
Novel Techniques to Improve Pathology Detection and Resection within the Lower Gastrointestinal Tract

Scarlet Nazarian

MBBS BSc MRCS

Supervisors:

Dr Nisha Patel

Mr Christopher Peters

Professor Daniel Elson

Professor Ara Darzi

Submitted to Imperial College London in partial fulfilment of the requirement for the degree of Doctor of Philosophy. Except for where indicated, it presents entirely my own work and describes the results of my own research.

2025

Declaration

The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution Non-Commercial 4.0 International Licence (CC BY-NC). Under this licence, you may copy and redistribute the material in any medium or format. You may also create and distribute modified versions of the work. This is on the condition that: you credit the author and do not use it, or any derivative works, for a commercial purpose. When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text. Where a work has been adapted, you should indicate that the work has been changed and describe those changes. Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law.

Acknowledgements

I am greatly indebted to all my supervisors for the guidance and unwavering support that they have provided me throughout my PhD and whose invaluable mentorship has been instrumental in the completion of this thesis.

Firstly, I would like to thank Nisha for being an incredible supervisor and for coming to the rescue when my initial PhD project fell through. Her extensive experience and dedication to research has been a constant source of inspiration for me. She has given me invaluable advice both for my project and on the importance of a work-life balance. Her support has not only allowed me to develop my own research ideas but has also facilitated meaningful collaborations that have helped advance my projects. Without her, I would truly not have been able to produce the work that I have nor be as proud as I am of the work that I have achieved.

Next, I'd like to thank Mr Peters and Professor Elson, my co-supervisors without whom this PhD would not have been possible. Their immense knowledge and academic excellence in the fields of upper GI surgery and medical imaging has enriched the work I have been able to produce. They have taught me so much and I will be eternally grateful to them for the wisdom they have imparted on me. I'd also like to thank Professor Darzi for giving me the opportunity and support to undertake this prestigious PhD at Imperial College. I am grateful for his guidance and impact he has had on my academic journey.

The work in this thesis is definitely made with love, hard work and lots of laughs along the way. And I couldn't have achieved any of it without my research partner and friend, Ioannis. I feel privileged to have worked with such a kind, supportive and dependable person. He has taught me so much and has been very patient, especially whilst I have navigated my way through the engineering aspects of this work. I am also grateful to Michal and Maxime, the BSc and PhD medical students who I have had the pleasure of working with. Their collaboration has made the process both productive and enjoyable.

I would like to extend my thanks to the teams at Hammersmith and Charing Cross Hospitals. The projects I have undertaken would not have run so smoothly without the support of the wards, theatres, histopathology department (especially the wonderful Martha), and the individuals who were kind enough to offer their time, expertise, and patience throughout the research process. Their willingness to facilitate data collection, provide technical assistance, and ensure seamless

coordination between departments has been invaluable, and I am truly grateful for their contributions.

Finally, I would like to thank my beautiful family for their unwavering support. A special thank you to my husband, Tom, who has had to endure the highs and lows of this rollercoaster and has done so with patience, constant encouragement and endless understanding.

I would like to dedicate this thesis to my two beautiful children – Lily and Hugo.

Abstract

Colorectal cancer (CRC) remains a major global health burden. Despite established screening tools such as colonoscopy and advanced imaging, many pre-cancerous or cancerous lesions are still missed or inadequately removed. Recent advances in machine learning (ML) have shown promise in improving lesion detection and characterisation in the lower gastrointestinal (GI) tract, with the aim of preventing post-colonoscopy CRC and reducing residual disease. This thesis explores novel screening tools and intra-operative techniques to enhance pathology detection and resection accuracy in the lower GI tract.

The effectiveness of colonoscopy is limited by operator dependency, leading to missed polyps. As the adenoma detection rate (ADR) inversely correlates with CRC risk, strategies to enhance colonoscopy quality were evaluated. The integration of artificial intelligence (AI) into endoscopy offers improved real-time diagnostic accuracy and procedural outcomes.

A systematic review and meta-analysis demonstrated high diagnostic accuracy of ML tools for colorectal polyp detection and characterisation, with potential to improve ADR and reduce CRC incidence. A second meta-analysis confirmed ML's ability to identify commonly missed lesions, including diminutive and non-pedunculated adenomas. A cross-sectional study of endoscopists revealed overall positive attitudes toward AI adoption, alongside key perceived challenges to clinical integration.

In surgical management, achieving clear margins is critical to reducing recurrence and improving survival. To address limitations of current intra-operative assessment tools, an ex-vivo study using diffuse reflectance spectroscopy (DRS) with ML classifiers successfully differentiated normal and tumour colonic tissue. This was followed by an in-vivo feasibility study confirming the DRS system's diagnostic accuracy and ergonomic suitability for real-time intra-operative use.

This thesis highlights the potential of AI and optical technologies to transform CRC detection and resection, offering promising avenues to enhance precision and improve patient outcomes.

Publications

Publications Directly Related to this Thesis

1. **Real-time classification of tumour and non-tumour tissue in colorectal cancer using diffuse reflectance spectroscopy and neural networks to aid margin assessment.**
Scarlet Nazarian, Ioannis Gkouzionis, Jamie Murphy, Ara Darzi, Nisha Patel, Christopher J. Peters, Daniel S. Elson
International Journal of Surgery 2024; doi: 10.1097/JS9.0000000000001102
2. **Real-time Tracking and Classification of Tumor and Nontumor Tissue in Upper Gastrointestinal Cancers Using Diffuse Reflectance Spectroscopy for Resection Margin Assessment.**
Scarlet Nazarian, Ioannis Gkouzionis, Michal Kawka, Marta Jamrozak, Josephine Lloyd, Ara Darzi, Nisha Patel, Daniel S. Elson, Christopher J. Peters
JAMA Surg. 2022; doi: 10.1001/jamasurg.2022.3899
3. **A YOLOv5-based network for the detection of a diffuse reflectance spectroscopy probe to aid surgical guidance in gastrointestinal cancer surgery.**
Ioannis Gkouzionis, Yican Zhong, Scarlet Nazarian, Ara Darzi, Nisha Patel, Christopher J. Peters, Daniel S. Elson
Int J Comput Assist Radiol Surg. 2024; doi: 10.1007/s11548-023-02944-9.
4. **Using Diffuse Reflectance Spectroscopy to Classify Tumor Tissue in Upper Gastrointestinal Cancers—Reply.**
Scarlet Nazarian, Ioannis Gkouzionis, Christopher J. Peters
JAMA Surg. 2023; doi:10.1001/jamasurg.2022.8433
5. **The future of endoscopy – what are the thoughts on artificial intelligence?**
Scarlet Nazarian, Hui Fen Koo, Emma Carrington, Ara Darzi, Nisha Patel
Journal of Experimental and Theoretical Artificial Intelligence 2023; doi: 10.1080/0952813X.2023.2178516
6. **Does computer-aided diagnostic endoscopy improve detection of commonly missed polyps? A meta-analysis.**
Scarlet Nazarian, Arun Sivananthan, Lakshmana Ayaru, Kinesh Patel, Hutan Ashrafian, Ara Darzi, Nisha Patel
Clinical Endoscopy 2022; doi: 10.5946/ce.2021.228
7. **Real-time tracking of a diffuse reflectance spectroscopy probe used to aid histological validation of margin assessment in upper gastrointestinal cancer resection surgery.**
Ioannis Gkouzionis, Scarlet Nazarian, Michal Kawka, Ara Darzi, Nisha Patel, Christopher J. Peters, Daniel S. Elson
Journal of Biomedical Optics 2022; doi: 10.1117/1.JBO.27.2.025001
8. **The diagnostic accuracy of artificial intelligence and computer-aided diagnosis for the detection and characterisation of colorectal polyps: A systematic review and meta-analysis.**
Scarlet Nazarian, Ben Glover, Hutan Ashrafian, Ara Darzi, Julian Teare
Journal of Medical Internet Research 2021; doi: 10.2196/27370
9. **Clinical validation benchmark dataset and expert performance baseline for colorectal polyp localization methods.**
Luisa F. Sánchez-Peralta, Ben Glover, Cristina L. Saratxaga, Juan Francisco Ortega-Morán, Scarlet Nazarian, Artzai Picón, J. Blas Pagador, Francisco M. Sánchez-Margallo
Journal of Imaging 2023; doi: 10.3390/jimaging9090167

Other Affiliated Publications

1. **Robotic versus laparoscopic cholecystectomy for difficult gallbladders: an observational study of tertiary centre cases.**
Michal Kawka, Zaynab Jawad, David Hakim, Madhava Pai, **Scarlet Nazarian**, Tamara Gall, Christopher Wadsworth, David Nicol, Long R. Jiao
Surgical Endoscopy 2025; doi: 10.1007/s00464-025-11586-8
2. **Development and validation of machine learning models to predict need for haemostatic therapy in acute upper gastrointestinal bleeding.**
Scarlet Nazarian, Frank Po Wen Lo, Jianing Qiu, Nisha Patel, Benny Lo, Lakshmana Ayaru
Therapeutic Advances in Gastrointestinal Endoscopy 2024; doi:10.1177/26317745241246899
3. **The influence of procedural volume on short-term outcomes for robotic pancreatoduodenectomy-a cohort study and a learning curve analysis.**
Michal Kawka, Tamara Gall, Fiona Hand, **Scarlet Nazarian**, David Cunningham, David Nicol, Long R. Jiao
Surgical Endoscopy 2023; doi: 10.1007/s00464-023-09941-8
4. **Patient Perspectives on Surgeon-Specific Outcome Reports in Bariatric Surgery**
Kyle Lam, **Scarlet Nazarian**, Nishita Gadi, Sherif Hakky, Krishna Moorthy, Christos Tsironis, Ahmed Ahmed, James M Kinross, Sanjay Purkayastha
Surgery for Obesity and Related Diseases 2022; doi: 10.1016/j.soard.2022.02.020
5. **A case for improved assessment of gut permeability: a meta-analysis quantifying the lactulose:mannitol ratio in coeliac and Crohn's disease.**
Jonathan Gan, **Scarlet Nazarian**, Julian Teare, Ara Darzi, Hutan Ashrafian, Alex J. Thompson
BMC Gastroenterology 2021; doi: 10.1186/s12876-021-02082-z
6. **The diagnostic accuracy of smartwatches for the detection of cardiac arrhythmia: A systematic review and meta-analysis.**
Scarlet Nazarian, Kyle Lam, Ara Darzi, Hutan Ashrafian
Journal of Medical Internet Research 2021; doi: 10.2196/28974

Abstract Publications

1. **Using Diffuse Reflectance Spectroscopy (DRS) to Identify Tumour and Non-tumour Tissue in Upper Gastrointestinal Specimens.**
Nazarian S, Gkouzionis I, Anandakumar A, Patel N, Elson D, Peters C. BJS Volume 108, Issue Supplement_7, October 2021, <https://doi.org/10.1093/bjs/znab361.142>. [Abstract from ASGBI 2021]
2. **Using Diffuse Reflectance Spectroscopy Probe Tracking to Identify Non-tumour and Tumour Tissue in Upper Gastrointestinal Specimens**
Gkouzionis I, **Nazarian S**, Anandakumar A, Darzi A, Patel N, Peters C, Elson D
European Conferences on Biomedical Optics 2021 (ECBO), OSA Technical Digest (Optica Publishing Group, 2021), paper ETu1A.1.
3. **Real-time tracking and classification of tumour and non-tumour tissue in upper gastrointestinal cancer specimens using diffuse reflectance spectroscopy**
Nazarian S, Gkouzionis I, Kawka M, Patel N, Darzi A, Elson D, Peters C. BJS Volume 108, Issue Supplement_9, 15 December 2021, <https://doi.org/10.1093/bjs/znab429.035>. [Abstract from UGI 2021]
4. **Using diffuse reflectance spectroscopy probe tracking to identify non-tumour and tumour tissue in upper gastrointestinal specimens**

Presentations

Oral Presentations

1. **In-vivo real-time assessment of resection margins using diffuse reflectance spectroscopy in upper gastrointestinal cancers**
Giot M, **Nazarian S**, Gkouzionis I, Ranjitkar R, Lloyd J, Elson DS, Peters CJ
Digestive Disease Week, San Diego 2025
2. **Diffuse Reflectance Spectroscopy for Real Time Differentiation of Tissue and Lymph Node Status in Upper Gastrointestinal Cancer Surgery.**
Giot M, Slim N, **Nazarian S**, Gkouzionis I, Ranjitkar R, Lloyd J, Elson DS, Peters CJ
UKIOG Annual Meeting, London, March 2025
3. **In vivo Differentiation of Tissue in Upper Gastrointestinal Cancers Using Diffuse Reflectance Spectroscopy for Resection Margin Assessment.**
Giot M, **Nazarian S**, Gkouzionis I, Ranjitkar R, Slim N, Lloyd J, Elson DS, Peters CJ
CRUK Clinical Researchers' Meeting, London, Feb 2025
4. **In-Vivo Classification of Stomach and Oesophageal Tissues During Upper Gastrointestinal Cancer Surgery Using Diffuse Reflectance Spectroscopy**
Gkouzionis I, **Nazarian S**, Giot M, Peters C, Elson D
SPIE Photonics Europe, Strasbourg, April 2024
5. **Real-time classification of tumour and non-tumour colorectal tissue using diffuse reflectance spectroscopy to aid resection margin assessment**
Nazarian S, Gkouzionis I, Patel N, Darzi A, Elson D, Peters C
Digestive Disease Week, Chicago 2023
6. **A YOLOv5 based network for the detection of a diffuse reflectance spectroscopy probe to aid surgical guidance in gastrointestinal cancer surgery**
Yican Zhong, Gkouzionis I, **Nazarian S**, Darzi A, Patel N, Peters C, Elson D
The 14th International Conference on Information Processing in Computer-Assisted Interventions, Munich, June 2023
7. **Diffuse reflectance spectroscopy for in-vivo stomach and oesophageal tissue classification during upper gastrointestinal cancer surgery**
Gkouzionis I, **Nazarian S**, Darzi A, Patel N, Peters C, Elson D
SPIE Photonics West, San Francisco, Jan 2023
8. **The application of deep learning when discriminating cancer and healthy tissue in upper gastrointestinal cancer surgery using diffuse reflectance spectroscopy**
Gkouzionis I, **Nazarian S**, Darzi A, Patel N, Elson D, Peters C
London Surgical Symposium, London, Nov 2022
9. **Using diffuse reflectance spectroscopy for real-time tissue assessment during upper gastrointestinal cancer surgery**
Gkouzionis I, **Nazarian S**, Darzi A, Patel N, Peters C, Elson D
IEEE International Conference on Biomedical & Health Informatics, Greece, Sept 2022

10. **Real-time tissue classification in stomach and oesophageal cancer based on diffuse reflectance spectroscopy**
Gkouzionis I, **Nazarian S**, Darzi A, Patel N, Peters C, Elson D
SPIE Photonics Europe, France, April 2022
11. **Real-time tissue classification in stomach and oesophageal cancer based on optical tracking of a diffuse reflectance spectroscopy probe**
Gkouzionis I, **Nazarian S**, Darzi A, Patel N, Peters C, Elson D
SPIE Photonics West, USA, January 2022
12. **Real-time Discrimination of Cancer Tissue from Healthy Tissue in Upper Gastrointestinal Cancers using a Diffuse Reflectance Spectroscopy Probe**
Nazarian S, Gkouzionis I, Patel N, Darzi A, Elson D, Peters C
Department of Surgery & Cancer Research Afternoon, Imperial College London, December 2021
13. **The use of machine learning for real-time detection of oesophageal and gastric cancer based on diffuse reflectance spectroscopy: a validation study.**
Nazarian S, Gkouzionis I, Kawka M, Elson D, Peters C
ESDE-IGCA international congress, Milan, November 2021
14. **Real-time tracking and classification of tumour and non-tumour tissue in upper gastrointestinal cancer specimens using diffuse reflectance spectroscopy**
Nazarian S, Gkouzionis I, Kawka M, Patel N, Darzi A, Elson D, Peters C
AUGIS Conference, Belfast, October 2021
15. **The Use of a Diffuse Reflectance Spectroscopy Probe and Tracking System to Classify Tumour and Non-tumour Tissue in Upper Gastrointestinal Cancer Specimens to Aid Margin Assessment**
Nazarian S, Gkouzionis I, Kawka M, Patel N, Darzi A, Elson D, Peters C
London Surgical Symposium, London, Sept 2021
16. **Using Diffuse Reflectance Spectroscopy Probe Tracking to Identify Non-tumour and Tumour Tissue in Upper Gastrointestinal Specimens**
Gkouzionis I, **Nazarian S**, Anandakumar A, Darzi A, Patel N, Peters C, Elson D
OSA European Biophotonics Conference, Virtual, June 2021
17. **Using Diffuse Reflectance Spectroscopy (DRS) to Identify Tumour and Non-tumour Tissue in Upper Gastrointestinal Specimens**
Nazarian S, Gkouzionis I, Anandakumar A, Patel N, Elson D, Peters C
ASGBI Conference, Virtual, May 2021
18. **Real-time optical tracking of a diffuse reflectance spectroscopy probe for gastrointestinal tissue analysis**
Gkouzionis I, **Nazarian S**, Avila-Rencoret F, Murphy J, Peters C, Elson D
SPIE Photonics West, Virtual, January 2021

Poster Presentations

1. **Differentiating tumour and normal tissue using diffuse reflectance spectroscopy *in vivo* in upper gastrointestinal cancer surgery**
Giot M, **Nazarian S**, Gkouzionis I, Darzi A, Patel N, Elson D, Peters C
London Surgical Symposium, London, March 2024
2. **Diffuse reflectance spectroscopy for tissue discrimination during *in vivo* upper gastrointestinal cancer surgery**
Gkouzionis I, **Nazarian S**, Darzi A, Patel N, Peters C, Elson D

European Conferences on Biomedical Optics, Munich, June 2023

3. **Real-time classification of colorectal tissue using diffuse reflectance spectroscopy to aid resection margin assessment**
Nazarian S, Gkouzionis I, Patel N, Darzi A, Elson D, Peters C
ESCP 17th Scientific & Annual Meeting, Dublin, Sept 2022
4. **Real-time tracking of a diffuse reflectance spectroscopy probe for tissue classification in colorectal cancer surgery**
Gkouzionis I, Nazarian S, Darzi A, Patel N, Peters C, Elson D
Hamlyn Symposium on Medical Robotics, London, Jun 2022
5. **The diagnostic accuracy of artificial intelligence and computer-aided diagnosis for the detection and characterisation of colorectal polyps: A systematic review and meta-analysis.**
Nazarian S, Glover B, Ashrafian H, Darzi A, Teare J
United European Gastroenterology (UEG), Virtual, Oct 2021
6. **Does computer aided diagnostic endoscopy improve detection of commonly missed polyps?: a meta-analysis.**
Nazarian S, Sivananthan A, Ayaru L, Patel K, Ashrafian H, Darzi A, Patel N
Digestive Disease Week (DDW), Virtual, May 2021

Table of Contents

Declaration.....	2
Acknowledgements	3
Abstract	5
Publications.....	6
Publications Directly Related to this Thesis	6
Other Affiliated Publications.....	7
Abstract Publications.....	7
Presentations	8
Oral Presentations.....	8
Poster Presentations	9
List of Figures	16
List of Tables	20
List of Acronyms	21
Chapter 1 Introduction	22
1.1 Colorectal Cancer	22
1.1.1 Post-Colonoscopy Colorectal Cancer	22
1.2 Development of Colorectal Cancer	23
1.3 Detection of Colorectal Cancer	24
1.3.1 Adenoma Detection Rate	25
1.3.2 Factors affecting ADR	25
1.3.2.1 Patient factors	26
1.3.2.2 Endoscopist factors	26
1.3.2.3 Endoscopy factors	27
1.3.3 Artificial Intelligence and Machine Learning	28
1.3.4 Machine Learning and Endoscopy	30
1.3.4.1 Ex vivo studies	30
1.3.4.2 In vivo studies.....	31
1.3.4.3 Commercial AI Devices	32
1.3.5 Barriers of AI Technology in Endoscopy	32
1.4 Management of Colorectal Lesions	33
1.4.1 Endoscopy Management Workflow for Colorectal Lesions	34
1.4.2 Endoscopic Resection Techniques.....	35
1.4.2.1 Biopsy Forceps Polypectomy.....	36
1.4.2.2 Snare Polypectomy.....	36
1.4.2.3 Endoscopic Mucosal Resection	38
1.4.3.4 Endoscopic Submucosal Dissection.....	39
1.4.3.5 Complex Lesions.....	41
1.4.3 Surgical Resection	41
1.4.3.1 Resection Margin Assessment	43
1.4.3.1.1 Frozen Section Analysis	43
1.4.3.2 Optical Techniques for Margin Assessment.....	44

1.5	Diffuse Reflectance Spectroscopy	45
1.5.1	Limitations of DRS.....	47
1.6	Commonalities between Upper and Lower Gastrointestinal Tract Cancers	48
1.7	Aims of this Thesis and Overarching Goal	50
Chapter 2 Lesion detection and characterisation with the use of computer-aided devices in endoscopy: systematic review and meta-analyses		51
2.1	The Diagnostic Accuracy of Artificial Intelligence and Computer-Aided Diagnosis for the Detection and Characterisation of Colorectal Polyps: A Systematic Review and Meta-Analysis. 51	
2.1.1	Background and Aims	51
2.1.1.1	Current Methods to Improve ADR	51
2.1.1.1.1	Visualisation of Mucosa Behind Folds	52
2.1.1.1.2	Extension of the Viewing Field	53
2.1.1.2	Training Initiatives to Improve ADR	53
2.1.1.3	AI Methods to Improve Detection and Characterisation of Polyps	54
2.1.1.4	Aims of the Study	56
2.1.2	Methods.....	56
2.1.2.1	Search Strategy.....	56
2.1.2.2	Inclusion and Exclusion Criteria.....	57
2.1.2.3	Study Selection	57
2.1.2.4	Data Extraction	57
2.1.2.5	Study Quality Assessment	58
2.1.2.6	Statistical Analysis	58
2.1.3	Results.....	58
2.1.3.1	Search Results and Characteristics.....	58
2.1.3.2	Detection or Localisation of a Polyp.....	64
2.1.3.3	Characterisation of a Detected Polyp	64
2.1.3.4	PDR and ADR for Polyp Detection (RCTs).....	66
2.1.3.5	Heterogeneity of Studies.....	67
2.1.3.6	Quality Assessment	68
2.1.4	Discussion	68
2.1.4.1	Limitations of the Study	69
2.1.4.2	Withdrawal times	69
2.1.4.3	Lack of Reporting Standards.....	70
2.1.4.4	Challenges of AI Integration into Healthcare	70
2.1.5	Conclusion.....	72
2.2	Does Computer-Aided Diagnostic Endoscopy Improve Detection of Commonly Missed Polyps? A Meta-Analysis.....	73
2.2.1	Background and Aims	73
2.2.1.1	Aims of this Study.....	74
2.2.2	Method	74
2.2.2.1	Search Strategy.....	74
2.2.2.2	Primary and Secondary Outcomes	74
2.2.2.3	Inclusion and Exclusion Criteria.....	74
2.2.2.4	Study Selection	75
2.2.2.5	Data Extraction	75
2.2.2.6	Study Quality Assessment	76
2.2.2.7	Statistical Analysis	76
2.2.3	Results.....	76
2.2.3.1	Eligible Studies.....	76
2.2.3.2	Study Characteristics	77
2.2.3.3	Commonly Missed Lesions	78
2.2.3.3.1	Diminutive Lesions	78
2.2.3.3.2	Non-Pedunculated Lesions.....	79

2.2.3.4	ADR and PDR	81
2.2.3.5	Heterogeneity of Studies.....	81
2.2.3.6	Quality Assessment	81
2.2.4	Discussion	81
2.2.4.1	Withdrawal Times	82
2.2.4.2	PCCRC Rates	82
2.2.4.3	Limitations of Study.....	82
2.2.4.4	Diminutive Lesions	83
2.2.5	Conclusion.....	83
Chapter 3 Views of endoscopists on the adoption of artificial intelligence: a cross-sectional questionnaire study		85
3.1	Background and Aims.....	85
3.1.1	Aims of the Study.....	87
3.2	Methods	87
3.2.1	Study Design	87
3.2.2	Questionnaire Distribution	88
3.2.3	Assessment of Competency.....	88
3.2.4	Statistical Analysis.....	88
3.2.5	Ethical Considerations	88
3.3	Results.....	89
3.3.1	Population.....	89
3.3.2	Effects of AI on Endoscopy	91
3.3.3	Effect of AI on ADR.....	92
3.3.4	Sentiments of Endoscopists based on Experience	93
3.3.5	Benefits and Barriers of AI in Endoscopy.....	94
3.3.6	Application of AI in Endoscopy	95
3.4	Discussion	96
3.5	Conclusion	99
Chapter 4 Improving detection of tumour tissue using diffuse reflectance spectroscopy to aid resection margin assessment in lower GI cancer surgery: an ex vivo validation study		100
4.1	Background and Aims.....	100
4.1.1	Intra-Operative Challenges of Margin Assessment	101
4.1.2	Intra-Operative Margin Mapping Techniques	102
4.1.2.1	Advanced Imaging and Spectroscopic Techniques	102
4.1.2.2	Optical Spectroscopy Techniques	104
4.1.2.3	Use of DRS for Tissue Discrimination	104
4.1.2.3.1	DRS Probe Tracking	106
4.1.3	Upper GI Tissue Discrimination	107
4.1.4	Aims of the Study.....	108
4.1.4.1	DRS Probe Tracking	108
4.1.4.2	Upper GI <i>ex vivo</i> Tissue Discrimination	108
4.1.4.3	Lower GI <i>ex vivo</i> Tissue Discrimination	109
4.2	Methods	109
4.2.1	Optical Marker-Based Tracking of DRS Probe	109
4.2.1.1	Marker Design, Detection and Segmentation.....	109
4.2.1.2	Calculation of DRS Tip Position & Probe Tracking.....	112
4.2.2	Graphical User Interface	113
4.2.3	DRS Instrumentation and Set-Up	114
4.2.4	Study Design	115
4.2.4.1	Upper GI Cancer	115

4.2.4.2	Lower GI Cancer	115
4.2.5	Data Collection & Processing	116
4.2.6	Histopathology Correlation	117
4.2.6.1	Challenges in Clinical Validation	117
4.2.6.2	Histopathological Analysis Protocol	118
4.2.7	Classification	120
4.2.7.1	Upper GI Tissue Classification	120
		122
4.2.7.2	Lower GI Tissue Classification	122
4.2.8	Statistical Analysis	124
4.3	Results	125
4.3.1	DRS Probe Tracking Evaluation	125
4.3.2	Patient Characteristics	127
4.3.2.1	Upper GI Cohort	127
4.3.2.2	Lower GI Cohort	129
4.3.3	Spectral Data Analysis	130
4.3.4	Classification Analysis	131
4.3.4.1	Upper GI Classification Results	131
4.3.4.1.1	Classification of Tumour Data for Pathological Staging	133
4.3.4.2	Lower GI Classification Results	135
4.4	Discussion	137
4.4.1	Study Findings	137
4.4.2	Strengths of Study	138
4.4.2.1	Data Acquisition Speed	138
4.4.2.2	Sample Size	138
4.4.2.3	Histopathology Correlation	139
4.4.2.4	Neural Network Classifier	139
4.4.2.5	Cost Benefit	140
4.4.3	Limitations	140
4.4.4	Future <i>In-Vivo</i> Use	142
4.5	Conclusion	143
Chapter 5 Classification of tumour and non-tumour tissue using diffuse reflectance spectroscopy in the gastrointestinal tract to aid resection margin assessment: an in vivo feasibility study		144
5.1	Background and Aims	144
5.1.1	Sterilisation and Infection Control	145
5.1.2	Safety	145
5.1.3	Durability	146
5.1.4	Ergonomics	146
5.1.5	Regulatory Compliance	146
5.1.6	Compatibility with Surgical Workflow	147
5.1.7	Aims of the Study	147
5.2	Methods	147
5.2.1	Patient and Public Involvement	148
5.2.2	Phantom Set-Up of DRS System	148
5.2.3	Study Design	149
5.2.4	DRS Set-Up	150
5.2.5	Graphical User Interface	152
5.2.6	Data Acquisition	153
5.2.7	Histopathology Correlation	154
5.2.8	Marker-less Probe Tracking	156
5.2.9	Spectral Classification	157
5.2.10	Usability Questionnaire Methodology	158

5.2.11	Statistical Analysis.....	159
5.3	Results	159
5.3.1	Patient Characteristics	159
5.3.2	Spectral Analysis	160
5.3.3	Marker-less Tracking Analysis.....	162
5.3.4	Usability Questionnaire Analysis	163
5.3.4.1	Respondent Characteristics.....	163
5.3.4.2	Ease of Use	164
5.3.4.3	Effectiveness and Ease of Learning	164
5.3.4.4	Satisfaction	165
5.3.4.5	Positive and Negative Impact.....	165
5.4	Discussion	166
5.4.1	Strengths of Study	166
5.4.2	Limitations of Study	167
5.4.3	Further Work	168
5.5	Conclusion	169
Chapter 6	Conclusions & Future Directions.....	171
6.1	Conclusions.....	171
6.2	Future Directions.....	172
6.2.1	Tumour Differentiation in Cancer Surgery	172
6.2.2	Virtual Reality & Autofluorescence Spectroscopy	174
6.2.3	Minimally Invasive Use of DRS.....	175
6.2.3.1	Laparoscopic Use of DRS	175
6.2.3.2	Endoscopic Use of DRS.....	175
6.2.4	Unmet Clinical Need & Commercial Opportunity	177
References.....		180
Appendices		202
Appendix 1	Search Strategy.....	202
Appendix 2	Quality Assessment	203
Appendix 3	Limitations of Study.....	205
Appendix 4	Database Searches	206
Appendix 5	Endoscopy Questionnaire	207
Appendix 6	Usability Questionnaire	210

List of Figures

Figure 1 Adenoma-carcinoma sequence. Loss of tumour-suppressor gene, APC, results in hyperproliferative epithelium. Further mutations of KRAS and DCC (deleted in colorectal cancer) lead to adenoma formation and eventually, together with loss of function of the p53 gene, lead to carcinoma.[17]	23
Figure 2 Deep learning is a subset of machine learning which is in turn a branch of AI	30
Figure 3 Endoscopy Management Workflow for Colorectal Lesions	34
Figure 4 Snare polypectomy. [Image via https://colorectalsurgeonsydney.com.au/conditions/polyps]	37
Figure 5 Endoscopic mucosal resection (EMR). (1) Injection of fluid to lift polyp. (2) Snared removal of polyp. (3) Retrieval with forceps. [Image via GastroMedicine & Endoscopy]	38
Figure 6 Endoscopic submucosal dissection (ESD). (1) Marking tissue margin. (2) Injection of fluid into submucosal space. (3) Submucosal dissection and resection of lesion. (4) Haemostasis. [Image via https://www.drpiyushranjan.com]	40
Figure 7 The three areas of interest when assessing intra-operative margins using real-time imaging techniques.[159]	44
Figure 8 Illustration of the basic principles of DRS with absorption and scattering of light within a sample of tissue.	46
Figure 9 Diffuse reflectance probe containing six peripheral illumination fibres around one light collection fibre.	46
Figure 10 DRS spectrum created through scattering and absorption of light measured over a range of wavelengths.	47
Figure 11 Three common colorectal lesion detection models.	55
Figure 12 PRISMA flow diagram for study selection. © Reproduced from [194] with permission.	59
Figure 13 Pooled analysis of specificity of polyp detection by the use of lesions or polyps within video frames or images. Effect sizes (ES) are shown with 95 per cent confidence intervals. A random-effects model was used. © Reproduced from [194] with permission.	65
Figure 14 Pooled analysis of sensitivity of polyp characterisation by the use of images. Effect sizes (ES) are shown with 95 per cent confidence intervals. A random-effects model was used. © Reproduced from [194] with permission.	65
Figure 15 Pooled analysis of polyp detection rate (PDR). Odds ratios (OR) are shown with 95 percent confidence intervals. A random-effects model was used for the meta-analysis. © Reproduced from [194] with permission.	66
Figure 16 Pooled analysis of adenoma detection rate (ADR). Odds ratios (OR) are shown with 95 per cent confidence intervals. A random-effects model was used for the meta-analysis. © Reproduced from [194] with permission.	67
Figure 17 PRISMA flow diagram for study selection in meta-analysis. © Reproduced from [257] with permission.	77
Figure 18 Pooled analysis of mean diminutive adenoma per patient in CADe vs. conventional colonoscopy groups. Effect sizes (ES) are shown with 95 per cent confidence intervals. A random-effects model was used. © Reproduced from [257] with permission.	80
Figure 19 Pooled analysis of mean non-pedunculated adenoma per patient in CADe vs. conventional colonoscopy groups. Effect sizes (ES) are shown with 95 per cent confidence intervals. A random-effects model was used. © Reproduced from [257] with permission.	80
Figure 20 Factors affecting the uptake of AI.	86
Figure 21 Summary of responses. © Reproduced from [96] with permission.	92
Figure 22 Key benefits of adoption of AI in endoscopy. © Reproduced from [96] with permission.	95
Figure 23 Key barriers of adoption of AI in endoscopy. © Reproduced from [96] with permission.	95
Figure 24 Best utilisation of AI within endoscopy. © Reproduced from [96] with permission.	96

Figure 25 Priority application of AI within endoscopy. © Reproduced from [96] with permission.	96
Figure 26 REIMS set-up ex-vivo and in operating theatres (in-vivo). Electrical current is applied to the tissue and the resultant charged particles are extracted through the custom-designed hand piece and drawn into the mass spectrometer to produce tissue-specific mass spectra. Image adapted from [374].....	103
Figure 27 Diagram to represent OCT. (A) Handheld single fibre (sf) OCT probe (upper inset: reference light derives from the probe tip; lower inset: a photo of the probe held by hand); (B) sfOCT probe is manually scanned across a region of interest to generate a 2D image. Image adapted from [375]..	103
Figure 28 Photo (A) and schematic diagram (B) of DRS set-up showing laptop connected to spectrometer (Ocean Optics, USB4000, Largo, USA) and light source (Ocean Optics, HL-2000-HP, Largo, USA), both of which connect to DRS probe.	105
Figure 29 Examples of surgical instruments with colour markers attached to aid in surgical tool detection and tracking.[316,317]	107
Figure 30 (a) Colour distribution of the biological tissue from laparoscopic images, with cyan colour circled as a potential colour for the marker. (b) DRS fibre probe with the green colour marker attached. © Reproduced from [330] with permission.	110
Figure 31 DRS probe detection and tracking flow diagram. © Reproduced from [330] with permission.	111
Figure 32 DRS probe detection workflow. (a) Example input video frame. (b) Blurred frame. (c) Hue channel of the blurred input frame. (d) Binary mask result of HSV segmentation. (e) Detected edge lines of the probe. (f) Detected tip point of the probe. © Reproduced from [330] with permission.	112
Figure 33 Python software for hardware system control, data acquisition and real-time classification. (Top) Graphical User Interface (GUI) highlighting function of each individual section. (Bottom) GUI in use during data collection. (a) Control panel, allowing for execution of data acquisition protocol, probe tracking, hardware calibration and light source control. (b) Augmented live video feed from the RGB camera. Real-time classification is displayed on the screen via text and a coloured overlay on the tissue, with a graduated colour scale indicating the probability of belonging to a particular class (Normal = 100% green – 100% pink = Tumour). (c) Reflectance data in the 420 – 1000 nm spectral range displayed in real-time. © Reproduced from [330] with permission.	114
Figure 34 DRS instrumentation for ex vivo data acquisition: (a) schematic diagram of set up; (b) photograph of set up. The DRS probe was connected to both the light source (Ocean Optics, HL-2000-HP, Largo, USA) and the spectrometer (Ocean Optics, USB4000, Largo, USA) to allow for data acquisition and sample illumination. All electronic devices communicated with proprietary software designed with Python version 3.6 (Python Software Foundation, Wilmington, USA) on the laptop.© Reproduced from [332] with permission.	115
Figure 35 Workflow for DRS data acquisition and processing. [GUI: graphical user interface] © Reproduced from [330] with permission.	117
Figure 36 Workflow for histopathological analysis of tissue for ground truth labelling.	118
Figure 37 Histopathological correlation workflow for processing of an oesophageal specimen. The correlation was performed using yellow tissue paint (marked with a yellow arrow across panels (a) – (e). (a) The ex vivo specimen with tracked optical biopsy sites of the suspected ‘tumour’ area. (b) Ex vivo specimen following painting of the suspected ‘tumour’ tissue with yellow paint. (c) Specimen, after having been placed in formalin, with yellow painted area visible. (d) Macroscopic slice of oesophageal tissue with the yellow painted area. (e) Microscopic H&E-stained slice with yellow paint (yellow arrow) and confirmed oesophageal adenocarcinoma (red outline). © Reproduced from [332] with permission.	119
Figure 38 Histopathological correlation workflow for processing of a colorectal specimen. (a) Photo of scanned colon tissue with tracked DRS locations over suspected tumour area indicated. (b) Photo of colon tissue with painted yellow dye over the suspected tumour tissue. (c) Photo of colon tissue after being in formalin with yellow tissue dye visible. (d) Photo of sliced colon tissue with yellow dye	

visible over the suspected tumour. (e) Scanned H&E slide showing yellow tissue dye over confirmed tumour tissue (marked in red). © Reproduced from [334] with permission.	120
Figure 39 Spectral data analysis workflow. © Reproduced from [330] with permission.	122
Figure 40 – The architecture of the proposed one-dimensional convolutional neural network (CNN) for colorectal tissue classification. The SpecNet is a five-layer CNN for learning visual representations in DRS data. It uses 20 kernels to filter the input data, and the output layer classifies spectra based on the spectral band size and number of output classes in the dataset. © Reproduced from [334] with permission.	124
Figure 41 (a)-(b) Illustration of probe tracking error with overlaid ground truth (blue line) and estimated (yellow line) probe tip position for two video sequences (600 and 450 frames each). (c)-(f) The ground truth and estimated projected DRS probe tip pixel position for x-axis and y-axis for each of the video sequences. Blue and red lines indicate the estimated and ground truth respectively. © Reproduced from [330] with permission.	126
Figure 42 Average (lines) and standard deviations (shaded areas) of the mean spectra for (a) normal and tumour stomach tissue and (b) normal and tumour oesophagus tissue.	131
Figure 43 Average (lines) and standard deviations (shaded areas) of the mean spectra for normal and tumour colorectal tissue.	131
Figure 44 ROC curves with AUC values for all classifiers for (a)stomach and (b)oesophagus datasets.	133
Figure 45 Three illustrations of real-time DRS probe tracking and classification at different frames during scanning of an oesophageal and gastric ex vivo tissue specimen. (a-c) Real-time tracking at each optical biopsy site coupled with the binary classification of each site on an oesophageal and gastric specimen. Tissue type highlighted on the screen in real-time (bottom left of each frame). © Reproduced from [330] with permission.	133
Figure 46 Distribution of spectra acquired for (a) gastric and (b) oesophageal cancer according to T staging.	134
Figure 47 Mean spectra of the different pathological tumour stages for (a) gastric and (b) oesophageal cancer.	134
Figure 48 ROC curves for one-vs-rest multi-class classification for (a) gastric and (b) oesophageal cancer pathological staging.	135
Figure 49 ROC curve with AUC values for all classifiers for colorectal tissue data.	136
Figure 50 Real-time tracking at each optical biopsy site coupled with the binary classification probability of each site as either normal (100% green) or tumour (100% tumour) tissue on a colorectal specimen.© Reproduced from [334] with permission.	137
Figure 51 Phantom feasibility study to evaluate the integration of the DRS system in vivo. (a) Phantom model for laparoscopic set-up. (b) Phantom model for open surgical set-up.	149
Figure 52 Specifications of the sterilisable DRS probe for in vivo use during surgery.	151
Figure 53 DRS system set-up for in vivo use.	152
Figure 54 In vivo data acquisition. (a) Spectral data and optical biopsy site highlighted live on screen with GUI software. (b) Open abdomen with DRS sampling in process using sterile DRS probe.	154
Figure 55 Histopathological correlation technique for suspected tumour area. (a) Oesophageal specimen immediately after resection with marked area of sampled tissue for suspected tumour using sterile marker pen. (b) Marked area on oesophagus painted with permanent yellow tissue dye for histopathological analysis.	154
Figure 56 A gastric specimen prior to fixation in formalin and processing by the pathologist. (The greater omentum is present and attached to the greater curvature of the stomach). The yellow painted region corresponds to the suspected tumour area. Line (1) corresponds to the distance from proximal margin to proximal border of yellow painted region. Line (2) corresponds to the length of the painted area. Line (3) corresponds to the 3 cm line drawn.	155
Figure 57 Data analysis framework for GI tissue classification.	158

Figure 58 Average (lines) and standard deviations (shaded areas) of the mean spectra for normal and tumour stomach tissue (left) and normal and tumour oesophagus tissue (right).....	161
Figure 59 ROC curves with AUC values for all classifiers for (a) stomach dataset and (b) oesophagus dataset.	162
Figure 60 Illustration of real-time DRS marker-less probe tracking at each optical biopsy site during an oesophagectomy. (a) – (c) Real-time visual correlation of scanned optical biopsy sites on tissue (green dots on each frame) using the DRS probe and marker-less tracking.....	163
Figure 61 Bar chart representing frequency of DRS probe use by respondents.....	163
Figure 62 Pie charts showing satisfaction rates with the use of the DRS probe.....	165
Figure 63 Potential designs for the DRS gripper that will attach onto the DRS ferrule to allow laparoscopic use of the probe in vivo.....	169
Figure 64 Schematic diagram of endoscopic optical probe in use during gastroscopy, adapted from paper by Waterhouse et al. [373] (A) Light is transferred using the illumination fibre and reflected light is relayed back to the spectrometer. (B) The spectral endoscope is inserted via the working channel of the gastroscope. (C) Light is delivered to the tissue area of interest to aid tissue characterisation.	176

List of Tables

Table 1 Factors affecting ADR.....	25
Table 2 Commercial AI devices.....	32
Table 3 Narrow-Band Imaging International Colorectal Endoscopic classification system [105].....	35
Table 4 Current methods to improve ADR	53
Table 5 Characteristics of included studies on polyp detection. © Reproduced from [194] with permission.	60
Table 6 Characteristics of included studies on polyp characterisation. © Reproduced from [194] with permission.	61
Table 7 Characteristics of RCTs into polyp detection. © Reproduced from [194] with permission.	63
Table 8 Study characteristics of included papers in meta-analysis. © Reproduced from [257] with permission.	78
Table 9 Summary of CADe versus control MPP and MAP (diminutive and non-pedunculated), percentage difference and effect size. © Reproduced from [257] with permission.	78
Table 10 Summary of CADe versus control for polyp and adenoma detection. (N/K – not known) © Reproduced from [257] with permission.	81
Table 11 Participant characteristics. © Reproduced from [96] with permission.....	90
Table 12 Comparing role, experience, ADR of endoscopists with views of AI on ADR. © Reproduced from [96] with permission.	92
Table 13 Sentiments of endoscopists based on experience. © Reproduced from [96] with permission.	93
Table 14 Projection error for DRS probe's tip location.	127
Table 15 Rotation angle and maximum detectable distance from the camera to track the DRS probe.	127
Table 16 Upper GI patient characteristics.	128
Table 17 Lower GI patient characteristics.	129
Table 18 Dataset summary for spectral acquisition for gastric, oesophageal and colorectal tissue.	130
Table 19 Performance metrics for the spectral data classification using supervised ML and permutation feature importance for the feature selection. Data showing the ability of multiple classifiers to detect normal vs. cancerous tissue presented as mean ± standard deviation. Overall accuracy is calculated as the proportion of correctly identified spectra over the total number of spectra.	132
Table 20 Performance metrics for the spectral data classification using supervised ML and permutation feature importance for the feature selection. Data showing the ability of multiple classifiers to detect normal vs. cancerous tissue presented as mean (95% confidence interval). Overall accuracy is calculated as the proportion of correctly identified spectra over the total number of spectra.	136
Table 21 Upper GI in vivo patient characteristics.	160
Table 22 Dataset summary for spectral acquisition for gastric and oesophageal tissue.....	161
Table 23 Performance metrics for all classifiers for stomach and oesophagus datasets. Data showing the ability of multiple classifiers to detect normal versus cancerous tissue presented as mean ± standard deviation. Overall accuracy is calculated as the proportion of correctly identified spectra over the total number of spectra.	162
Table 24 Usability questionnaire results on ease of use of the DRS probe, ease of learning and future practice.	164
Table 25 Assessment of bias for each included study in seven domains according to the QUADAS-2 tool. 😊 = low risk of bias; 😞 = high risk of bias; ? = unclear risk of bias.	204
Table 26 Assessment of bias for RCTs using the Jadad scale; score ≥ 3 = good quality.....	204
Table 27 Limitations of current studies assessing polyp detection or characterisation.....	205

List of Acronyms

ADR	Adenoma detection rate	IBD	Inflammatory bowel disease
ANN	Artificial neural networks	IP	Intellectual property
APC	Adenomatous polyposis coli	IPC	Industry Partnership and Commercialisation
ASCRS	American Society of Colon and Rectal Surgeons	JAG	Joint Advisory Group on Gastrointestinal Endoscopy
ASGE	American Society of Gastrointestinal Endoscopy	KF	Kalman filter
AUC	Area under the curve	LGBM	Light Gradient Boosting Machine
BCSP	Bowel Cancer Screening Programme	LRM	Longitudinal resection margin
BSG	British Society of Gastroenterology	LSL	Laterally spreading lesion
CAC	Cap-assisted colonoscopy	MAD	Median absolute deviation
CIMP	CpG island methylator phenotype	MHRA	Medicines and Healthcare products Regulatory Agency
CIR	Caecal intubation rate	ML	Machine learning
CNN	Convolutional neural networks	MLP	Multi-layer Perceptron
CFP	Cold forceps polypectomy	MRI	Magnetic resonance imaging
CRC	Colorectal Cancer	NBCSP	National bowel cancer screening programme
CRM	Circumferential resection margin	NBI	Narrow band imaging
CSP	Cold snare polypectomy	NICE	NBI international colorectal endoscopic
CT	Computerised tomography	OCT	Optical coherence tomography
CV	Cross-validation	PCCRC	Post-colonoscopy colorectal cancer
DL	Deep learning	PIVI	Preservation and Incorporation of Valuable endoscopic Innovations
DNN	Deep neural network	PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
DRS	Diffuse reflectance spectroscopy	QUADAS	Quality Assessment of Diagnostic Accuracy Studies
EMR	Endoscopic mucosal resection	RCT	Randomised controlled trial
ESD	Endoscopic submucosal dissection	REIMS	Rapid evaporative ionisation mass spectrometry
ESGE	European Society of Gastrointestinal Endoscopy	RMSE	Root mean squared error
ESS	Elastic scattering spectroscopy	ROC	Receiver-operator characteristic
FGS	Fluorescence-guided surgery	SNV	Standard normal variate
FIT	Faecal immunochemical test	SSP	Sessile serrated polyps
FSA	Frozen section analysis	SVM	Support vector machines
FUSE	Full Spectrum Endoscopy	TAM	Technology Acceptance Model
GI	Gastrointestinal	TER	Third eye retroscope
GUI	Graphical user interface	TERRACE	Third Eye Retroscope Randomized Clinical Evaluation
HDI	Human Development Index	UTAUT	Unified Theory of Acceptance and Use of Technology
HFP	Hot forceps polypectomy	XGB	Extreme Gradient Boosting
HSP	Hot snare polypectomy		
HSV	Hue-saturation value		

Chapter 1 Introduction

1.1 Colorectal Cancer

Colorectal cancer (CRC) presents a significant global health burden. It ranks as the third most commonly diagnosed malignancy worldwide, accounting for nearly 10% of all cancer cases. Moreover, it is the second leading cause of mortality, contributing to approximately 9% of all cancer-related deaths globally. [1] Large geographical variations in incidence and mortality rates are observed, with rates highest in Europe. Incidence rates are shown to be four times higher in countries of very high Human Development Index (HDI) compared to countries with low HDI levels. Furthermore, the burden of CRC is projected to increase to 3.2 million new cases and 1.6 million deaths by 2040, with most cases predicted to occur in high or very high HDI regions. [2]

In the UK, over 40,000 new cases are diagnosed annually and CRC leads to around 16,000 deaths every year. [3] The overall five-year relative survival of colorectal cancer in England is 50.7%. [4] The best survival rates are observed amongst those who are diagnosed with early-stage tumours. The five-year survival of patients diagnosed with the earliest stage of the disease (Dukes A) is 93.2%, compared to only 6.6% for those with advanced disease with metastasis at diagnosis (Duke D). Therefore, optimising diagnostic and screening services is crucial.

1.1.1 Post-Colonoscopy Colorectal Cancer

Unfortunately, CRC can appear after a colonoscopy that is negative for cancer. Post-colonoscopy colorectal cancer (PCCRC) may be the result of the development of a rapidly growing new tumour that was not present at initial colonoscopy, but, more often, it is due to missed or inadequately excised pre-cancerous lesions or the CRC simply being missed by the endoscopist. [5,6] As a result, the British Society of Gastroenterology has proposed that the PCCRC rate should be used as a quality indicator for colonoscopy. Wide variations exist in PCCRC rates, with studies reporting prevalence rates that range from approximately 2% to 10%. [5,7–9] This is in part due to variations in the methods used to define PCCRC and the exclusions applied to the dataset. Thus, international benchmarking on an agreed method for calculating PCCRC rates is required.

1.2 Development of Colorectal Cancer

The development of CRC typically starts from a series of genetic alterations that accumulate over time in the cells lining the colon and rectum and, as a result of focal changes within benign, precancerous polyps. There are two main types of polyps with malignant potential, adenomas and sessile serrated polyps (SSPs), each of which is associated with different genetic pathways for developing into CRC. [10,11]

The adenoma-carcinoma sequence, as illustrated in Figure 1, proposed by Vogelstein and colleagues, outlines the step-wise progression from normal colonic epithelial cells to adenomatous polyps and eventually, invasive carcinomas. [12] Typically, mutations develop within the APC (adenomatous polyposis coli) gene, which affects chromosome segregation during cell division. Subsequent mutations in the KRAS oncogene affect cell growth, differentiation, motility and survival. These mutations eventually lead to loss of function of the p53 gene, allowing cells to evade apoptosis and undergo uncontrolled proliferation, thus leading to carcinogenesis. [13,14] In contrast, development of SSPs often begins with mutations in the BRAF gene, resulting in altered growth signalling and loss of apoptosis. [15,16]

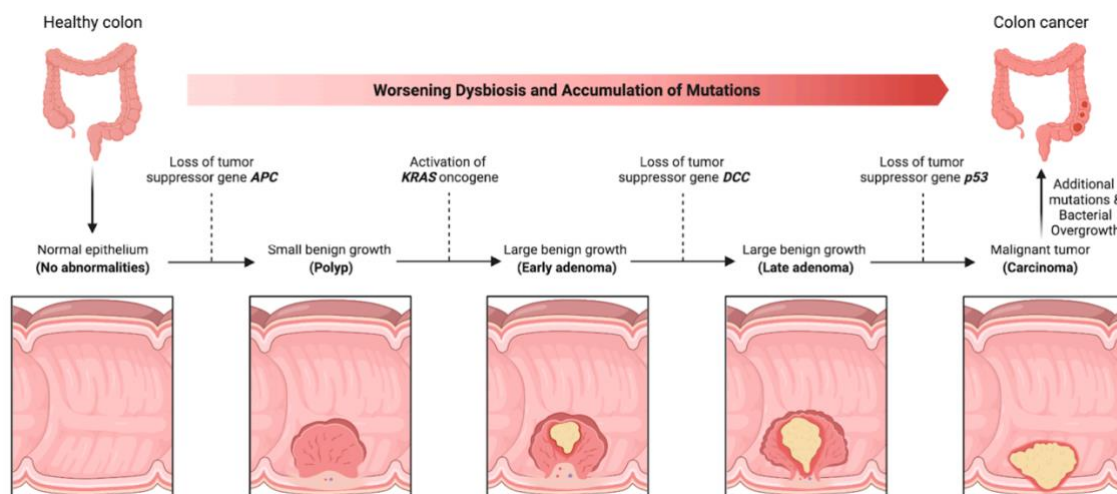


Figure 1 Adenoma-carcinoma sequence. Loss of tumour-suppressor gene, *APC*, results in hyperproliferative epithelium. Further mutations of *KRAS* and *DCC* (deleted in colorectal cancer) lead to adenoma formation and eventually, together with loss of function of the *p53* gene, lead to carcinoma. [17]

In addition, epigenetic alterations also play a significant role in CRC development. [18] DNA methylation and histone modifications can silence tumour suppressor genes and activate

oncogenes, contributing to cancer initiation and progression. The CpG island methylator phenotype (CIMP) has been associated with a subset of CRCs characterised by widespread DNA hypermethylation. [19]

As the genetic and epigenetic alterations accumulate, adenomatous polyps may undergo further changes, leading to the development of invasive carcinoma. Invasive CRCs can invade through the colonic wall and metastasize to distant organs, significantly reducing patient survival rates. [20]

1.3 Detection of Colorectal Cancer

The National Bowel Cancer Screening Programme in the UK aims to detect CRC at an early stage, when treatment is more likely to be effective. The programme currently invites individuals between the ages of 52 and 74 to participate in routine screening every two years. More recently, the age for starting bowel cancer screening has begun to lower, with plans to extend the invitation to those aged 45 and above. [21] Screening is achieved using a home-based faecal immunochemical test (FIT). The FIT is a non-invasive screening tool that quantitatively measures human haemoglobin levels in stool samples, thereby identifying occult gastrointestinal bleeding. If the FIT result is positive, individuals are offered a follow-up colonoscopy to further investigate any abnormalities. This screening approach has been shown to reduce CRC mortality. [22]

Colonoscopy remains the gold standard screening tool for CRC. [23,24] Whilst imaging modalities, such as computerised tomography (CT) colonography, have a high sensitivity for CRC, the diagnosis can only be confirmed through a biopsy. [25] As such, colonoscopy is a technique that allows both diagnostic and therapeutic effect. Magnified mucosal visualisation in high definition is complemented with the ability to perform biopsies and resect detected lesions or polyps. It has been shown that colonoscopy with polypectomy reduces the incidence of CRC by up to 90%. [26–28]

Colonoscopy is relatively safe with studies suggesting a risk of perforation of less than 1/1000, most often as a result of polypectomy. In an asymptomatic screened population, colonoscopy has been shown to have a pooled sensitivity of 92.5% and a specificity of 73.2% per patient for different polyp sizes. [29] The overall success of screening colonoscopy depends upon several parameters such as bowel preparation, caecal intubation rate, withdrawal time, and adenoma detection rate. [30] Despite these factors, mucosal visualisation with standard colonoscopes is between 87-92% in a

clean colon, highlighting the limitations of colonoscopy to adequately visualise the entire colon surface. [31]

1.3.1 Adenoma Detection Rate

The adenoma detection rate (ADR) is the primary colonoscopy quality indicator. It is defined as the proportion of average-risk screening colonoscopies with at least one adenoma. In the UK, the British Society of Gastroenterology, the UK Joint Advisory Group on GI Endoscopy, and the Association of Coloproctology of Great Britain and Ireland have set the ADR target as 15% with an aspirational target at 20%. [23]

The ADR is invaluable to the quality process of a colonoscopy procedure, but it is not readily available since it relies upon correlation with pathology reports. Its importance has been highlighted through a study by Corley et al. which states that for every 1% increase in ADR, there is a 3% decrease in the subsequent risk of CRC (hazard ratio 0.97; 95% Confidence Interval (CI), 0.96–0.98), and a 5% decrease in risk of fatal interval CRC (hazard ratio 0.95; 95% CI, 0.94–0.97).[32]

1.3.2 Factors affecting ADR

The ADR is influenced by a variety of factors (Table 1). These variables can be divided into patient factors, endoscopist factors and endoscopy factors. [33] It is not yet fully understood to what extent these factors affect the quality of the colonoscopy.

Table 1 Factors affecting ADR

Factors affecting ADR	
Patient Factors	Age
	Sex
	Personal history of adenoma or colorectal cancer
	Colonoscopy completion rate
	Adequate bowel preparation
	Patient positioning
Endoscopist Factors	Expertise or individual ADR
	Training and education
	Withdrawal time

Endoscopy Factors	Techniques to improve colour/contrast of adenomas
	Extension of the viewing field

1.3.2.1 Patient factors

Patient-related factors contribute to ADR variability. Studies have shown that the ADR is affected by several patient characteristics including age, sex and personal history of adenoma or CRC. [34,35] It has been shown that, relative to women, men are at one-and-a-half to two-fold increased age-specific risk for advanced adenomas. [36] Patients with a history of adenomas or CRC, as well as those with a family history of CRC, are associated with higher chances of adenoma presence, affecting the overall ADR. [37]

Some patient factors are also associated with colonoscopy completion rates. The completeness of the examination of the colon during colonoscopy is assured by intubation of the caecum and remains a UK key performance indicator. In large studies, a higher caecal intubation rate (CIR) was associated with a significantly increased ADR. [38,39] Lower colonoscopy completion rates have been reported in patients with advanced age, lower BMI, female gender and prior abdominal or pelvic surgery. [40–42]

Another UK key performance indicator for colonoscopy is adequate bowel preparation. Excellent bowel preparation has been associated with superior detection of advanced adenomas. [43] Several papers have examined the patient's role in bowel preparation and colonoscopy quality. Lower socioeconomic status, lack of social support and limited English proficiency, have been linked to worse bowel preparation quality. [44,45]

Patient positioning during a colonoscopy procedure also affects the ADR. In a single-centre study of 120 patients, dynamic position changes during colonoscopy withdrawal significantly improved ADR in the mid and distal colon when compared to examination of the patient in the left lateral position. [46]

1.3.2.2 Endoscopist factors

The ADR of endoscopists is associated with the knowledge, skills and motivation that they possess when performing a colonoscopy procedure. An individual endoscopist's ADR is an independent predictor for interval cancer after screening colonoscopy. [32,47] Endoscopists with more years of experience and high-volume practice are associated with a higher ADR. [48] Some studies have suggested that gastroenterologists are more proficient at colonoscopy than other physicians, including general surgeons, by showing that patients had a significant increased risk of subsequent incidence of CRC when having their procedure performed by the latter group. [49] In a study analysing differences in detected polyps by endoscopists with either a low or high ADR, it was shown that expert detectors more frequently diagnosed not only adenomas, but also high-risk adenomas, compared with low-ADR endoscopists. [50]

Adenoma detection is dependent on skill, training and experience, as shown by the variability in ADR between endoscopists. [51,52] Several studies have sought to improve ADR by increasing training and education for endoscopists. A large-scale randomised trial involving 40 colonoscopy screening centres demonstrated an improvement in ADR through dedicated training of screening centre leaders, resulting in sustained change of their own practice as well as the performance of the centre as a whole. [53] Another study randomised endoscopists to receiving educational intervention in the form of two 1-hour training sessions, and showed that those in the educational group had a significant increase in ADR (47% vs. 36%). [54] The endoscopists were followed up for five months and those in the educational group showed a sustained increase in ADR, highlight the durability of the educational training. [55]

Withdrawal time is another important factor in the detection of adenomas, and a marker of colonoscopy quality that is associated with operator technique. A mean withdrawal time of less than six minutes has been associated with lower ADR and higher AMR. [56–58]

1.3.2.3 Endoscopy factors

Endoscopy factors associated with ADR are based upon two areas: 1) identification of adenomas through techniques to improve their colour/contrast compared to surrounding colonic mucosa, and 2) extension of the viewing field through visualisation of mucosa behind colonic folds. Wide-angle, high-definition endoscopes combined with widescreen, high-definition monitors have been implemented for performing screening colonoscopy to aid polyp detection.

Technologies surrounding image-enhanced endoscopy (IEE) have been developed with an aim to improve ADR. Through the use of dye, optical or electronic strategies to create contrast enhancement during endoscopy, IEE aims to improve detection and characterisation of adenomas. Some previous studies demonstrated that certain image enhancement techniques, including chromoendoscopy, were not superior to standard white light endoscopy in detecting adenomas. [59–62] A systematic review and meta-analysis in 2014 showed that neither autofluorescence imaging (AFI), flexible spectral imaging colour enhancement (FICE), i-scan nor narrow band imaging (NBI) improved the ADR. [63]

However, a more recent meta-analysis in 2019 of 40 studies concluded that NBI and AFI were significantly better at identifying patients with adenomas than standard white light endoscopy. [64] The authors stated that although chromoendoscopy was demonstrated to be superior to standard and high-definition white light colonoscopies in detecting adenomas, it would not be advised to apply this technique to screening colonoscopies of average-risk individuals, given the time-consuming nature of the technique requiring endoscopy and the spraying of dye.

There are several technologies that are available to aid an increase in adenoma detection by allowing examination behind mucosal folds. These endoscopic devices have highlighted some promising results in increasing visualisation of the mucosa, leading to a reduced AMR or an increase in the ADR. [65,66] Endoscopes providing a 330-degree angle of view of the colon demonstrated an AMR of 7% compared to 41% by a standard forward-viewing endoscope ($p < 0.0001$) in an international, multicentre, randomised, tandem colonoscopy trial. [67] This will be discussed further later in this thesis.

1.3.3 Artificial Intelligence and Machine Learning

In recent years, artificial intelligence (AI) has rapidly developed, and its application has penetrated most industries, including the medical field. This technology is able to provide large-scale, data-driven solutions to previously time and labour-intensive problems. AI is characterised by its perception and cognitive ability in computer vision, including object classification, detection, segmentation, and generation. [68] It is based on algorithms that perform tasks that would otherwise require human intelligence. [69]

Within healthcare, AI tools have been developed to aid with tasks ranging from visual identification and classification of disease to clinical decision-making. [70] In the field of radiology, studies have highlighted improved cancer detection by radiologists when supported by AI systems. [71,72] In ophthalmology, AI-enabled retinal imaging has been shown to accurately predict cardiovascular disease and death. [73] These AI-assisted systems have been built using machine learning techniques. Machine learning (ML) is a branch of AI which involves a set of algorithms being trained to perform certain tasks based on training data that is inputted. [74] The more data that is inputted, the more the system learns and the better it becomes at its function. It learns and adapts accordingly so that it is eventually able to correctly apply the learned pattern recognition on new data.

Deep learning (DL) is a subset of ML (Figure 2) and utilises algorithms that are based on neural networks loosely modelled on the biological interconnected neurons of animals. [68] One of the most popular deep learning models is based on artificial neural networks (ANN). ANNs are created by simulating a network of model neurons in a computer system and “teaching” the network how to solve certain problems. [75] This allows large datasets to be analysed and can be used as a foundation to predict future outcomes. Deep neural networks (DNNs) are a specific form of ANNs, built by stacking multiple layers of neural networks to create a “deep” structure. DNNs are able to learn more powerful features directly from training data. [76] For image data analysis, convolutional neural networks (CNNs), which are a class of DNNs, are used since they convolve the input and pass its result to the next layer. Convolution is an important aspect of image processing tasks. When analysing image data, CNNs are more accurate, and less computation power is required, compared to DNNs. Additionally, deep CNN (D-CNN) architectures have been developed. The architecture of these algorithms includes multiple layers of processing between the input and output layers, allowing analysis of complex data with efficient performance. These have formed the basis of recent computer-aided detection (CAD/CADe) and diagnosis (CADx) systems. [77–79]

As such, ML and DL have the potential to reduce human factors and optimise lesion detection, making it a popular novel technology in the field of endoscopy.

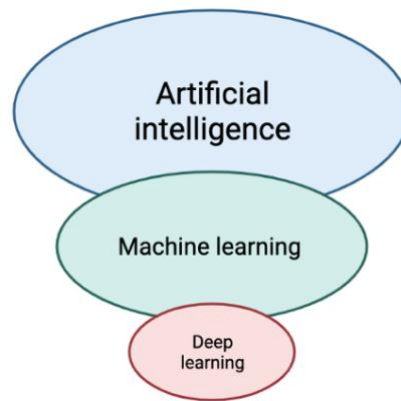


Figure 2 *Deep learning is a subset of machine learning which is in turn a branch of AI*

1.3.4 Machine Learning and Endoscopy

Traditional endoscopic techniques rely heavily on the expertise of the endoscopist to identify abnormalities and made clinical decisions. ML algorithms have introduced a new dimension to endoscopy by augmenting human expertise with the ability to process vast amounts of data, identify patterns and enhance diagnostic accuracy. This technology has emerged as a promising tool in detection and classification of lesions within the field of endoscopy. [80,81] Algorithms have been developed and optimised using still images or videos from colonoscopy procedures and integrated into CAD systems. [82] Systems, including those using CNNs, have shown promising results in identifying lesions in endoscopy. [83] Furthermore, ML has also been used to enhance the quality of endoscopic imaging. [84] Techniques for image enhancement can improve the visibility of lesions and fine structures within the gastrointestinal (GI) tract. ML-driven image enhancement systems can aid endoscopists obtain clearer images and make more accurate diagnoses.

An ideal CAD system in endoscopy should have the following characteristics: (1) high sensitivity for detection of polyps; (2) low rate of false positives; and (3) low latency, so that polyps are detected and flagged in real-time during the procedure. [80] The strive for perfection has led to the emergence of numerous ML systems in the form of novel devices that are intended to assist endoscopists in the detection of polyps in real-time. [82]

1.3.4.1 *Ex vivo* studies

Initially, *ex vivo* studies showed promising outcomes for the use of algorithms in detecting colorectal polyps. Karkanis et al. examined 1200 polyp frames from 60 videos using an algorithm based on

colour analysis and found a sensitivity of 93.6% and specificity of 99.3%. [85] However, these results were limited by long processing time and still images. These limitations were tackled by other studies using methods focussing on edge detection, [86] elliptical shape features, [87] and window median depth of valleys accumulation (WM-DOVA) energy maps. [88] These studies were able to present near real-time video analysis speeds for polyps.

The emergence of deep learning allowed CAD systems to be trained to recognise polyps without a continuous external input. Zhang et al. analysed a total of 180 images as testing datasets to compare a deep learning CNN structure to expert endoscopists. [89] They showed that the algorithm was able to outperform endoscopists in polyp detection accuracy (86% vs 74%) and highlighted a sensitivity of the ML system of 98%. Misawa et al. analysed the per-polyp sensitivity and found it to be 94% when using CNN algorithms on a dataset of 1.8 million still frames evaluated by expert endoscopists. [90] Similarly, Hasan et al. used 1.5 million still frames from 338 videos to assess the per-polyp sensitivity of a CNN algorithm and found it to be 99.7%. [91] For this study, Hasan and colleagues used a system called the GI-Genius (Medtronic) which was able to anticipate the detection against endoscopists in 82% of cases.

1.3.4.2 *In vivo* studies

More recently, *in vivo* randomised controlled trials (RCT) have been published on the use of AI in colonoscopy. The first RCT was published in 2019 by Wang et al. in China. [92] This was a prospective trial evaluating the ADR as the primary outcome when using a real-time automatic polyp detection system which provided a simultaneous visual notice and sound alarm on polyp detection. Of the 1058 patients included, 522 were randomised to colonoscopy with CAD. The study showed a significantly increased ADR in the CAD group versus control group (29.1% vs. 20.3%; $p < 0.001$), primarily due to increase in detection of diminutive adenomas (185 vs. 102; $p < 0.001$) without a significant difference in large adenomas (77 vs. 58, $p = 0.075$).

In 2020, Gong et al. set out to assess the quality indicators for colonoscopy with the development of the ENDOANGEL system. [93] They used this real-time quality improvement system based on D-CNN models to monitor real-time withdrawal speed during colonoscopy. In the intention-to-treat population, they reported a significant increase in ADR when using the ENDOANGEL system compared to the control group (16% vs 8%). However, a major limitation of the study was the lack of

external validation, with the novel system only having been validated at one centre in China, which may have resulted in the very low ADR rates presented in both groups.

The performance of a CAD system in a multi-centre western setting was tested by Repici et al. using the GI-Genius (Medtronic). [94] The ADR was 54.8% when using the CAD system compared to 40.4% in the control group. Diminutive adenomas (≤ 5 mm) were detected in a significantly higher proportion of patients in the CAD group (33.7%) than in the control group (26.5%), as were adenomas of 6–9 mm (10.6% vs. 5.8%), regardless of morphology or location.

1.3.4.3 Commercial AI Devices

These RCTs were quickly followed by the introduction of commercial devices by leading endoscopy manufacturers, despite a lack in studies reporting results of implementation or integration of CAD systems into the clinical healthcare setting within the context of endoscopy workflow. [95] These manufacturers have opted to protect their intellectual property and initially introduce their devices as “second readers” supporting endoscopists in the decision-making process until improvements are made and stakeholder backing is sought. A list of commercial AI devices is shown in Table 2.

Table 2 Commercial AI devices

Commercial AI device	Manufacturer	Country approval
EndoBrain- EYE	Olympus	Japan
GI Genius	Medtronic	European Union, Australia, Israel, South Arabia
CAD Eye	Fujifilm	European Union
Discovery AI	Pentax	European Union

1.3.5 Barriers of AI Technology in Endoscopy

Given the speed at which AI technology in endoscopy is developing, considerations should focus on the real-world implementation of ML in routine clinical practice. The future success of AI in endoscopy will rely on the ability to overcome key ethical and regulatory issues around deployment

of the systems. These challenges must be addressed with full cooperation of stakeholders including healthcare professionals and patients. [96]

Ethical dilemmas include data sharing and privacy, as well as practical issues such as issues pertaining to potential patient harm, liability, decision-making bias and legal frameworks. Initially, healthcare professionals, using AI systems as “second readers” for detecting lesions, could be held responsible for any errors. However, as AI technology progresses and ML systems become more advanced, they are likely to gain more autonomy, especially given their potential to adapt and continuously learn in real-time. As such, strategies in relation to legal responsibility must be put in place. Although AI systems are designed and built by humans, with the aim of doing no harm and achieving their goal in a safe manner, it will be challenging to find responsible parties amongst software developers, engineers, companies and healthcare providers in the case of a medical error. [97]

In addition to legal frameworks, the regulatory bodies for novel devices and technology must also put in place standardised protocols for ML systems within endoscopy to provide assurances of safety and effectiveness. It is clear that the regulatory pathways for AI are currently evolving. Only a few endoscopy AI systems have been approved internationally, such as EndoBRAIN®, for which regulatory approval was obtained by Japan’s Pharmaceuticals and Medical Devices Agency (PMDA). [98] Although this may appear to represent a major barrier to progression, a balance must be sought between protecting patients and promoting innovation.

In this thesis, I will be exploring the sentiments of endoscopists towards ML and its emergence in endoscopy, including their views on the main barriers to the introduction of this technology within the clinical setting. To gain the confidence of stakeholders, as well as to address certain challenges within this sector, more RCTs are needed to report the applicability of ML systems in endoscopy. Once sufficient evidence is gathered, clear evidence-based guidelines can be created for use of AI endoscopic systems in a healthcare setting. [81]

1.4 Management of Colorectal Lesions

The management approach for colorectal lesions can vary significantly based on the nature and characteristics of the lesions. The primary objectives of management include accurate identification of lesions, risk stratification, ensuring appropriate treatment decisions, and minimising the risk of

disease progression. Within the Bowel Cancer Screening Programme (BCSP) in the UK, colorectal polyps are found in 60% of colonoscopies, of which almost 80% are found to be adenomatous polyps. [99] The National Polyp Study is a pivotal study which provided strong evidence that the removal of polyps prevents CRC. [100] The study followed 1,418 patients, who had at least one adenoma resected during colonoscopy, for a mean of 6 years. It was found that the incidence of CRC in the study cohort was significantly lower (76%) than expected when compared to the Surveillance Epidemiology and End Results group. In addition, no CRC deaths were reported.

1.4.1 Endoscopy Management Workflow for Colorectal Lesions

A systematic and thorough approach to the detection, evaluation, resection and histopathological assessment of colorectal polyps and lesions is important in aiding early detection and effective management of CRC. Figure 3 shows the endoscopy management workflow for colorectal lesions.

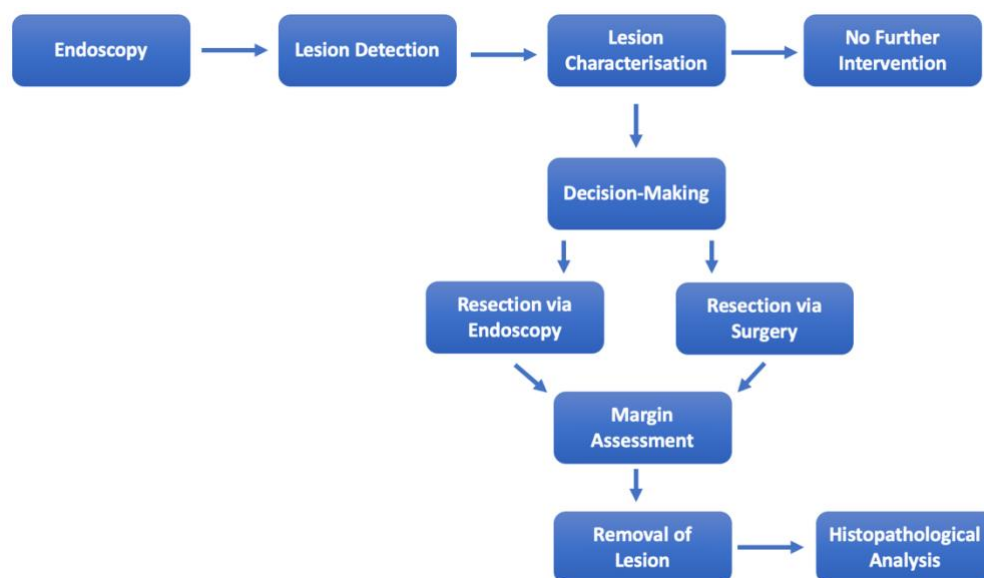
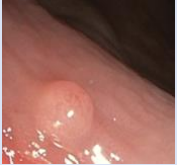
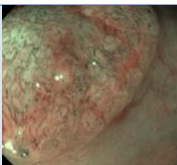


Figure 3 Endoscopy Management Workflow for Colorectal Lesions

Once detected, the first step in establishing whether a polyp is benign or has malignant potential, is through macroscopic characterisation, which is also known as “optical biopsy”. This provides information to predict the histology, biological behaviour and prognosis of the lesion, and is based on its location, size, morphology, suspected histopathology and estimated depth of invasion. The size of polyps can be stratified as diminutive (≤ 5 mm), small (6-9 mm), intermediate (10-19 mm), and

large (≥ 20 mm). [101] In terms of their characteristics, the Paris classification is a consensus system widely used to describe colorectal polyp morphology [102]: protruding types (I) contain pedunculated (Ip) and sessile (Is) lesions; non-protruding and non-excavated types (II) include slightly elevated (IIa), completely flat (IIb), and slightly depressed (IIc) lesions; an ulcer is seen in excavated type (III). Lesions assigned to Is/IIa/IIb/IIc with a size of >10 mm and extending laterally (in contrast to vertically) along the colonic wall are also known as laterally spreading lesions (LSLs). The careful examination of colorectal polyps through high-resolution narrow band imaging (NBI) endoscopy allows for the identification of 1) lesion colour, 2) microvascular architecture, and 3) surface pattern (or pit pattern). [103] These three characteristics serve as the basis for the NBI international colorectal endoscopic (NICE) classification, aiding in the differentiation of hyperplastic polyps (Type 1), adenomas (Type 2), and deeply invasive submucosal cancers (Type 3). [104] This is highlighted further in Table 3.

Table 3 *Narrow-Band Imaging International Colorectal Endoscopic classification system* [105]

	Colour	Vessels	Pits	Association	Endoscopic Image
Type 1	Same or lighter than background	None or lacy vessels	Dark or white spots of uniform size	Hyperplastic or serrated polyps	
Type 2	Browner than background	Brown vessels	Oval or tubular white spots	Adenomatous polyps	
Type 3	Dark brown	Disrupted or missing vessels	Amorphous or absent pits	Deep submucosal invasion	

Once detected and carefully evaluated, the endoscopist must decide the best management option for the colorectal lesions, whether through endoscopic approaches or more invasive surgical methods. These techniques will be discussed further below.

1.4.2 Endoscopic Resection Techniques

Endoscopic removal of colorectal polyps has been shown to reduce the incidence and mortality from CRC. [106] More than 90% of polyps detected during screening colonoscopies are small (<10 mm in size), mostly benign, and do not contain advanced disease. These lesions can be easily managed by conventional endoscopic resection techniques, such as cold biopsy or snare polypectomy. Around 10-15% of colorectal polyps are considered “complex”, and conventional methods are inappropriate for their removal due to their size, location and/or their morphology. These require more advanced techniques or invasive procedures. [107]

1.4.2.1 Biopsy Forceps Polypectomy

Biopsy forceps polypectomy removes polyps through the action of the biopsy forceps which “bite” off the polypoid tissue with the use of its jaws. This can be repeated with further “bites” if there is any residual tissue. The specimen is retrieved directly by using the biopsy forceps and is delivered through the working channel of the endoscope. Biopsy forceps polypectomy is also known as cold forceps polypectomy or hot forceps polypectomy depending on the use of electrocautery or not. Cold forceps polypectomy (CFP) has advantages including widespread device availability and ease of use. However, with an increase in lesion size, number of bites and obscured bleeding field after the initial bite, it has been shown that CFP is associated with incomplete resection rates and inadequate treatment for the removal of diminutive polyps. [108,109] Hot forceps polypectomy (HFP) may be employed as a method to manage bleeding. Nevertheless, the utilization of HFP for polypectomy is discouraged due to its higher likelihood of incomplete resection, insufficient tissue sampling for histopathological examination, and elevated risks of adverse events when compared to snare excision. [101] Biopsy forceps can also be categorised based on their jaw volume, with distinctions made between large-capacity forceps and jumbo forceps. The latter featuring a larger cup size. Jumbo forceps have demonstrated superiority over standard forceps in achieving comprehensive visual endoscopic resection of sessile polyps measuring ≤ 6 mm and in reducing procedure duration. [110]

1.4.2.2 Snare Polypectomy

Snare polypectomy utilises a wire loop which can be extended from a plastic catheter. The catheter is placed down the biopsy channel of the endoscope and placed in position before the snare is opened. The wire loop is carefully placed around the base of the polyp and the snare is then closed to guillotine the polyp.

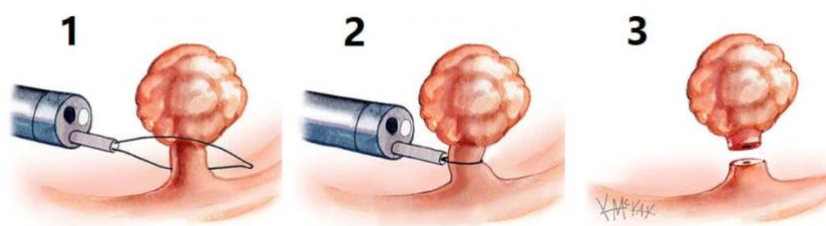


Figure 4 Snare polypectomy. [Image via <https://colorectalsurgeonssydney.com.au/conditions/polyps>]

Cold snare polypectomy (CSP) is the preferred technique for all diminutive (≤ 5 mm) colorectal polyps, as per the recommendations set out by the current European Society of Gastrointestinal Endoscopy (ESGE) guidelines. [101] Its advantages in shortening procedure time and lowering treatment costs contribute to the wide use of CSP in clinical practice. [111] For small non-pedunculated polyps (6-9 mm), CSP is also preferred due to a better safety profile when compared with hot snare polypectomy (HSP). CSP has lower rates of delayed bleeding, perforation, and post-polypectomy syndrome, and it has the advantage of shorter procedure time when compared with HSP. [112]

HSP is defined as snare polypectomy requiring electrocautery. This technique is recommended for all pedunculated polyps. Similar to CSP, the snare is closed around the base of the polyp, and in HSP, electric current is applied allowing polyp removal through coagulation and tissue cutting. In this situation, a base plate is attached to the patient. There is limited research on the optimal diathermy setting for HSP, and as such, this varies widely in clinical practice. [113] The use of electrocautery has advantages and disadvantages. A pure cutting current is advised against by the ESGE since, although it provides high-quality incision and reduced thermal injury, it leads to higher rates of immediate post-polypectomy bleeding. Similarly, a pure coagulation current is advised against by the British Society of Gastroenterology (BSG) due to increased risk of delayed post-polypectomy bleeding and thermal tissue injury. [114] Automatic microprocessor systems can be used to achieve a blend of coagulation and cutting energy, which automatically balances the need for haemostasis and the prevention of thermal injury.

The side-effects of electrocautery may be alleviated through submucosal injection. [115] The injection of a solution into the submucosal space lifts the polyp and may reduce the risk of thermal injury and can improve polyp resection rates. [116]

1.4.2.3 Endoscopic Mucosal Resection

Endoscopic mucosal resection (EMR) is a well-established technique used for the minimally invasive removal of precancerous or early-stage malignant lesions (Figure 5). [117] EMR is primarily employed in the management of colorectal polyps and early-stage neoplasms, providing an effective alternative to surgical resection. EMR is particularly suitable for the removal of lesions larger than 20 mm and those with specific characteristics suggesting low risk for deep submucosal invasion.

The procedure involves the injection of a lifting solution (usually normal saline or a viscous solution) into the submucosal layer beneath the lesion, creating a cushion and elevating the lesion from the surrounding tissue. Subsequently, a snare-like device is employed to capture and resect the lesion while maintaining haemostasis by electrocautery. The resected specimen is then collected for histopathological analysis to determine the nature and extent of the lesion. Lesions of <20 mm can usually be removed in a single piece, whereas lesions of ≥ 20 mm often require piecemeal resection, namely piecemeal EMR. In piecemeal EMR, a lesion is first cut into a small piece to accurately carry out histological diagnosis and the residual part is then resected piece by piece. [118] Attempts to en-bloc resect polyps >20 mm with EMR is associated with a higher risk of potential complications and failure. A recent meta-analysis showed that the pooled proportion of successful en-bloc resection for polyps >20 mm with either conventional or underwater EMR was unacceptably low (49.7%-58.7%). [119]

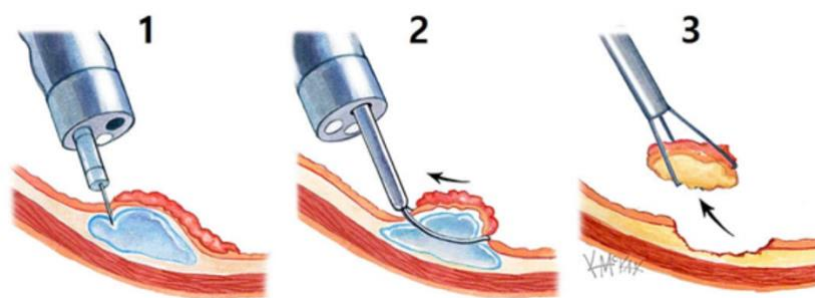


Figure 5 Endoscopic mucosal resection (EMR). (1) Injection of fluid to lift polyp. (2) Snared removal of polyp. (3) Retrieval with forceps. [Image via GastroMedicine & Endoscopy]

There are various solutions available for submucosal injections, usually with the addition of adrenaline to minimise the risk of bleeding. Vital dyes like methylene blue or non-vital dyes like indigo carmine help identify a deep muscular layer injury or perforation. A meta-analysis of five RCTs

showed a significant increase in en-bloc resection (OR 1.91, 95% CI: 1.11–3.29, $p = 0.02$) and fewer residual lesions (OR 0.54, 95% CI: 0.32–0.91, $p = 0.02$) with viscous solutions compared to normal saline used for submucosal injection for EMR. [120]

Several variations of EMR exist, including cap-assisted EMR (EMR-C), band ligation-assisted EMR (EMR-L), and underwater EMR (UEMR), each with its own set of indications and advantages. EMR has been shown to be effective and safe, with high rates of complete resection for appropriate lesions, thus reducing the need for surgery.

A meta-analysis of 33 studies showed the overall recurrence risk for EMR resection to be 15% (95% CI 12–19%). [121] The recurrence rate was higher after piecemeal resection (20%) than en-bloc resection. A multi-centre prospective study of 1,000 successful EMR procedures for sessile or laterally spreading colonic lesions ≥ 20 mm in size showed an early recurrent/residual adenoma rate of 16% (95% CI: 13.6–18.7%). [122] Out of the total, 71.7% of these were diminutive.

1.4.3.4 Endoscopic Submucosal Dissection

Endoscopic submucosal dissection (ESD) represents a revolutionary advancement in the field of therapeutic endoscopy, particularly in the management of neoplastic lesions. This minimally invasive technique allows en-bloc resection of large, early-stage tumours and pre-cancerous lesions. It differs from traditional EMR by enabling precise dissection of lesions from the submucosal layer, rather than relying solely on snare excision.

The ESD technique starts by demarcating the tissue margin, typically extending 3–5 mm beyond the lesion's boundaries. Following this, an injection procedure, similar to that used in EMR, is undertaken. The submucosal space is expanded with a fluid injection, effectively creating separation between tissue layers. This expansion serves to provide precise control over the depth and lateral extent of resection. Next, an endoscopic knife with electrical cutting current is used to incise the lesion's circumference. Subsequently, the submucosal layer is dissected. This technique is able to resect lesions en-bloc in a single piece regardless of the size of the lesion. A meta-analysis evaluating 66 studies comparing EMR and ESD for colorectal lesions showed a higher en-bloc resection rate of 90.5% with ESD compared to 62.8% with EMR (OR 0.18; 95% CI 0.16–0.2).

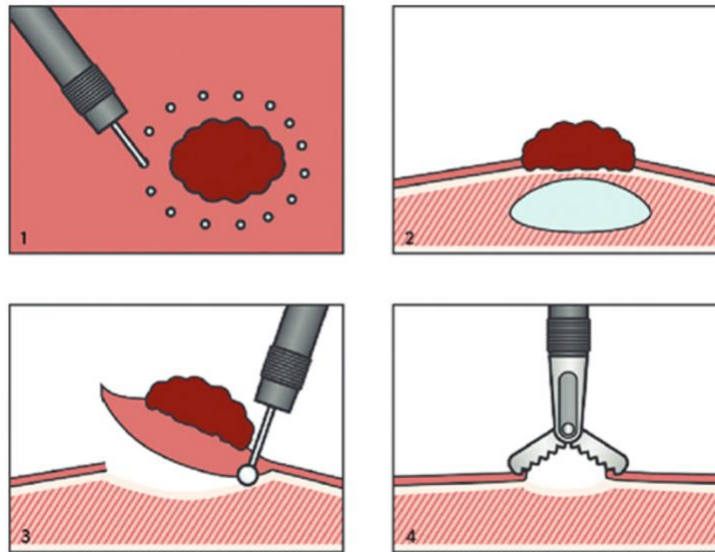


Figure 6 Endoscopic submucosal dissection (ESD). (1) Marking tissue margin. (2) Injection of fluid into submucosal space. (3) Submucosal dissection and resection of lesion. (4) Haemostasis. [Image via <https://www.drpiyushranjan.com>]

Advanced endoscopic skills, knowledge of the technique and specialised endoscopic equipment are needed for ESD. Therefore, careful case selection for this procedure is crucial to achieve clear margins and reduce the risk of recurrence. The aim of ESD is for curative management of cases in which there is a low risk of lymph node metastasis. En-bloc resection by colorectal ESD has shown high rates of curative resection and low rates of recurrence. [123,124] A recent meta-analysis highlighted that, during a 2-year follow-up period, the local recurrence rate after colorectal ESD was 1%. [125] Moreover, the intact specimens produced by ESD facilitate more accurate assessment from both pathological and oncological perspectives. [126,127]

Endoscopic tools have been developed to achieve homeostasis and manage a perforation. [128] However, the technical complexity, risk of adverse events and long duration of the procedure, still limit the use of ESD. [126,129] Studies have shown that ESD is associated with a steep learning curve and that intense skill training is needed by expert endoscopists to reduce the risk of perforation. [130,131] Furthermore, the risk of perforation with ESD is high. A study showed that the perforation rate was 4.8% with ESD compared to 0.9% in EMR ($p < 0.0001$). [132] A meta-analysis of 97 studies with colorectal lesions removed by standard ESD showed a perforation rate of 5.2% (95% CI: 4.4–6.1%). [126] The risk of perforation has been shown to be associated with tumour size, location, submucosal fibrosis and operators with limited experience. [133]

1.4.3.5 Complex Lesions

Lesions which are characterised as technically challenging at endoscopic resection, carry an increased risk of CRC, high recurrence rates in the range of 10-20% after piecemeal resection, risk of adverse events with resection, increased risk of interval cancer after incomplete resection and potential for increased medicolegal issues. [134] For this reason, it is extremely important to identify complex lesions or polyps prior to endoscopic resection and refer these for surgical management, when appropriate. Complex polyps can be identified through their size (>20 mm), shape (flat/bulky), extent (polyps crossing two haustral folds), location (right side, ileocaecal valve, dentate line), or due to previous attempts of endoscopic resection. [135–137] Identification also requires access to clinical expertise. Complex polyps should be managed by expert endoscopists with training in advanced polypectomy, and the confidence in referring challenging lesions to achieve optimal oncological resection and patient outcome in case of malignant polyps. [138]

The prognosis of CRC varies depending on depth of tumour invasion. The tumour, nodes and metastasis (TNM) model is a clinic-pathological method of staging CRC, with T1 tumours indicating invasion into the submucosa. The Kikuchi classification system describes submucosal invasion in sessile and flat malignant colorectal polyps by dividing the submucosa into three levels: sm1 describes invasion into the upper third of submucosa, sm2 describes invasion into the middle third of submucosa and sm3 describes invasion into the lower third of submucosa. Lesions with suspicion for deep submucosal invasion, defined as tumour involvement ≥ 1 mm or sm3, are associated with a high risk of lymph node metastasis and residual recurrence (10-18%), and should be referred for surgical management. [139]

Diagnosis of the depth of tumour invasion can be difficult, and relies on endoscopic findings. The use of NBI, staining, observation of surface and vascular structures with magnification endoscopes, and the use of endoscopic ultrasonography have improved the accuracy of determining invasion depth, however, studies have reported that the accuracy remains between 59% and 84%. [140,141]

1.4.3 Surgical Resection

Despite advanced endoscopic techniques and improving skills, some patients still require subsequent surgical resection of colorectal lesions, due to a high risk of lymph node metastases or positive resected margins. Oncological resection of colorectal lesions is performed via segmental resection of the affected portion of the bowel and its draining lymph node basin. Harvesting these nodes can

achieve local control and stage the disease. The intraoperative localisation of small lesions or a previous polypectomy site is often challenging, especially during the laparoscopic approach. Despite many well-known advantages, the laparoscopic resection of colorectal cancer still has significant limitations, including a limited operative field of view and lack of tactile sensation. Therefore, without precise pre-operative localisation, several studies have reported inadvertent excision of healthy bowel where the primary tumour was not adequately located.

Various methods are widely used for pre-operative localisation of colorectal lesions. Radiological techniques, such as CT or magnetic resonance imaging (MRI), are commonly used to identify and precisely locate colorectal lesions. These imaging modalities are particularly useful for assessing potential involvement of adjacent structures, provide a great deal of information regarding vascular anatomy and studies report a high accuracy for MRI in T staging. However, limitations include difficulty in differentiating fibrosis from tumour infiltration, especially following chemoradiation therapy. This has been shown to compromise the ability to distinguish early stage T3 tumours from T2 tumours in rectal cancer. [142]

Double contrast barium enema is another effective method for identifying large tumours, however, small lesions are frequently missed. [143] Given that approximately 10-20% of tumour locations identified from colonoscopy are inconsistent with the intraoperative tumour site, [144,145] preoperative endoscopic tattooing is a safe, highly effective and common approach for localisation. The accuracy of endoscopic tattooing for localisation has been shown to vary between 70% to 100%. [146] Several studies have highlighted endoscopic tattooing to be a safe technique with relatively small number of complications. [147,148]

One of the crucial pillars of surgical quality is achieving negative resection margins. Two critical margins that are often assessed are the circumferential resection margin (CRM) and the longitudinal resection margin (LRM). Negative CRM and LRM can be considered indicative of a successful oncological resection. [149] A positive CRM, which refers to the presence of tumour cells at the cut edge of the surrounding tissue, is associated with a higher risk of local recurrence and a poor prognosis amongst patients with colon, as well as rectal, cancer. [150–152] Similarly, a positive LRM, indicating that tumour cells are present at the proximal or distal surgical margin, has also been shown to be a predictor for local recurrence, development of distant metastasis and decreased disease-free survival. [153]

Ensuring clear margins in both the circumferential and longitudinal dimensions is essential in optimising the curative potential of surgery for CRC. However, despite adhering to established surgical protocols aimed at achieving macroscopically negative resection margins, positive resection margins are still detected microscopically on post-operative histopathological examination. [154]

1.4.3.1 Resection Margin Assessment

Intra-operative margin assessment in colorectal surgery plays a pivotal role in ensuring the adequacy of tumour resection and, consequently, patient outcomes. Currently, the assessment of surgical margins primarily relies on conventional techniques, such as visual and tactile feedback, as well as frozen section analysis. These techniques are essential components of the surgical workflow, allowing surgeons to gauge the completeness of tumour removal. Visual inspection involves the meticulous examination of the surgical field under optimal lighting conditions to identify any residual tumour tissue, whilst tactile feedback complements this by enabling surgeons to physically assess, via manual palpation, the resection area for any indistinct or suspicious areas. Visual and tactile feedback alone are not sufficiently reliable in determining the margin status, and lack the sensitivity required to detect subtle residual tumour tissue, particularly when dealing with lesions that are very small or deeply embedded within the tissue. The subjective nature of these techniques can introduce variability in interpretation and will likely result in false-negative assessments.

1.4.3.1.1 Frozen Section Analysis

Frozen section analysis (FSA) is an intra-operative diagnostic technique, that involves the examination of a tissue sample during surgery. The procedure involves the surgeon removing a portion of the tissue mass and sending this sample to a pathologist. The pathologist freezes the tissue in a cryostat machine, cuts it with a microtome, then stains it with various dyes so that it can be examined under the microscope. The pathologist then provides a diagnosis to guide the surgeon during the ongoing surgical procedure. The whole process can take between 10-45 minutes.

Although this technique has high accuracy, with reports within GI cancers of up to 99.4%, [155] its diagnostic value in the detection of inadequate margins is limited due to sampling error and finite number of sampling points. [156] Furthermore, it is time consuming, costly and requires the use of multiple workforces. In a study by DiNardo et al., the estimated cost of intraoperative frozen section averaged as much as \$3,123 per patient. [157] In addition, the potential for freezing artifacts can

introduce challenges to the process. One study found the reasons for inappropriate quality of frozen sections to be freezing artifacts, insufficient staining and small volume of tissue. [158]

Definitive histopathological analysis by a pathologist is the gold standard method for margin assessment. However, this usually takes several days or up to a week. Therefore, definitive histopathological assessment cannot aid intra-operative decision making, such as delineating the tumour margin within the surgical field during the resection of the tissue, assessing the surgical cavity to aid immediate resection of any residual tumour tissue, nor allowing assessment of the surgical specimen within the operating room to guide whether immediate re-resection is necessary. The ideal intra-operative tool for margin assessment should meet the following clinical criteria: capable of diagnosing with high accuracy, quick to perform, easy to operate, real-time feedback, applicable to a wide patient group, cost-effective and most importantly, leading to lower reoperation rates.

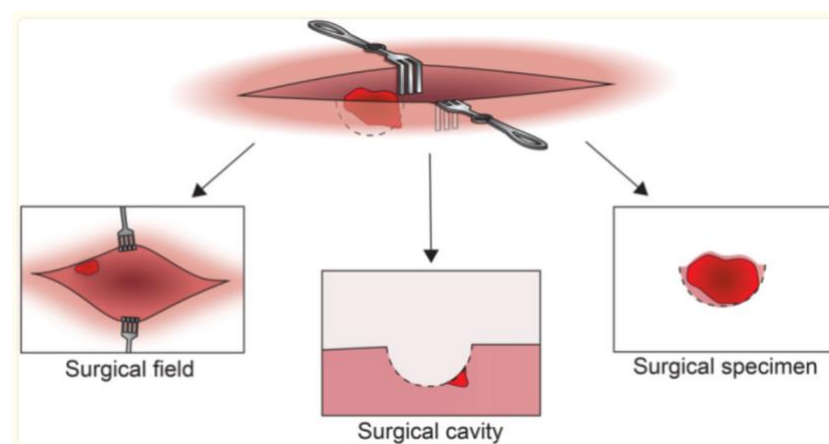


Figure 7 The three areas of interest when assessing intra-operative margins using real-time imaging techniques.[159]

1.4.3.2 Optical Techniques for Margin Assessment

To reduce the associated consequences of inadequate margins, there has been a growing interest in the use of optical techniques for intra-operative margin assessment. [159] Optical technologies offer the opportunity to extract structural and morphological information from biological tissues using light-tissue interactions. The key advantages of optical techniques, including non-invasive and non-ionising properties, make them ideal for use intra-operatively for discrimination of tissue type. Light delivered to biological tissues undergoes several processes, such as reflection and refraction at the

surface, as well as scattering from cellular structures. [160] Wavelengths of light may be absorbed by molecules within tissue known as chromophores, such as water, lipids and haemoglobin. [161] Moreover, light may be emitted by structures creating a further interaction known as fluorescence. As a result of these properties, several optical techniques are currently being explored for tissue discrimination and margin assessment.

In this thesis, I will be highlighting the development of an optical technique, known as diffuse reflectance spectroscopy, through application of ML for intra-operative margin assessment, and exploring its technical properties, diagnostic accuracy and its feasibility in clinical practice.

1.5 Diffuse Reflectance Spectroscopy

Diffuse reflectance spectroscopy (DRS) is an optical technique which measures the intensity of diffusely reflected light as a function of wavelength. [162] The DRS probe contains several fibres, one of which transmits light onto the tissue and others which collect the reflected light back within the probe for measurement by a spectrometer (Figure 9). Typically, light is delivered to the sample through an optical fibre. Upon entering the tissue, the light may undergo several interactions with the tissue, including absorption and scattering (Figure 8). Scattering refers to the change in direction of individual light particles as they interact with various structures within the tissue. Different tissues have different morphology and biochemical composition which affects the amount of light which is absorbed and scattered. This is highlighted in the changes seen in the DRS spectrum. Light which is detected has generally undergone multiple scattering events, and so the change in the recorded spectrum also infers information about the absorbers within the tissue. Spectral analysis enables extraction of information about the optical properties of biological tissue, and as such, can be used to differentiate tissue type.

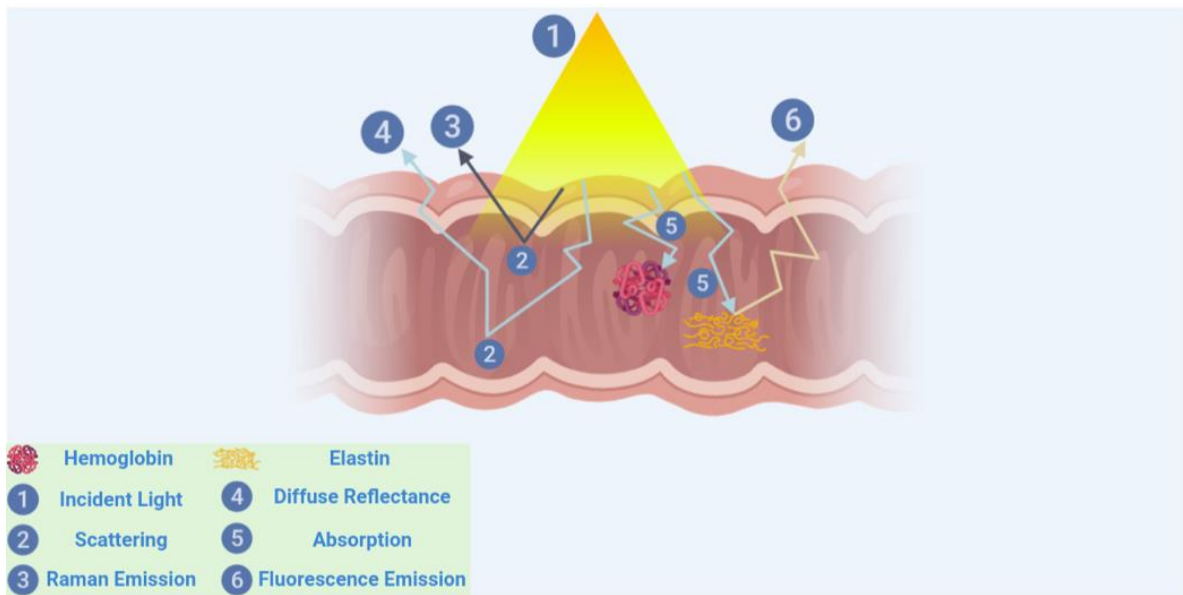


Figure 8 Illustration of the basic principles of DRS with absorption and scattering of light within a sample of tissue.

The distance between the emitting and receiving fibres within a DRS probe is important in determining the depth at which the tissue is primarily assessed. By adjusting the separation between these fibres, one can control the average path that the detected light traverses within the tissue. Increasing the distance between these fibres enables the average path of the detected light to penetrate deeper into the tissue. This occurs because the increased separation allows the light to travel a more extended and varied route before reaching the detector. Consequently, a larger volume of tissue is probed, enabling assessment at greater depths. Although the fibre distance primarily dictates the average measurement depth, it is concomitantly influenced by the tissue's composition and the wavelength of the incident light.

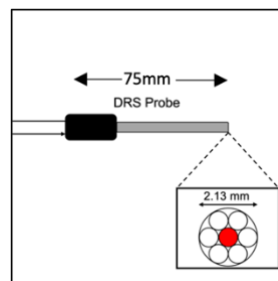


Figure 9 Diffuse reflectance probe containing six peripheral illumination fibres around one light collection fibre.

The intensity of the received signal is influenced by the interaction of light with tissue components. When the wavelength of the incident light corresponds to a range that is primarily absorbed by the tissue, the intensity of the received signal tends to be low. This is because a significant portion of the incident light is absorbed by chromophores present in the tissue. Conversely, when the wavelength of the incident light aligns with a wavelength predominantly scattered by the tissue, the received signal tends to exhibit heightened intensity. This elevation is attributable to scattering events that cause the light to deviate from its initial trajectory, resulting in a greater proportion of scattered light reaching the detector. The absorption and scattering properties of tissue exhibit wavelength dependency. Therefore, the diffuse reflectance is measured over a range of wavelengths to generate a complete diffuse reflectance spectrum (Figure 10).

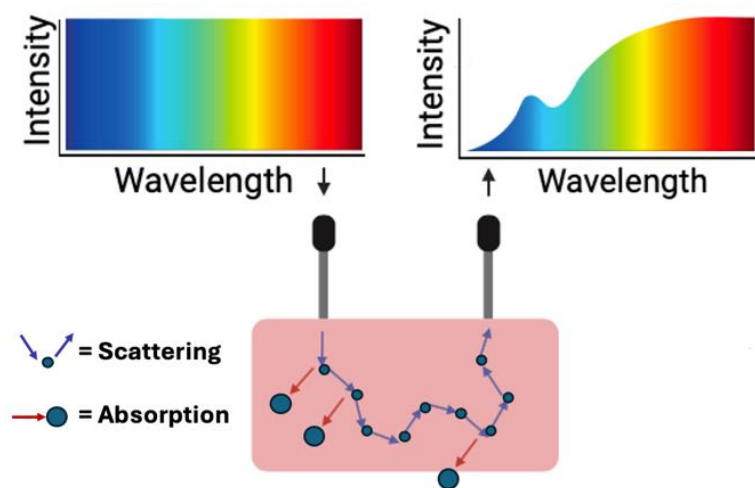


Figure 10 DRS spectrum created through scattering and absorption of light measured over a range of wavelengths.

DRS has been applied as a diagnostic tool in differentiating tissue type in a large number of organs including lung, skin, breast, cervix and brain. [163–167] Recent studies have explored the use of DRS in discriminating benign and malignant tissue. Cancer leads to significant structural changes at the cellular and sub-cellular levels within tissue which can be evidenced by changes in spectral measurements of the diffuse reflected light.

1.5.1 Limitations of DRS

The progression of DRS as a tool for lesion detection and margin assessment has been challenging due to its limitations. Firstly, it is a point-based system meaning that it yields spectral plots of single point measurements of diffuse reflectance. As a result, when scanning tissue, it is not able to collect continuous data and studies have only been able to obtain a small number of spectra in a given time

period. [168] Secondly, DRS primarily measures light absorption and scattering when applied to the surface of tissue. Therefore, the penetration depth of DRS is not fully understood and has been the focus of multiple studies. [169] And lastly, biological variability, such as tissue perfusion based on changes in blood flow or oxygenation levels, can impact the interpretation of DRS data. This leads to difficulty in fully assessing DRS results when studying tissue *ex vivo*. [170]

Addressing these limitations requires combining DRS with other techniques and advanced data analysis methods. Incorporating ML into DRS enhances its capabilities and allows development into a clinically applicable tool. This is achieved through enhanced data analysis, quality control, reduced human error and automation. ML-driven improvements can lead to more accurate and reliable tissue discrimination, ultimately benefiting patient care. [171]

Further advancements in probe tracking, data collection, and histological correlation are also essential. Firstly, precise tracking of the DRS probe is crucial for real-time, accurate tissue mapping. Without tracking, it is challenging to identify and record the exact locations from which spectra are obtained, limiting spatial accuracy during *in vivo* applications. Developing integrated probe-tracking mechanisms would enhance the ability to map spectral data onto tissue accurately, enabling more precise assessments during surgical procedures.

Secondly, improving data collection methods will be key to making DRS a clinically reliable tool. Traditional DRS data acquisition is often limited in speed and spatial continuity, but innovations in hardware and software could allow for more efficient collection of spectral data across larger tissue areas. This could transform DRS from a point-based to a more comprehensive imaging modality, allowing for continuous data capture and ultimately providing a more holistic view of tissue characteristics.

Finally, correlating DRS data with histological findings remains critical to validate the accuracy of DRS in tissue differentiation. Establishing a robust correlation between spectral patterns and histopathological results will enhance the specificity and sensitivity of DRS for clinical use. With these developments, DRS has the potential to evolve from an experimental tool into a valuable aspect of intra-operative decision-making in cancer surgery.

1.6 Commonalities between Upper and Lower Gastrointestinal Tract Cancers

Although this thesis primarily focuses on the lower GI tract, several components of the research were undertaken in the context of upper GI cancers. It is, therefore, important to acknowledge the shared features between upper and lower GI malignancies which support the translation of findings across both domains.

From an anatomical perspective, the upper and lower GI tracts form a continuous hollow organ system lined by mucosa and submucosa. Both regions demonstrate complex tissue architecture that influences light-tissue interactions, making optical diagnostic techniques, such as DRS, broadly applicable. Similarly, both upper and lower GI organs are subject to minimally invasive approaches, including laparoscopic and endoscopic procedures, which provide common ground for intraoperative optical guidance tools.

In regard to cancer biology, upper and lower GI tumours share several hallmarks of carcinogenesis, including dysregulated epithelial cell turnover, accumulation of genetic and epigenetic alterations and progression through premalignant stages – for example, Barrett’s oesophagus in the upper GI tract and adenomatous polyps in the colon. Angiogenesis, altered cellular metabolism and changes in tissue scattering and absorption properties are observed across GI malignancies.

Staging and management of both upper and lower GI cancers are also similar with classification systems focussing on depth of invasion, nodal involvement and metastatic spread. Surgical resection and adequate margin assessment remains key in curative treatment, supported by multimodal strategies including chemotherapy and radiotherapy. The accuracy of tumour margin assessment remains a major determinant of surgical outcome in both upper and lower GI cancers which is precisely the challenge that this thesis seeks to address.

Given these commonalities in anatomy, cancer biology, staging and surgical principles, it is reasonable to draw on findings from both upper and lower GI cancer studies. The inclusion of upper GI work in this thesis reflects these shared characteristics and demonstrates that methods developed and validated in one part of the GI tract can be applied to another. However, the principal focus of this thesis remains the lower GI tract, with the overarching aim of advancing techniques to improve detection of lesions, surgical precision and resection margins and ultimately survival and outcomes for patients with colorectal cancer.

1.7 Aims of this Thesis and Overarching Goal

This thesis will address the challenges in CRC detection and resection through a series of complementary studies from exploring ML technologies to enhance lesion detection during colonoscopy, assessing endoscopists' perspectives on AI integration, to validating a DRS probe system for real-time tumour margin assessment in CRC. Together these studies demonstrate a pathway for improved detection of polyps, as well as improved intra-operative margin evaluation, with the ultimate goal of integrating DRS technology into endoscopy to support both detection and resection of complex lesions through techniques such as ESD. In doing so, this work highlights the potential of innovative optical and AI-based technologies to transform CRC diagnosis and treatment, contributing to improved patient outcomes.

Chapter 2 Lesion detection and characterisation with the use of computer-aided devices in endoscopy: systematic review and meta-analyses

Work from this chapter was published in:

- **Diagnostic accuracy of artificial intelligence and computer-aided diagnosis for the detection and characterisation of colorectal polyps: A systematic review and meta-analysis.** Scarlet Nazarian, Ben Glover, Hutan Ashrafian, Ara Darzi, Julian Teare. *Journal of Medical Internet Research* 2021; doi: 10.2196/27370
- **Does computer-aided diagnostic endoscopy improve detection of commonly missed polyps? A meta-analysis.** Scarlet Nazarian, Arun Sivananthan, Lakshmana Ayaru, Kinesh Patel, Hutan Ashrafian, Ara Darzi, Nisha Patel. *Clinical Endoscopy* 2022; doi: 10.5946/ce.2021.228

2.1 The Diagnostic Accuracy of Artificial Intelligence and Computer-Aided Diagnosis for the Detection and Characterisation of Colorectal Polyps: A Systematic Review and Meta-Analysis.

2.1.1 Background and Aims

As previously stated, colorectal cancer poses a significant global health concern, ranking as the third most commonly diagnosed malignancy and a leading cause of mortality. [172] It typically develops from sporadic colorectal adenomatous polyps. These lesions are detected and resected by colonoscopy, significantly reducing the incidence and mortality of CRC. [106] However, the effectiveness of colonoscopy procedures is hindered by operator-dependent limitations, resulting in potential missed polyps. Studies have highlighted that the miss rate for adenomas may be as high as 26%. [173] The ADR, the quality indicator for colonoscopy, has been shown to inversely correlate with the PCCRC risk, with each 1% increase in ADR corresponding to a 3% reduction in subsequent risk of cancer. [32] This has emphasised the need for strategies to enhance endoscopy quality.

2.1.1.1 Current Methods to Improve ADR

The PCCRC rate is associated with the endoscopists' ability to evaluate the colonic mucosa and resect lesions. As much as 86% of CRC cases following colonoscopy are associated with inadequate colon visualisation or incomplete removal of pre-cancerous lesions. [5] The clinical challenges

presented by endoscopes in detecting adenomas have been met by the production of new technology or modifications to standard endoscopic devices. These include cap-assisted colonoscopy (CAC), Endocuff Vision (Arc Medical, Leeds, England), EndoRings (EndoAid, Caesarea, Israel), FUSE Full Spectrum Endoscopy (EndoChoice, Alpharetta, GA), Third-eye panoramic device (Avantis Medical Systems, San Jose, CA), and G-EYE system (Smart Medical Systems, Israel). These approaches are aimed at improving visualisation of mucosa behind folds or helping to extend the viewing field (**Table 4**).

2.1.1.1.1 Visualisation of Mucosa Behind Folds

Endoscopic caps are disposable, transparent attachments that are applied to the distal end of an endoscope. They are used as a means of manipulating and flattening bowel folds, thereby allowing visualisation of their proximal aspects. In a large randomised controlled trial, CAC had no benefit in ADR. [174]

The Endocuff is a single-use device made up of a radio-opaque sleeve with flexible finger-like projections which is mounted at the tip of the colonoscope. The projections sit flush in two rows and extend during withdrawal to help flatten the colonic mucosa. This allows assessment of otherwise unsighted areas of the colon. In a randomised controlled trial involving 492 patients, EndoCuff-assisted colonoscopy was shown to increase the ADR by 14.7% ($p<0.0001$) when compared to standard colonoscopy. [66]

Similarly, EndoRings is a distal add-on device that consists of clear silicone discs or rings which are able to deflect in opposite directions during insertion or withdrawal to flatten folds. The AMR was demonstrated to be 10.4% in the device group compared to 48.3% in the standard colonoscopy group ($p<0.001$) in a study assessing the value of EndoRings. [65]

The G-EYE system is composed of a balloon that is integrated into the distal part of the colonoscope. An accompanying inflation system (NaviAid Spark2C Inflation System; Smart Medical Systems Ltd) is operated using a foot pedal. The balloon is deflated on insertion of the scope and inflated on withdrawal from the caecum to straighten out folds. The G-EYE balloon endoscope was compared to standard colonoscopy in a randomised tandem study with 126 patients and shown to significantly lower AMR (7.5% vs 44.7%; $p=0.0002$). [175]

2.1.1.1.2 Extension of the Viewing Field

The Full Spectrum Endoscopy (FUSE) system provides high-resolution, 330° field of view, achieved by three imagers and LED groups positioned on the left and right side of the scope's distal tip. Two multicentre, prospective randomised tandem studies have shown a reduction in the AMR when using the FUSE system versus standard colonoscopy. [67,176] With the wider field of view provided by the FUSE system, the ADR was 34% compared to 28% in the standard forward-viewing colonoscopy group. [67]

The Third Eye Retroscope (TER) is an auxiliary device which extends beyond the tip of the scope to bend 180 degrees in a J-shape form allowing retrograde view of the colonic lumen during withdrawal. However, concerns over cost and inability to use the working channel of the scope led to development of the Third Eye Panoramic (Avantis). This device attaches to the distal tip of the colonoscope and provides a 330-degree field of view whilst keeping the working channel free for use. The Third Eye Retroscope Randomized Clinical Evaluation (TERRACE) study demonstrated a 23.2% increase in detection of adenomas by the TER device compared to standard colonoscopy. [177] This was confirmed by a subset analysis of the results involving 349 patients. [178]

Table 4 *Current methods to improve ADR*

Aim	Method/device	Overall effect (vs. standard colonoscopy)
Visualisation of mucosa behind folds	Endoscopic caps	No benefit in ADR [174]
	EndoCuff	14.7% increase in ADR (p<0.0001) [66]
	EndoRings	AMR 10.4% vs. 48.3% (p<0.001) [65]
	G-Eye	AMR 7.5% vs 44.7% (p=0.0002) [175]
Extension of the viewing field	FUSE	Reduction in AMR [67,176] ADR 34% vs. 28% [67]
	Third Eye	23.2% increase in ADR [177]

2.1.1.2 Training Initiatives to Improve ADR

Despite advances in modern endoscopy, it remains provider-dependent, leading to differences in outcomes. [179] Human and technical factors lead to a small yet significant proportion of undetected lesions during colonoscopy. Rex et al. observed different endoscopists performing back-to-back colonoscopies on the same patient and noted large differences in the adenoma detection miss rates, ranging from 17% to 48%. [180] In addition to colonoscopy quality assessment and ADR,

training is another measure recommended to improve endoscopists' performance and reduce the burden of PCCRC. Numerous studies propose that enhancing the educational and behavioural skills of endoscopists can increase the ADR. [181] Training initiatives, involving practical instruction and consistent feedback, yielded positive outcomes in improving ADR in clinical trials. [53,181] A systematic review highlighted that a training intervention compared to no intervention was associated with an improved ADR, supporting brief, focussed training for continuing professional development purposes. [182] Nevertheless, the rise in detection rate observed in these limited trials was marginal, with high heterogeneity between the training interventions, highlighting the ongoing challenge for endoscopists to identify small, subtle or flat lesions.

2.1.1.3 AI Methods to Improve Detection and Characterisation of Polyps

In recent years, there has been a notable integration of AI technologies in healthcare, in particular in the field of diagnostics. [183] The capacity of AI and ML methods to detect and characterise various pathologies, has the potential to enhance and outperform traditional CAD systems. Within endoscopy, ML algorithms can be specifically trained to identify or characterise polyps in real-time. Research has explored two primary endoscopic methodologies: those analysing non-magnified endoscopic images and those focussing on cellular imaging at a microscopic level, known as optical biopsy.

Colorectal polyp detection can be defined as the process of training a ML model to learn a set of features representing a polyp from a given image or live video stream. In some cases, these models are further developed by adding a classification component to be able to distinguish different classes of polyps.

When building an automated polyp detection system using ML, a common structure is followed. This includes sample normalisation, augmentation, model fine-tuning and evaluation as the standard first steps. Firstly, a dataset consisting of either static images or video frames is needed to train the model. Next, the samples are pre-processed and augmented using various techniques such as flips, random crops and rotations. Augmentation is required when public datasets are small, and as a result, do not cover all possible scenarios, such as variations in illumination, the visual field and polyp sizes.

Colorectal lesion detection models are built with one of three objectives: (1) polyp segmentation, (2) polyp detection, and (3) polyp classification. Some models are trained to carry out multiple objectives simultaneously. Once a model is trained to detect a lesion or polyp, training images with their corresponding labels are fed to a network which learns how to localise a lesion in a given image. The format for this labelling is usually a combination of four coordinates representing the bounding box, a variable representing whether a polyp is present, and sometimes, the class of the polyp is added, provided that such information exists in the dataset. In contrast, segmentation models are trained to draw an image segmentation mask around a detected polyp. Segmentation models are trained on the colonoscopy images, and the corresponding image masks are used as the input labels. Finally, classification models are trained to classify polyps based on the group a polyp belongs to without identifying the location. [184]

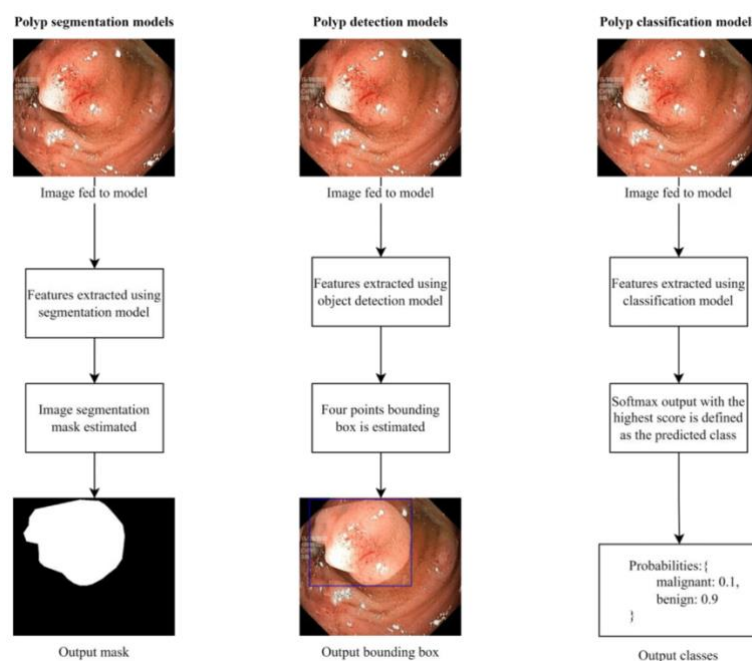


Figure 11 Three common colorectal lesion detection models.

The primary advantage of these methods lies in their potential to elevate the ADR, subsequently reducing the occurrence of CRC. These technologies are able to do this by detecting polyps even at the very limit of the field of view, by detecting multiple polyps at the same time and by distinguishing mucus caps or rims of debris around polyps from residual stool.

Additionally, computer-aided characterisation of detected polyps may alleviate the financial strain on the healthcare system associated with the histological analysis of low-risk colorectal polyps. Most of the polyps detected during colonoscopy are diminutive in size, with a negligible risk of progression

to cancer. [185] Unnecessary resection and histopathology evaluation of these non-neoplastic lesions are associated with increased costs and risk of significant morbidity. [186] The Preservation and Incorporation of Valuable endoscopic Innovations (PIVI) initiative, established by the American Society of Gastrointestinal Endoscopy (ASGE), has outlined specific benchmarks for the integration of new endoscopic technologies, including the optical diagnosis of small colorectal polyps. [187] The “resect and discard” approach is recommended when the optical biopsy aligns by 90% with histopathological assessment for post-polypectomy surveillance intervals, and the “diagnose and leave” strategy is recommended for hyperplastic polyps when the negative predictive value (NPV) for diminutive rectosigmoid adenomas is 90% or more. Several studies, predominantly conducted at single centres, have met the PIVI criteria highlighting that a “resect and discard” strategy or a “diagnose and leave” strategy could be adopted. [188,189] However, a recent multi-centre study emphasised that the precision of optical diagnosis necessitates imaging advancements before it can be relied upon to determine surveillance without histology. [190]

2.1.1.4 Aims of the Study

The aims of this study were to systematically review and meta-analyse the diagnostic accuracy of AI-based technologies in the detection and characterisation of colorectal polyps.

2.1.2 Methods

This review was carried out and reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement. [191] This review was registered on PROSPERO (registration number: CRD42020169786).

2.1.2.1 Search Strategy

A comprehensive literature search was undertaken using the databases of EMBASE, Medline and the Cochrane Library. All published articles up until October 2020 were included. Search terms used in EMBASE and Medline included “colon*”, “polyp”, “artificial intelligence OR machine learning” and “computer aided or assisted and diagnos* OR detect*”. Studies in the Cochrane Library were identified with the terms “colonic polyp”, “artificial intelligence” and “diagnosis, computer-assisted” (Appendix 1).

2.1.2.2 Inclusion and Exclusion Criteria

Inclusion criteria were as follows:

- Studies reporting computer-aided detection of colorectal polyps retrospectively using endoscopic images or videos.
- Studies reporting computer-aided classification of colorectal polyps retrospectively using endoscopic images or videos.
- Studies reporting the use of computer-aided diagnosis of colorectal polyps during colonoscopy.
- Studies reporting ADR, PDR, sensitivity, specificity, and diagnostic accuracy data; or studies with adequate information to calculate these data.
- Studies published or translated into English.

Exclusion criteria were as follows:

- Studies with no original data present (e.g. review article, letter).
- Studies with no full text available.
- Studies conducted in patients with inflammatory bowel disease (IBD).
- Studies greater than 20 years old.
- Studies without adequate data to calculate sensitivity, specificity and diagnostic accuracy data, PDR and ADR, adenoma miss rate (AMR), or mean adenomas per patient, or not reporting these data.

2.1.2.3 Study Selection

Two reviewers (SN and BG) conducted a screening of all retrieved articles to eliminate duplicates. Titles and abstracts were then screened for relevance by the two reviewers independently, and irrelevant studies were excluded. Next, full-text reviews of remaining studies were completed. The reference lists of identified review articles and included papers were scrutinised for relevant studies. Disagreements about eligibility were settled by consensus, both after screening and following full-text review. All final articles met the predetermined inclusion and exclusion criteria.

2.1.2.4 Data Extraction

Data was gathered from studies onto a standard spreadsheet template. For each study, we extracted the following data: study details (first author, year of publication, journal), primary outcome (polyp detection vs. characterisation), study design (type of study, method of AI, exclusion criteria), information on type of imaging modality (images or videos, images for training, images for validation), and information regarding diagnostic accuracy characteristics (sensitivity, specificity, accuracy, ADR, PDR).

2.1.2.5 Study Quality Assessment

Study quality was independently assessed using the Quality Assessment of Diagnostic Accuracy Studies (QUADAS 2) tool. [192] Each domain was classified as low risk, high risk or unclear risk of bias. For randomised controlled studies, the Jadad scale was used for quality scoring. [193] Studies with a Jadad score of 3 or more were considered “good” quality.

2.1.2.6 Statistical Analysis

The individual proportions and their variances were independently calculated and combined using the DerSimonian and Laird random-effects model. This method accounted for both between-study and within-study variances, influencing the weighting of each study. The combined values and their 95% confidence intervals (CI) were computed and visually represented on funnel plots. Statistical heterogeneity was assessed using the I^2 statistic, categorising values as low (<30%), moderate (30-60%), or high (>60%). Analyses were performed using Stata version 15 (StataCorp LP, College Station, TX). Probability values (p-value) ≤ 0.05 were considered statistically significant.

2.1.3 Results

2.1.3.1 Search Results and Characteristics

A total of 899 articles were initially sourced from database searches. Following the removal of duplicates, 575 records underwent screening based on their titles and abstracts. Subsequently, 141 articles were considered for full-text review. Further assessment, including the application of the exclusion criteria, led to the inclusion of 48 studies in this systematic review and meta-analysis. Figure 12 illustrates the PRISMA flow diagram of screening and selecting the studies.

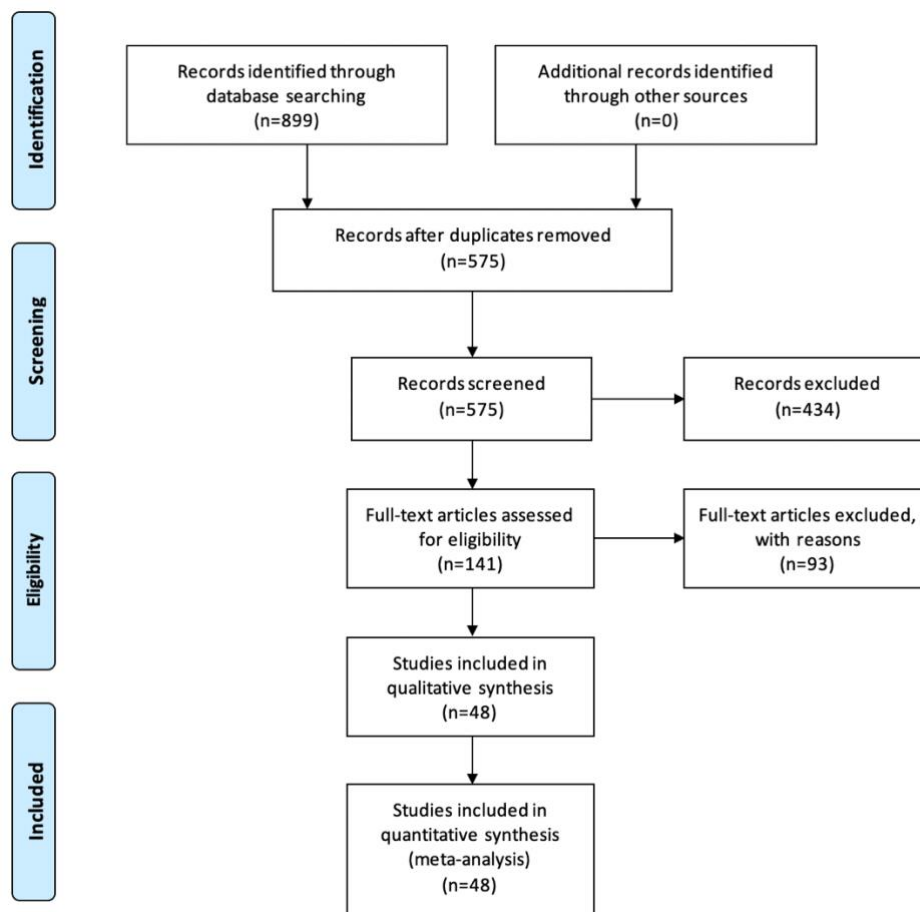


Figure 12 PRISMA flow diagram for study selection. © Reproduced from [194] with permission.

Studies in this systematic review included pre-clinical studies for polyp detection (Table 5), pre-clinical studies for polyp characterisation (Table 6) and recent RCTs (Table 7). All studies were published between 2003 and 2020. The outcome measures included polyp detection in 18 studies, polyp characterisation in 22 studies, and polyp detection rates in eight studies. The studies examining sensitivity, specificity, and accuracy of each AI system reported results on a per-patient, per-polyp, and/or per-image basis, whereas RCTs evaluating the ADR and PDR consistently presented results on a per-patient basis.

Studies focusing on polyp detection primarily employed CNN or DNN as their ML approach. Among the 14 studies on polyp detection conducted retrospectively, the number of images used to train or validate the AI systems varied considerably, ranging from 8 images [86] to 5545 images [195] in different studies.

Table 5 Characteristics of included studies on polyp detection. ©Reproduced from [194] with permission.

Authors	Year	Primary outcome	Recruitment	Machine learning approach	Imaging modality	No. of patients	No. of polyps	Total number of images	Images for training	Images for validation
Karkanis et al [85]	2003	Detection	Retrospective	CWC	RGB- colour frame grabber	66	95	1380	180	1200
Fu et al [196]	2014	Detection	Retrospective	SFFS with SVM classifier	Still image enhanced by PCT	100	N/K	365	292	73
Wang et al [86]	2015	Detection	Retrospective	Polyp edge detection - ECSP	Video clip	N/K	43	N/K	8	53
Tajbakhsh et al (a) [197]	2015	Detection	Retrospective	Hybrid context-shape approach	CVC-Colon DB	N/K	15	300	300	N/K
Tajbakhsh et al (b) [197]	2015	Detection	Retrospective	Hybrid context-shape approach	ASU- Mayo DB	N/K	10	19,400	300	N/K
Fernández-Esparrach et al [88]	2016	Detection	Retrospective	WM-DOVA maps	White light colonoscope	N/K	31	612	N/K	N/K
Park & Sargent [198]	2016	Detection	Retrospective	CNN- CRF model	White light and NBI	N/K	92	11802	N/K	N/K
Urban et al [199]	2018	Detection	Retrospective	Deep CNNs	NBI images	>2000	N/K	8641	N/K	N/K
Wang et al [195]	2018	Detection	Retrospective	Deep learning, built on SegNet Architecture	Still images	2428	N/K	N/K	5545	27113

Misawa et al [200]	2018	Detection	Retrospective	CNN	White light images	73	155	546	411	135
Figueiredo et al [201]	2019	Detection	Retrospective	SVM binary classifier	White light images	42	42	N/K	N/K	N/K
Yamada et al [202]	2019	Detection	Retrospective	Faster R-CNN	Still images	N/K	752	N/K	>4000	4840
Wang et al [82]	2019	Detection	Real-time / prospective	Deep learning, built on SegNet Architecture	Video stream	1058	767	-	-	-

Table 6 Characteristics of included studies on polyp characterisation. ©Reproduced from [194] with permission.

Authors	Year	Primary outcome	Recruitment	Machine learning approach	Image modality	No. of patients	No. of polyps/ lesions	Total number of images	Images for training	Images for validation
Tischendorf et al [203]	2010	Characterisation	Prospective pilot	SVM classifier	Magnification NBI	223	209	N/K	208	N/K
Gross et al [204]	2011	Characterisation	Prospective	SVM classifier	Magnification NBI	214	434	N/K	433	N/K
Ganz et al [205]	2012	Characterisation	Retrospective	Shape-UCM	NBI	N/K	N/K	N/K	58	87
Takemura et al [206]	2012	Characterisation	Retrospective	SVM classifier	Magnification NBI	N/K	371	N/K	1519	371
Mori et al [207]	2015	Characterisation	Retrospective	EC- CAD	EC	152	176	N/K	N/K	N/K
Kominami et al [189]	2016	Characterisation	Retrospective	SVM classifier	Magnification NBI	41	118	N/K	2247	N/K
Misawa et al [208]	2016	Characterisation	Retrospective	EndoBRAIN	NBI & EC	N/K	85	1079	979	100
Mesejo et al [209]	2016	Characterisation	Retrospective	SfM	White light & NBI	N/K	76	N/K	N/K	N/K

Mori et al [210]	2016	Characterisation	Retrospective	SVM classifier	EC-CAD	123	205	N/K	N/K	6051
Takeda et al [211]	2017	Characterisation	Retrospective	SVM classifier	EC-CAD	242	375	5843	5643	200
Byrne et al [212]	2017	Characterisation	Retrospective	DCNN	NBI	N/K	125	N/K	60089	N/K
Komeda et al [213]	2017	Characterisation	Retrospective	CNN-CAD	Endoscopic images	N/K	N/K	1200	N/K	N/K
Misawa et al [214]	2017	Characterisation	Retrospective	EndoBRAIN/ECV-CAD	NBI	100	124	1834	173	1661
Mori et al [215]	2018	Characterisation	Retrospective	N/K	EC	N/K	144	N/K	N/K	N/K
Chen et al [216]	2018	Characterisation	Prospective	DNN- CAD	NBI	193	284	2441	2157	284
Renner et al [217]	2018	Characterisation	Retrospective	DNN	NBI & HDWL	250	231	788	602	186
Mori et al (a) [218]	2018	Characterisation	Prospective	SVM classifier	NBI & EC	325	466	N/K	61925	450
Mori et al (b) [218]	2018	Characterisation	Prospective	SVM classifier	NBI & EC	325	466	N/K	61925	450
Kudo et al (a) [219]	2019	Characterisation	Retrospective	EndoBRAIN system	White light, NBI & EC	89	100	N/K	69142	5065
Kudo et al (b) [219]	2019	Characterisation	Retrospective	EndoBRAIN system	White light, NBI & EC	89	100	N/K	69142	5065
Figueiredo et al [220]	2019	Characterisation	Retrospective	Segmentation algorithm	NBI	10	11	86	43	43

Table 7 Characteristics of RCTs into polyp detection. © Reproduced from [194] with permission.

Authors	Year	Primary outcome	Recruitment	Machine learning approach	Imaging modality	No. of patients	No. of polyps	PDR – AI [%]	PDR – control [%]	ADR – AI [%]	ADR – control [%]	Withdrawal time (excl. biopsy) AI vs. control [mins]
Wang et al [82]	2019	Detection	Real-time / prospective	ANN- SegNet Architecture	Video stream	1058	767	45.02	29.1	29.12	20.34	6.18 vs. 6.07 (p=0.15)
Wang et al [221]	2020	Detection	Prospective	ANN- SegNet Architecture	Video stream	962	809	52	37	34	28	6.48 vs. 6.37 (p=0.14)
Su et al [222]	2020	Detection	Prospective	DCNN	Video stream	623	273	38.31	25.4	28.9	16.5	7.03 vs. 5.68 (p<0.001)
Gong et al [93]	2020	Detection	Prospective	DCNN	Video stream	704	N/K	47	34	16	8	6.38 vs. 4.76 (p<0.0001)
Liu et al [223]	2020	Detection	Prospective	ANN	Video stream	1026	734	43.65	27.81	39.1	23.89	6.82 vs 6.74 (p<0.001)
Luo et al [224]	2020	Detection	Prospective	CNN - YOLO	Video stream	150	185	38.7	34	N/K	N/K	6.22 vs 6.17 (p = 0.102)
Repici et al [94]	2020	Detection	Prospective	CNN - GI GENIUS	Video stream	685	596	N/K	N/K	54.8	40.4	6.95 vs 7.25 (p = 0.1)
Wang et al [225]	2020	Detection	Prospective	ANN - Endoscreener	Video stream	369	N/K	63.59	55.14	42.39	35.68	6.55 vs 6.51 (p = 0.745)

Abbreviations: N/K – not known, CNN – convolutional neural network, CWC – colour wavelet covariance, WM-DOVA - window Median Depth of Valleys Accumulation, DB – database, CRF – conditional random field, NBI – narrow band imaging, DNN – deep neural network, DCNN – deep convolutional neural network, ANN – artificial neural network, EC – endocytoscopy, SFFS - sequential floating forward selection, SVM - support vector machines, SfM – Structure-from-Motion, HDWL – High definition white light, PCT – principal components

In most instances, NBI or endocytoscopy served as the preferred imaging method for characterising polyps, except for one study where the imaging modality was unspecified. [215] Data for polyp characterisation was gathered retrospectively in 18 studies, while in three studies that collected data prospectively, a support vector machine (SVM) classifier was utilised as the ML approach. Similar to studies focusing on polyp detection, those investigating polyp characterisation exhibited substantial variation in the number of images used for training or validating the AI system. However, studies on polyp characterisation focused more on the number of polyps used than overall images, as evident in Table 6.

2.1.3.2 Detection or Localisation of a Polyp

Machine learning systems used for polyp detection were evaluated, in regard to their diagnostic accuracy, using 103,049 still images across 10 studies. The findings indicated a pooled sensitivity of 0.84 (95% CI 0.74-0.93), specificity of 0.87 (95% CI 0.83-0.90), and accuracy of 0.89 (95% CI 0.81-0.97). In 14 studies, diagnostic performance was based on lesions within video frames or images, revealing a sensitivity of 0.92 (95% CI 0.89-0.95), specificity of 0.89 (95% CI 0.84-0.94; see Figure 13), and accuracy of 0.87 (95% CI 0.76-0.97). A set of 11 studies explored the accuracy of polyp detection using images or video clips from over 17,401 patients, resulting in a sensitivity of 0.92 (95% CI 0.90-0.94), specificity of 0.93 (95% CI 0.91-0.96), and accuracy of 0.92 (95% CI 0.87-0.98).

2.1.3.3 Characterisation of a Detected Polyp

Nine studies outlined the diagnostic accuracy of computer analysis for individual image frames. This analysis involved 22,862 images and revealed a sensitivity of 0.92 (95% CI 0.90-0.95; see Figure 14), a specificity of 0.79 (95% CI 0.68-0.91), and an accuracy of 0.87 (95% CI 0.83-0.91). An additional 20 studies investigated the diagnostic accuracy of methods predicting the histological diagnosis of polyps, yielding a sensitivity of 0.94 (95% CI 0.92-0.95), a specificity of 0.87 (95% CI 0.83-0.90), and an accuracy of 0.91 (95% CI 0.88-0.93). In total, 16 studies examined diagnostic accuracy using images or video clips from a cohort of 4,001 patients who had undergone colonoscopy, demonstrating a sensitivity of 0.94 (95% CI 0.92-0.95), a specificity of 0.82 (95% CI 0.73-0.91), and an accuracy of 0.90 (95% CI 0.86-0.94).

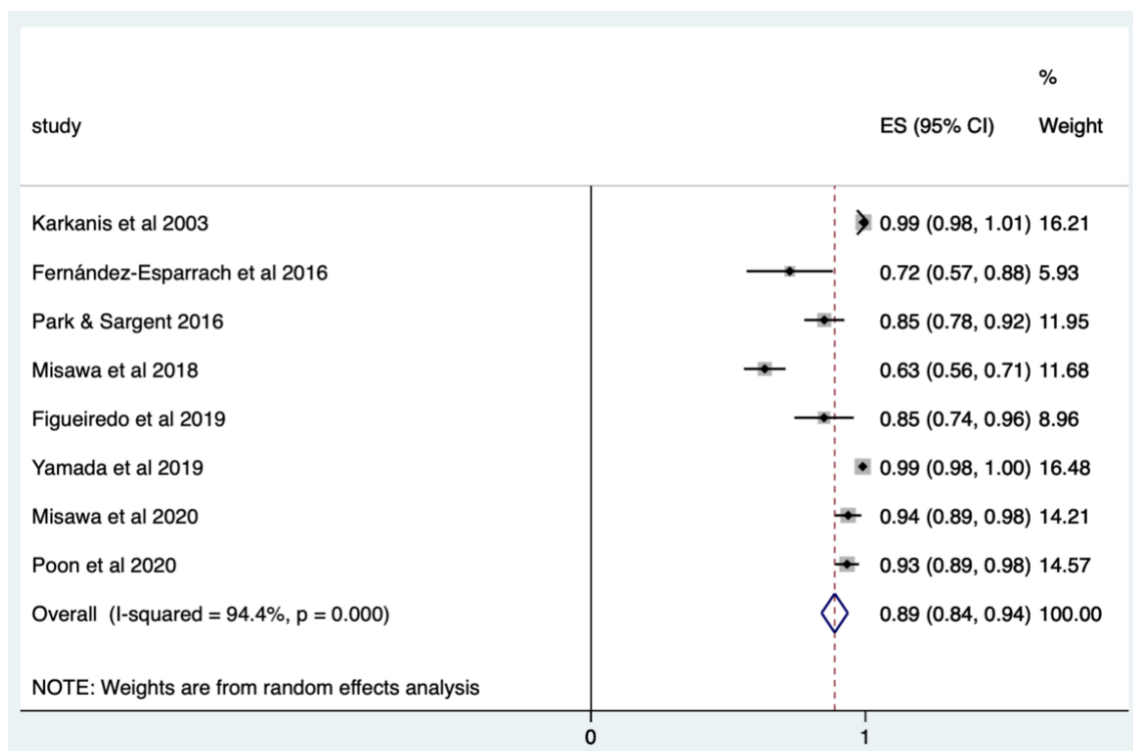


Figure 13 Pooled analysis of specificity of polyp detection by the use of lesions or polyps within video frames or images. Effect sizes (ES) are shown with 95 per cent confidence intervals. A random-effects model was used. © Reproduced from [194] with permission.

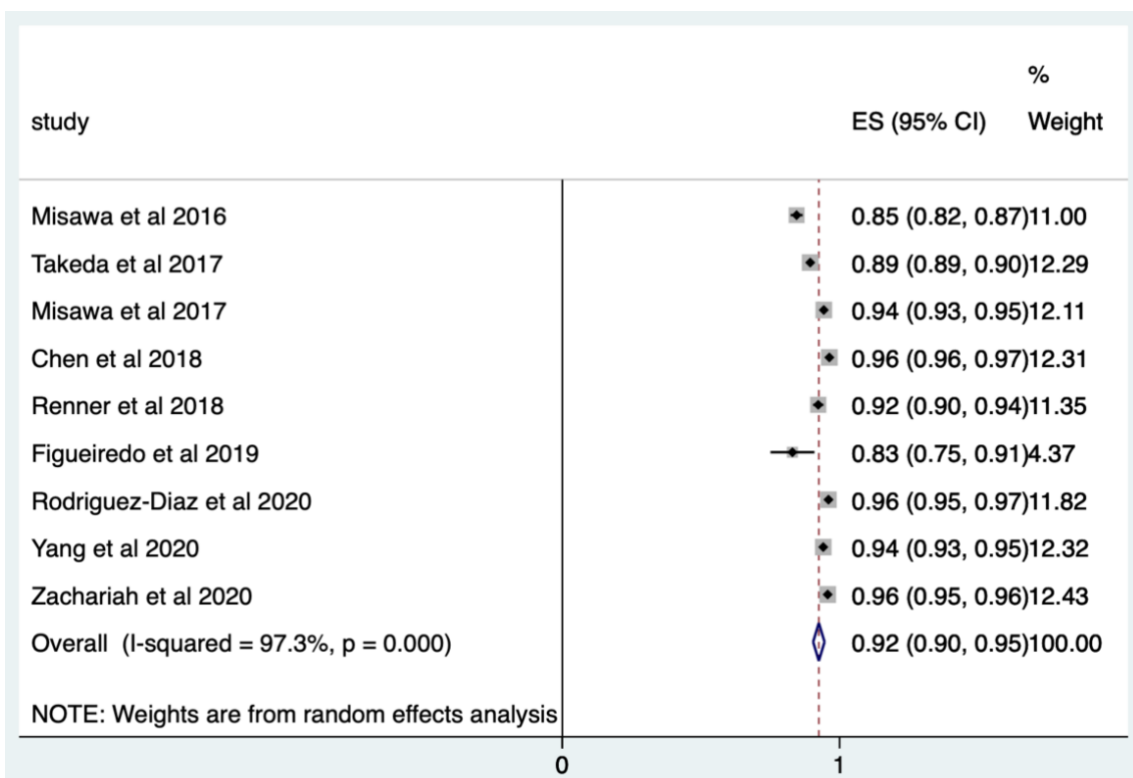


Figure 14 Pooled analysis of sensitivity of polyp characterisation by the use of images. Effect sizes (ES) are shown with 95 per cent confidence intervals. A random-effects model was used. © Reproduced from [194] with permission.

2.1.3.4 PDR and ADR for Polyp Detection (RCTs)

A total of 5,577 patients were included in the eight RCTs assessing PDR and ADR, with 2,438 in the AI group and 2,463 in the control group, undergoing standard colonoscopy. [82,93,221,222] These trials prospectively collected data using deep learning methods on real-time video streams from colonoscopy procedures.

The meta-analysis revealed a significant increase in the pooled PDR when AI was utilised for polyp detection during colonoscopy, compared to patients undergoing standard colonoscopy (OR 1.75; 95% CI 1.56-1.96; $p = 0.0005$; Figure 15). The PDR varied between 38-64%, with a median of 45%, when employing AI during colonoscopy. Furthermore, a significant increase in the ADR was observed when comparing patients undergoing colonoscopy with AI to those with standard colonoscopy (OR 1.53; 95% CI 1.32-1.77; $p = 0.0005$; Figure 16). The ADR ranged between 16-55%, with a median of 34%, when AI technology was employed compared to standard colonoscopy.

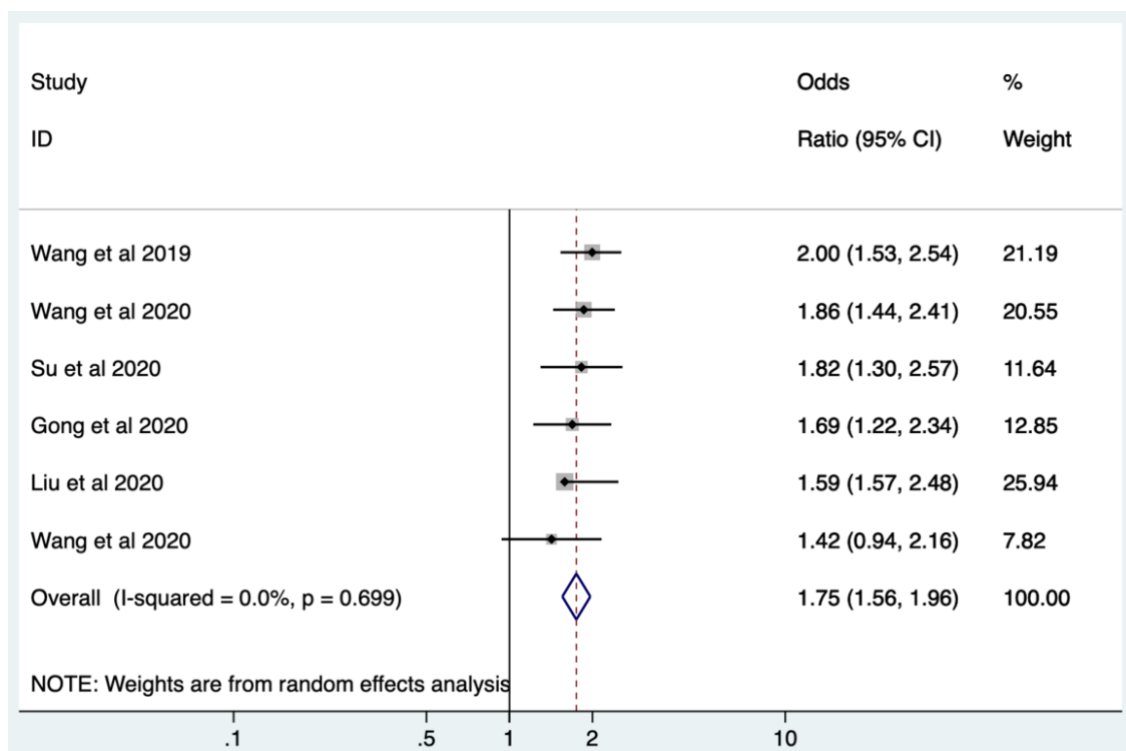


Figure 15 Pooled analysis of polyp detection rate (PDR). Odds ratios (OR) are shown with 95 percent confidence intervals. A random-effects model was used for the meta-analysis. © Reproduced from [194] with permission.

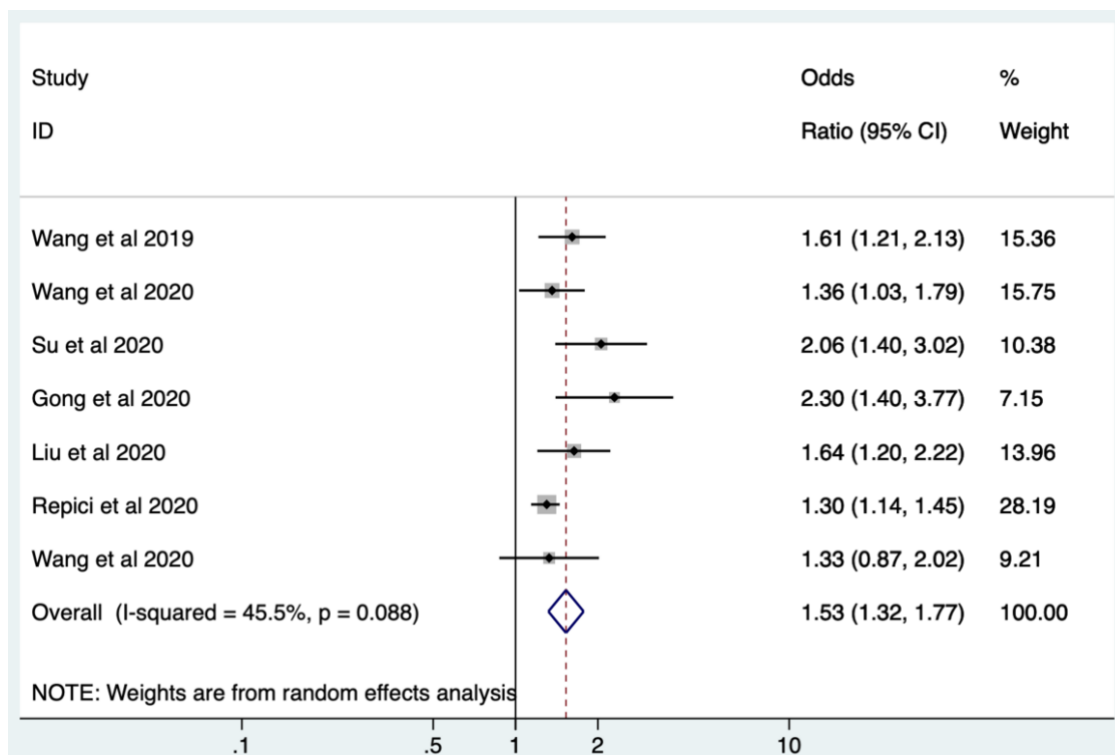


Figure 16 Pooled analysis of adenoma detection rate (ADR). Odds ratios (OR) are shown with 95 per cent confidence intervals. A random-effects model was used for the meta-analysis. © Reproduced from [194] with permission.

2.1.3.5 Heterogeneity of Studies

A high degree of variability was present between studies, with statistically significant heterogeneity observed when comparing studies related to polyp detection and characterisation in terms of sensitivity, specificity, and accuracy ($p < 0.05$). The least variability in polyp detection studies occurred in those evaluating accuracy with polyp data ($I^2 = 86.3\%$), whilst the highest variability was evident in studies analysing the sensitivity of ML systems using image datasets ($I^2 = 99.9\%$). In the context of studies focusing on polyp characterisation, the lowest heterogeneity was observed in those assessing sensitivity using patient datasets ($I^2 = 51.1\%$), and the highest heterogeneity was noted when evaluating specificity using image datasets ($I^2 = 99.9\%$).

When assessing RCTs, a low degree of heterogeneity was identified for the PDR ($I^2 = 0\%$, $p = 0.70$), and a moderate degree of heterogeneity was found for the ADR ($I^2 = 45.5\%$, $p = 0.09$). These results were not statistically significant.

2.1.3.6 Quality Assessment

The evaluation of bias in the studies using the QUADAS-2 tool is illustrated in Appendix 2. The majority of the RCTs collected a score of 3 or higher on the Jadad scale, indicating that their quality was good (Appendix 2; Table 26). One study received a score of 2, suggesting poor quality. However, upon a comprehensive examination of the paper and its evidence, it was included in the final analysis. [226] This decision was made because, despite the absence of explicit mention of blinding, the study's participant selection process was justified, with consecutive patients having been enrolled, and there were no concerns about applicability. The paper adhered to our study's selection criteria and was consistent with other included studies.

2.1.4 Discussion

This systematic review and meta-analysis aimed to evaluate the diagnostic accuracy of AI technologies in detecting and characterising colorectal polyps. Studies employed a wide range of ML systems for these purposes. The overall diagnostic accuracy of AI systems for polyp detection and characterisation was found to be high, with sensitivities, specificities, and accuracies generally exceeding 84% and 82%, respectively. These results highlight the success of current ML systems and algorithms to detect and characterise polyps, and indirectly, regarding the rate of false-positives.

This meta-analysis underscores a significant increase in both PDR and ADR when employing AI systems concurrently with colonoscopy in real-time for identifying colonic and rectal polyps. The overall odds ratios were 1.75 (95% CI 1.56-1.96, $p < 0.05$) for PDR and 1.53 (95% CI 1.32-1.77, $p < 0.05$) for ADR. According to the UK key performance indicators and quality assurance standards for colonoscopy, the recommended minimum ADR is 15%, with an aspirational target of 20%. [23] Previous findings indicate that endoscopists with an ADR below 20% have a hazard ratio for interval cancer ten times higher than those exceeding 20%. [47] All RCTs reviewed demonstrated ADR surpassing 15% when utilising AI systems for polyp detection, with the majority exceeding an ADR of 25%. [82,221,222] These results represent a promising advancement in using AI for the early detection of missed polyps and potentially contributing to a reduction in CRC incidence.

The assessment of quality of the diagnostic accuracy studies in this paper revealed an overall low risk of bias, substantiating the credibility of the study outcomes and suggesting their potential

applicability to clinical practice. The primary source of bias in the RCTs was identified in the blinding process. This may have influenced an overestimation of the effects of AI in polyp detection.

2.1.4.1 Limitations of the Study

The published studies exhibit multiple limitations (Appendix 3; Table 27). The factors contributing to the miss-rate of polyps are diverse, encompassing patient-related, polyp-related, and image-related factors. [227,228] The studies in this review employed a range of imaging modalities, enhancing their potential applicability in a clinical setting. However, it is noteworthy that many studies employing image enhancement techniques predominantly utilised NBI. Therefore, it may be important to validate the effectiveness of AI systems in endoscopy by incorporating image enhancement approaches from alternative manufacturers such as i-scan (Pentax Medical) and blue laser imaging (Fujifilm). Certain studies focussing on polyp characterisation utilised magnification NBI, which is not widely used in Western endoscopic practice, making it less applicable to healthcare settings in the Western world. [189,203,206,229] Despite notable advancements in computer-assisted technologies aimed at increasing the ADR, challenges related to image quality persist. Several studies in this review opted to exclude images that were either blurred or of poor quality when evaluating the diagnostic accuracy of ML systems. [201,208,210,219] Recent RCTs have made attempts to address the issue of image quality by developing models capable of recognising blurry frames. [93,222] Additionally, some studies excluded images with poor bowel preparation. [201,203,216] Although adequate bowel cleansing has been emphasised for thorough mucosal inspection, a meta-analysis has indicated that low-quality preparation might not significantly impact ADR, as these patients often undergo repeat colonoscopy. [230] Most RCTs included in this review relied on the Boston Bowel Preparation Scale to evaluate the adequacy of bowel preparation. [231]

2.1.4.2 Withdrawal times

Adequate withdrawal time allows for thorough mucosal inspection, ensuring a careful examination of all folds and flexures, to minimise the risk of missing any polyps. Research has demonstrated a positive association between increased withdrawal time and higher ADR, supporting the consideration of withdrawal times as a quality indicator for screening colonoscopy. [232] Assessing withdrawal time in pre-clinical studies poses challenges due to the utilisation of still images and video clips. In the evaluated RCTs, withdrawal times (excluding biopsy time) were generally longer when AI-based technology was employed, although statistical significance varied among studies

(Table 7). However, it is the ability to record withdrawal time which is equally as important, suggesting that the use of ML may contribute to maintaining quality control during colonoscopy examinations. [93,222]

2.1.4.3 Lack of Reporting Standards

AI is a relatively recent and evolving domain in medical practice, and as such, the absence of evidence-based standards for its development is apparent. This is underscored by the inconsistencies in validating ML systems across different studies. The diversity in the data used to train algorithms, such as static images from colonoscopies [213,214] or images of polyps, [197,215] along with variations in sample sizes, including very small sizes in some studies, [197,220] contributes to the observed high degree of heterogeneity in the included studies. While this heterogeneity may suggest that our findings are applicable to a broad spectrum of study settings and outcomes, it also underscores the challenge of inconsistencies in AI system development, potentially compromising their design and impeding implementation in clinical settings. In response to this issue, we are actively developing a new, multi-disciplinary, consensus-based reporting standards statement known as STARD-AI. This initiative aims to establish stringent guidelines for all clinical trials involving AI, particularly those reporting diagnostic accuracy. [233,234]

The absence of standardised practices in these studies introduces a form of selection bias. In traditional computer programming, intelligent systems were constructed by manually designing models, allowing for a clear understanding of the rules guiding conclusions. However, neural networks and DL methods face criticism for their "black box" issue, as they often fail to provide a clear description of the generated results. This dichotomy between the need for explanations and the pursuit of efficiency poses a challenge. Many studies in this systematic review did not disclose their algorithms, raising questions about whether they exclusively employed algorithms that yielded desired outcomes without a comprehensive understanding of the underlying processes.

2.1.4.4 Challenges of AI Integration into Healthcare

Various factors contribute to the limited applicability of these studies in clinical practice. Many trials into polyp detection and characterisation have predominantly taken place in Japan or China. [214,218,219][82,93,196] Disparities in polyp biology and tumorigenesis in these regions may restrict generalisability to Western endoscopic practices. [235] Additionally, for real-time detection to be

effective, the operation of the AI system to detect and characterise polyps must be timely, practical, and non-disruptive to workflow. However, most existing studies are undertaken in a non-clinical setting and conducted retrospectively, with only a few recent RCTs. More RCTs are imperative to generate prospective data, evaluating ML systems in the context of ongoing colonoscopy procedures.

The financial considerations associated with the implementation of AI systems in endoscopy must be reviewed. The studies in this systematic review lack sufficient evidence in addressing the cost-effectiveness of AI systems. Prior to clinical implementation, research should highlight the potential savings on healthcare systems and histopathology departments, encompassing both workload and costs. A recent study exploring the integration of AI with the diagnose-and-leave strategy for diminutive polyps has shown promising results, indicating significant cost reductions in colonoscopy based on prospective data. [236] However, additional studies are warranted to further substantiate these findings.

Currently, real-time detection systems in colonoscopy lack the ability to function independently without human guidance. Understanding the evolving roles of endoscopists, nurses and other healthcare workers is essential for the future. Moreover, addressing the skills gap to prepare the workforce for AI is crucial. As ML systems for polyp detection continue to advance, AI is poised to become an integral part of routine colonoscopy procedures. This evolution will enable the procedure to be conducted by staff without the need for extensive training or accreditation. [237] In such a scenario, expert endoscopists may only be required for patients with complex polyps necessitating advanced management.

The purpose of AI in detecting and characterising polyps is to serve as a "checker system" for endoscopists. Consequently, it is advocated to integrate AI into endoscopy as a supplementary tool, not a substitute for clinicians. Ensuring a high level of accuracy is essential for AI systems, ideally operating with 100% sensitivity and a minimal false-positive rate. Despite these expectations, AI is still prone to biases and errors, with the risk of automation bias when endoscopists consistently follow its predictions. [238] Also, ML systems may inadvertently perpetuate biases from their training datasets, exacerbating health disparities. [239] Notably, several studies in this meta-analysis excluded patients with IBD or sessile serrated polyps, limiting applicability to these populations. [82,207,211] Acknowledging the importance of including diverse patient cohorts, including those with benign colonic pathologies beyond polyps, is recognized. However, given the early stage of this

technology, creating validated AI algorithms for all patient cohorts is challenging and not entirely feasible until the technology matures and demonstrates effectiveness across various patient groups.

Despite highlighting satisfactory performance of AI systems in this systematic review, most studies are pre-clinical trials that haven't adequately addressed clinical needs. Consequently, there is a lack of confidence among endoscopists and patients regarding complete adoption of the system. Clinical expectations surpass the objectives of ML algorithms. For seamless integration of an AI system into routine practice, diagnostic accuracy for polyp detection and characterisation must meet desired thresholds whilst instilling confidence in fulfilling quality requirements.

Two additional challenges pose a threat to the successful integration of AI in healthcare: patient confidentiality and accountability. The absence of rigorous policies governing the use of training data in AI results in weak methods for de-identifying patient information. We advocate for standardised guidance to address the need for consent and proper utilisation of patient data in AI training. As algorithm-based healthcare systems become operational, questions of accountability emerge. In scenarios where a ML system, in collaboration with an endoscopist, misidentifies a polyp as hyperplastic instead of adenomatous, determining liability becomes crucial. Establishing a robust legal framework, in collaboration with national and international endoscopy representative groups like the Joint Advisory Group on Gastrointestinal Endoscopy (JAG) in the UK and the ASGE in the US, is essential to safeguard both endoscopists and patients. Addressing these critical concerns is paramount for instilling confidence and trust in the utilisation of ML in healthcare delivery.

2.1.5 Conclusion

This systematic review and meta-analysis underscores the increasing interest in employing AI for polyp detection and characterisation during colonoscopy. The accuracy for ML systems in this role is high, offering the potential to improve ADR, and subsequently reduce CRC incidence. Nevertheless, AI and ML systems are still in a phase of evolution. To advance this field, there is a need for high quality research with contemporary trial designs, emphasising the use of larger datasets and prospective validation of AI systems in clinical settings. Additionally, these systems must establish quality assurance measures within a robust ethical and legal framework before garnering full acceptance from clinicians and patients in the future.

2.2 Does Computer-Aided Diagnostic Endoscopy Improve Detection of Commonly Missed Polyps? A Meta-Analysis.

2.2.1 Background and Aims

Cancer of the colon and rectum typically develops from sporadic adenomatous polyps. More than 90% of adenomas do not progress to cancer. [240–244] However, since it is not possible to reliably identify those that will progress, complete removal of an adenoma is expected in order to eliminate the risk of cancer. [245] National colonoscopy screening programmes have demonstrated effectiveness in lowering the 10-year risk of CRC. [246,247] Despite this, achieving complete mucosal visualisation in a clean colon with a standard colonoscope is not yet possible. This presents several limitations, including the potential of missed polyps.

Studies have estimated the PCCRC rate to account for 2-10% of all CRCs, and most of these cases are attributed to missed or incompletely resected lesions. [248] It has been shown that up to 86% of PCCRC cases are associated with deficient visualisation of the colon or incomplete resection of pre-neoplastic lesions. [249] Regardless of interventions, such as training sessions, or add-on devices, the ability of endoscopists to detect very small, subtle or flat lesions remains a limiting factor. A recent meta-analysis indicated significantly increased miss rates for diminutive ($\leq 5\text{mm}$) and non-pedunculated adenomas compared to larger and pedunculated adenomas. [173]

Non-polypoid (flat or depressed) adenomas have been shown to be more easily overlooked in routine practice, more challenging to resect and a subset of them have been shown to have the potential to progress more rapidly to cancer. [250–252] Therefore, techniques to improve detection of these commonly missed lesions are extremely important.

More recently, CAD systems in endoscopy, based on CNNs, have been validated and made commercially available for use during colonoscopy procedures. This novel technology has been shown to detect polyps in real-time with greater diagnostic accuracy, when compared to standard colonoscopy, and lead to a higher ADR. [194,253] However, despite meta-analyses assessing the impact of CAD systems on ADR in real-time, the ability of these systems to detect the most commonly missed polyps, including diminutive and non-pedunculated lesions, has not been fully determined. [194,254,255]

2.2.1.1 Aims of this Study

The aim of this meta-analysis was to provide an up-to-date analysis of the impact of real-time CAD systems in endoscopy on the detection of the most commonly missed polyps, including non-pedunculated and diminutive lesions.

2.2.2 Method

This review was carried out and reported in accordance with the PRISMA statement.

2.2.2.1 Search Strategy

A thorough review of published articles up to October 2020 was conducted, utilising databases such as EMBASE, Medline, and the Cochrane Library. The search terms employed in EMBASE and Medline encompassed "colon*," "polyp," "adenoma," "artificial intelligence OR machine learning," and "computer aided or assisted and diagnosis* OR detect*." Meanwhile, studies in the Cochrane Library were identified using terms like "colonic polyp," "artificial intelligence," and "diagnosis, computer-assisted" (Appendix 4).

2.2.2.2 Primary and Secondary Outcomes

The primary outcomes for this study were mean adenoma per patient (MAP) and mean polyp per patient (MPP) for diminutive ($\leq 5\text{mm}$) and non-pedunculated (flat or sessile) lesions.

In place of detection rates, mean lesions per patient were calculated in this study. This was due to the data available in the studies analysed. It is important to note that there is good evidence suggesting that mean lesion per patient rates correlate with PDR and ADR. [256]

2.2.2.3 Inclusion and Exclusion Criteria

Inclusion criteria were as follows:

- RCTs reporting CAD of colorectal polyps during real-time colonoscopy.
- Studies reporting mean polyp per patient, mean adenoma per patient; or studies with adequate information to calculate this data.

- Studies reporting the detection rate of polyps or adenomas according to morphology, size and location.
- Studies reporting the ADR or PDR of CAD systems during real-time colonoscopy.
- Studies published or translated into English.

Exclusion criteria were as follows:

- Studies with no original data present (e.g. review article, letter).
- Studies with no full text available.
- Studies conducted in patients with IBD.
- Studies greater than 20 years old.
- Studies without adequate data to calculate PDR and ADR, mean polyp per patient, or mean adenomas per patient, or not reporting this data.
- Studies with overlapping data to the largest and latest data set from centres publishing multiple studies.

2.2.2.4 Study Selection

Two authors, SN and AS, conducted a thorough review of the articles. Duplicate articles were identified and excluded during the initial screening. Titles and abstracts were independently reviewed for relevance, leading to the exclusion of studies unrelated to the research question. The remaining articles underwent a comprehensive review, and reference lists were scrutinised for pertinent studies. Any disagreements in article selection were resolved through consensus between SN and AS, both during the screening process and after a full-text review. All included articles adhered to the established inclusion and exclusion criteria.

2.2.2.5 Data Extraction

Data was gathered from studies onto a standard spreadsheet template. For each study, the following data was extracted: study details (first author, year of publication, journal), study design (type of study, method of artificial intelligence [AI], exclusion criteria), and information regarding polyp characteristics (mean polyp per patient, mean adenoma per patient, morphology, size, location). Data was also gathered for the overall ADR and PDR in each study.

2.2.2.6 Study Quality Assessment

A quality assessment was carried out by independently evaluating each study using the Jadad scale for RCTs. Studies achieving a Jadad score of 3 or higher were classified as being of "good" quality.

2.2.2.7 Statistical Analysis

The mean polyp and adenoma per patient, categorised by morphology and size, were calculated by dividing the total lesions by the corresponding number of patients for each group.

Independent proportion and their differences were calculated and pooled through DerSimonian and Laird random-effects modelling. This considered both between-study and within-study variances which contributed to study weighting. Pooled values and 95% CIs were computed and represented on forest plots. Statistical heterogeneity was determined by the I^2 statistic; where <30% was low, 30-60% was moderate and >60% was high. Analyses were performed using Stata version 15 (StataCorp LP, College Station, TX). A p-value ≤ 0.05 was considered statistically significant.

2.2.3 Results

2.2.3.1 Eligible Studies

From the initial database searches, 899 articles were recognised. Following the elimination of duplicates, a screening process based on titles and abstracts was applied to 575 records. Subsequently, 141 articles were considered for full-text review. After further evaluation and application of exclusion criteria, seven studies were ultimately included in this meta-analysis. The detailed screening and selection process is illustrated in Figure 17.

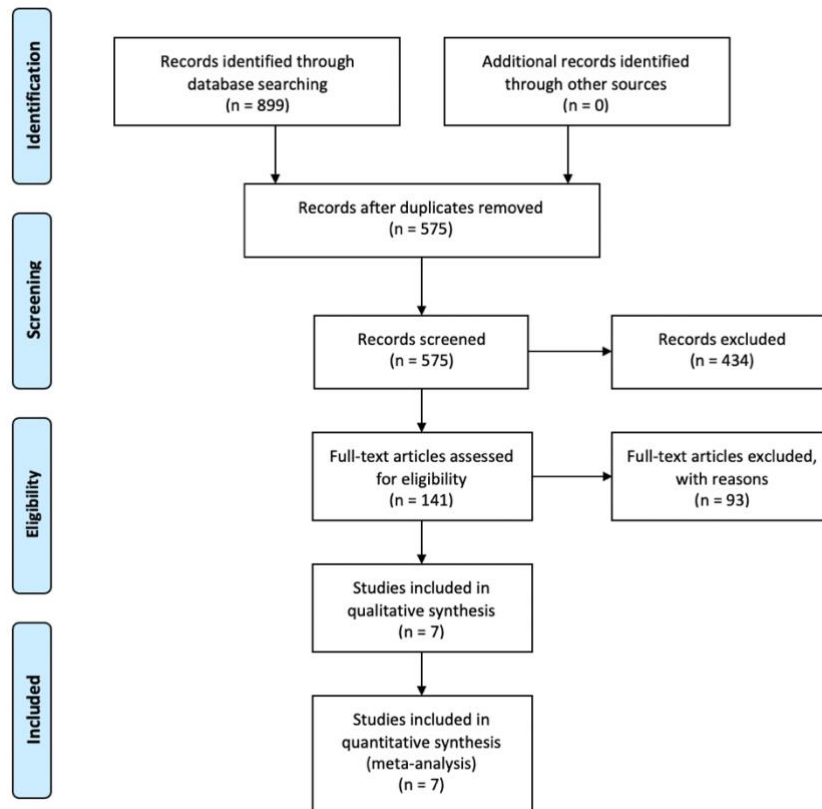


Figure 17 PRISMA flow diagram for study selection in meta-analysis. ©Reproduced from [257] with permission.

2.2.3.2 Study Characteristics

Each of the seven studies included in this meta-analysis utilised CAdE systems in real-time. The analysis of diminutive MPP was available from six studies, and from five studies for diminutive MAP. Four studies evaluated MPP for non-pedunculated polyps, and MAP for non-pedunculated lesions was reported in six studies. One study exclusively analysed polyp detection without reporting on adenoma detection. [224]

Sufficient data was provided in all studies to calculate MPP, MAP, mean diminutive lesions (polyps and adenomas) per patient, and mean non-pedunculated lesions per patient. The studies included a total of 5,028 patients, in which 3,052 polyps and 1,847 adenomas were analysed. Study characteristics are shown in Table 8.

Table 8 Study characteristics of included papers in meta-analysis. ©Reproduced from [257] with permission.

Author and Year	Machine Learning Approach	No. of Patients	No. of Patients - CADe	No. of Patients - Control	Reported Polyps (Total)	Reported Adenomas (Total)	Withdrawal Time (excl. biopsy) CADe vs. Control [mins]
Wang et al 2019 [82]	ANN- SegNet Architecture	1058	522	536	✓ (767)	✓ (424)	6.18 vs. 6.07 (p=0.15)
Wang et al 2020 [221]	ANN- SegNet Architecture	962	484	478	✓ (809)	✓ (462)	6.48 vs. 6.37 (p=0.14)
Su et al 2020 [222]	DCNN	623	308	315	✓ (273)	✓ (169)	7.03 vs. 5.68 (p<0.001)
Gong et al 2020 [93]	DCNN	704	355	349	✓ (284)	✓ (85)	6.38 vs. 4.76 (p<0.0001)
Liu et al 2020 [223]	ANN	1026	508	518	✓ (734)	✓ (394)	6.82 vs 6.74 (p<0.001)
Luo et al 2020 [224]	CNN - YOLO	150	77	82	✓ (185)	X	6.22 vs 6.17 (p = 0.102)
Repici et al 2020 [94]	CNN - GI GENIUS	685	341	344	x	✓ (313)	6.95 vs 7.25 (p = 0.1)

Table 9 Summary of CADe versus control MPP and MAP (diminutive and non-pedunculated), percentage difference and effect size. ©Reproduced from [257] with permission.

		CADe versus Control	Percentage Difference	Effect Size
Diminutive	MPP	0.69 vs 0.37	85.3 %	0.30 (95% CI 0.19 – 0.42)
	MAP	0.31 vs 0.17	80.0 %	0.13 (95% CI 0.09-0.18)
Non-Pedunculated	MPP	0.84 vs 0.43	91.7 %	0.39 (95% CI 0.35-0.44)
	MAP	0.32 vs 0.19	71.5 %	0.05 (95% CI 0.02-0.07)

2.2.3.3 Commonly Missed Lesions

A summary of the results can be found in Table 9.

2.2.3.3.1 Diminutive Lesions

In the CADe group, six studies detected 1,557 diminutive polyps among 2,254 patients, while the control group identified 849 diminutive polyps among 2,278 patients. This resulted in 0.691

diminutive polyps per patient in the CAdE group compared to 0.373 in the control group, with a pooled effect size of 0.30 (95% CI 0.19 – 0.42). This indicates an 85.3% increase in the detection rate of diminutive polyps with CAdE.

In regard to adenomas, five studies in the CAdE group detected 680 diminutive adenomas among 2,177 patients, while the control group identified 381 diminutive adenomas among 2,196 patients. This translated to 0.312 diminutive adenomas per patient in the CAdE group and 0.173 in the conventional colonoscopy group, yielding a pooled effect size of 0.13 (95% CI 0.09-0.18; Figure 18). The CAdE group exhibited an 80.0% increase in the detection of diminutive adenomas compared to conventional colonoscopy.

2.2.3.3.2 Non-Pedunculated Lesions

Across four studies, a total of 1,769 non-pedunculated lesions were identified, with 1,158 in the CAdE group (1,337 patients) and 611 in the control group (1,393 patients). This resulted in 0.841 non-pedunculated polyps per patient in the CAdE group and 0.439 non-pedunculated polyps in the conventional colonoscopy group. The use of CAdE demonstrated a 91.7% increase in the detection of non-pedunculated polyps in the studies. Analysing non-pedunculated lesions in the CAdE versus conventional group generated a pooled effect size of 0.39 (95% CI 0.35-0.44).

In six studies, 818 non-pedunculated adenomas were documented in the CAdE group (2,518 patients) and 477 in the control group (2,540 patients). This resulted in 0.325 non-pedunculated adenomas per patient in the CAdE group compared to 0.189 in the control group, yielding a pooled effect size of 0.05 (95% CI 0.02-0.07; Figure 19). The CAdE system demonstrated a 71.5% increased detection rate of non-pedunculated adenomas compared to conventional colonoscopy.

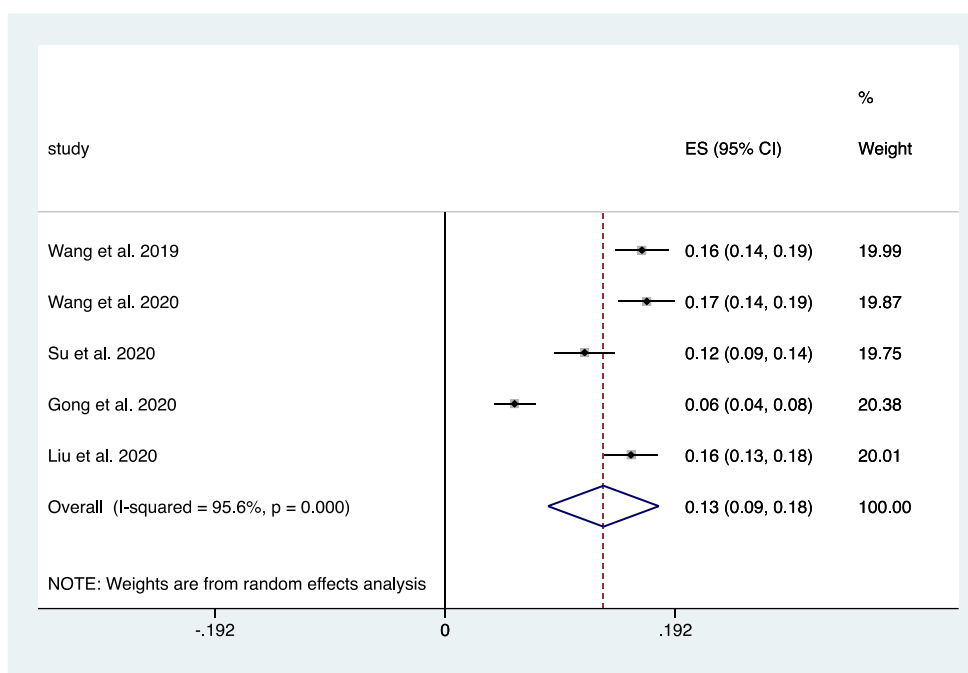


Figure 18 Pooled analysis of mean diminutive adenoma per patient in CAde vs. conventional colonoscopy groups. Effect sizes (ES) are shown with 95 per cent confidence intervals. A random-effects model was used. © Reproduced from [257] with permission.

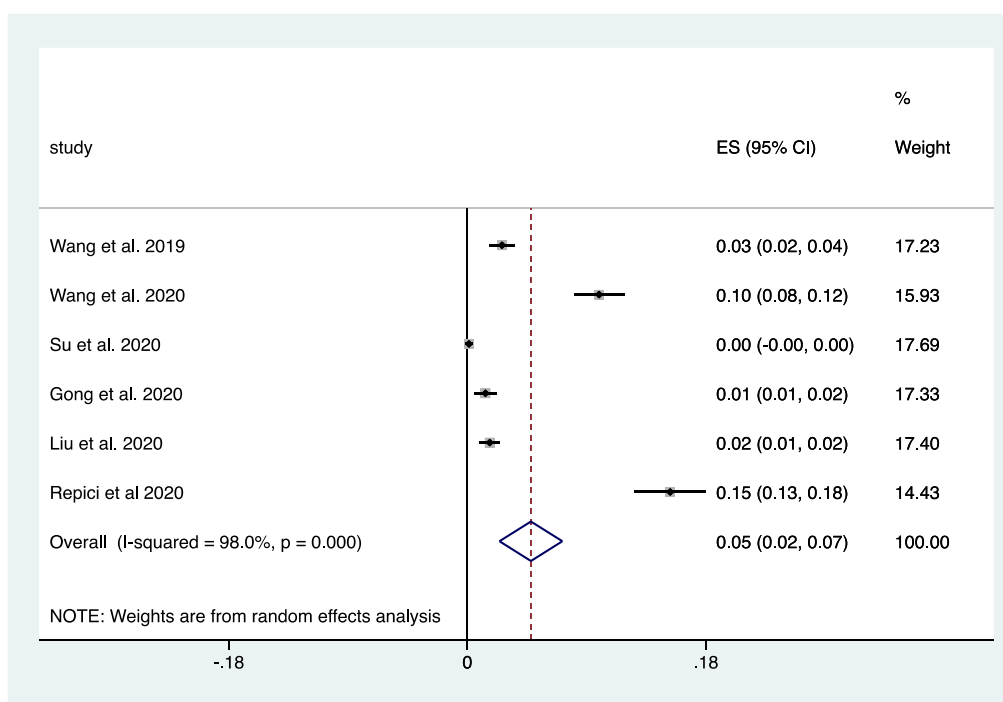


Figure 19 Pooled analysis of mean non-pedunculated adenoma per patient in CAde vs. conventional colonoscopy groups. Effect sizes (ES) are shown with 95 per cent confidence intervals. A random-effects model was used. © Reproduced from [257] with permission.

2.2.3.4 ADR and PDR

In total, six studies evaluated the PDR and ADR for CADe systems compared to standard colonoscopy (Table 10). In all cases, the utilisation of CADe systems for real-time polyp detection demonstrated an improvement in both ADR and PDR. The mean ADR for CADe versus standard colonoscopy was 33.65% and 22.85%, respectively, with CADe ranging from 16% to 54.8% and standard colonoscopy from 8% to 40.4%.

Table 10 Summary of CADe versus control for polyp and adenoma detection. (N/K – not known) © Reproduced from [257] with permission.

Author & Year	PDR – CADe [%]	PDR – control [%]	ADR – CADe [%]	ADR – control [%]
Wang et al 2019 [82]	45.02	29.1	29.12	20.34
Wang et al 2020 [221]	52	37	34	28
Su et al 2020 [222]	38.31	25.4	28.9	16.5
Gong et al 2020 [93]	47	34	16	8
Liu et al 2020 [223]	43.65	27.81	39.1	23.89
Luo et al 2020 [224]	38.7	34	N/K	N/K
Repici et al 2020 [94]	N/K	N/K	54.8	40.4

2.2.3.5 Heterogeneity of Studies

There was a high degree of heterogeneity amongst studies. The heterogeneity was statistically significant when comparing the studies ($p < 0.05$). The variation amongst studies ranged between 95.6% and 98.7%.

2.2.3.6 Quality Assessment

The assessment of bias for the studies using the Jadad scale is shown in Appendix 2; Table 26. Most of the RCTs scored 3 or more on the Jadad scale and were, therefore, considered to be of good quality. One study scored 2, suggesting poor quality.

2.2.4 Discussion

This meta-analysis evaluates the real-time performance of CADe systems in identifying lesions most commonly missed by endoscopists, expanding on prior evidence that these systems improve both ADR and PDR overall. Findings from this review indicate superior performance of CADe systems

compared to standard colonoscopy in the detection of diminutive and non-pedunculated polyps, including adenomas. The CAdE systems exhibited a percentage increase in detection ranging from 72% to 92%, with a more pronounced effect size observed in polyp detection compared to adenoma detection.

2.2.4.1 Withdrawal Times

Overall, variance in withdrawal times noted across the majority of RCTs was not statistically significant with the integration of CAdE systems. This implies that the endoscopy workflow is unlikely to be hugely affected by the integration of a CAdE system, and the duration of the procedure may not be prolonged for the detection of challenging lesions. However, three studies did report a significantly extended withdrawal time for CAdE systems, [93,222,223] making it difficult to assess the genuine influence of CAdE systems on withdrawal time due to the diversity in system types and the relatively early stage of their implementation in real-time healthcare settings.

2.2.4.2 PCCRC Rates

Enhancing adenoma detection is recognised as a pivotal factor in lowering CRC risk, considering that the majority of PCCRCs stem from missed lesions. [258] Therefore, a reasonable assumption can be concluded to state that the significant increase in detection of the most commonly missed adenomas would translate into reduced PCCRC rates.

Furthermore, the increased detection of polyps often encompasses benign lesions like hyperplastic polyps. To enhance the ADR, AI-based technology designed for polyp characterisation during endoscopy is undergoing refinement. Once effective, this technology holds the promise of enabling strategies such as "resect and discard" or "diagnose and leave" for small polyps—a goal that has, thus far, encountered challenges in widespread adoption in routine clinical practice. [259,260]

2.2.4.3 Limitations of Study

There were several limitations in this study which should be considered. Although the CAdE systems in the papers included in this meta-analysis were based on neural networks, the algorithms used for each differed and the type of architecture were not the same. This likely contributed to the large heterogeneity in the studies. While our results provide an overview of the collective impact of CAdE

systems in identifying this specific category of lesions, it's important to note that this approach might diminish and, consequently, underestimate the influence of individual systems. Additionally, variations in endoscopic experience and withdrawal times, which were not standardised across studies, could contribute to the heterogeneity observed in this meta-analysis.

Previous meta-analyses have excluded the paper by Gong et al. as the system used in the study focussed on quality of withdrawal rather than direct lesion detection. [93] Our meta-analysis focuses on the general application of CAdE systems and its impact on lesion detection, and therefore, we included this paper in our analysis. This likely also contributed to the large heterogeneity. Conducting additional real-time studies will enable targeted analyses of various CAdE systems and their specific applications.

Moreover, the predominant setting for the studies was in Chinese institutions, potentially limiting the generalisability of the findings to Western healthcare systems. The range of endoscopist experience varied from 1-3 years in some studies [93] to 4-6 years in others, [222] along with differences in overall procedure numbers. Previous research has established a significant link between years of endoscopy experience and ADR. [51] Experienced endoscopists tend to exhibit higher skill levels, impacting the detection of diminutive or non-pedunculated polyps. It is crucial to explore whether the effectiveness of CAdE systems aligns with the endoscopists' experience levels for cost-effective deployment.

2.2.4.4 Diminutive Lesions

The malignant potential of diminutive adenomas remains uncertain. A study by Vleugels et al. noted that the risk of diminutive lesions containing advanced pathology and cancer was low, with the risk being 5.6% and 0.07%, respectively. [261] However, Kim et al. demonstrated an increased risk of metachronous cancer when three or more non-advanced diminutive adenomas were present, along with advanced pathology. [262] The quantity of diminutive lesions is a factor considered in certain international screening programmes. Utilising optical diagnosis, based on AI systems, has the potential to enhance the overall detection of colonic lesions, and may lead to a 'resect and discard' approach for diminutive pathology.

2.2.5 Conclusion

The use of CAdE systems has the potential to enhance colonoscopy procedures and improve early detection of CRC. This meta-analysis illustrates the ability of CAdE systems to improve detection of commonly missed adenomas, and presents the opportunity to improve the effectiveness of colonoscopy procedures whilst reducing the risk of PCCRC. Nevertheless, the existing evidence supporting the use of CAdE systems in real-time in colonoscopy is limited, necessitating further trials. These trials should encompass comparisons between various systems and varying levels of endoscopic expertise before considering broader integration into routine clinical practice.

Chapter 3 Views of endoscopists on the adoption of artificial intelligence: a cross-sectional questionnaire study

Work from this chapter was published in:

- **The future of endoscopy – what are the thoughts on artificial intelligence?**

Scarlet Nazarian, Hui Fen Koo, Emma Carrington, Ara Darzi, Nisha Patel

Journal of Experimental and Theoretical Artificial Intelligence 2023; doi:

10.1080/0952813X.2023.2178516

3.1 Background and Aims

The integration of AI into the realm of endoscopy represents a significant stride towards enhancing diagnostic precision and therapeutic outcomes in the management of GI lesions. Over the past decade, advancements in ML algorithms, coupled with the availability of vast datasets, have fuelled the development of AI systems tailored for real-time image analysis during endoscopic procedures. These innovations promise to revolutionise the way that colorectal lesions are detected and characterised, providing a potential paradigm shift in the field of gastroenterology.

As demonstrated in previous chapters, potential benefits of AI in endoscopy are manifold. One of the primary advantages lies in the improved detection of abnormalities, particularly in the form of polyps and lesions that may be subtle or easily missed. [92] AI algorithms, when trained on diverse datasets, have demonstrated the ability to differentiate between normal mucosa and pathological features, thereby also improving the characterisation of lesions. [217] Moreover, the integration of AI technologies holds the promise of reducing miss rates, ensuring a more comprehensive evaluation of the GI mucosa and enhancing early detection of lesions, which is crucial for effective intervention and improved patient outcomes. [93,94]

Despite the promising prospects, the assimilation of AI into endoscopic practice has encountered a notable degree of scepticism among physicians and surgeons. This scepticism stems from a complex interplay of factors, including concerns about the reliability of AI algorithms, the potential for over-reliance on automated systems, and the need for continuous validation and refinement of these technologies in real-world clinical settings. The literature lacks studies evaluating the perception of endoscopists about the adoption of this novel technology into clinical practice. As we evaluate these

sentiments, it becomes imperative to explore and address these reservations, acknowledging the critical role of human expertise while navigating the evolving landscape of AI applications in endoscopy.

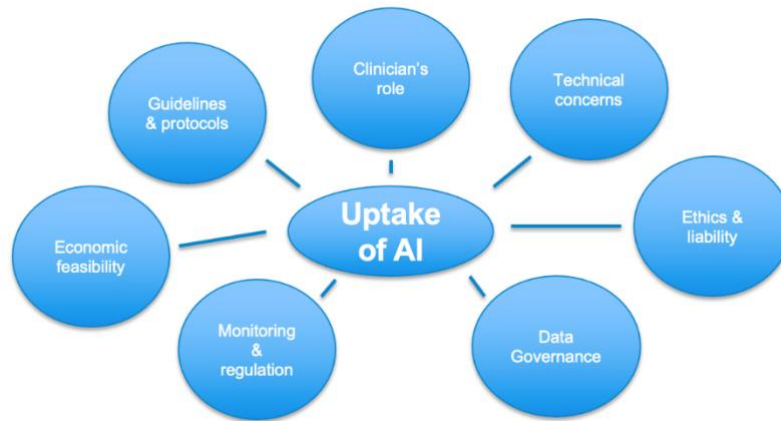


Figure 20 Factors affecting the uptake of AI.

Many limitations of AI systems require addressing before widespread use becomes possible. Figure 20 highlights some of the factors affecting the uptake of AI. Currently, a significant variation exists between AI models and the datasets used for their training, emphasising the need for standardised regulations to be implemented. [263,264] This heterogeneity leads to several challenges, such as potential increase in false-positive results, extended procedural times, increased dependence of endoscopists on the software, and an associated risk in costs. [265] Addressing these concerns is pivotal for establishing a framework that ensures the reliability, efficiency, and cost-effectiveness of AI applications in endoscopy.

The successful integration of any novel technology necessitates the endorsement of both institutional providers, such as hospitals, and the end users within the clinical domain, namely endoscopists. Examining models that originated outside the healthcare context but have achieved widespread adoption within it, provide valuable frameworks encompassing the factors that influence technology acceptance. The Technology Acceptance Model (TAM) and the Unified Theory of Acceptance and Use of Technology (UTAUT) are two such models. [266] The UTAUT, derived from the original TAM, introduces four key variables: perceived usefulness, effort expectancy, social influence (reflecting the perception of the degree to which influential individuals endorse the system's use), and facilitating conditions (pertaining to the perception of organisational and

technical infrastructure supporting the system). These variables collectively influence the intention to use, often considered a surrogate for actual usage. [263,264]

AI endoscopy is still in its infancy but assessment and understanding of the current attitude of endoscopists towards AI in their clinical practice is an important part of the evolution process. Presently, the literature predominantly concentrates on the technical aspects, such as the diagnostic accuracy of AI endoscopy; however, there is a paucity of research examining the qualitative aspects of its integration into the healthcare setting, in particular on its effect on the end-users.

3.1.1 Aims of the Study

The primary objective of this study was to establish the knowledge levels among physicians and surgeons regarding AI, and to gain insights into their perspectives on the benefits and obstacles associated with the integration of AI systems into the clinical setting.

3.2 Methods

3.2.1 Study Design

A cross-sectional study in the form of an anonymous online questionnaire was carried out. The questionnaire was hosted online. Given that there was no existing validated questionnaire available to capture the perspectives of clinicians towards AI in endoscopy, we developed a bespoke survey. The content of the questionnaire was generated through an expert consensus process. Four clinicians (a mixture of gastroenterologists and surgeons) with experience in endoscopy and an academic interest in AI, were consulted. Proposed questions were collated, refined and reviewed until consensus was reached on the final set of questions.

The survey encompassed 18 questions featuring a range of question types, such as polar (yes/no) questions, a 4-point Likert scale, a ranking question, and multiple-response questions. To mitigate acquiescence response bias, questions were phrased in both positive and negative formulations. Participants were asked about their specialty, training level, endoscopy experience, familiarity with AI, and perspectives on AI technology in endoscopy. A full copy of the questionnaire can be found in Appendix 5.

3.2.2 Questionnaire Distribution

All endoscopists practicing in the UK were eligible to participate. Inclusion criteria consisted of endoscopists in surgical or gastroenterology departments, including consultants, trainees, fellows and endoscopy nurses.

Participants were approached to take part in the study through email communication. The online questionnaire was available between March and April 2021. A reminder email was sent 2 weeks after the initial email. Participation was voluntary. The confidentiality of the survey and the anonymity of the respondents were ensured.

The Qualtrics survey software was used for the questionnaire due to its ease of use and the fact that data could be easily aggregated and exported for analysis. The questionnaire was divided into two parts; part 1 gave information about the study with access to the participant information sheet and consent form, and part 2 was the questionnaire itself. Data was gathered onto a standard spreadsheet.

3.2.3 Assessment of Competency

The UK Key Performance Indicators & Quality Assurance Standards[23] were used to evaluate competency amongst endoscopists and minimal ADR. Those with >200 colonoscopies were considered competent, and an ADR of 15% was defined as the minimum.

3.2.4 Statistical Analysis

Descriptive statistics were performed to summarise the questionnaire findings. All statistical analysis was carried out using SPSS version 27 (IMB, Armonk, NY). Chi-squared test (or Fisher's exact test when appropriate) was used to assess statistically significant differences between endoscopist characteristics and views on AI in gastroenterology. Statistical significance was defined as $p < 0.05$.

3.2.5 Ethical Considerations

Full ethical review and approval was obtained through the Science Engineering Technology Research Ethics Committee (SETREC number: 20ICHH6530), and informed consent was obtained from all participants prior to starting the questionnaire.

3.3 Results

3.3.1 Population

A total of 186 endoscopists practicing in the UK were invited to participate in the questionnaire study via email communication. Among them, 98 (52.7%) agreed to participate in the study, and 75 (40.3%) fully completed the questionnaire. Of these respondents, 58.7% held consultant positions, whilst 41.3% were registrars, fellows or nurse endoscopists. Regarding experience, 68% had completed over 200 colonoscopy procedures to achieve competency, and 58.7% demonstrated an ADR exceeding 15%. The majority of participants (60%) were located in the London region. Participant characteristics are presented in Table 11.

Table 11 Participant characteristics. ©Reproduced from [96] with permission.

Region of Work	n	%
North-East England	2	2.67
North-West England	2	2.67
Yorkshire and the Humber	2	2.67
West Midlands	2	2.67
East Midlands	5	6.67
East of England	4	5.33
London	45	60.00
South-East England	9	12.00
South-West England	4	5.33
Role		
Consultant gastroenterologist	30	40.00
Consultant surgeon	14	18.67
Gastroenterology registrar	20	26.67
Surgical registrar	4	5.33
Clinical fellow in gastroenterology	5	6.67
Clinical fellow in surgery	1	1.33
Nurse endoscopist	1	1.33
Consultant	44	58.67
Non-consultant	31	41.33
Level of Colonoscopy Experience		
0-200 procedures	24	32.00
200-500 procedures	12	16.00
500-1000 procedures	4	5.33
More than 1000 procedures	35	46.67
Competent	51	68.00
Non-competent	24	32.00
ADR		
0-15%	7	9.33
15-25%	13	17.33
25-40%	15	20.00
40-60%	11	14.67
>60%	5	6.67
Unsure	24	32.00

3.3.2 Effects of AI on Endoscopy

Responses from all participants in this study are highlighted in Figure 21. The survey findings indicate varying degrees of familiarity among endoscopists regarding AI, with the highest proportion (52.8%) perceiving themselves to be slightly familiar, while 22.6% admitted to having no familiarity with AI. Optimism was observed regarding the anticipated impact of AI on endoscopic practices. The majority of endoscopists expressed positive expectations, with 72.8% foreseeing AI's potential to enhance ADR, 80% anticipating improvements in PDR, 72.8% envisioning elevated endoscopy quality, and 78.6% believing in AI's utility for lesion characterisation.

A significant portion of endoscopists (62.89%) expressed the belief that AI would seamlessly integrate into routine endoscopy practice within the next five years. Notably, apprehensions about AI posing a threat to clinical practice were minimal, with fewer than 3% of respondents viewing AI as a potential threat. Instead, a substantial majority (81.4%) perceived AI as a valuable tool in their clinical toolkit.

However, differing opinions emerged regarding the economic implications of AI adoption, as 55.7% of respondents believed that AI would increase the overall cost of endoscopy. Furthermore, no consensus was reached on the perceived impact of AI on the efficiency of endoscopy practices, with an equal proportion of 37.1% expressing "yes" and "unsure". Similarly, respondents were divided on the necessity of AI in endoscopy, with 34.3% asserting its indispensability and an equivalent percentage negating its essential role, showcasing a nuanced and varied perspective within the endoscopic community.

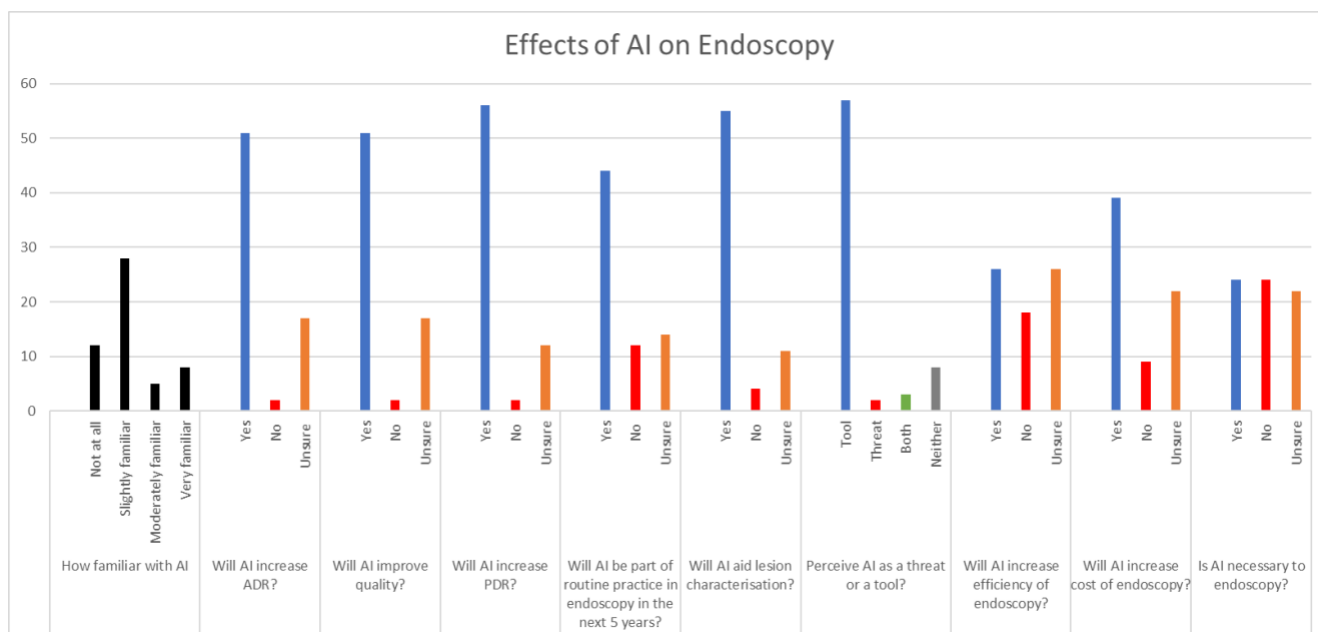


Figure 21 Summary of responses. ©Reproduced from [96] with permission.

3.3.3 Effect of AI on ADR

In the evaluation of endoscopists' perspectives concerning the potential impact of AI on ADR, a thorough analysis of roles, experience levels, and ADR performance was undertaken. The statistical analysis revealed that the experience of endoscopists, quantified by the number of colonoscopies conducted, exerted the most significant influence on their perceptions regarding the efficacy of AI in improving ADR ($p=0.01$). Specifically, among those with fewer than 200 completed colonoscopies, 52.4% expressed a belief that AI could increase ADR, while 47.6% remained uncertain about its potential impact. In contrast, a substantially larger proportion of experienced endoscopists (81.6%) exhibited confidence in the capacity of AI to improve ADR, with only 14.3% expressing uncertainty (Table 12).

Table 12 Comparing role, experience, ADR of endoscopists with views of AI on ADR. ©Reproduced from [96] with permission.

Improvement in ADR with AI?					
		No (%)	Yes (%)	Unsure (%)	P-value
Role	Consultant	2.60	76.90	20.50	0.69
	Non-consultant	3.20	67.70	29.00	
Experience (no. of colonoscopies)	<200	0.00	52.40	47.60	

	>200	4.10	81.60	14.30	0.01
ADR of endoscopist	<15%	0.00	66.70	33.30	
	>15%	2.40	73.80	23.80	
	Unsure	4.50	72.70	22.70	0.95

3.3.4 Sentiments of Endoscopists based on Experience

Upon analysing the perceptions of endoscopists with varying levels of experience, notable distinctions emerged in their attitudes towards the potential impact of AI on endoscopic quality and outcomes. Particularly, endoscopists with greater experience exhibited a higher likelihood of believing in the positive influence of AI on the quality of endoscopy (83.7% vs. 47.6%; $p=0.005$) and its capacity to PDR (87.8% vs. 61.9%; $p=0.008$).

Conversely, when comparing the sentiments of endoscopists with higher and lower experience levels, no statistically significant differences were discerned regarding the perceived impact of AI on efficiency ($p=0.085$), cost ($p=0.690$), or necessity ($p=0.068$) (Table 13).

Table 13 Sentiments of endoscopists based on experience. © Reproduced from [96] with permission.

		Competence (no. of colonoscopies)		p-value
		<200	>200	
Improve quality	No	2 (9.5%)	3 (6.1%)	0.005
	Yes	10 (47.6%)	41 (83.7%)	
	Unsure	9 (42.9%)	5 (10.2%)	
Increase PDR	No	0 (0)	2 (4.1%)	0.008
	Yes	13 (61.9%)	43 (87.8%)	
	Unsure	8 (38.1%)	4 (8.2%)	
Increase efficiency	No	2 (9.5%)	16 (32.7%)	0.085
	Yes	8 (38.1%)	18 (36.7%)	
	Unsure	11 (52.4%)	15 (30.6%)	
Increase cost	No	2 (9.5%)	7 (14.3%)	0.690
	Yes	11 (52.4%)	28 (57.1%)	
	Unsure	8 (38.1%)	14 (28.6%)	
Routine in 5 years	No	4 (19%)	8 (16.3%)	
	Yes	10 (47.6%)	34 (69.4%)	

	Unsure	7 (33.3%)	7 (14.3%)	0.147
Necessary	No	9 (42.9%)	15 (30.6%)	
	Yes	3 (14.3%)	21 (42.9%)	
	Unsure	9 (42.9%)	13 (26.5%)	0.068
Lesion characterisation	No	1 (4.8%)	3 (6.1%)	
	Yes	14 (66.7%)	41 (83.7%)	
	Unsure	6 (28.6%)	5 (10.2%)	0.154
Threat/Tool	Threat	2 (9.5%)	0 (0)	
	Tool	15 (71.4%)	42 (85.7%)	
	Both	0 (0)	3 (6.1%)	
	Neither	4 (19%)	4 (8.2%)	0.049

3.3.5 Benefits and Barriers of AI in Endoscopy

In exploring the perspectives of participants regarding the benefits and barriers associated with the integration of AI in endoscopy, notable patterns emerged. The participants identified the three primary advantages of AI in endoscopy as its potential to improve ADR, improve PDR, and instil greater confidence during procedural performance. Conversely, key impediments hindering the widespread adoption of AI in endoscopy were perceived to be the associated costs, issues related to accessibility, and a notable absence of clear guidelines. Strikingly, when participants were tasked with ranking barriers, the ethical implications of AI in endoscopy emerged as the least prominent barrier among the other factors (Figure 22 and Figure 23).

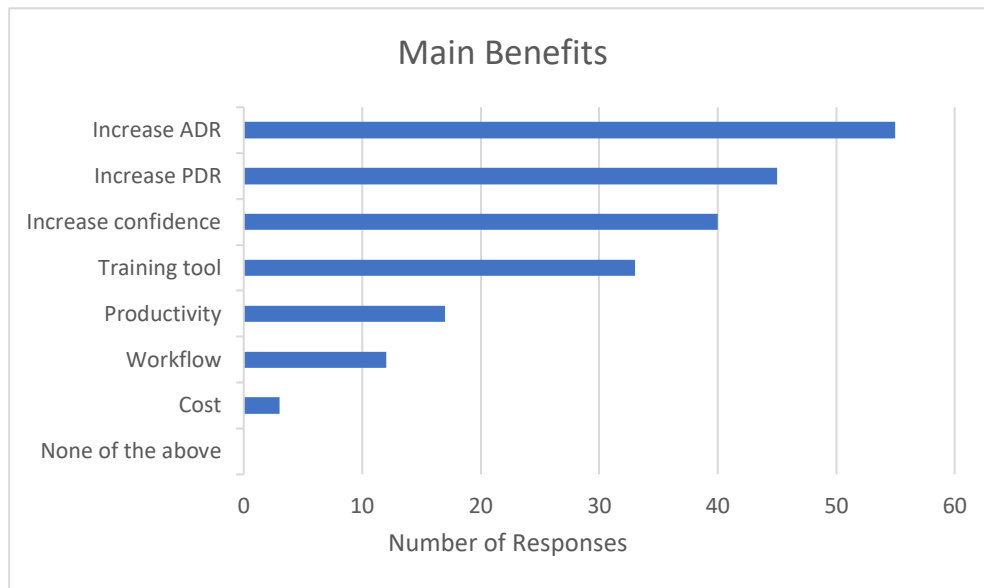


Figure 22 Key benefits of adoption of AI in endoscopy. © Reproduced from [96] with permission.

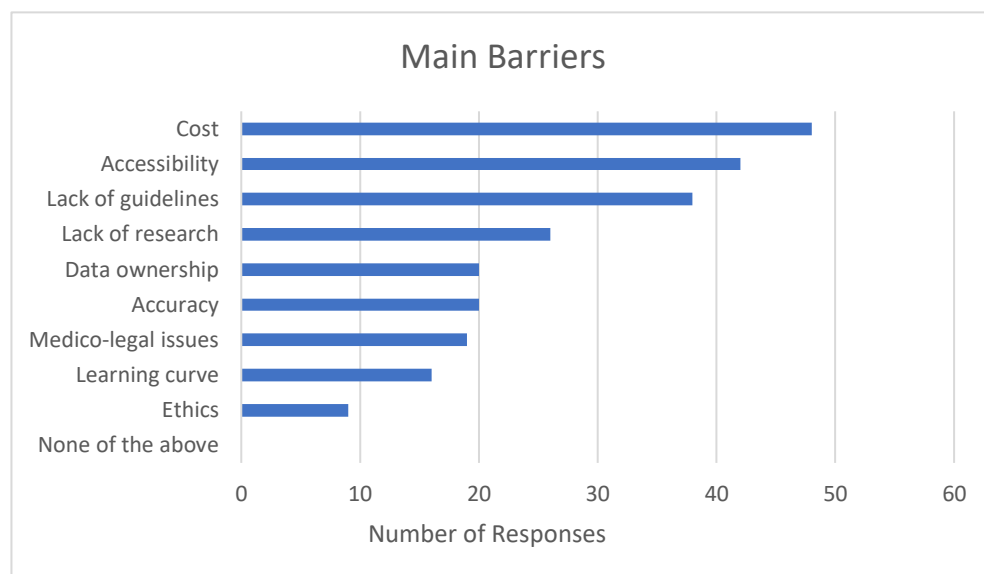


Figure 23 Key barriers of adoption of AI in endoscopy. © Reproduced from [96] with permission.

3.3.6 Application of AI in Endoscopy

Participants' perspectives on the optimal applications of AI in endoscopy were revealed through their responses. A large proportion of responders (36.7%) considered the most promising utilisation of AI to be in *in-vivo* lesion detection and characterisation, closely followed by its application in assessing the quality of endoscopy, as indicated by 27.8% of the participants (Figure 24).

Furthermore, when prioritising the areas deemed most crucial for AI integration, endoscopists

underscored video capsule endoscopy as the highest priority, followed in descending order by hepatobiliary procedures (e.g. ERCP), upper GI examinations, endoscopic ultrasound, and lastly, lower GI procedures (Figure 25).

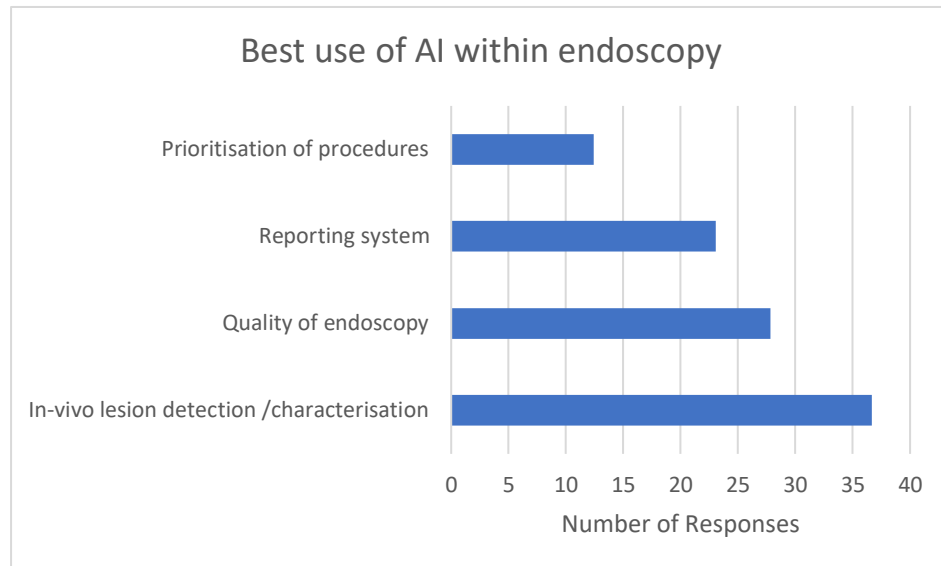


Figure 24 Best utilisation of AI within endoscopy. © Reproduced from [96] with permission.



Figure 25 Priority application of AI within endoscopy. © Reproduced from [96] with permission.

3.4 Discussion

The integration of AI into endoscopic practice has already begun, promising enhanced diagnostic precision and procedural outcomes. This study delves into the nuanced perspectives of endoscopists, providing valuable insights into their views on the prospective role of AI within the field of endoscopy. The results highlight that the majority of endoscopists felt that the adoption of AI in endoscopy would increase the ADR (72.8%), PDR (80%) and improve the quality of the endoscopy

(72.8%). Endoscopists also believed it would have a positive impact on their practice and perceived AI as a tool rather than a threat.

In this study, endoscopists with greater experience demonstrated a higher likelihood to endorse the positive impact of AI on the quality of endoscopy. Notably, a substantial majority of experienced endoscopists, comprising 81.6%, expressed confidence in the capability of AI to improve ADR. This nuanced examination underscores the significance of experience as a pivotal factor informing and influencing endoscopists' perceptions regarding the role of AI in ADR enhancement.

The optimism among endoscopists regarding the adoption of AI, as evidenced by their anticipation of improvements in ADR, PDR, and overall endoscopy quality, underscores the growing recognition of the potential impact of AI. However, a pivotal observation surfaces from this optimism—endoscopists, to a considerable extent, exhibit a gap in their understanding of AI systems. This knowledge gap serves as a crucial backdrop, potentially contributing to the lack of consensus on the necessity of AI in endoscopy. The revelation that over 75% of respondents indicated only slight or no familiarity with AI shows an important need for targeted education and training initiatives within the endoscopic community.

Despite the promising outlook, the trajectory from the demonstrated positive impact of AI on polyp detection, [194] to the seamless incorporation of these systems into clinical practice remains a challenge. The rapid development of commercial AI devices by endoscopy manufacturers outpaces the empirical scrutiny of real-time efficacy in healthcare settings, emphasising the critical need for robust evidence from large RCTs. [95,267] These trials, coupled with evidence-based guidelines, are essential for instilling genuine confidence in the clinical utility of AI.

Lesion characterisation emerges as a focal point in the envisaged AI-driven paradigm, with nearly 80% of endoscopists foreseeing the aid of AI in this aspect. The prospect of an AI-enabled "resect and discard" strategy for lesions, effectively reducing the need for additional resources, aligns synergistically with initiatives like the PIVI set forth by the ASGE. [187,268] However, the ethical and medico-legal considerations surrounding such strategies necessitate meticulous deliberation and clear guidelines to navigate these complex areas.

The financial implications of the adoption of AI technology in endoscopy is still undetermined. Endoscopists in this study expressed associated costs as the main barrier to adoption of AI in

endoscopy. While recent cost-effectiveness analyses provide compelling support for the potential reduction in polypectomy-related costs through AI implementation, the overarching economic implications of AI-driven polyp detection systems remain unknown. [236,269] The cost of a missed polyp and a diagnosis of a PCCRC may completely outweigh the cost of an AI system able to accurately detect adenomas, but without studies to show this, clinicians and healthcare systems cannot allocate their funding based on such assumptions. [270] In order to persuade essential stakeholders to accept AI and provide reimbursement for its use, more studies based on large clinical trials are required to highlight the cost-benefit related outcomes of this technology.

This study has many limitations that should be considered. Firstly, the cross-sectional design of this study introduces susceptibility to misclassification, primarily attributed to recall bias concerning the experiences of endoscopists with AI. The inherent nature of cross-sectional studies, reliant on the participants' ability to accurately recall past events, may introduce potential inaccuracies in their responses. However, it is important to note that the study's questionnaire strategically focussed on gauging the prevailing sentiments and perspectives of endoscopists, rather than relying on the precise recollection of specific events. This deliberate approach aims to mitigate the impact of recall bias to a certain extent.

Secondly, a noteworthy limitation stems from the observed low response rate among participants. While more than half of the initially invited endoscopists expressed consent to participate, only 40% completed the comprehensive questionnaire. This lower-than-desired response rate could potentially be attributed to the demanding and time-constrained schedules inherent in the practice of endoscopy. To enhance the response rate, alternative data collection methods such as telephone or face-to-face interviews could have been explored. These approaches may have facilitated a more engaged participation, allowing for a more in-depth exploration of the nuanced perspectives held by endoscopists regarding AI integration in their clinical practices.

Moreover, the study's questions lacked an open-ended format, restricting participants from elaborating on their thoughts or exploring tangential topics not directly related to the primary question. This limitation could have led to an incomplete understanding of endoscopists' views and potentially skewed results. However, it is essential to note that open-ended questions, while fostering depth, pose challenges in terms of analysis and data quality. Furthermore, the study's reliance on descriptive statistics rather than inferential statistics in the analysis represents an additional limitation. While descriptive statistics provide a succinct summary of the collected data,

they fall short of offering detailed insights into the strength of relationships between factors, such as endoscopist experience and their opinions on the anticipated impact of AI on ADR and PDR. The adoption of inferential statistics would have allowed for a more nuanced exploration of these relationships, enriching the depth of the study's findings.

Lastly, this was a national study, taking into consideration the views of endoscopists in the UK. This introduces a potential limitation since stakeholders in the UK largely work in a public healthcare setting, and may have differing views and priorities to those in other countries, particularly those operating within private healthcare systems. A comparative study conducted in the US revealed substantial interest among gastroenterologists in the application of AI to colonoscopy. [271] However, this was once again a national study. Thus, a more expansive international questionnaire study is needed for a broader perspective to ensure a comprehensive understanding of global views on AI adoption in endoscopy, particularly the priorities for stakeholders, across diverse healthcare landscapes.

3.5 Conclusion

In view of the growing interest and advancements in AI systems within the realm of endoscopy, a comprehensive examination of stakeholder and endoscopist perspectives is imperative. This study highlights that endoscopists harbour optimistic views regarding the transformative impact of AI systems on endoscopy, categorising the technology as a valuable tool rather than a perceived threat to their clinical practice. Concurrently, the study shows significant challenges and barriers inherent in the application of AI in endoscopy, demonstrating the necessity for their systematic resolution to achieve broader acceptance among end-users and, subsequently, solidify patient trust. Despite the promising outlook, an absence of large randomised clinical trials persists, which are crucial to establish evidence-based guidelines and fully understand the nuanced cost-benefit outcomes in this rapidly growing technological paradigm. This underscores the importance for future research endeavours to delve into the intricacies of AI integration in endoscopy, paving the way for a more informed and resilient adoption of this novel technology within the clinical landscape.

Chapter 4 Improving detection of tumour tissue using diffuse reflectance spectroscopy to aid resection margin assessment in lower GI cancer surgery: an *ex vivo* validation study

Work from this chapter was published in:

- **Real-time classification of tumour and non-tumour tissue in colorectal cancer using diffuse reflectance spectroscopy and neural networks to aid margin assessment.**
Scarlet Nazarian, Ioannis Gkouzionis, Jamie Murphy, Ara Darzi, Nisha Patel, Christopher J. Peters, Daniel S. Elson
International Journal of Surgery 2024; doi: 10.1097/JS9.0000000000001102
- **Real-time Tracking and Classification of Tumor and Nontumor Tissue in Upper Gastrointestinal Cancers Using Diffuse Reflectance Spectroscopy for Resection Margin Assessment.**
Scarlet Nazarian, Ioannis Gkouzionis, Michal Kawka, Marta Jamrozak, Josephine Lloyd, Ara Darzi, Nisha Patel, Daniel S. Elson, Christopher J. Peters
JAMA Surg. 2022; doi: 10.1001/jamasurg.2022.3899
- **Using Diffuse Reflectance Spectroscopy to Classify Tumor Tissue in Upper Gastrointestinal Cancers—Reply.**
Scarlet Nazarian, Ioannis Gkouzionis, Christopher J. Peters
JAMA Surg. 2023; doi:10.1001/jamasurg.2022.8433
- **Real-time tracking of a diffuse reflectance spectroscopy probe used to aid histological validation of margin assessment in upper gastrointestinal cancer resection surgery.**
Ioannis Gkouzionis, Scarlet Nazarian, Michal Kawka, Ara Darzi, Nisha Patel, Christopher J. Peters, Daniel S. Elson
Journal of Biomedical Optics 2022; doi: 10.1117/1.JBO.27.2.025001

4.1 Background and Aims

The definitive treatment for CRC is surgical resection of the tumour. The aim of surgery is to resect all cancerous tissue, both on a macroscopic and microscopic level. In doing so, a R0 resection should be achieved, which refers to residual tumour at the primary site after treatment. The Association of Coloproctology of Great Britain & Ireland (ACPGBI) recommend that a R0 resection is achieved in

>90% of colorectal cancers predicted to be resectable on appropriate staging. [272] However, the need for maximal resection should be balanced with the aim of preserving as much healthy tissue as possible, in order to minimise patient morbidity and functional impairment. In this way, the accurate mapping of tumour margins is key, particularly given that positive circumferential and longitudinal resection margins following colorectal cancer surgery have been shown to significantly increase rates of recurrence and reduce long-term survival. [151,273]

4.1.1 Intra-Operative Challenges of Margin Assessment

The oncological outcomes of patients who undergo curative colectomy are greatly influenced by the quality of surgery, and the meticulous attention to resection margins. [274] Currently, there is a lack of consensus on the optimal length of bowel that should be resected in colon cancer surgery to achieve negative resection margins. The American Society of Colon and Rectal Surgeons (ASCRS) recommends that colon resection includes proximal and distal margins of 5–7 cm to ensure the adequate removal of pericolic nodes at risk of metastasis. [275] However, other studies have demonstrated that pericolic lymph nodes as far as 8-10 cm from the tumour may harbour tumour metastases. [276]

Patients with CRC who have undergone neo-adjuvant therapy pose another challenge in the assessment of resection margins intra-operatively. The administration of pre-operative treatments, while invaluable in shrinking tumours and potentially improving surgical outcomes, introduce complexities that impact the tactile and visual feedback crucial for effective tumour margin assessment. Neo-adjuvant therapy often induces alterations in tissue consistency and vascularity, leading to the formation of scar tissue that can complicate the surgeon's ability to discern the tumour boundaries by touch alone. The conventional reliance on tactile feedback for identifying the transition from malignant to healthy tissue becomes complex, potentially increasing the risk of incomplete resection. [277,278]

These challenges emphasise the pressing need for innovative intra-operative techniques that can augment surgeons' capabilities, providing real-time, objective information to improve tumour margin assessment and ultimately contribute to improved oncological outcomes, particularly in patients who have undergone neo-adjuvant therapy.

4.1.2 Intra-Operative Margin Mapping Techniques

Frozen section analysis remains the gold standard technique for intra-operative tumour margin assessment. This technique has been shown to have a diagnostic accuracy of 93%, a sensitivity of 67%, and a specificity of 100% in gastric and oesophageal adenocarcinoma.[279] However, FSA is limited by many factors. The technique often involves finite sampling, introducing the risk of sampling error and inadequate representation of the entire tumour. In addition, the time required for processing frozen sections may prolong the duration of surgery, potentially compromising the timely decision-making essential for optimal oncological outcomes. Furthermore, the cost implications associated with the labour-intensive process and the need for specialised personnel contribute to the overall constraints of FSA as an ideal method for intra-operative tumour margin assessment in real-time. [280]

4.1.2.1 Advanced Imaging and Spectroscopic Techniques

Many studies have therefore focussed on developing advanced imaging and spectroscopic techniques for improved tumour visualisation and margin delineation. [281–283] Fluorescence-guided surgery (FGS) utilises fluorescent markers, such as indocyanine green, to selectively highlight tumour tissue under near-infrared light, enabling real-time visual guidance. [284] Whilst this approach enhances intraoperative visualisation, challenges include the lack of binding of the fluorescent agents specifically to tumour cells, compromising tumour specificity, and the limited penetration depth of fluorescence. [285] In addition, quantifiable fluorescence imaging information is difficult to obtain using open-field devices due to uncontrollable parameters, such as variable light intensity, working distance and influence of ambient light on the instruments in the operating room. [286]

Rapid evaporative ionisation mass spectrometry (REIMS) is another imaging technique which allows characterisation of tissue *in vivo* by rapid analysis of the aerosol (“smoke”) released during electrosurgical dissection, providing molecular information for accurate tumour identification. [287] Nevertheless, this technique is invasive and faces challenges in distinguishing between normal and cancerous tissue with subtle molecular differences. In addition, the REIMS technique is highly complex, requiring specialised equipment and skilled operators which adds significant cost and makes widespread clinical adoption challenging.

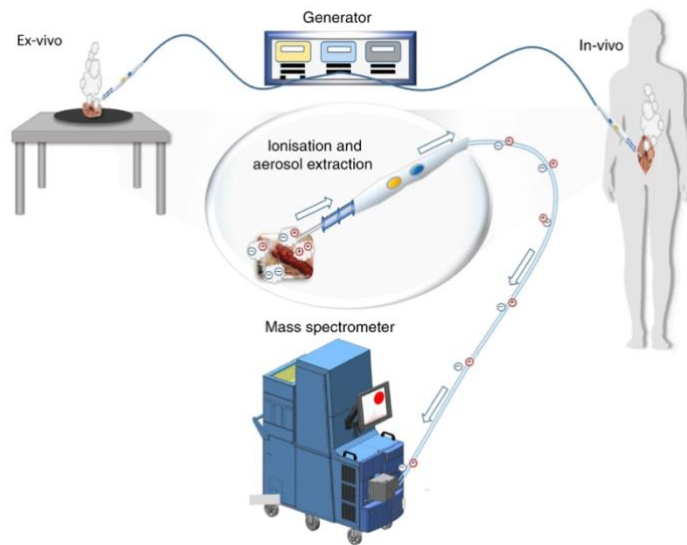


Figure 26 REIMS set-up ex-vivo and in operating theatres (in-vivo). Electrical current is applied to the tissue and the resultant charged particles are extracted through the custom-designed hand piece and drawn into the mass spectrometer to produce tissue-specific mass spectra. Image adapted from [374].

Optical coherence tomography (OCT) provides high-resolution, real-time imaging of tissue microstructure. [288] While OCT excels in visualising surface layers, its penetration depth is limited, hindering comprehensive assessment of deeper tumour margins. [289]

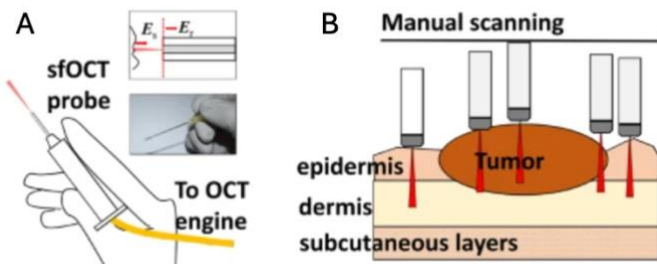


Figure 27 Diagram to represent OCT. (A) Handheld single fibre (sf) OCT probe (upper inset: reference light derives from the probe tip; lower inset: a photo of the probe held by hand); (B) sfOCT probe is manually scanned across a region of interest to generate a 2D image. Image adapted from [375].

These advanced techniques represent promising avenues, however, they have either failed to or have not yet translated into routine clinical practice. This is because they either cannot provide real-time data, do not adapt into the surgical workflow, are not cost-effective, or have been designed as bench-top adjuncts only. [290–294]

4.1.2.2 Optical Spectroscopy Techniques

These challenges can potentially be addressed by utilising optical spectroscopy techniques, which have been shown to exhibit a high sensitivity and specificity (>90%) for discriminating between normal and tumour tissue, such as Raman spectroscopy and DRS. [295] These methods allow the delivery of light in a highly localised manner to tissue via optical fibres. Analysis of light interactions with the tissue provide insight into the metabolic and pathological state of tissue. [296] A significant advantage of optical spectroscopy techniques is that they do not rely on pigments, dyes, or external fluorophores. This minimises the requirement for supplementary agents, streamlining the detection process.

Additionally, optical spectroscopy methods have the potential to visualise tissue and cellular changes at a micro-scale that is otherwise invisible to the naked eye. This ability enables detection of early-stage dysplasia and cell transformations without the need for costly microscopic imaging devices. In this way, optical spectroscopy has potential to advance early detection of cancers.

Raman spectroscopy, in particular, has enabled highly accurate *in vivo* diagnosis of a range of cancers including breast, skin and colon. [281,297,298] Although Raman spectroscopy has shown much promise in the accurate diagnosis of cancerous tissue, its application for tumour margin delineation is limited. This is due to the infrequency of Raman scattering events, which, despite being faster in single-point imaging compared to DRS, still necessitate long acquisition times to cover larger tissue areas. Furthermore, Raman spectroscopy is more complex and costly than DRS, limiting its accessibility. In addition, Raman spectroscopy lacks accuracy in depth perception for both probe tracking and margin delineation. [299]

4.1.2.3 Use of DRS for Tissue Discrimination

Diffuse reflectance spectroscopy is a point-based spectroscopy technique which holds considerable promise in cancer margin assessment due to its ability to provide real-time, in-situ information about tissue composition and structure. DRS operates by measuring the scattering and absorption of light within tissues, offering valuable insights into the optical properties of different cellular and subcellular components. This technique facilitates the discrimination between normal and tumour tissue based on variations in their optical signatures, present at both early and late stages of malignant transformation. [300]

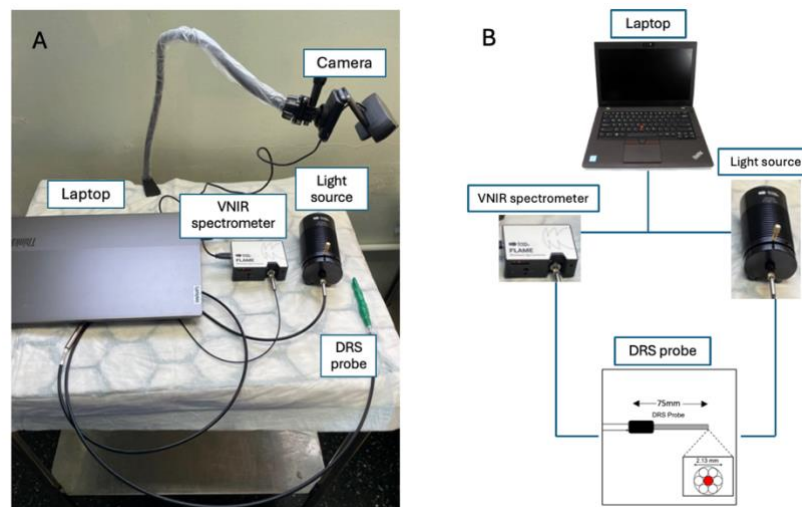


Figure 28 Photo (A) and schematic diagram (B) of DRS set-up showing laptop connected to spectrometer (Ocean Optics, USB4000, Largo, USA) and light source (Ocean Optics, HL-2000-HP, Largo, USA), both of which connect to DRS probe.

What makes DRS particularly promising is its non-invasive nature, allowing for real-time assessment during surgery without the need for tissue excision or disruption. The DRS probe is applied to the surface of tissue and emits light via optical fibres. Additionally, DRS can provide information at multiple wavelengths, enabling a comprehensive analysis of the tissue's diffuse reflectance spectrum. Furthermore, the equipment needed for performing DRS measurements is known for its simplicity and cost-effectiveness, generally consisting of a broadband light source, a spectrometer and the DRS probe. Overall, DRS shows promise as a cost-effective, non-invasive, real-time, and objective technique for margin delineation. By leveraging these capabilities, DRS has the potential to enhance intra-operative decision-making, aiding surgeons in accurately identifying tumour margins and minimising the risk of incomplete resection, ultimately contributing to improved oncological outcomes in cancer patients.

The DRS system has been utilised for *ex vivo* differentiation between healthy and tumour tissue across various malignancy types, including skin, lung, and breast cancers. [164,301,302] In addition, there has been significant interest in the use of DRS in CRC. [303–309] These studies have been able to discriminate tumour from normal colon tissue with high overall accuracy of 0.95[303], 0.94[304] and 0.91.[305] However, the success thus far in the use of DRS technology for assessing colorectal specimens has lacked approaches which would make it clinically relevant and integrate smoothly into the operating field. For example, in the majority of studies, the use of the DRS probe on the

tissue has been point-based with only a small number of spectra being acquired per sample point on the tissue. [306]

Here lies the major limitation with the DRS technique. Point-wise application of this spectroscopy technique provides diagnostic information at discrete locations, rather than a diagnostic image for a large area. The localised nature of point measurements may result in oversights, especially when tumours exhibit heterogeneity or when the probe fails to capture regions with subtle variations in tissue composition. Although this approach can provide highly accurate diagnosis of a given point, tumour margin assessment is unrealistic unless surgeons can visualise and record data for many points simultaneously. This is particularly challenging when DRS is used endoscopically or during minimally invasive surgery, where the ergonomics of scanning and viewing the DRS probe site are even more demanding. The absence of visible marks on the tissue when scanning with the DRS probe presents a challenge for tracking the probed sites during tumour margin assessment. In order to track the DRS probe, an approach must be established in which the tip of the probe which is in direct contact with the tissue during spectra acquisition, is utilised as a reference point for localisation. Additionally, another limitation with current uses of DRS, is the disruption of the serosal layer of GI tissue when obtaining data. [305] This may jeopardise the quality of the specimen being sent for processing and hinder CRM measurement, leading to inaccuracies.

4.1.2.3.1 DRS Probe Tracking

To overcome the challenges of a point-based DRS system, the tip of the DRS probe should ideally be tracked to enable localisation of the biopsy site. [310] Various surgical tool tracking approaches have been explored, with optical tracking emerging as a widely evaluated method in surgical research. [311,312] This approach involves optical markers reflecting or emitting a non-visible wavelength of light to determine a tool's pose. By measuring the location of several of these markers using a camera system, rigid transformations between the marker co-ordinate systems and the camera co-ordinate system can be estimated. This method has been commonly used in minimally invasive surgery. [313] Many optical tracking systems employ markers such as fiducials or colour labels to facilitate the recognition and tracking of the tool's tip. In one study, a dual-pattern hybrid marker, incorporating both circular dots and chessboard vertices, demonstrated high detection rates and accurate pose estimation of a laparoscopic gamma probe. [314] In another study, biocompatible colour markers appropriate for real-time use during surgery achieved high sensitivity and specificity for a multiple instrument tracking system. [315]

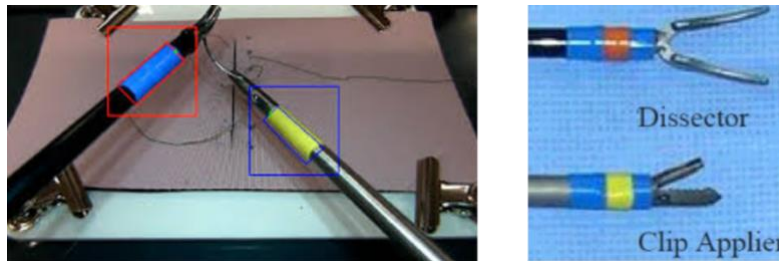


Figure 29 Examples of surgical instruments with colour markers attached to aid in surgical tool detection and tracking.[316,317]

The combination of DRS with image guidance allows the spatial relationship between the probe and the tissue to be continuously monitored by tracking the position and orientation of the DRS probe's tip. The process of localising the area of tissue being scanned through DRS probe tip tracking involves several key steps:

- The DRS probe is affixed with optical markers which allow real-time localisation of the probe's tip in relation to the anatomy.
- Optical imaging using a camera system allows visualisation of the surgical field and the DRS probe tip's position within it. This imaging data serves as a reference for tracking the probe's location and activity.
- The information from the DRS probe and the image guidance system is fused, allowing for a comprehensive understanding of the probe's tip relative to the tissue.
- The position of the DRS probe tip is displayed to the surgical team on a user interface, providing real-time feedback during the sampling process.

These steps allow real-time tracking of the probe's tip to aid continuous monitoring of the tissue being scanned without the need for any interruption of the tissue surface nor any visible marks.

4.1.3 Upper GI Tissue Discrimination

Although DRS has been applied to tissue discrimination across several malignancy types and utilised in the context of pre-malignant upper GI disease, [318,319] to our knowledge, its use in oesophageal and gastric cancer has not been fully explored.

Oesophageal and gastric cancers are associated with poor prognosis, with 5-year overall survival of patients estimated to be between 15-25%. [320,321] This is despite an improvement in survival over

the last half-century, as a result of advances in modern neoadjuvant therapies and the introduction of new surgical techniques. [322–324] The presence of residual disease has been demonstrated to nearly double the risk of mortality in oesophageal cancer. [325,326] Likewise, a positive resection margin in gastric cancer has been identified as an independent risk factor associated with reduced overall and recurrence-free survival. [327]

The lack of accurate tumour mapping tools combined with anatomical limitations of the upper abdomen and thorax lead to difficulties in achieving negative resection margins. [328] This is compounded by the lack of consensus on optimal margin length and attempts to preserve as much healthy tissue as possible. [329]

4.1.4 Aims of the Study

This component of the thesis had three broad aims:

1. Develop a system for DRS probe tracking.
2. Validate DRS in *ex vivo* specimens from gastric and oesophageal cancer surgery.
3. Validate DRS in *ex vivo* specimens from colorectal cancer surgery.

4.1.4.1 DRS Probe Tracking

This study presents a tracking system to enable localisation of the tip of a handheld DRS probe. The system allowed tracking of the two-dimensional (2D) position and orientation (x, y-axes, angle) of the spectroscopic probe using a colour marker to aid real-time assessment of tumour margins when used on *ex vivo* cancer specimens.

The aim of this study was to support real-time differentiation of tumour and non-tumour tissue, and thereby improve existing methods for intra-operative margin assessment, by introducing a system to continuously track a DRS probe during surgery.

4.1.4.2 Upper GI *ex vivo* Tissue Discrimination

We explored the use of DRS in discriminating oesophageal and gastric tissue as part of this validation study. The aim was to initially assess the system to evaluate whether DRS was able to accurately differentiate cancer tissue from normal tissue in *ex vivo* gastric and oesophageal specimens and to

develop a real-time classification system able to provide live feedback about tissue type to the surgeon intraoperatively.

4.1.4.3 Lower GI *ex vivo* Tissue Discrimination

The primary aim of this study was to evaluate the diagnostic accuracy of a DRS probe and tracking system to distinguish cancer and non-cancer tissue in *ex vivo* colorectal specimens, with the potential to aid margin assessment and provide augmented visual maps to the surgeon in real-time.

4.2 Methods

The technique for development of a DRS probe tracking system will be described followed by the methodology for *ex vivo* experiments in later sections. The role of data analysis and system development in this chapter was carried out by Dr Ioannis Gkouzionis, with input from the author of this thesis.

4.2.1 Optical Marker-Based Tracking of DRS Probe

4.2.1.1 Marker Design, Detection and Segmentation

Accurate detection and segmentation are necessary for DRS probe tracking. For calculation of the 2D probe tip position, the location of the probe in the image plane was identified. The midline of the probe could be calculated using the two edges of the probe, allowing recognition of the tip point of the probe as lying somewhere along this line.

A distinct and easily distinguishable marker colour was chosen given the unremarkable probe colour. The aim of a colour marker was to allow effective and accurate tracking during surgical procedures. Whilst patterns can also be employed as markers, a colour marker is more effective in minimising inaccuracies. This is because a patterned marker, such as a checkerboard, may be susceptible to lighting changes which affect its unique pattern, leading to tracking errors. Similarly, the marker may be obscured by surgical instruments or tissue, hindering identification of its entire pattern and therefore leading to unreliable results.

To determine the optimal marker colour for the probe, the colour distribution of surgical laparoscopic images, containing all possible colours that could be present during the course of a procedure, was analysed. The hue-saturation value (HSV) colour space was used since it directly related to the colour signature of the images (Figure 30). The colour distribution of the biological tissue ranged mainly from yellow to purple. For the colour marker, a colour that was most contrasting was needed. For this reason, the colours cyan and green were selected. In this study, a colour marker in the form of green plastic tape was used. This was wrapped around the shaft of the probe ensuring it extended to the tip of the cylindrical probe to enhance tracking precision. The use of the HSV colour space meant that the marker's distinctive colour signature would be maintained regardless of changes in illumination during surgical procedures.

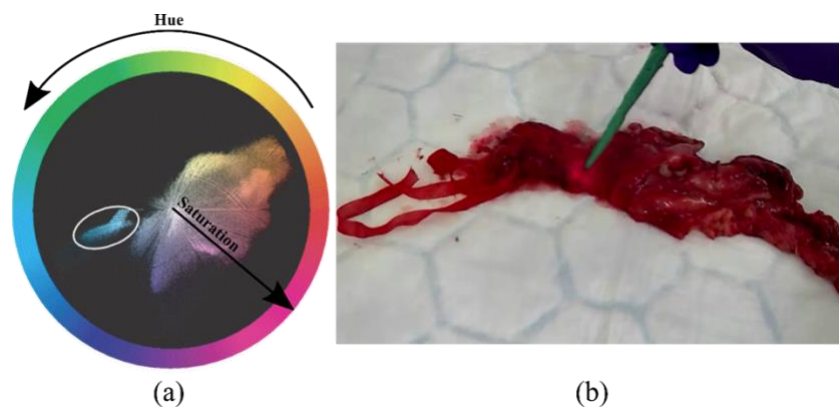


Figure 30 (a) Colour distribution of the biological tissue from laparoscopic images, with cyan colour circled as a potential colour for the marker. (b) DRS fibre probe with the green colour marker attached. © Reproduced from [330] with permission.

The use of the green colour allowed it to be automatically segmented from the optical images. This segmentation process is vital for accurately identifying the marker's position in the image frames and determining the tool's pose.

The comprehensive approach for employing a monocular red/green/blue (RGB) camera (C920, Logitech International S.A., Lausanne, Switzerland) within a 2D colour-based optical marker tracking system is outlined in Figure 31.

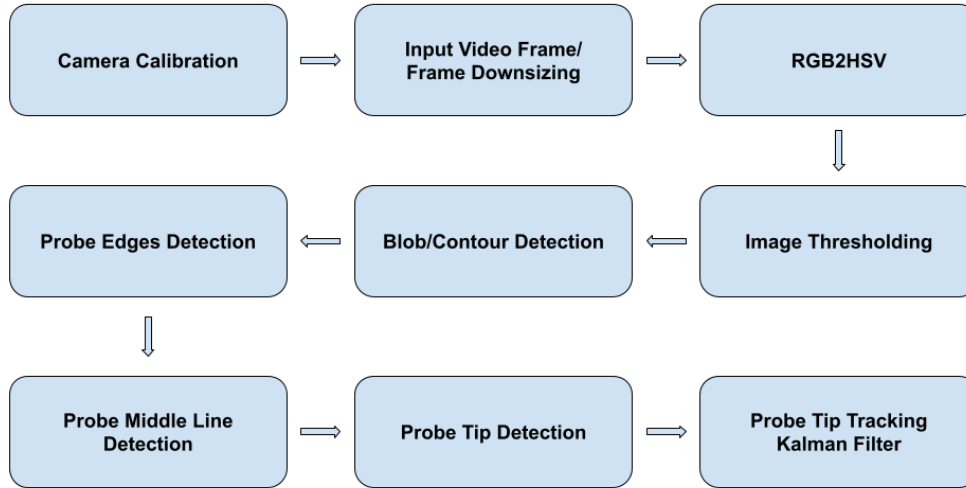


Figure 31 DRS probe detection and tracking flow diagram. © Reproduced from [330] with permission.

A 2D image calibration method was employed. [331] This calibration approach necessitated the observation of a planar pattern, specifically a checkerboard, from various orientations and positions using monocular RGB camera. A series of previously captured images from the camera were then scrutinised to deduce the 3×3 intrinsic matrix K . This matrix encompassed essential intrinsic parameters such as focal length, optical centre co-ordinates, and radial distortion coefficients associated with the lens. Additionally, the extrinsic matrix was determined, comprising the 3×3 rotation matrix, denoted as R , and the 3×1 translation vector, t . The extrinsic matrix represented the camera's position and orientation with respect to the 3D co-ordinate system. This meticulous calibration process ensured that the tracking system was able to accurately convert 2D image points of the colour marker into 3D co-ordinates, ensuring precise position and orientation of the probe based on its appearance in the camera view. As a result, accurate and reliable tracking information can be determined enabling effective surgical navigation.

To expedite processing and enhance frames-per-second, the input video frames underwent resizing from 1920×1080 to 640×480 . Subsequently, the frame underwent a blurring process to reduce high spatial frequency information while preserving visibility of the flat-textured DRS probe. The image was then transformed from RGB to HSV colour space and subjected to thresholding based on predetermined upper and lower boundaries for hue (h) and saturation (s), corresponding to either cyan or green colour. The resulting segmented frame was converted back to RGB, and morphological operations, including erosion and dilation, were implemented. These operations served to eliminate negative impulsive noise and positive noise while addressing potential gaps in the mask. This series

of image processing steps aimed to optimise the representation of the DRS probe within the video frames.

Diverse contours within the mask, indicative of segmented objects, were recognised and systematically arranged. The largest contour was singled out as the likely location of the identified colour marker. Subsequently, the corner points of the bounding box enclosing the contour were extracted. These corner points, along with the Hough transform, were employed to discern the edges of the probe. Utilising a conventional description of a line in the $x - y$ plane ($y = mx + c$), the midline of the probe was computed based on the detected edges.

4.2.1.2 Calculation of DRS Tip Position & Probe Tracking

As depicted in Figure 32, there existed two potential locations corresponding to the tip position—one situated at the bottom left of the white segmented probe and the other at the top right. To ascertain the direction for locating the probe tip along the midline, an assumption was made that the probe's movement in the yaw axis during *ex vivo* benchtop sampling would be confined to the range of $0 - 180^\circ$, with zero degrees defined as parallel to the hard surface.

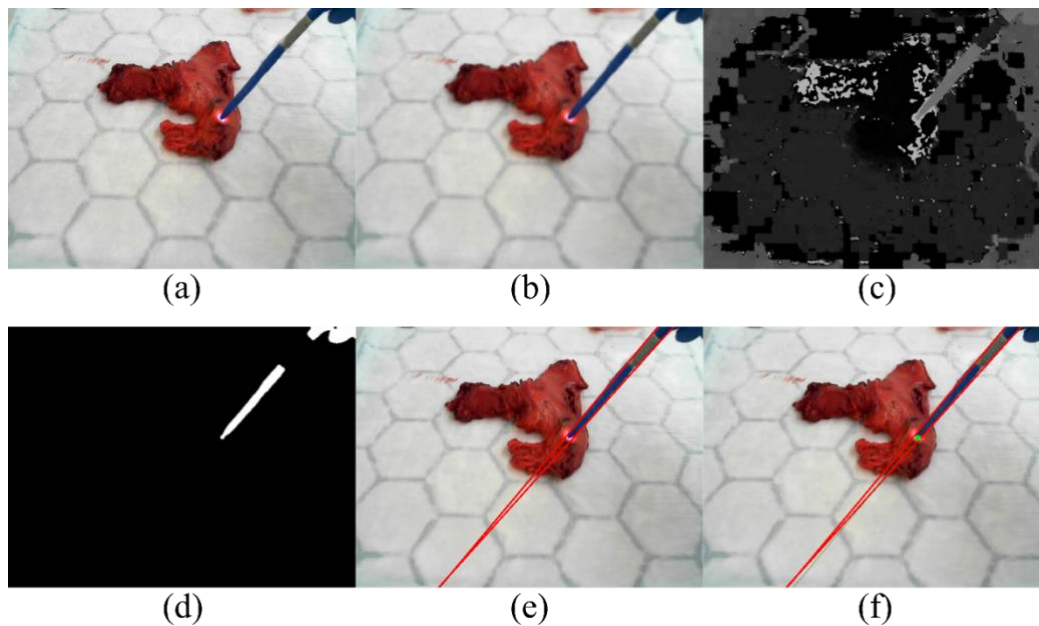


Figure 32 DRS probe detection workflow. (a) Example input video frame. (b) Blurred frame. (c) Hue channel of the blurred input frame. (d) Binary mask result of HSV segmentation. (e) Detected edge lines of the probe. (f) Detected tip point of the probe. © Reproduced from [330] with permission.

To address real-world scenarios within the operating room such as partial occlusion, varying illumination, and changes in probe velocity, a standard Kalman filter (KF) was introduced to estimate the tip location of the probe based on its previous positions. The KF's responsibilities were divided into two main steps: prediction and correction. In the prediction step, the KF tracked the state matrix, encompassing the current probe tip location, while the process covariance matrix contained the predictive error associated with those measurements. With each new time frame, the state transition matrix adjusted the state and process matrices based on the current position, generating estimations of both the new position and covariance. Simultaneously, the Kalman gain was computed alongside the detected probe tip positions. The correction step involved updating the current state using the Kalman filter in conjunction with the measurement obtained in that specific time frame.

4.2.2 Graphical User Interface

A graphical user interface (GUI) was developed using Python 3.6 (Python Software Foundation, Wilmington, USA) and Qt5 toolkit to incorporate all functionalities on a standard PC (Intel i3 processor at 3.30 GHz with 8 GB RAM). The GUI, highlighted in Figure 33, consisted of a variety of functions, such as “open camera” to capture the live video feed on the left side of the screen and “start tracking” for the initiation of the probe tracking. Spectral data was collected and highlighted live on the right side of the screen by pressing the “start acquisition” function. Before starting the acquisition, the “experimental parameters” function allowed entry of patient metadata, spectrometer integration time, spectral range, tissue type and DRS calibration into a dialog box. A single image frame was captured of the tissue sample being examined without the DRS probe. On completion of the acquisition process, this frame was annotated with all the probed locations. The live video feedback from the RGB camera was annotated in real-time with the probed positions, and the current spectrum at each location was displayed in the main GUI window. In addition, a colourmap was utilised to visually indicate the probability of a data point belonging to a class of tissue type (green for normal, pink for tumour). Real-time classification was added to the GUI following the interim analysis of the data. When the experiment was complete, all data was saved for further correlation with the histopathology reports.

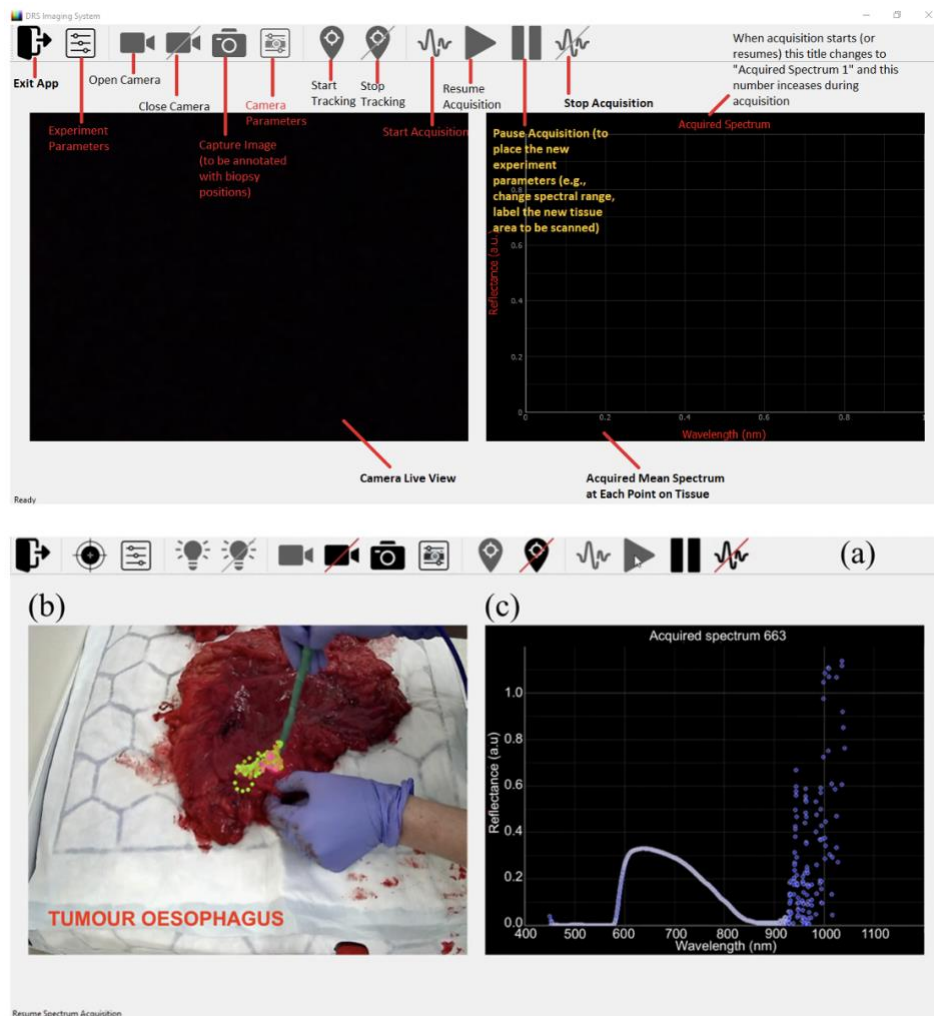


Figure 33 Python software for hardware system control, data acquisition and real-time classification. (Top) Graphical User Interface (GUI) highlighting function of each individual section. (Bottom) GUI in use during data collection. (a) Control panel, allowing for execution of data acquisition protocol, probe tracking, hardware calibration and light source control. (b) Augmented live video feed from the RGB camera. Real-time classification is displayed on the screen via text and a coloured overlay on the tissue, with a graduated colour scale indicating the probability of belonging to a particular class (Normal = 100% green – 100% pink = Tumour). (c) Reflectance data in the 420 – 1000 nm spectral range displayed in real-time. © Reproduced from [330] with permission.

4.2.3 DRS Instrumentation and Set-Up

The DRS system consisted of a fibre-optic probe (Ocean Optics Inc., QR600-7-SR-125F) containing six peripheral illumination fibres around one light collection fibre (600 μm diameter) within a 3.175mm stainless steel ferrule. The total probe length was exactly 76.2 mm. A tungsten halogen light source (360 – 2,400 nm, typical output power 8.8 W, Ocean Optics Inc., HL-2000-HP) was connected to the six illumination fibres, whilst the central fibre was connected to a spectrometer (Ocean Optics Inc., USB4000). The set-up is shown in Figure 34. The spectrometer was controlled by the developed GUI. The spectral data acquired by the spectrometer was in the spectral range of 400-1,000 nm. To

protect the DRS probe tip from tissue contamination a sterile plastic camera drape (365 Healthcare Ltd.) was chosen due to the low optical interference, low cost, and sterility.

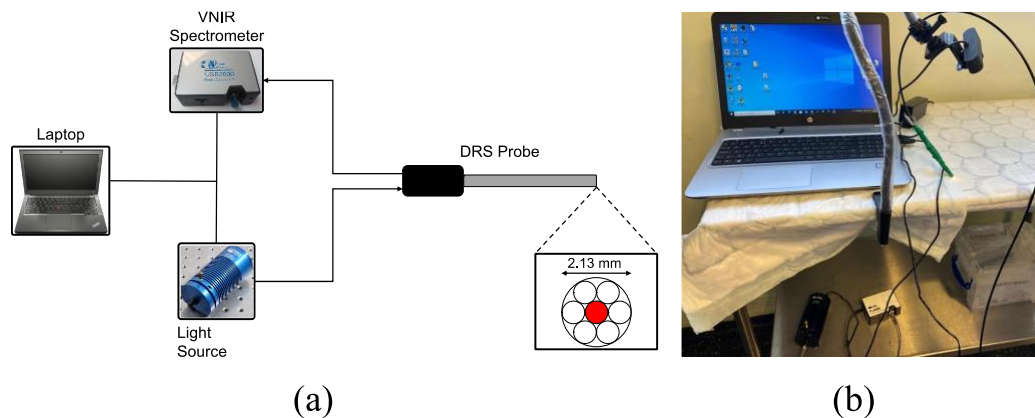


Figure 34 DRS instrumentation for *ex vivo* data acquisition: (a) schematic diagram of set up; (b) photograph of set up. The DRS probe was connected to both the light source (Ocean Optics, HL-2000-HP, Largo, USA) and the spectrometer (Ocean Optics, USB4000, Largo, USA) to allow for data acquisition and sample illumination. All electronic devices communicated with proprietary software designed with Python version 3.6 (Python Software Foundation, Wilmington, USA) on the laptop. © Reproduced from [332] with permission.

4.2.4 Study Design

4.2.4.1 Upper GI Cancer

This was an *ex vivo* validation study to explore the feasibility of using the DRS probe and tracking system on gastrointestinal tissue. Consecutive adult patients, over the age of 18, undergoing oesophageal or gastric cancer resection were prospectively recruited into the study between July 2020 and July 2021 at Hammersmith Hospital in London, United Kingdom. All patients were given a patient information sheet. Only patients who were competent to consent were included. *Ex vivo* tissue specimens were included for patients undergoing elective surgery for either oesophageal carcinoma (adenocarcinoma or squamous cell carcinoma) or gastric adenocarcinoma. The study was performed with approval from the Harrow Research Ethics Committee (ref. no. 08/H0719/37).

4.2.4.2 Lower GI Cancer

This was a prospective validation study covered under the same ethical approval (ref. no. 08/H0719/37). Adult patients, over the age of 18, undergoing elective colorectal cancer resection surgery at Imperial College NHS Trust in London were prospectively recruited between April 2021 and July 2022. Abdominoperineal excision of rectum (APER) cases were excluded from this study, since we only collected colonic tissue data. This was due to the presence of mesorectal fat

surrounding the rectum. Normal tissue data was included for analysis for cases confirmed as T0 on formal histology.

4.2.5 Data Collection & Processing

The DRS system was set-up inside the operating theatre. The specimens were sampled within 15 minutes of termination of the blood supply using the hand-held DRS probe. Freshly resected specimens were placed on a clear white disposable absorbent pad immediately following removal from the patient. Initially, an overview image was taken of the specimen, followed by a video recording using the RGB camera. The live video feedback allowed tracking of the tip of the DRS probe and mapping of sampling locations in real-time. Data acquisition lasted approximately 10 minutes.

Spectral data was obtained by probing the outer surface of the oesophageal, gastric or colorectal specimens over a non-tumour, or normal, tissue area and over the suspected tumour area. These areas were identified by visual and haptic inspection by a surgeon. Normal tissue areas were chosen as far away from the suspected tumour area as possible, and based on an area with minimal overlying fat tissue, and no staples or sutures. Staples and clips would be expected to affect the spectral data due to their reflectance properties. Probing the surface of the specimen left no mark on the tissue and did not disrupt the outer layer in any way.

The DRS probe acquired around 80 spectra per second. At each sampling point, 20 measurements, equating to 20 spectra, were acquired and averaged. The mean spectrum was displayed in real-time. A minimum of 200 spectral measurements were acquired for each tissue type (normal or tumour) per specimen. This was dependent on the size of the sampled region. Where possible, DRS measurements were taken for both oesophageal and gastric tissue from the same specimen.

Data were acquired in the spectral range of 420 – 720 nm, with a wavelength separation of 0.5 nm for the oesophageal and gastric specimens, and in the spectral range of 450 – 1100 nm, with a wavelength separation of 0.34 nm for the colorectal specimens. Wavelength, or λ , separation denotes the interval between successive measurement points in a spectrum. At each point measurement location, the acquired spectra were averaged to calculate a mean spectrum of each DRS sampling location. Subsequently, to account for inter-patient comparability, background light,

and signal quality variability, spectral normalisation and noise reduction was performed for each acquired spectrum. A workflow for the DRS acquisition and processing is highlighted in Figure 35.

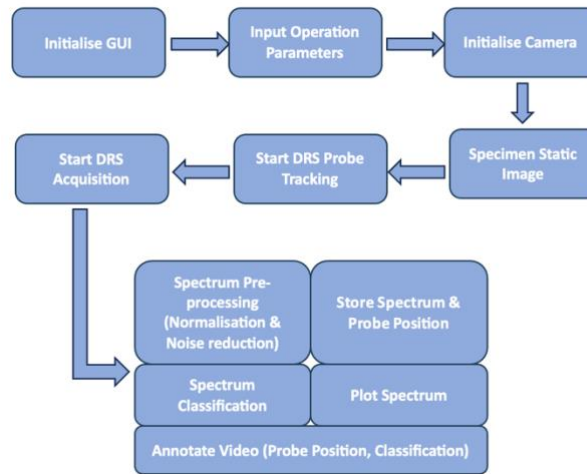


Figure 35 Workflow for DRS data acquisition and processing. [GUI: graphical user interface] © Reproduced from [330] with permission.

4.2.6 Histopathology Correlation

Establishing a robust ground truth is paramount for the training of machine learning classifiers in medical applications, particularly in the context of cancer detection. Accurate delineation of tumour and non-tumour areas is crucial for developing models that can reliably identify malignancies during surgical procedures. The gold standard technique for understanding the true areas occupied by the tumour is histopathological analysis. By aligning digital data from suspected tumour locations with histopathological examination, we aimed to create a comprehensive and reliable reference dataset that could serve as the foundation for training and validating machine learning algorithms.

4.2.6.1 Challenges in Clinical Validation

Achieving accurate correlation between optical measurements and histopathological analysis poses a significant challenge in the clinical validation of optical technologies. [333] Previous efforts in achieving histopathological correlation as part of the surgical workflow have encountered issues. Many studies have resorted to sampling suspected tumour sites and processing each sample individually, thereby lacking a holistic representation of the entire specimen, and leading to inaccurate and unstable fit results. [305] This approach falls short of capturing the spatial intricacies of the tumour across the entire surgical field. In addition, other studies have attempted to visually

correlate annotated pathology slides with the RGB images taken from the measurement locations to identify the histological tissue type at each measured location. [303] However, it is essential to acknowledge the limitations associated with comparing digitalised images of tissue captured during surgery to the specimen once dissected and assessed. Firstly, this method requires disturbing the tissue surface and potentially affecting the precise calculation of the cancer margins. Secondly, in a clinical setting, rather than a controlled or experimental environment, the shrinkage and structural changes that occur upon formalin preservation of the tissue following resection can introduce discrepancies between the real-time digital data and the post-processing assessment. This underscores the need for a meticulous and adaptable approach to histopathological correlation, considering the dynamic nature of tissue characteristics throughout the surgical workflow. This study endeavoured to overcome this limitation by correlating digital data with histopathological analysis on a larger scale, encompassing the entirety of the specimen and taking into consideration translation of the methodology in a real-world scenario.

4.2.6.2 Histopathological Analysis Protocol

Following the acquisition of all spectra, the suspected tumour area was painted using yellow tissue dye (Cancer Diagnostics Inc., Durham, USA) for subsequent correlation with histopathological analysis, serving as the reference standard. An image was taken at this stage. The entire tissue sample was then immersed in formalin and dispatched to the histopathology department, where it underwent processing in accordance with established protocols. Comprehensive macroscopic images of the entire specimen, along with cross-sectional slices, were captured by the dissector in the histopathology department. These images included areas of the tissue painted with yellow dye. This process of histopathological analysis using yellow tissue dye is highlighted in Figure 36.

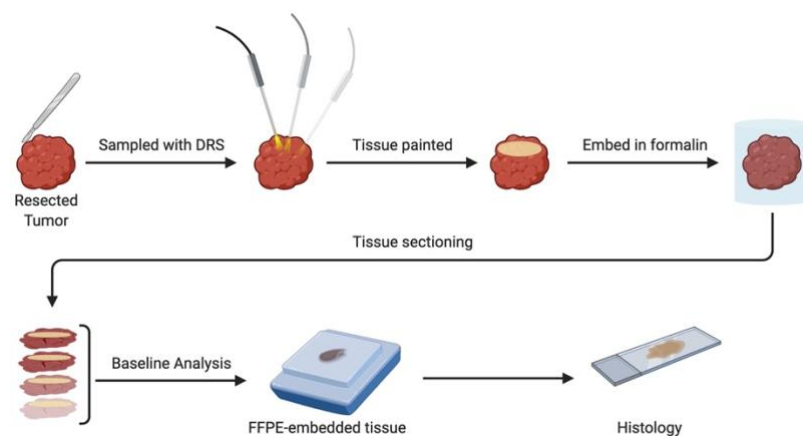


Figure 36 Workflow for histopathological analysis of tissue for ground truth labelling.

Once the areas marked with yellow dye were confirmed by the pathologist as tumour, the corresponding video frames and acquired spectra were correlated with the histopathological information, as shown in Figure 37 for an oesophageal specimen and Figure 38 for a colorectal specimen. These regions were labelled as tumour, whilst the remaining spectra were labelled as non-tumour tissue. These labels were considered ground truth labels for the training of the ML classifiers. For specimens in which the suspected “tumour” area was confirmed by the pathologist as fully regressed tumour following neoadjuvant therapy (T0), only normal tissue data collected for the specimen was analysed. Tracked locations on the specimen, which could not be confirmed by the histopathological analysis, were excluded.

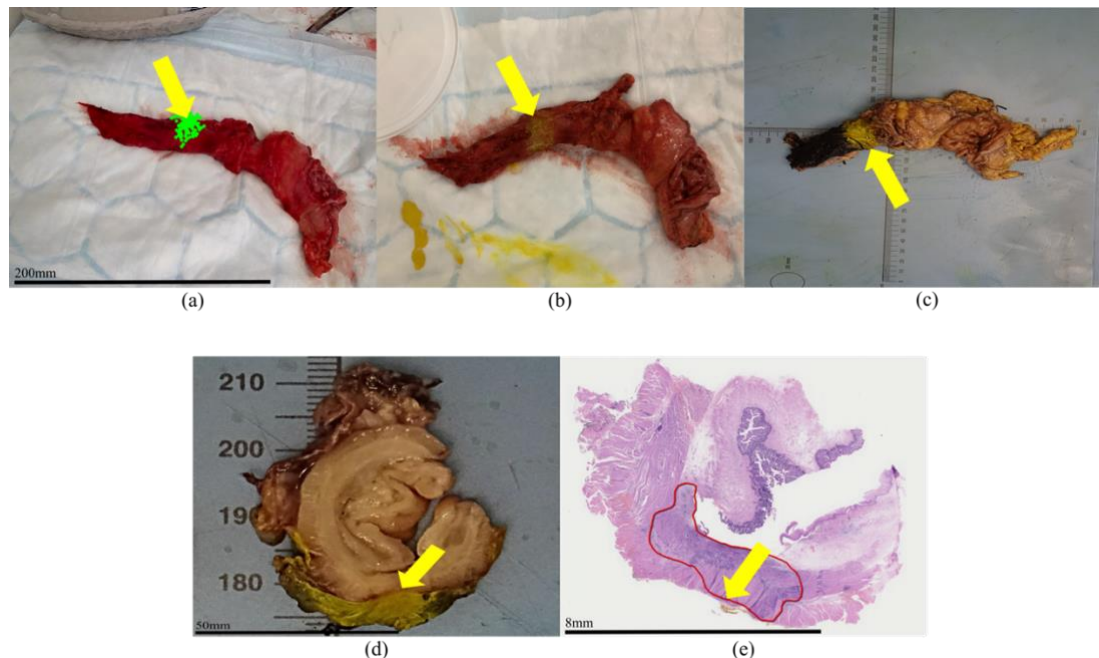


Figure 37 Histopathological correlation workflow for processing of an oesophageal specimen. The correlation was performed using yellow tissue paint (marked with a yellow arrow across panels (a) – (e)). (a) The ex vivo specimen with tracked optical biopsy sites of the suspected ‘tumour’ area. (b) Ex vivo specimen following painting of the suspected ‘tumour’ tissue with yellow paint. (c) Specimen, after having been placed in formalin, with yellow painted area visible. (d) Macroscopic slice of oesophageal tissue with the yellow painted area. (e) Microscopic H&E-stained slice with yellow paint (yellow arrow) and confirmed oesophageal adenocarcinoma (red outline). © Reproduced from [332] with permission.

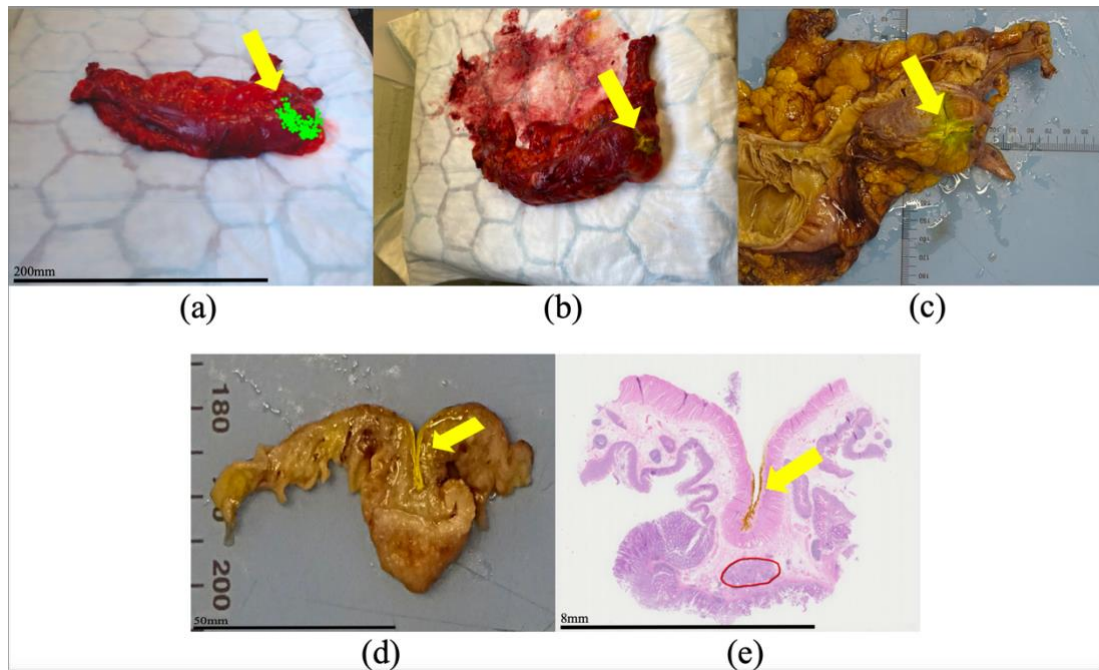


Figure 38 Histopathological correlation workflow for processing of a colorectal specimen. (a) Photo of scanned colon tissue with tracked DRS locations over suspected tumour area indicated. (b) Photo of colon tissue with painted yellow dye over the suspected tumour tissue. (c) Photo of colon tissue after being in formalin with yellow tissue dye visible. (d) Photo of sliced colon tissue with yellow dye visible over the suspected tumour. (e) Scanned H&E slide showing yellow tissue dye over confirmed tumour tissue (marked in red). © Reproduced from [334] with permission.

4.2.7 Classification

4.2.7.1 Upper GI Tissue Classification

Data from all patients were amalgamated into distinct datasets for either oesophageal or gastric tissue. Instances of *ex vivo* labelling errors of spectral data were identified and subsequently excluded from further analysis. This was carried out through retrospective review of videos depicting real-time data acquisition, coupled with the analysis of the DRS probe's position over the tissue areas. Probe contact artefacts (e.g. lack of probe contact with the tissue) were automatically removed. To address inter-patient variations, background light fluctuations, and signal quality disparities, spectral normalisation and noise reduction were performed. These procedures incorporated the use of white reflectance standard and dark field readings.

Moreover, an automated method was implemented for outlier (e.g. erroneous measurements or air interference) detection and removal. For this method, the median of the residuals was calculated, along with the 25th and 75th percentiles. The difference between each historical value and the

residual mean was then analysed. If the historical value was 1.5 times greater than the median absolute deviation (MAD) away from the median of the residuals, that value was classified as an outlier. Additionally, the ground truth provided by histopathology correlation was used for the validation of the method.

To mitigate overfitting and enhance classification accuracy, feature extraction was conducted on the oesophageal and gastric datasets. The focus was on identifying the two most prominent peaks within a spectrum. The mean intensity values in the spectral range corresponding to these peaks, along with the mean intensity value across the overall spectral range were derived. Additionally, another three feature extraction techniques underwent evaluation, namely the permutation feature importance, [335] the recursive feature elimination, [336] and the Boruta method. [337]

The dataset was subsequently divided into the training and testing sets through the repeated stratified k-fold cross-validation (CV) technique, employing five folds and five repeats. This specific cross-validation method was chosen after a thorough literature review and after consideration of the nature of the data used in this study. The stratified five-fold CV method ensures a more stable and reliable result since training and testing is performed on several different sections of the dataset and deals with imbalanced data. Using this CV method, the dataset is split into five folds such that each fold contains approximately the same percentage of samples of each target class as the complete set. When stratified five-fold CV is compared to leave-one-out CV, the latter requires building n models instead of five, where n stands for the number of samples in the dataset. This means that the leave-one-out CV method is more computationally expensive than the stratified five-fold CV technique. The spectral data analysis process is depicted in Figure 39.

Binary classification into normal and tumour tissue was executed using distinct supervised ML classifiers. These included linear Support Vector Machine (SVM), Multi-layer Perceptron (MLP), Light Gradient Boosting Machine (LGBM) and Extreme Gradient Boosting (XGB).

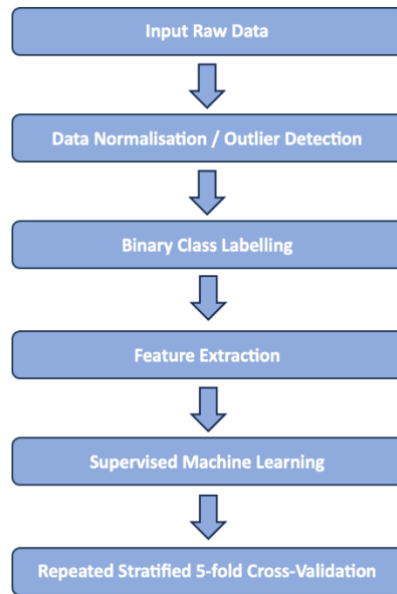


Figure 39 Spectral data analysis workflow. ©Reproduced from [330] with permission.

4.2.7.2 Lower GI Tissue Classification

Data for colorectal tissue from all patients were collected. Upon data collection, spectral normalisation and noise reduction were performed, using a white reflectance standard and dark field readings to account for inter-patient, background light, and signal quality variability. After that, the spectral data were normalised by removing the mean and scaling to unit variance, using the standard normal variate (SNV) method. [338] An automated method was implemented for outlier (e.g., erroneous measurements or air interference) detection and removal.

In addition to the intensity measurements, spectral shape-based features were extracted and used for classification, comprising the intensity value and wavelength of the greatest peak for each spectrum, the intensity and wavelength of the secondary peak, and the mean intensity in three spectral regions, namely 480-550 nm, 580-700 nm, and 450-720 nm. The haemoglobin saturation coefficient was also derived, using Monte Carlo simulations, to quantify the effect of changing haemoglobin saturation on the curve features. To reduce overfitting and improve discrimination accuracy, we performed the Boruta feature selection method on the data. [339]

To ensure unbiased and reliable performance evaluation, the dataset was methodically divided into distinct training and testing sets, adhering to a patient-wise grouping framework. This strategy ensured that data originating from each patient was collectively allocated to either the training or

testing set, circumventing any inadvertent intermixing that might lead to data leakage or skewed assessments. The rationale underpinning this patient-wise grouping strategy was to reflect the patient-specific variability encountered in real-world clinical scenarios.

A stratified five-fold cross-validation methodology was employed exclusively on the training subset. This five-fold approach entailed the subdivision of the training data into five equitably balanced subsets, characterised by an equitable representation of both tumorous and healthy tissue samples representing the real-world distribution. Within this stratified five-fold cross-validation framework, the training process was iteratively executed four times, with one subset reserved for validation, effectively mitigating the potential for overfitting and ensuring a comprehensive model assessment.

The stratified five-fold cross-validation was performed using tree-based classifiers like random forest (RF) and gradient boosting (XGB, LGBM) as they are effective with highly complex data. [340–342] Conventional ML algorithms were also used to differentiate tumour and non-tumour tissue, such as linear Support Vector Machine (SVM), and a shallow neural network (MLP). In addition, a one-dimensional CNN, named SpecNet, was developed and used.

The hierarchical structure of CNNs is emerging as the most effective and successful method for acquiring visual representations. Figure 40 highlights the hierarchical architecture of the proposed CNN-based SpecNet for spectral classification. The network contains five layers with weights, including the input layer, convolutional layer C1, max pooling layer M2, fully connected layer F3, and output layer. Each spectral measurement in DRS can be regarded as a 1D image, and the size of the SpecNet input layer is $(n_1, 1)$, where n_1 is the number of spectral bands (1,922 in this study). This treatment acknowledges the inherent multidimensionality of the data, condensing the spectral information across the bands into a single-channel representation akin to an image with a single height dimension. The first hidden convolutional layer C1 filters the input data with 20 kernels of size $k_1 \times 1$. Layer C1 contains $20 \times n_2 \times 1$ nodes, and there are $20 \times (k_1 + 1)$ trainable parameters between layers C1 and the input layer. The second hidden layer is the max pooling layer M2 with a kernel size of $(k_2, 1)$ and contains $20 \times n_3 \times 1$ nodes, where $n_3 = n_2 / k_2$. The fully connected layer F3 has n_4 nodes, and there are $(20 \times n_3 + 1) \times n_4$ trainable parameters between this layer and layer M2. The output layer has n_5 nodes, and there are $(n_4 + 1) \times n_5$ trainable parameters between this layer and the F3 layer.

The output data of the neural network comprises binary class predictions that decisively assign each input spectrum sample to either of the two classes: healthy tissue or cancerous tissue. These binary class predictions serve as a categorical affirmation of the network's evaluation of each individual pixel's spectral attributes. Upon traversing the network's layers, which include convolutional, pooling, and fully connected components, the model collectively aggregates and processes spectral patterns representative of each tissue type. This culminates in the final layer, where the network furnishes a definitive declaration regarding the classification label for each pixel—either "healthy" or "tumour."

Python 3.6 (Python Software Foundation, Wilmington, USA) was used for feature engineering and machine learning classification, and MATLAB R2020b (MathWorks, Natick, USA) was used for data processing. For the training and validation of the SpecNet, TensorFlow deep learning framework was employed.

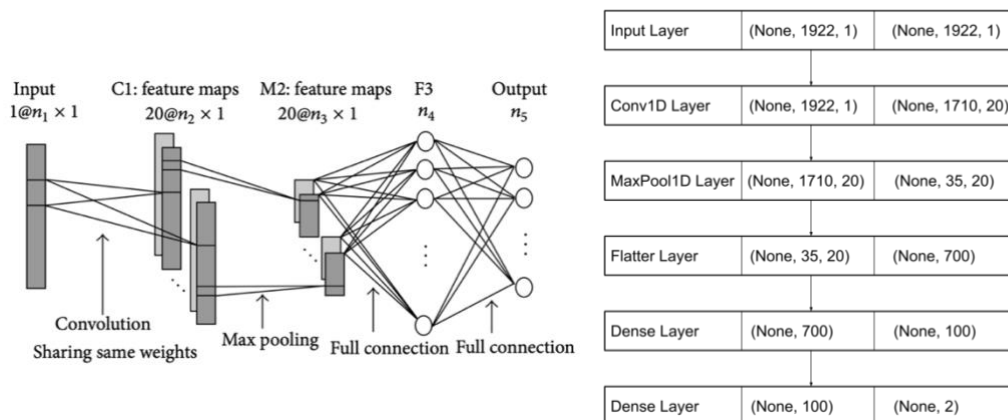


Figure 40 – The architecture of the proposed one-dimensional convolutional neural network (CNN) for colorectal tissue classification. The SpecNet is a five-layer CNN for learning visual representations in DRS data. It uses 20 kernels to filter the input data, and the output layer classifies spectra based on the spectral band size and number of output classes in the dataset. © Reproduced from [334] with permission.

4.2.8 Statistical Analysis

Machine learning classifiers were evaluated in terms of sensitivity, specificity, overall accuracy, and the area under the curve (AUC). Overall accuracy was calculated as the proportion of correctly identified spectra over a total number of spectra. Receiver-operator characteristic (ROC) curves were plotted. Python 3.6 (Python Software Foundation, Wilmington, USA) was used for ML classification, MATLAB R2020b (MathWorks, Natick, USA) was used for data processing, and R 4.0 (R Foundation, Vienna, Austria) was used for visualisation and statistical analysis.

4.3 Results

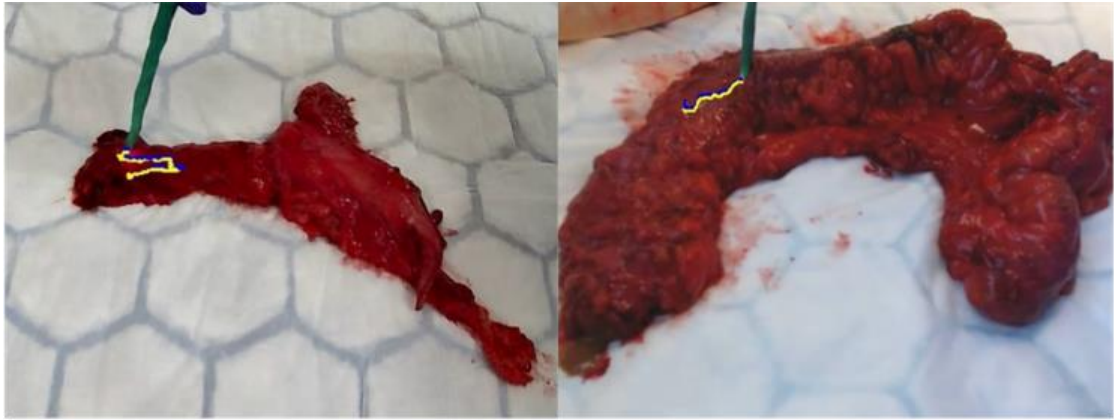
4.3.1 DRS Probe Tracking Evaluation

To assess tracking accuracy, two videos from each tissue specimen underwent analysis through the manual labelling of the 2D tip position (x, y) by two distinct individuals (data labelling tool via Labelbox Inc., California, USA). Subsequently, the manually labelled positions were compared to the estimated tip locations determined by the proposed algorithm (

Figure 41).

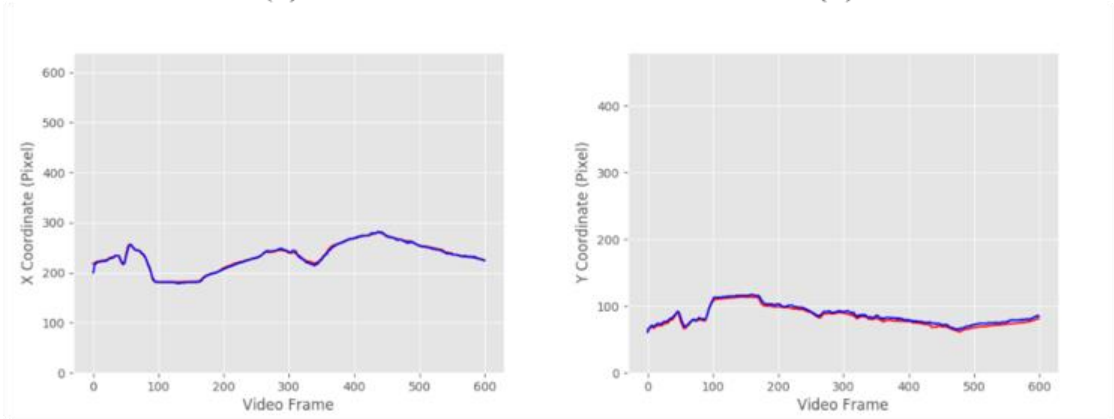
The root mean squared error (RMSE) between the estimated and manually labelled tip position was 1.18 ± 0.58 mm (4.21 ± 2.07 pixels) and 1.05 ± 0.28 mm (3.75 ± 0.97 pixels) for the x and y directions, respectively, as shown in Table 14. The maximum measured error observed across all 1050 frames from both videos was 1.76 mm (6.28 pixels) at a working distance of 30 cm.

Figure 41 illustrates frames from each video with overlays of the manually labelled and estimated positions of the probe's tip, distinguished by different colours.



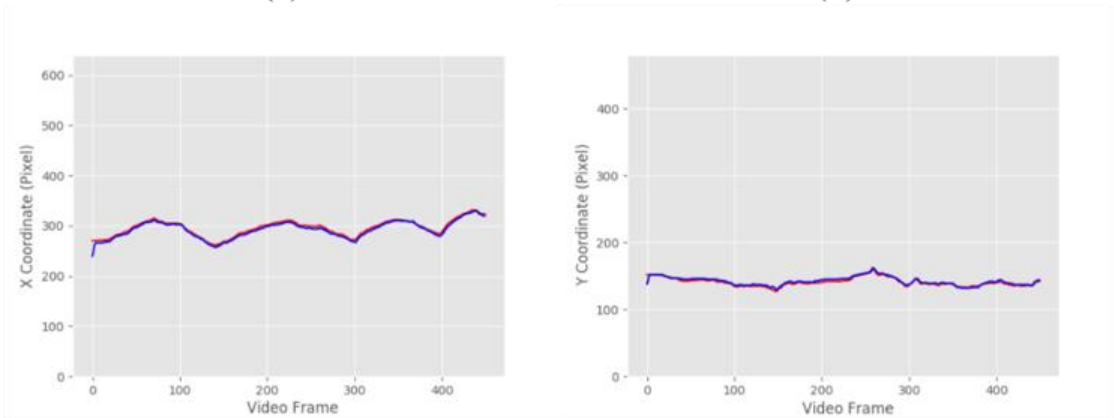
(a)

(b)



(c)

(d)



(e)

(f)

Figure 41 (a)-(b) Illustration of probe tracking error with overlaid ground truth (blue line) and estimated (yellow line) probe tip position for two video sequences (600 and 450 frames each). (c)-(f) The ground truth and estimated projected DRS probe tip pixel position for x-axis and y-axis for each of the video sequences. Blue and red lines indicate the estimated and ground truth respectively. © Reproduced from [330] with permission.

Table 14 Projection error for DRS probe's tip location.

Dimension	RMSE (mm)	Maximum error (mm)	Minimum error (mm)
x	1.18±0.58	1.76	0.00
y	1.05±0.28	1.39	0.00

For further validation, the detection limits and rates were calculated by recording the maximal experimental distance and rotation angle of the probe (Table 15). The recorded distance was measured from the camera to the probe and the rotation limits were established around the probe's local coordinate axes (roll, pitch, and yaw). When testing the distance limits, the probe was translated along the axis of the camera until the detection failed. To assess the rotational motion limits, the probe was placed at 30 cm from the camera, a typical distance for practical *ex vivo* tissue scanning.

The image-based tracking framework was used to map all spatio-temporally tracked biopsy sites and re-project them back onto the image plane to provide a live augmented view of the *ex vivo* spectral acquisition procedure. The biopsy site was defined by the probe tip location, taken from the tracking procedure, and the geometrical characteristics of the probe.

Table 15 Rotation angle and maximum detectable distance from the camera to track the DRS probe.

Rotation Axis	Rotation Angle & Distance
Roll (°)	360°
Pitch (°)	360°
Yaw (°)	0° - 180°
Distance to Camera (cm)	10 – 100

4.3.2 Patient Characteristics

4.3.2.1 Upper GI Cohort

This study assessing upper GI tissue recruited a total of 37 patients, and their characteristics are detailed in Table 16. Three patients were excluded due to incomplete data acquisition, resulting in

34 patients being included in the final analysis. The median age of the participants was 68, with the majority being male (n=22, 64.7%), in line with the typical demographics of oesophageal and gastric cancer. Among the patients, 19 (55.9%) underwent oesophagectomy, whilst 15 (44.1%) underwent a gastrectomy procedure. The predominant histological finding was adenocarcinoma, accounting for 79.4% of cases, with 5.9% being squamous cell carcinoma. The most common tumour stage was T1 disease for gastric cancer patients (33.3%), and T3 disease in oesophageal cancer patients (57.1%). Pathological staging revealed that 14.7% of tumours had regressed or showed no tumour presence. With regard to neoadjuvant therapy, 60% of patients with gastric cancer underwent chemotherapy, radiotherapy or both, compared to 89% of patients with oesophageal cancer.

Table 16 Upper GI patient characteristics.

Patient Characteristic	No. of patients (%)	Gastric Cancer No. (%)	Oesophageal Cancer No. (%)
Age, median (range)	68 (35 – 89)	61 (43 – 89)	68 (35 – 80)
Sex			
Female	12 (35.3)	4 (26.7)	8 (42.1)
Male	22 (64.7)	11 (73.3)	11 (57.9)
Procedure			
Oesophagectomy	19 (55.9)		
Gastrectomy	15 (44.1)		
Tumour histology			
Adenocarcinoma	27 (79.4)	12 (80.0)	15 (78.9)
Squamous Cell Carcinoma	2 (5.9)	2 (13.3)	0 (0.0)
Tumour stage (TNM staging)			
T0	5 (14.7)	1 (6.7)	4 (21.1)
T1	6 (17.7)	5 (33.3)	1 (5.2)
T2	4 (11.8)	3 (20.0)	1 (5.2)
T3	15 (44.1)	4 (26.7)	11 (57.9)
T4	4 (11.8)	2 (13.3)	2 (10.6)
Neoadjuvant therapy			
Chemotherapy	24 (70.6)	8 (53.3)	16 (84.2)
Chemoradiotherapy	1 (2.9)	0 (0.0)	1 (5.3)
Radiotherapy	1 (2.9)	1 (6.7)	0 (0.0)

None	8 (23.5)		
------	----------	--	--

4.3.2.2 Lower GI Cohort

A total of 23 patients were recruited into the study assessing colorectal cancer specimens. The age of patients ranged from 50 to 88 years old, with a median age of 70. The majority of patients undergoing colorectal cancer surgery were male (74%). Most patients (n= 14; 61%) underwent an anterior resection procedure, whilst 8 (35%) patients underwent a right hemicolectomy and only 1 (4%) patient underwent a sigmoid colectomy. The majority of patients (47.8%) had T3 disease on histopathological analysis. There were 3 patients with T0 disease on histology. These were found to be benign adenomatous polypoid lesions. Normal tissue data was collected for these patients. Table 17 outlines the patient characteristics.

Table 17 *Lower GI patient characteristics.*

Patient Characteristic	No. of patients (%)
Age, median (range)	70 (50 – 88)
Sex	
Female	6 (26.0)
Male	17 (74.0)
Procedure	
Anterior Resection	14 (61.0)
Right Hemicolectomy	8 (35.0)
Sigmoid Colectomy	1 (4.0)
Tumour histology	
Adenocarcinoma	20 (87.0)
Benign Adenomatous Polyp	3 (13.0)
Tumour stage (TNM staging)	
T0	3 (13.0)
T1	5 (21.7)
T2	2 (8.7)
T3	11 (47.8)
T4	2 (8.7)
Neoadjuvant therapy	

Chemotherapy	1 (4.3)
Chemoradiotherapy	0
Radiotherapy	0
None	22 (95.7)

4.3.3 Spectral Data Analysis

Overall, a total of 23 distinct sets of normal stomach data, 16 of normal oesophagus data and 23 of normal colorectal data were recorded. In regard to cancerous tissue, 10 sets of stomach data were collected, together with 10 for oesophagus and 19 for colorectal. This is highlighted in Table 18.

In total, 5,496 mean spectra were collected for normal stomach, 2,441 mean spectra for gastric cancer, 3,677 mean spectra for normal oesophagus and 2,483 mean spectra for oesophageal cancer. Each processed mean spectrum for upper GI tissue contained 505 equally spaced intensity measurements in the 468-720 nm spectral range (resolution 0.5 nm), with data from 420-468 nm and 720-1,000 nm spectral ranges excluded following interim analysis. The means of all spectra for upper GI tissue classes are depicted in Figure 42.

For colorectal tissue, a total of 7,692 mean spectra were acquired. This included 3,632 mean spectra for normal tissue and 4,330 mean spectra for tumour tissue. Each processed mean spectrum contained 1,922 equally spaced intensity measurements in the 450-1000 nm spectral range. The means of all spectra for each of the colorectal tissue classes are highlighted in Figure 43.

Table 18 Dataset summary for spectral acquisition for gastric, oesophageal and colorectal tissue.

	Normal Gastric Tissue	Cancer Gastric Tissue	Normal Oesophageal Tissue	Cancer Oesophageal Tissue	Normal Colorectal Tissue	Cancer Colorectal Tissue
Distinct sets of data	23	10	16	10	23	19
Mean spectra collected	5,496	2,441	3,677	2,483	3,632	4,330

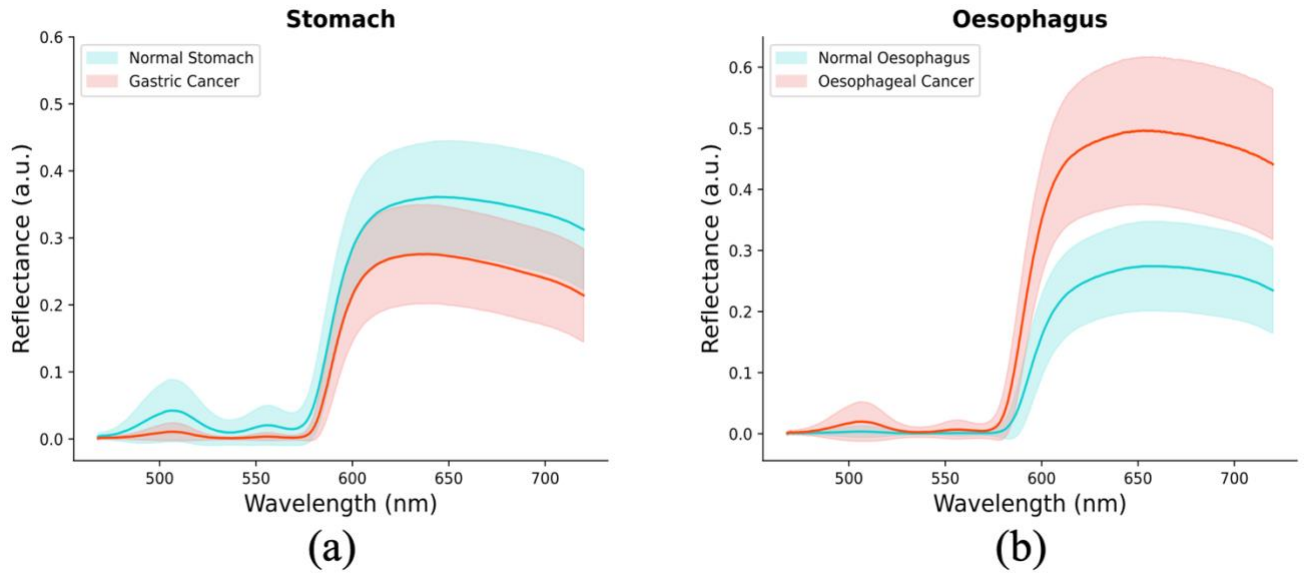


Figure 42 Average (lines) and standard deviations (shaded areas) of the mean spectra for (a) normal and tumour stomach tissue and (b) normal and tumour oesophagus tissue.

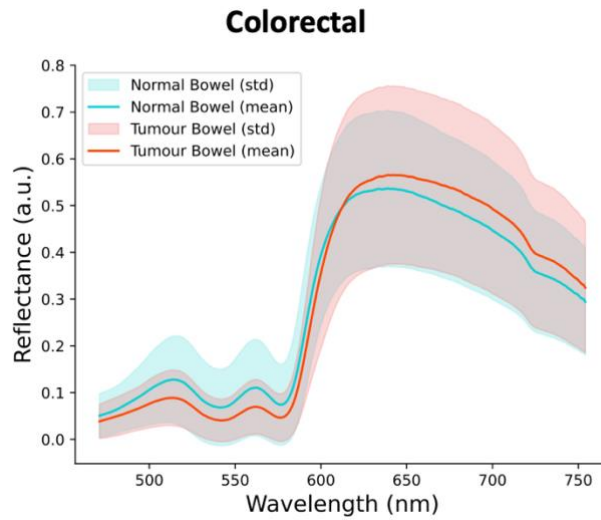


Figure 43 Average (lines) and standard deviations (shaded areas) of the mean spectra for normal and tumour colorectal tissue.

4.3.4 Classification Analysis

4.3.4.1 Upper GI Classification Results

The outcomes of the classification analysis for spectral data from the stomach and oesophagus are presented in Table 19. Notably, the XGB classifier emerged as the top-performing machine learning

algorithm for both stomach and oesophagus datasets. In a comparative evaluation against the array of algorithms assessed in this study, the XGB model consistently demonstrated superior performance across key metrics such as accuracy, sensitivity, and specificity. It achieved a normal versus cancer diagnostic accuracy of 93.86 ± 0.66 for stomach tissue and 96.22 ± 0.50 for oesophageal tissue.

Table 19 Performance metrics for the spectral data classification using supervised ML and permutation feature importance for the feature selection. Data showing the ability of multiple classifiers to detect normal vs. cancerous tissue presented as mean $\pm$ standard deviation. Overall accuracy is calculated as the proportion of correctly identified spectra over the total number of spectra.

[LGBM: light gradient boosting machine; MLP: multi-layer perceptron; SVM: support vector machine; XGB: extreme gradient boosting]

Classifier	Stomach				Oesophagus			
	Accuracy	Sensitivity	Specificity	AUC	Accuracy	Sensitivity	Specificity	AUC
LGBM	0.9363 $\pm$ 0.0	0.910 $\pm$ 0.01	0.949 $\pm$ 0.00	0.973 $\pm$ 0.00	0.962 $\pm$ 0.02	0.943 $\pm$ 0.00	0.974 $\pm$ 0.00	0.990 $\pm$ 0.00
MLP	0.900 $\pm$ 0.00	0.862 $\pm$ 0.04	0.917 $\pm$ 0.01	0.961 $\pm$ 0.00	0.919 $\pm$ 0.02	0.851 $\pm$ 0.04	0.959 $\pm$ 0.02	0.960 $\pm$ 0.00
SVM	0.876 $\pm$ 0.00	0.834 $\pm$ 0.01	0.895 $\pm$ 0.01	0.943 $\pm$ 0.00	0.883 $\pm$ 0.00	0.821 $\pm$ 0.04	0.923 $\pm$ 0.02	0.957 $\pm$ 0.00
XGB	0.938 $\pm$ 0.00	0.913 $\pm$ 0.01	0.951 $\pm$ 0.00	0.974 $\pm$ 0.00	0.962 $\pm$ 0.00	0.946 $\pm$ 0.00	0.972 $\pm$ 0.00	0.989 $\pm$ 0.00

The XGB classifier exhibited sensitivity and specificity values for stomach tissue measured at 91.31% and 95.13%, respectively. Meanwhile, for oesophageal tissue, the classifier exhibited even higher sensitivity and specificity values, reaching 94.60% and 97.28%, respectively. Reflecting the robustness of the XGB model, the AUC values were 0.974 for stomach and 0.989 for oesophagus. The ROC curves for all the classifiers are shown in Figure 44.

In addition, the XGB model also outperformed all other algorithms used for classification in computational efficiency. Benefiting from its intrinsic parallel processing capabilities, the XGB model exhibited a significantly shorter computation time of only 3.5 seconds, encompassing both the training and validation phases, setting it apart as an efficient and effective tool in the realm of spectral data classification.

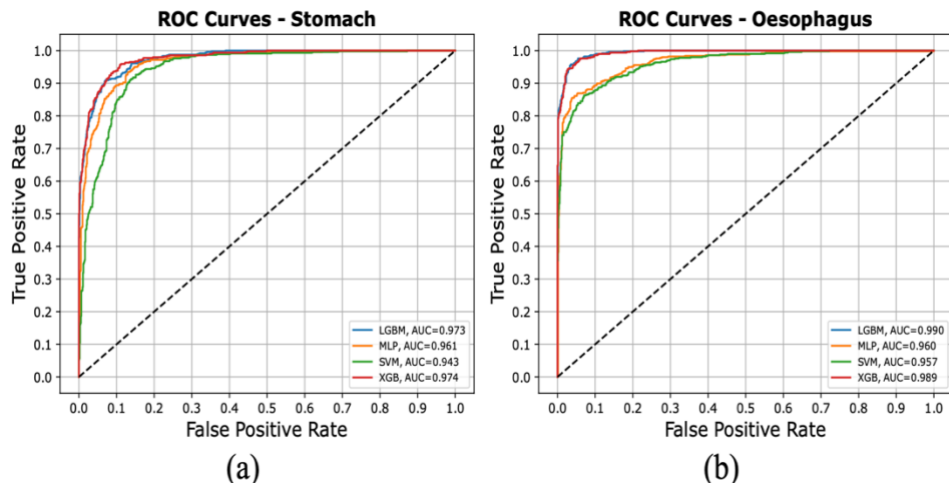


Figure 44 ROC curves with AUC values for all classifiers for (a)stomach and (b)oesophagus datasets.

Real-time tissue classification was achieved and presented on the user interface when using the DRS probe. Real-time tracking at each optical biopsy site coupled with the binary classification of each site was visualised as either normal (100% green) or tumour tissue (100% pink) using a graduated colourmap. Tissue type was highlighted on the screen in real-time during sampling, as highlighted in Figure 45.

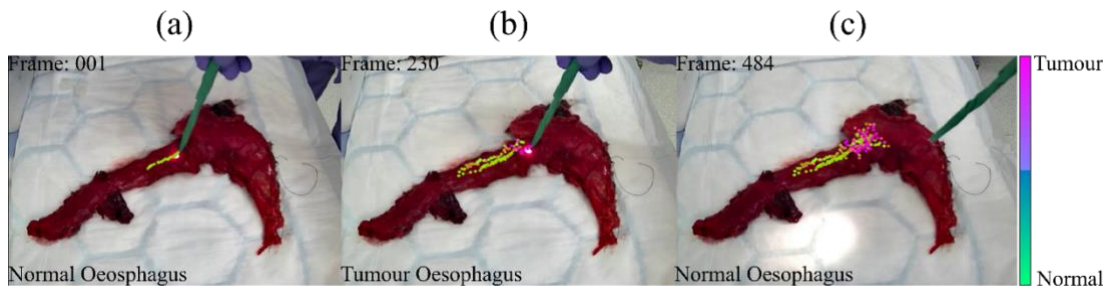


Figure 45 Three illustrations of real-time DRS probe tracking and classification at different frames during scanning of an oesophageal and gastric ex vivo tissue specimen. (a-c) Real-time tracking at each optical biopsy site coupled with the binary classification of each site on an oesophageal and gastric specimen. Tissue type highlighted on the screen in real-time (bottom left of each frame). © Reproduced from [330] with permission.

4.3.4.1.1 Classification of Tumour Data for Pathological Staging

Whilst binary classification provided valuable results, the complex nature of tumour heterogeneity necessitated a more nuanced approach. Therefore, a multi-class classification was carried out on T1, T2, T3 and T4 tumours, with the aim of determining subtler differentiations within tumour

specimens. Establishing pathological staging in cancer surgery is crucial for prognosis and in guiding management, thereby contributing to improved patient care.

Figure 46 depicts the number of spectra acquired for each stage of disease as per the TNM staging system for gastric and oesophageal tumours. For gastric cancer, a total of 1,070 spectra were recorded for T1 tumours, 646 spectra for T2, 550 spectra for T3 and 917 spectra for T4. For oesophageal cancer, there were 291 spectra acquired for T1 tumours, 166 for T2, 1,638 spectra for T3 and 421 spectra for T4. The means of all spectra for each of the tissue classes is represented in Figure 47.

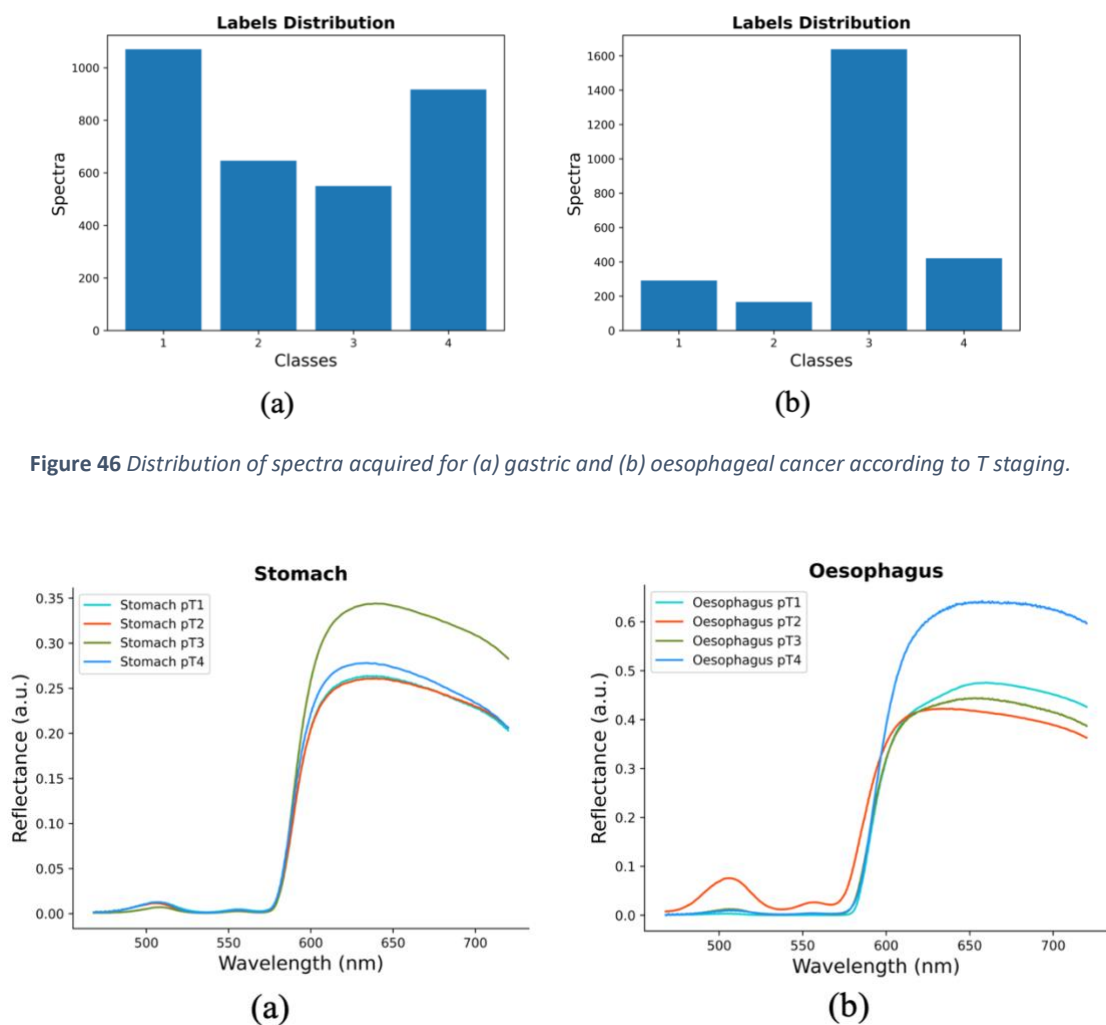


Figure 47 Mean spectra of the different pathological tumour stages for (a) gastric and (b) oesophageal cancer.

Analysis using ROC curves, as represented in Figure 48, were carried out showing high AUC rates for the differentiation of T-staging across all stages of tumour using ML classifiers in a one-vs-rest multi-class technique.

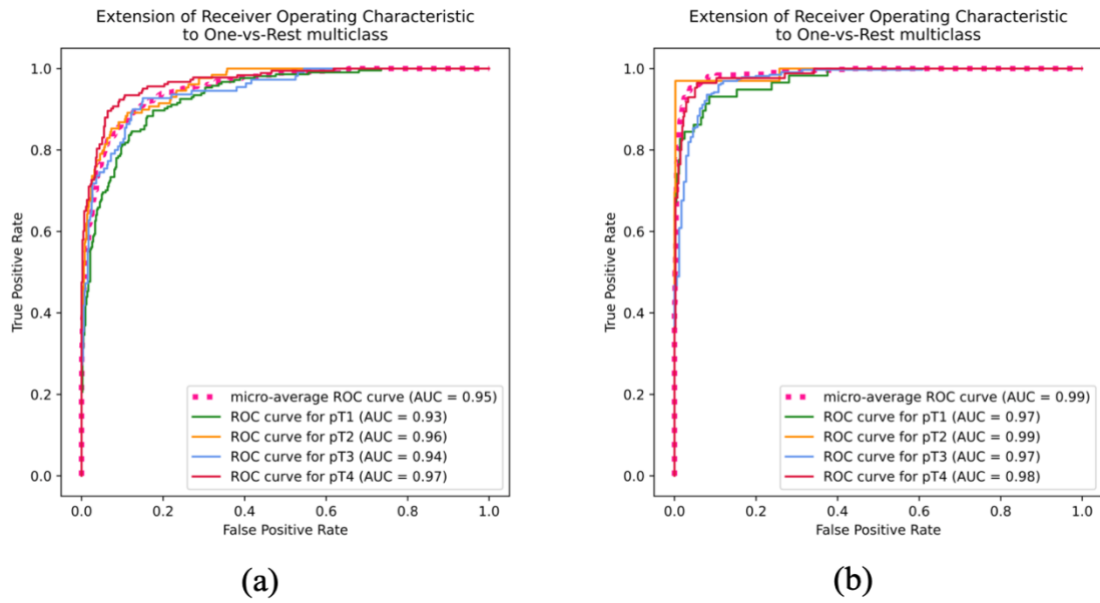


Figure 48 ROC curves for one-vs-rest multi-class classification for (a) gastric and (b) oesophageal cancer pathological staging.

4.3.4.2 Lower GI Classification Results

The findings from the classification analysis of colorectal spectral data are outlined in Table 20. Notably, the 1D CNN and tree-based classifiers (RF, XGB, LGBM) demonstrated superior performance compared to SVM, exhibiting accuracies of over 88% and AUC rates of over 94%. Although no significant difference was observed between SpecNet and tree-based classifiers, the SpecNet classifier generally exhibited higher accuracy (90.8%), sensitivity (93%), and specificity (88%). Additionally, the AUC of the SpecNet classifier was 96.8%, underscoring its robust discrimination ability.

As depicted in Figure 49, the ROC curves provide a graphical representation of classifier performance. The improved performance of the SpecNet classifier over the traditional ML classifiers was modest. One potential explanation for this could be attributed to the relatively limited amount of spectral data utilised for training the SpecNet classifier. Deep neural networks, such as SpecNet, typically necessitate larger datasets to fully leverage their capabilities and demonstrate optimal performance. This highlights the importance of dataset size in maximising the efficacy of deep learning algorithms.

Table 20 Performance metrics for the spectral data classification using supervised ML and permutation feature importance for the feature selection. Data showing the ability of multiple classifiers to detect normal vs. cancerous tissue presented as mean (95% confidence interval). Overall accuracy is calculated as the proportion of correctly identified spectra over the total number of spectra.

Colorectal Tissue				
Classifier	Accuracy (%)	Sensitivity (%)	Specificity (%)	AUC (%)
SpecNet	90.8 (88.7 – 92.9)	93.0 (90.7 – 95.3)	88.0 (84.1 – 91.9)	96.8 (95.5 – 98.1)
RF	89.8 (87.5 – 92.1)	90.0 (87.9 – 92.1)	89.6 (86.4 – 92.8)	96.2 (94.9 – 97.5)
LGBM	89.3 (87.5 – 91.1)	89.9 (87.9 – 91.9)	88.6 (85.6 – 91.6)	96.2 (95.1 – 97.3)
XGB	88.7 (87.0 – 90.4)	89.2 (87.2 – 91.2)	88.2 (85.2 – 91.2)	95.8 (94.7 – 96.9)
MLP	88.8 (86.2 – 91.4)	89.9 (85.6 – 94.2)	87.4 (82.8 – 92.0)	94.5 (92.5 – 96.5)
SVM	80.5 (75.9 – 85.1)	80.5 (72.5 – 88.5)	80.5 (72.3 – 88.7)	86.5 (83.3 – 89.7)

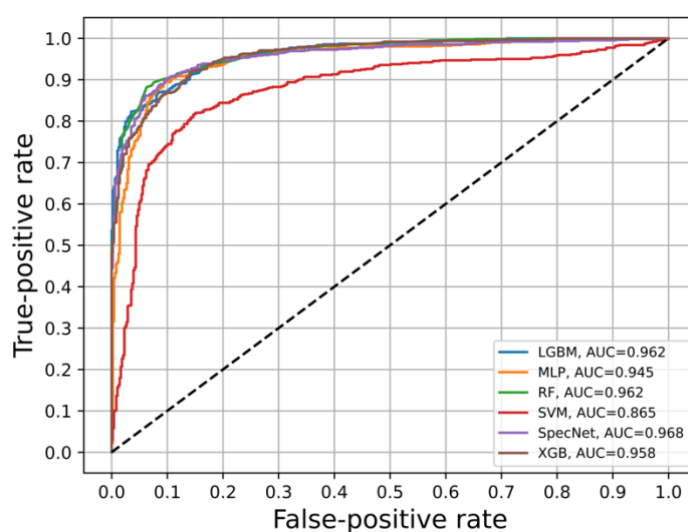


Figure 49 ROC curve with AUC values for all classifiers for colorectal tissue data.

Table 20 illustrates that the specificity of the RF classifier was higher than that for SpecNet. This observation may stem from various factors, including differences in feature extraction and dimensionality, given that SpecNet is a one-dimensional CNN whilst RF utilises an ensemble learning approach that inherently accommodates the complex relationships and interactions within the dataset. Furthermore, the performance variation could be attributed to differences in data availability and sample size, since RF exhibits robustness in scenarios with limited data points, making it particularly suitable for situations involving relatively small sample sizes.

Real-time tissue classification was achieved and presented on the user interface when the DRS probe was used to scan the tissue. Real-time tracking at each optical biopsy site together with the binary classification probability of each site was visualised as either normal (100% green) or tumour tissue

(100% pink) using a graduated colourmap. Tissue type was highlighted on the screen in real-time during sampling, as illustrated in Figure 50.

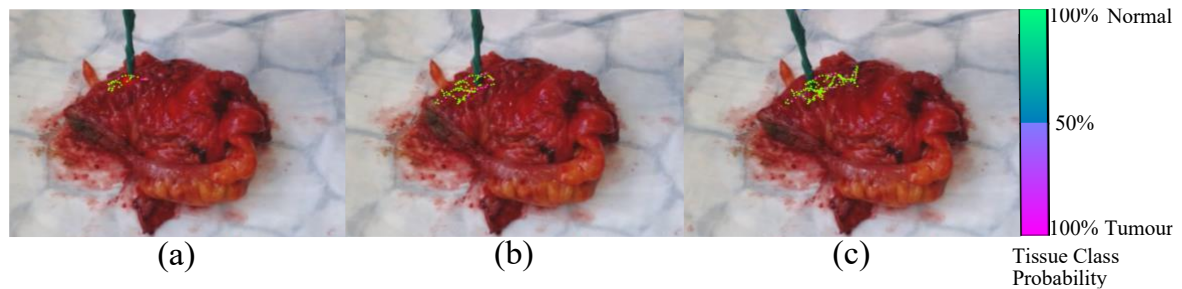


Figure 50 Real-time tracking at each optical biopsy site coupled with the binary classification probability of each site as either normal (100% green) or tumour (100% tumour) tissue on a colorectal specimen. © Reproduced from [334] with permission.

This study analysed 4,330 mean spectra for tumour tissue, increasing the accuracy of results. The classification in this study was run on a subset of 13 patients (3,117 mean spectra of tumour tissue), and then on the remaining 10 patients, resulting in a classification error of ± 0.05 . This ensured the statistical reproducibility of the proposed machine learning models by following the aforementioned data transformations and model training hyperparameters.

4.4 Discussion

4.4.1 Study Findings

In this study, we evaluated the development of a real-time tracking system and the diagnostic accuracy of a DRS probe in discriminating normal and tumour tissue in stomach, oesophageal and colorectal specimens. We highlighted the successful integration of a real-time tracking system for a DRS probe able to scan tissue intra-operatively and to present the optical tracking sites to the surgeon on a video feed, overcoming the point-based limitation that other studies using DRS have faced. The tracking system showed a measured root mean square error of 1.18 mm and 1.05 mm for the x and y directions, respectively, and a maximum measured error of 1.76 mm when used. The colour green achieved an excellent performance when used as a marker for probe tracking.

This tracking system allowed classification of spectral data for tumour and non-tumour tissue in real-time. Machine learning classifiers and a novel 1D CNN-based classifier were able to differentiate tumour and normal tissue with high diagnostic accuracy, achieving an overall accuracy of 94%, 96% and 91% for stomach, oesophageal and colorectal cancers, respectively. The graphical user interface

presented live feedback regarding the tissue type being sampled to the surgeon, to aid assessment of resection margins of tumour. In this way, a smaller proportion of healthy tissue could be resected, leading to a lower risk of morbidity, and a more accurate and sufficient margin could be identified, leading to lower risk of recurrence of disease. As a result, the DRS probe and tracking system has the potential to influence intra-operative decision-making and lead to better patient outcomes in colorectal cancer care.

4.4.2 Strengths of Study

4.4.2.1 Data Acquisition Speed

The DRS system presented in this study shows suitability for use intra-operatively to aid margin assessment in cancer surgery. We developed a system with high spectral resolution and high data acquisition speed of approximately 80 spectra per second. As a result, we were able to scan a large area on the tissue, and able to differentiate tissue in real-time, whilst creating minimal disruption to the surgical workflow.

A limitation of *ex vivo* sampling in previous studies has been that following resection of the tissue and the blood supply being cut off, there is potential for tissue degradation impacting the spectra that are acquired. [343] In this prospective validation study, a total of more than 7,500 mean spectra were acquired immediately after resection of colorectal specimens intra-operatively, minimising the disruption to perfusion of the specimens. Data collection was complete within 15 minutes of specimens having been resected. This process was notably faster compared to previous studies in which data had been collected within an average time of 60 minutes after surgical resection. [303,344] It also ensured that the results were mirroring the *in vivo* setting as much as possible. In addition to these findings, we have been able to show the possibility of real-time tissue discrimination live on-screen in the operating theatre during colorectal cancer procedures. We have been able to achieve this using a developed DRS tracking system and a graduated colour map for tissue classification represented on an easy-to-use graphical user interface. [330,332]

4.4.2.2 Sample Size

This study has been able to show results that set it apart from previous studies exploring the use of DRS for colorectal cancer detection. [303,305,306,345,346] Dhar et al. [346] showed an accuracy of 80% for differentiation of colonic cancer from normal tissue, which was based on 31 spectra from

colorectal cancer sites. Baltussen et al. [303] achieved a mean accuracy of 95% in differentiating colorectal tumour tissue from healthy bowel wall. However, the classifier used in their study was trained on 448 locations, with only 164 locations for tumour tissue. Langhout et al. [305] achieved an accuracy of 91% by analysing 603 separate tissue measurements, with 276 on tumour tissue, and Schols et al. [306] included 60 spectra from colonic tumour tissue. The small sample sizes in these studies limits their reliability and may lead to overfitting.

4.2.2.3 Histopathology Correlation

The national guidelines by the United Kingdom Royal College of Pathologists recommend preserving the integrity of the bowel at the tumour site to maintain the circumferential margin, the assessment of which would be otherwise compromised if the specimen was disturbed. [347,348] Various methodologies have been employed in studies exploring the use of DRS for colorectal cancer tissue assessment when aiming to correlate findings with histopathological analyses. In the study by Langhout et al. [305], biopsies were taken from macroscopically recognised tumour tissue and sent off for histopathological assessment, whilst the gold standard used by Baltussen et al.[303] involved samples of tissue being taken from the specimen and put into cassettes before being sent off to the histopathology department. These techniques would have disrupted the integrity of the specimens, deeming them inappropriate for CRM assessment, and are not a true representation of the surgical workflow during colorectal cancer procedures. In this study, we were able to use yellow tissue paint to mark the area of suspected tumour tissue following collection of spectral data and follow this through to the H&E slide confirming a tumour diagnosis. This technique was quick, taking only a few minutes, simple and meant that the circumferential margin of the tissue was preserved.

In addition, the light emitted by the DRS probe was able to penetrate colonic tissue allowing the system to detect tumour tissue in deeper muscular layers away from the serosal surface on which it was used. Histopathological analysis allowed insight into the systems' success in differentiating normal and tumour tissue in T2 or T3 tumours. However, we plan on further assessing this in future studies through sub-set analyses.

4.2.2.4 Neural Network Classifier

This study was further strengthened by the development and use of a CNN classifier which was able to achieve an accuracy when discriminating colorectal tissue of over 90%. Neural networks and deep

learning have been used in combination with optical techniques for intra-operative detection of cancers previously.[349,350] However, other studies exploring the use of neural networks for CRC detection have either analysed mucosal tissue,[351] or have evaluated neural networks in view of using datasets to determine an optimal classification strategy.[352] Our study, on the other hand, was able to utilise a CNN classifier in combination with DRS to accurately classify each spectral measurement as normal or cancerous in real-time.

4.2.2.5 Cost Benefit

Compared to current gold standard frozen sections, which can cost more than \$3,000 per patient to perform,[157] our system has much lower costs, with a portable set-up system and no additional workforce required. The system we have developed has a one-off cost of £8,000 and given that it can be used for multiple sampling events, the per-patient costs would be much less. Our DRS system is also completely non-invasive and does not disrupt the tissue itself or interfere with its structure, allowing for a full analysis when sent to the histopathology department, especially important for small tumours. Nevertheless, the aim is for the DRS system to be integrated into the surgical workflow in the future as a device to assess tumour margins, with the intention to help reduce the burden on histopathology departments once the system has been fully developed, with less requirement for frozen section analysis.

4.4.3 Limitations

There are several limitations to this study that must be taken into consideration. Firstly, these were single-centre studies which would have hindered the generalisability of our findings to broader patient populations and diverse clinical settings. Moreover, this may have exhibited inherent biases and limitations fundamental to the specific institutional practices and patient cohorts of a tertiary teaching hospital in which the studies were based. The lack of variation in surgical expertise, pathological assessment protocols, and perioperative care practices may have influenced the interpretation and applicability of study findings. Consequently, extrapolating the results of a single-centre study to a wider clinical context requires careful consideration of these contextual factors. In addition, the limited geographic scope of these single-centre studies may have restricted the diversity of patient populations recruited, potentially overlooking important demographic and clinical variations that could affect the performance outcomes of the DRS system. Collaborative multi-centre studies involving diverse patient cohorts and clinical settings will be essential for

validating the efficacy and utility of the DRS system as an intraoperative margin assessment technique.

Secondly, the colorectal specimens were sampled immediately after resection without any cleaning techniques being applied, resulting in differing quantities of stool or blood in the bowel lumen. Although this may have caused some inconsistencies with the data, the process reflects what would be expected if the system was applied to the *in vivo* clinical setting. Furthermore, we did not collect data on fat tissue and chose to collect normal tissue data from areas of the bowel without overlying fat to ensure consistent spectra. However, when collecting data for tumour, fat tissue could not always be avoided. In future studies, we aim to acquire spectral data for fat tissue to be able to apply this information to the results to ensure accurate and robust results.

Cases undergoing neoadjuvant therapy may have been affected by the presence of fibrosis tissue. We excluded samples which were confirmed after histopathological analysis as fibrosis tissue, and the sparsity of the data did not allow further analysis. However, we would benefit from analysing fibrosis tissue in future studies, especially since this type of tissue is difficult to differentiate intra-operatively by visual and haptic feedback by the surgeon when making decisions about tumour margin assessment.

In this study, specimens were placed on a white background, whilst colours in the surrounding field could be controlled. However, this may not necessarily mirror translation into an *in vivo* setting where differences in colour and artefacts, such as bleeding or other instruments, exist. This creates the challenge of tracking the colour marker accurately. In addition, the tracking system used provided 2D visual feedback only. We are currently collecting 3D stereoscopic data which will offer enhanced spatial awareness and depth perception during tissue scanning and probe localisation, potentially facilitating more accurate margin assessment.

Additionally, DRS measurements were performed from the exterior surface of the colorectal specimen, specifically the serosa, without necessitating opening of the lumen. An inherent challenge in such an approach lies in the potential variability in light penetration due to the heterogeneous nature of the tissue surface and the diverse underlying tissue properties. The depth to which light penetrates and the variability of this depth in manipulated tissues is critical for accurate DRS measurements. Studies were conducted to determine the penetration depth of light in different tissue types, and the results indicated certain predictable patterns in DRS changes associated with

various tissue conditions. [344,353] It is important to understand that the chemical and physical characteristics of the tissue above the tumour may not exhibit homogeneity across the entire specimen's outer surface. This variability can stem from factors like surgical manipulation, bleeding, or innate tissue characteristics. Such heterogeneities can potentially mask or alter the DRS signatures of underlying tumour tissue, posing challenges for real-time surgical applications. We are actively exploring calibration techniques that adjust for these variabilities, ensuring consistent and reliable measurements.

Finally, the absence of external validation on a separate cohort of patients, ideally sourced from another hospital or healthcare facility, presents a further limitation. External validation is crucial for assessing the generalisability and robustness of our findings beyond the confines of our single-centre study and ensures applicability of the DRS system when applied to broader patient populations. However, the use of the DRS system initially as a feasibility study for upper GI specimens followed by the validation of this technology and device on a lower GI patient cohort allowed for a comprehensive assessment of its applicability across different anatomical regions, thereby contributing to a more robust understanding of its potential clinical utility and performance. With this in mind, external validation of our classification model on a separate cohort of patients is a crucial step towards the clinical translation of this technology. It is also vital that decision analysis studies and key stakeholder interviews are conducted to identify the optimal use-cases for the technology, the barriers and facilitators to its adoption in the future. [354,355] The integration of the DRS technology into the surgical workflow to allow tumour mapping can be facilitated via augmented reality or an optional heat-map overlay added onto a laparoscopic or robotic camera feed, which constitutes only a minimal deviation from the standard plan of care. [356]

4.4.4 Future *In-Vivo* Use

A concurrent *in vivo* validation study of the real-time tracking technology and classification is planned, for which adaptation of the DRS probe system will be needed. As such, we have developed a sterilisable probe to allow *in vivo* spectral measurements. An *in vivo* setting will allow application of the DRS probe directly onto the surface of colorectal tissue prior to tissue resection and cessation of blood supply. Intra-operatively, the DRS system will be able to facilitate real-time tissue classification with live on-screen categorisation of “tumour” or “normal” tissue. This approach aims to enhance surgical decision-making by providing instantaneous feedback to the surgeon during the critical stages of the procedure, particularly in guiding margin assessment. The system will be able to

be used both during open and laparoscopic surgery. For the latter, the DRS probe will be inserted through a port site and be used to scan the tissue with the assistance of laparoscopic instruments.

Clinical application of the *in vivo* technique would be particularly useful in scenarios in which a high risk of incomplete resection is anticipated. It is vital that DRS is applied with high accuracy, but we have shown that it is also important for the system to be incorporated with standard surgical protocols and easily integrate into the surgical workflow in a way that will best aid surgeons and improve outcomes. This will be explored further in chapter 5.

4.5 Conclusion

We have highlighted that the use of a DRS probe and tracking system can be successfully adopted intra-operatively for real-time tumour margin assessment in colorectal cancer with the aim of potentially aiding surgeons in decision making processes. Given the success in our *ex vivo* findings, we intend on expanding the study to include rectal cancer resections, since a positive CRM in this cohort has been shown to be an independent, poor prognostic factor for local recurrence and survival.[357–359] We plan to use DRS for assessment of tissue infiltration of the peritoneum and other organs in future studies, to be able to assess for metastasis of tumour in real-time during cancer procedures. The development of advanced deep learning neural networks to achieve accurate probe tracking, as well as for classification of spectral data, will be undertaken. Moreover, correlation of the real-time tissue classification with clinical outcomes, such as recurrence, positive resection margin rate, or overall survival, will be required to show the clinical utility of this technology before its potential clinical adoption. [360]

Chapter 5 Classification of tumour and non-tumour tissue using diffuse reflectance spectroscopy in the gastrointestinal tract to aid resection margin assessment: an *in vivo* feasibility study

Work from this chapter was published in:

- **A YOLOv5-based network for the detection of a diffuse reflectance spectroscopy probe to aid surgical guidance in gastrointestinal cancer surgery.**

Ioannis Gkouzionis, Yican Zhong, **Scarlet Nazarian**, Ara Darzi, Nisha Patel, Christopher J. Peters, Daniel S. Elson

Int J Comput Assist Radiol Surg. 2024; doi: 10.1007/s11548-023-02944-9.

5.1 Background and Aims

In chapter 4 we highlighted the success of a DRS probe and tracking system in differentiating tumour and non-tumour tissue in real-time to aid margin assessment through *ex vivo* validation studies.

[332,334] Machine learning classifiers were able to differentiate tumour and normal tissue with high diagnostic accuracy, achieving an overall accuracy of 94%, 96% and 91% for stomach, oesophageal and colorectal cancers, respectively.

Despite the promising outcomes observed in the *ex vivo* studies, challenges regarding the applicability and reliability of the data persist, particularly considering the controlled nature of the experimental environment. The DRS configuration utilised in these studies may not accurately replicate real-world conditions encountered during real-world procedures, where factors such as variations in tissue colour and the presence of potential artefacts, like bleeding or other instruments, may influence the efficacy of the technique. Demonstrating the seamless integration of the DRS probe and tracking system with established surgical protocols is crucial. This ensures its compatibility with existing surgical practices and facilitates smooth integration into the surgical workflow, ultimately enhancing surgical precision and patient outcomes.

Previous research has evaluated the role of DRS in differentiating tissue type in an *in vivo* setting. Several studies have explored the feasibility of utilising DRS in margin assessment for breast surgery. [361–363] However, a recurring challenge in these studies has been the difficulty in accurately correlating the tissue volume measured by DRS *in vivo* with its precise location on histopathology

slides obtained post-operatively. This alignment is critical as histopathological assessment serves as the gold standard for evaluating surgical margins. Consequently, discrepancies between optical measurements and histopathological findings impede the development of reliable classification algorithms and validation of the technology. In addition, many studies prematurely disrupt the tissue via penetration during the optical biopsy technique. [165] This disruption can alter the tissue architecture and introduce artefacts, complicating the interpretation of histopathological findings. This, coupled with the slow acquisition speeds of around one spectra per second, further challenges the seamless integration of DRS into clinical practice.

Some research has also explored the use of DRS for colorectal cancer with greater success in histopathological correlation. However, these studies have highlighted small sample sizes which limits the reliability of the results. [364] The study by Baltussen et al. analysed 835 spectra in total, with only 151 spectra from colorectal tumour tissue. [304] Studies with larger sample sizes are imperative to validate the efficacy of DRS in colorectal cancer assessment and improve its clinical applicability.

The *in vivo* application of a DRS system is dependent on many factors. These include sterilisation and infection control, safety, durability, ergonomics, regulatory compliance and compatibility with surgical workflow. Each of these factors will be discussed in further detail.

5.1.1 Sterilisation and Infection Control

For utilisation of the DRS probe intra-operatively, standards and protocols must be followed regarding sterility of surgical instruments. The prevention and control of infection is an essential part of the management of patients. Infection control teams within hospitals ensure these stringent standards are met. The DRS probe must withstand a variety of sterilisation methods commonly used including autoclaving, in which steam sterilisation is utilised at very high temperatures, dry heat or chemical detergents. These methods should not lead to degradation, discolouration or damage of the DRS probe, rather the probe should be able to maintain structural integrity and functionality after repeated sterilisation cycles.

5.1.2 Safety

The design and construct of the DRS probe must be biocompatible to ensure it does not elicit adverse reactions or tissue damage when in contact with biological tissues. Biocompatibility testing ensures that materials used in the probe construction are safe for prolonged intraoperative use without causing harm to the patient. A risk assessment should be completed and risk mitigation actions should be put in place to avoid potential risks during intra-operative use of the probe. This may include time limits for the scanning procedure to avoid overheating of the probe, handling of the probe by experienced clinicians and researchers and use of a separate electrical transformer to avoid disruption to the surgical equipment in case of electrical faults.

5.1.3 Durability

The DRS probe should be durable and maintain consistent performance throughout the surgical procedure without malfunction or failure. It should be designed to withstand the rigors of surgical procedures, including repeated use and exposure to bodily fluids and mechanical stress.

5.1.4 Ergonomics

An *in vivo* DRS probe should seamlessly integrate into the surgical workflow. As a result, the design of the probe should prioritise ease of use. Ergonomic features such as comfortable grip, lightweight design and the probe's manoeuvrability can enhance user experience and ensure secure handling, even in challenging surgical environments. Steady and precise manipulation of the probe promotes accurate tissue assessment and contributes to overall surgical efficiency and safety.

5.1.5 Regulatory Compliance

All medical devices in the UK are regulated by the Medicines and Healthcare products Regulatory Agency (MHRA). They ensure that all regulatory standards and guidelines are adhered to during a clinical investigation for a medical device. Compliance with regulatory bodies ensures the safety, efficacy and quality of a medical device for clinical use.

In this study, a notification of a clinical investigation to MHRA Devices was not required for the following reasons:

- The study is not sponsored or supported by the manufacturer, and the device has been purchased for use in the study.
- The study will only be conducted at clinical sites within Imperial NHS Trust.

- There is no intention to commercialise the product using the data generated from this study.

5.1.6 Compatibility with Surgical Workflow

Ensuring the seamless integration of a new device like the DRS probe into the existing surgical workflow is paramount for its successful adoption and utilisation. The introduction of such a tool should complement established practices and procedures without imposing significant burdens on surgical staff. This means that the device should fit seamlessly into the workflow, requiring minimal adjustments or additional steps to incorporate it into routine surgical protocols. Compatibility with existing equipment and protocols streamlines the adoption process, reduces disruption to surgical routines, and ultimately facilitates efficient and effective patient care. Moreover, a device that seamlessly integrates into the surgical workflow enhances user acceptance, promotes confidence in its use, and maximises its potential clinical benefits. Therefore, prioritising compatibility with the surgical environment ensures smooth implementation and optimised performance of the DRS probe in clinical practice.

5.1.7 Aims of the Study

The aims of this feasibility study were multifarious:

1. To determine the feasibility of the DRS system to be utilised in an *in vivo* setting intra-operatively.
2. To evaluate the diagnostic accuracy of the DRS system to differentiate tumour and non-tumour GI tissue *in vivo*, with the potential to aid margin assessment.
3. To assess the ergonomics of the DRS probe within the surgical workflow with a usability questionnaire.

5.2 Methods

The role of data analysis and system development in this chapter was carried out by Dr Ioannis Gkouzionis and Mr Maxime Giot, with input from the author of this thesis. Data collection and development of histopathological correlation was conducted by the author of this thesis in close collaboration with Mr Maxime Giot.

5.2.1 Patient and Public Involvement

Patient and public involvement (PPI) was incorporated as an important aspect in the designing and planning of this study. This was carried out to improve the relevance, clarity and feasibility of the research design and outcome measures, and to ensure a diverse range of perspectives. We organised a focus group, comprising of members of the public and patients, who would be able to provide insight into key areas of our research. The focus group was organised in two parts. Initially, we invited the group to review our research protocol, consent form and information sheet via email. We provided adequate time for the information to be read and critiqued. Subsequently, a follow-up email was sent to gather feedback about the research proposal and design. The feedback received was reflected upon and used to implement necessary changes to the study's design and information pack.

The second part of the PPI focus group was arranged as a face-to-face meeting with the group to engage in insightful discussion about the study's benefits and challenges. A presentation was given to outline the study's aims more clearly. Following this, we discussed the study protocol in detail, and ways of ensuring successful recruitment of patients through the use of clear images, translators and open dialogue with a next of kin when appropriate. This was invaluable in helping to shape the way the study would be conducted and to guide the ethical application. Public involvement has allowed our project to foster transparency and inclusivity.

5.2.2 Phantom Set-Up of DRS System

We carried out a phantom feasibility study, in which we embarked on a crucial exploration of the potential for utilising the DRS system *in vivo*, focusing on the laparoscopic and open surgical scenarios. Employing a laparoscopic stack and a phantom surgical abdomen model, our objective was twofold: to assess the feasibility of integrating the DRS system into real-time surgical procedures and to identify the challenges that must be addressed to transition from an *ex vivo* setup to an *in vivo* environment (Figure 51).

Throughout our study, we encountered several notable challenges that shed light on the complexities of implementing the DRS system in surgical practice. One such obstacle revolved around maintaining sterility, a fundamental requirement in surgical settings. While attempting to employ a sterile probe cover for the DRS probe, we encountered difficulties since the sterile plastic

cover was too bulky and did not allow access through the laparoscopic port, posing logistical hurdles. Furthermore, the utilisation of the probe cover may have inadvertently impacted the intensity of light emitted and reabsorbed by the DRS probe, potentially compromising the accuracy of spectral measurements. We concluded that a sterilisable DRS probe, which would preserve both the size of the probe and the accuracy of spectral data acquisition, would be required for *in vivo* use.

Additionally, the unique characteristics of laparoscopic procedures presented distinct challenges. The smooth and flexible nature of the DRS probe made it challenging to grasp securely within the abdomen, highlighting the importance of ergonomics and instrument design in laparoscopic settings. Moreover, for open procedures, we recognised the necessity for innovative solutions to stabilise the camera, such as employing a long tripod to arch over the patient, ensuring optimal visualisation and data acquisition. These findings underscore the critical need for meticulous planning and innovation to overcome technical and logistical challenges and successfully integrate the DRS system into clinical practice for real-time tissue assessment during surgical interventions.

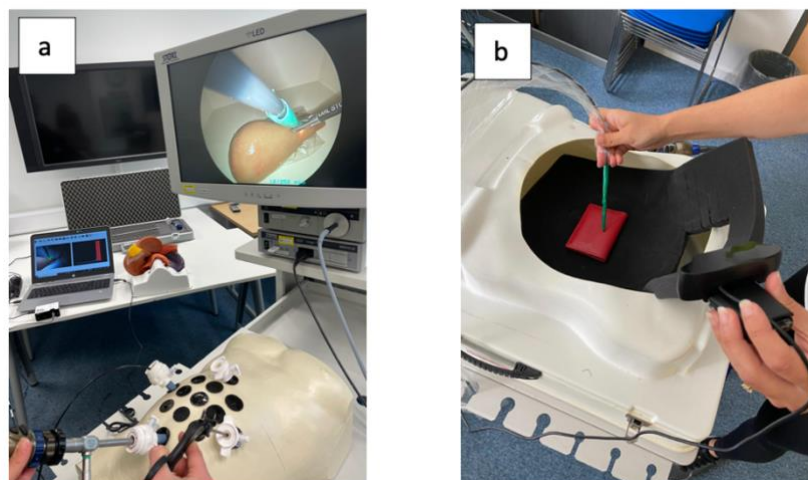


Figure 51 Phantom feasibility study to evaluate the integration of the DRS system *in vivo*. (a) Phantom model for laparoscopic set-up. (b) Phantom model for open surgical set-up.

5.2.3 Study Design

This was an *in vivo* feasibility study evaluating the use of a DRS system in differentiating tissue types. Patients undergoing oesophageal or gastric cancer resection were prospectively recruited into the study between May 2022 and August 2024 at Hammersmith Hospital in London, United Kingdom. Inclusion criteria included patients over the age of 18 undergoing oesophageal or gastric cancer

resection surgery. Open cases were included only for this study due to logistical challenges for laparoscopic use of the DRS probe. All patients gave written consent to partake in the study. For any cases confirmed as regressed or T0 on formal histology, normal tissue data was analysed only, whilst data labelled as suspected tumour was excluded. Ethical approval was gained from the HRA and Health and Care Research Wales (HCRW) - reference 22/SC/0048.

5.2.4 DRS Set-Up

A sterilisable DRS probe was designed and manufactured for *in vivo* use. This was a custom-made probe (Bundle UVNSS 600, CeramOptec GmbH) with six peripheral illumination fibres around one light collection fibre, 600 µm core diameter, 0.22 numerical aperture within a 3.175 mm ferrule (Figure 52). The sterilisable probe was able to withstand temperatures up to 200 degrees Celsius and was approved to undergo standard hospital sterilisation protocols. The probe was presented in theatre within a sterile surgical tray.

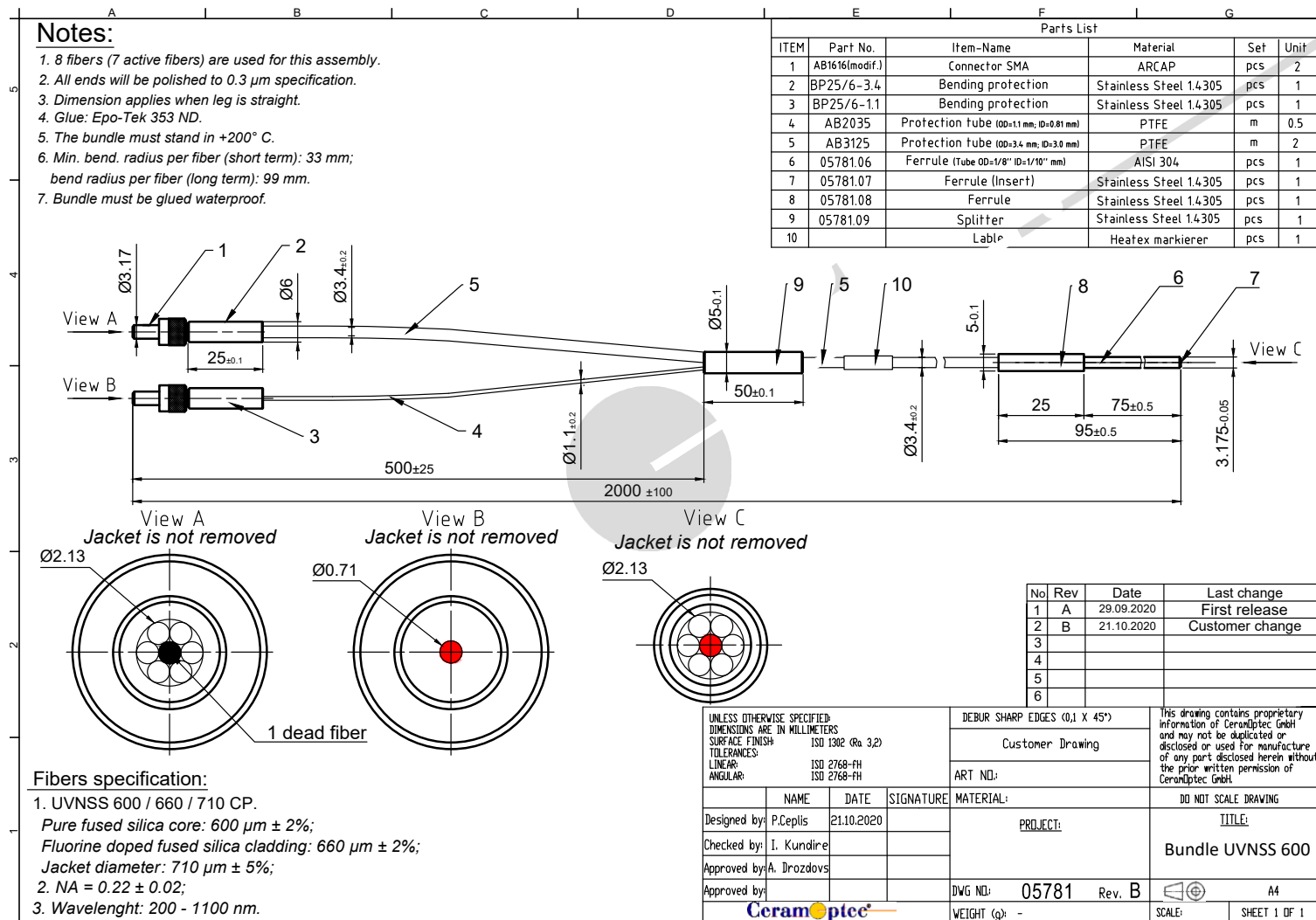


Figure 52 Specifications of the sterilisable DRS probe for in vivo use during surgery.

The probe was connected to the remaining set-up consisting of a tungsten halogen light source (Ocean Optics Inc., HL-2000-HP), a spectrometer (Ocean Optics Inc., USB4000), a camera (C920, Logitech International S.A., Lausanne, Switzerland) and a laptop computer.

In addition to the set-up used for *ex vivo* data collection, certain adjustments were made to accommodate the system in a theatre setting. An extendable tripod was used to allow overhead recording of colour videos. This, together with the camera, were encased in a sterile, disposable cover to ensure sterility. Sterile, single-use drapes were used for the trolley housing all the equipment. The trolley was wheeled next to the operating table when acquiring data. Moreover, to avoid any disruption to the operating theatre electrical system, a separate electrical transformer was utilised. The set-up is illustrated in Figure 53.



Figure 53 DRS system set-up for *in vivo* use.

5.2.5 Graphical User Interface

A GUI was used, as described previously in Chapter 4. This was developed using Python 3.6 (Python Software Foundation, Wilmington, USA) and Qt5 toolkit to incorporate all functionalities on a standard PC (Intel i3 processor at 3.30 GHz with 8 GB RAM). Before starting data acquisition, the experimental parameters were inputted into the system. This included patient demographics, type of procedure, tissue type, spectrometer integration time, spectral range and DRS calibration. The GUI allowed the user to directly visualise optical biopsy sites live on-screen in real-time.

5.2.6 Data Acquisition

The DRS system was set up as described above, and spectral data was collected intra-operatively once appropriate dissection and access had been achieved. The patient demographics, the procedure, the tissue type and all other information was inputted into the GUI software. The sterile DRS probe was removed from the surgical tray and the ends of the probe were connected to the spectrometer and light source. Initially, a single image frame was captured of the tissue where sampling would occur. Following this, spectral data was acquired. Measurements were taken from the stomach and/or the oesophagus before termination of blood supply to these tissues. (Figure 54)

The DRS probe was applied along the surface of the tissue to obtain normal tissue data and suspected tumour tissue data. These areas were identified by visual and haptic inspection by a surgeon. Normal tissue was chosen as far away from the suspected tumour area as possible, and based on an area with minimal overlying fat tissue, and no staples or sutures. The sampling process left no marks on the tissue and did not disrupt the surface layer in any way. Data collection was completed within 15 minutes. The DRS probe was sent off for sterilisation at the end of each procedure.

The DRS probe acquired around 80 spectra per second. At each sampling point, 20 measurements, equating to 20 spectra, were acquired and averaged. The mean spectrum was displayed in real-time. A minimum of 200 spectral measurements were acquired for each tissue type (normal or tumour) per organ. This was dependent on the size of the sampled region and access to the tissue area. Where possible, DRS measurements were taken for both oesophageal and gastric tissue.

Data were acquired in the spectral range of 420 – 720 nm, with a wavelength separation of 0.5 nm for the oesophageal and gastric tissues. Subsequently, to account for inter-patient comparability, background light, and signal quality variability, spectral normalisation and noise reduction was performed for each acquired spectrum.

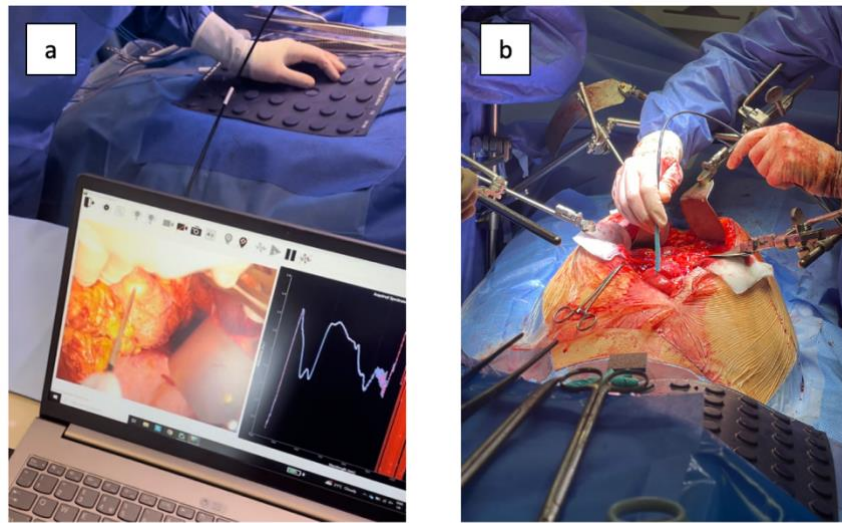


Figure 54 *In vivo* data acquisition. (a) Spectral data and optical biopsy site highlighted live on screen with GUI software. (b) Open abdomen with DRS sampling in process using sterile DRS probe.

5.2.7 Histopathology Correlation

At the end of the *in vivo* sampling process, a sterile marker was used to demarcate the suspected tumour area on the tissue where spectral measurements had been taken. This was non-traumatic to the tissue. Once the specimen was resected, a RGB photograph was taken with a ruler next to the specimen. This acted as a reference for pixel information in the 2D images and assisted in correlating measurements from the specimen. The marked area was then painted with yellow tissue dye (Cancer Diagnostics Inc) to allow histopathological correlation. (Figure 55)

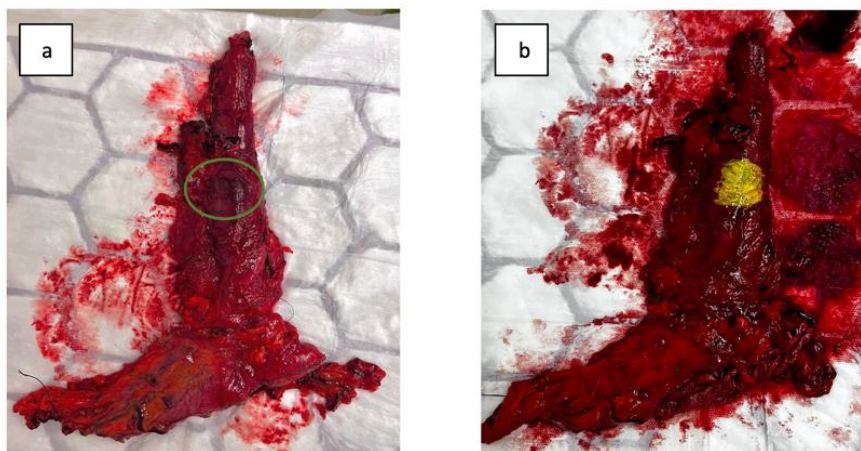


Figure 55 *Histopathological correlation technique for suspected tumour area. (a) Oesophageal specimen immediately after resection with marked area of sampled tissue for suspected tumour using sterile marker pen. (b) Marked area on oesophagus painted with permanent yellow tissue dye for histopathological analysis.*

To enhance the histopathological correlation technique used in chapter 4, the methodology was refined to account for shrinkage of the specimen after placement in formalin and to ensure a more precise delineation of the relationship between sampling location and confirmed tumour tissue. The steps below were followed:

1. The distance from the proximal margin of the specimen to the proximal border of the yellow tissue dye was measured.
2. The length of the painted area in the proximal to distal plane was measured.
3. A 3 cm line was drawn on the specimen before placement in formalin, as a reference test.

These steps are illustrated in Figure 56. Once these steps were completed, the specimen was placed in formalin and sent to the histopathology department to be processed according to standard protocols. Macroscopic photos of the whole specimen and slices of the specimen were taken.

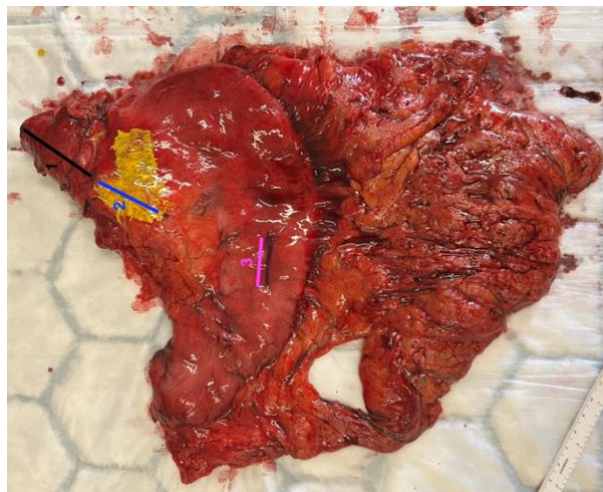


Figure 56 A gastric specimen prior to fixation in formalin and processing by the pathologist. (The greater omentum is present and attached to the greater curvature of the stomach). The yellow painted region corresponds to the suspected tumour area. Line (1) corresponds to the distance from proximal margin to proximal border of yellow painted region. Line (2) corresponds to the length of the painted area. Line (3) corresponds to the 3 cm line drawn.

During processing of the specimen by the dissector, the three measurements were taken again following time in formalin to allow calculation of the shrinkage coefficient. Slices of approximately 4 mm width were taken from the yellow painted region and labelled (e.g. A1, A2, A3 etc.). These slices were placed in individual cassettes and underwent standardised procedures for analysis. The histopathology report noted, if any, which blocks contained tumour. This information was then used to define whether the painted area on the specimen contained confirmed tumour. Once the area on the painted region was confirmed as tumour, the corresponding images were correlated with the

video frames and recordings containing the tracked positions of the suspected tumour. The video frames were then manually labelled as tumour or non-tumour tissue. These labels were considered ground truth labels for the training of machine learning classifiers.

Tracked locations on the specimen, which could not be confirmed by the histopathological analysis, were excluded. For specimens in which the suspected tumour area was confirmed by the pathologist as fully regressed tumour following neoadjuvant therapy, only normal tissue data collected for the specimen were analysed.

5.2.8 Marker-less Probe Tracking

Whilst a colour marker for DRS probe tracking was effective in providing real-time feedback on optical biopsy sites in *ex vivo* studies, the *in vivo* set-up for this study presented challenges that limited its effectiveness. Limitations included issues with sterility of the colour marker, with the possibility of increasing infection risk, disruption of the surgical workflow presented by having to attach and detach the colour marker onto the probe each time, and issues with lighting conditions and occlusion of the probe, caused by other tissues in the abdomen or surgical instruments, affecting the detection and tracking of the probe.

To address these limitations and ensure robustness of results, DL methodologies and CNNs were utilised to learn and extract relevant features directly from image data to enable a marker-less approach to detection and tracking of the DRS probe *in vivo*. These ML models would allow continuous tracking of the probe regardless of lighting variations, motion or challenging surgical conditions. Marker-less tracking did not require any artificial markers or modifications to the probe. Instead, it relied on natural, local features of the probe to characterise and track it within the image sequence. In this way, any potential disruption to the surgical procedure or workflow was minimised.

To carry out marker-less probe tracking, the original videos of the DRS probe sampling the tissue were taken and split into multiple frames. This led to a total of 1942 frames which were then labelled using two annotation types: labelling the boundary box around the metal shaft of the probe and labelling the tip of the DRS probe. In this way, labelled datasets were created. A model was then designed using the You Only Look Once (YOLO) v5 network architecture. [365] YOLO approaches object detection as a single, end-to-end regression problem, offering a fast and versatile detection

framework. To enhance the accuracy of DRS probe tip detection, a four-point tip regression was added to the network to reduce errors associated with probe localisation.[366]

5.2.9 Spectral Classification

Upon data collection, spectral normalisation and noise reduction were performed, using a white reflectance standard and dark field readings to account for inter-patient, background light, and signal quality variability. After that, the spectral data were normalised by removing the mean and scaling to unit variance, using the standard normal variate (SNV) method. An automated method was implemented for outlier (e.g., erroneous measurements or air interference) detection and removal. For this method, the median of the residuals was calculated, along with the 25th and 75th percentiles. The difference between each historical value and the residual mean was then calculated. If the historical value was 1.5 times the median absolute deviation (MAD) away from the median of the residuals, that value was classified as an outlier. In addition to the intensity measurements, spectral shape-based features were extracted and used for classification, comprising the intensity value and wavelength of the greatest peak for each spectrum, the intensity and wavelength of the secondary peak, and the mean intensity in three spectral regions, namely 480-550 nm, 580-700 nm, and 450-720 nm.

The proposed framework, illustrated in Figure 57, consisted of two key components: 1) stratified group k-fold cross-validation (SGKCV) and 2) machine learning classifiers. This framework encompassed seven main stages: data preparation, pre-processing, feature extraction and selection, training/building of three classifier models, implementing SGKCV techniques, and evaluating performance using overall accuracy, sensitivity, specificity, and the ROC curve.

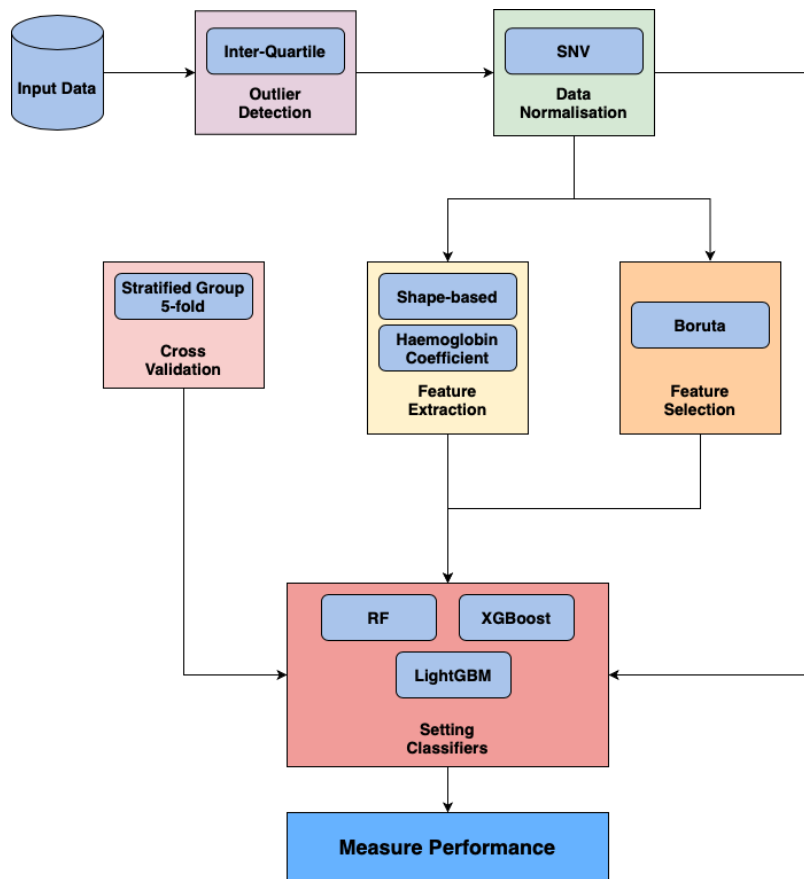


Figure 57 Data analysis framework for GI tissue classification.

5.2.10 Usability Questionnaire Methodology

A usability questionnaire was developed based on existing validated questionnaires, to determine whether the DRS probe was compatible with the surgical workflow. [367,368] We identified key users or stakeholders with different levels of expertise and knowledge within the theatre setting who would be handling or using the DRS probe.

An anonymous online questionnaire was carried out. The survey contained 22 questions featuring a range of question types, including multiple choice, 6-point Likert scale and open questions.

Participants were asked questions about their profession, the usefulness of the probe, the ease of use, ease of learning and satisfaction. A full copy of the questionnaire can be found in *Appendix 6 Usability Questionnaire*.

All end-users of the probe were asked to participate in the study. The questionnaire was hosted online. The Qualtrics survey software was used for the questionnaire since data could be easily aggregated and exported for analysis. The questionnaire was divided into two parts; part 1 gave information about the study, and part 2 was the questionnaire itself. Data was gathered onto a standard spreadsheet.

5.2.11 Statistical Analysis

For spectral data analysis, a stratified five-fold cross-validation was performed using tree-based classifiers, including random forest (RF) and gradient boosting (XGB, LGBM), as well as the CNN classifier, SpecNet. These classifiers were evaluated in terms of sensitivity, specificity, overall accuracy, and the area under the curve (AUC). Overall accuracy was calculated as the proportion of correctly identified spectra over a total number of spectra. Receiver-operator characteristic (ROC) curves were plotted.

For the usability questionnaire, descriptive statistics were performed to summarise the questionnaire findings. Graphs, including bar graphs, were used to represent the findings.

5.3 Results

5.3.1 Patient Characteristics

A total of 39 patients were recruited into this *in vivo* study, as detailed in Table 21. Seven cases were excluded due to one being a neuroendocrine tumour on histology, and the rest due to incomplete data analysis (system issues, flat spectra), resulting in 32 patients being included in the final analysis. The median age of patients was 67.5, with the youngest patient being 37 years of age and the oldest 84 years. The majority of patients were male (71.9%). An oesophagectomy was performed on 17 (53.1%) patients and a gastrectomy was performed on 15 (46.9%) patients. On tumour histology, all patients were diagnosed with adenocarcinoma. The most common tumour stage was T3 disease in both gastric and oesophageal cancers (56.2%). A total of 53.3% of patients with gastric cancer had T3 disease and 58.8% of patients with oesophageal cancer had T3 disease. Predominantly, patients with gastric and oesophageal cancer underwent some form of neoadjuvant therapy. For patients with gastric cancer, 93.3% underwent chemotherapy. For patients with oesophageal cancer, 82.4% underwent chemotherapy and 11.8% underwent chemoradiotherapy. Overall, 2 (6.3%) patients did

not undergo any form of neoadjuvant therapy; one patient with gastric cancer and the other with oesophageal cancer.

Table 21 Upper GI in vivo patient characteristics.

Patient Characteristic	No. of patients (%)	Gastric Cancer No. (%)	Oesophageal Cancer No. (%)
Age, median [range]	67.5 [37-84]	61 [37-84]	71 [52-80]
Sex			
Female	9 (28.1)	5 (33.3)	4 (23.5)
Male	23 (71.9)	10 (66.7)	13 (76.5)
Procedure			
Oesophagectomy	17 (53.1)	0	17 (100)
Gastrectomy	15 (46.9)	15 (100)	0
Tumour histology			
Adenocarcinoma	32 (100)	15 (100)	17 (100)
Tumour stage (TNM staging)			
T0	5 (15.6)	3 (20)	2 (11.8)
T1	5 (15.6)	1 (6.7)	4 (23.5)
T2	2 (6.3)	1 (6.7)	1 (5.9)
T3	18 (56.2)	8 (53.3)	10 (58.8)
T4	2 (6.3)	2 (13.3)	0
Neoadjuvant therapy			
Chemotherapy	28 (87.5)	14 (93.3)	14 (82.4)
Chemoradiotherapy	2 (6.3)	0	2 (11.8)
Radiotherapy	0	0	0
None	2 (6.3)	1 (6.7)	1 (5.9)

5.3.2 Spectral Analysis

In total, the analysis included 29 unique datasets for healthy stomach tissue, 11 for stomach cancer, 15 for healthy oesophagus tissue, and 12 for oesophageal cancer, as shown in Table 22. The healthy stomach data comprised of 4,810 spectra, while the stomach cancer data included 1,660 spectra. For

the oesophagus, there were 1,518 spectra associated with healthy tissue and 1,598 spectra corresponding to cancerous tissue.

Table 22 Dataset summary for spectral acquisition for gastric and oesophageal tissue.

	Normal Gastric Tissue	Cancer Gastric Tissue	Normal Oesophageal Tissue	Cancer Oesophageal Tissue
Distinct sets of data	29	11	15	12
Mean spectra collected	4,810	1,660	1,518	1,598

The means of all spectra with standard deviation for each of the tissue classes are shown in Figure 58.

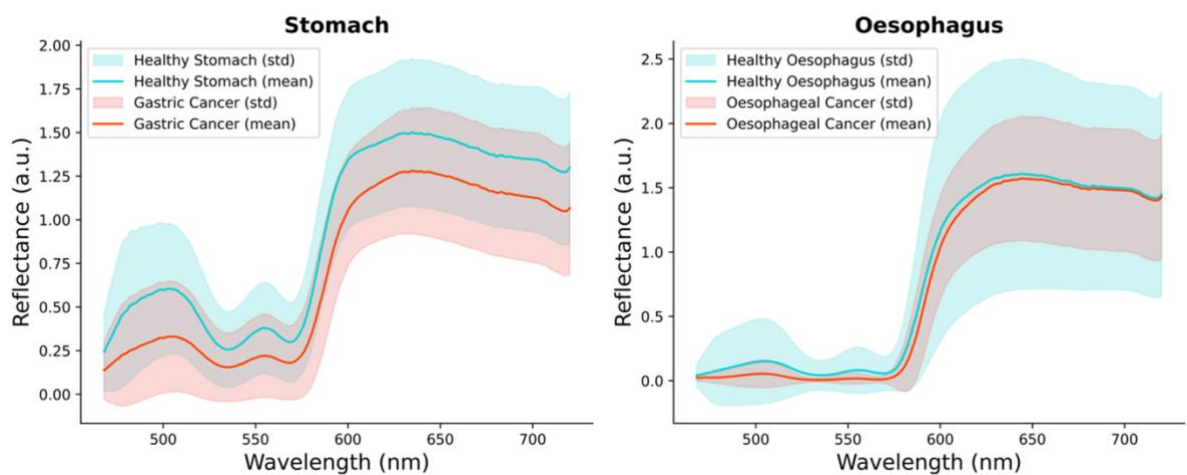


Figure 58 Average (lines) and standard deviations (shaded areas) of the mean spectra for normal and tumour stomach tissue (left) and normal and tumour oesophagus tissue (right).

The outcomes of the classification analysis for spectral datasets from stomach and oesophagus tissue are highlighted in Table 23. The XGB classifier performed the best for stomach tissue data compared to the RF and LGBM classifiers across key metrics including accuracy, sensitivity and specificity. The XGB classifier achieved a normal versus cancer diagnostic accuracy of 90.58 ± 0.85 , and an AUC of 0.957 (Figure 59) for stomach tissue.

For the oesophagus dataset, the LGBM classifier demonstrated superior performance when compared to the other classifiers overall. The LGBM classifier achieved a diagnostic accuracy for normal versus cancer oesophageal tissue of 83.81 ± 1.40 , and an AUC of 0.924 (Figure 59).

Table 23 Performance metrics for all classifiers for stomach and oesophagus datasets. Data showing the ability of multiple classifiers to detect normal versus cancerous tissue presented as mean $\pm$ standard deviation. Overall accuracy is calculated as the proportion of correctly identified spectra over the total number of spectra.

Classifier	Mean (SD) (%)							
	Stomach				Oesophagus			
	Accuracy	Sensitivity	Specificity	AUC	Accuracy	Sensitivity	Specificity	AUC
RF	89.71 (0.79)	73.24 (2.39)	95.40 (0.67)	95.04 (0.60)	83.78 (1.18)	85.99 (1.83)	81.46 (2.68)	92.28 (0.73)
XGB	90.58 (0.85)	77.36 (2.38)	95.14 (0.69)	95.72 (0.54)	83.40 (1.11)	85.24 (1.96)	81.47 (2.46)	92.02 (0.74)
LGBM	90.31 (0.74)	76.87 (2.23)	94.95 (0.62)	95.62 (0.54)	83.81 (1.40)	85.69 (1.89)	81.83 (2.81)	92.44 (0.82)

[RF: random forest; XGB: extreme gradient boosting; LGBM: light gradient boosting machine].

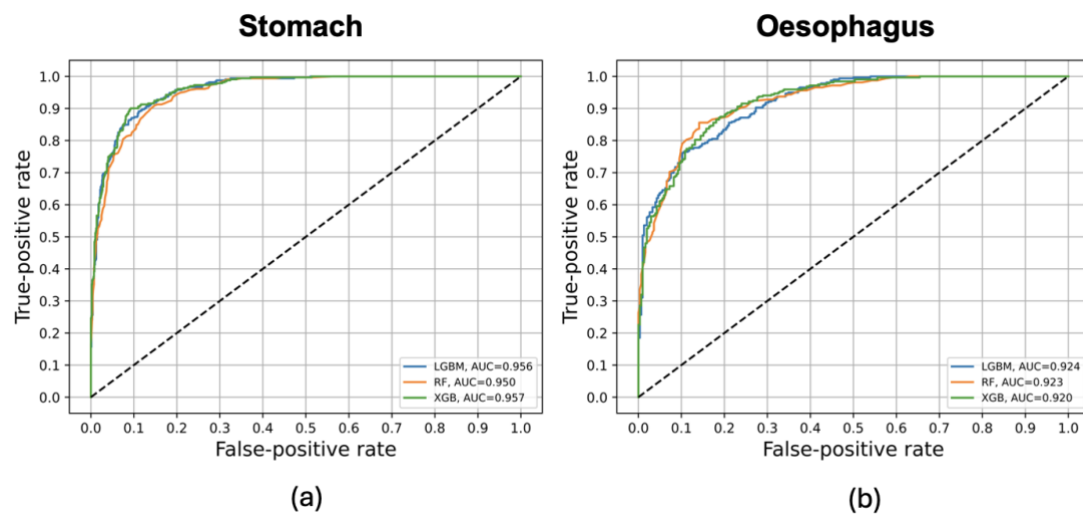


Figure 59 ROC curves with AUC values for all classifiers for (a) stomach dataset and (b) oesophagus dataset.

5.3.3 Marker-less Tracking Analysis

Overall, an average error of 4.90 pixels was achieved. The network was able to detect and track the tip of the probe with a prediction precision of 93.67%. In the inference mode, the developed model was able to detect and track the DRS probe in 23 frames per second.

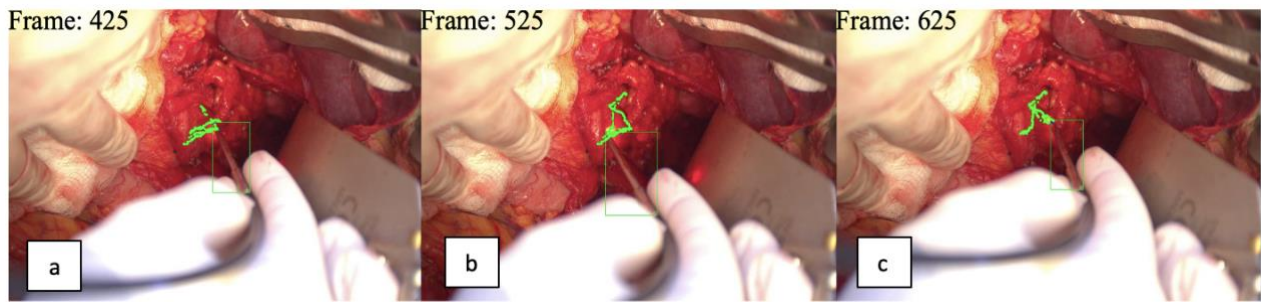


Figure 60 Illustration of real-time DRS marker-less probe tracking at each optical biopsy site during an oesophagectomy. (a) – (c) Real-time visual correlation of scanned optical biopsy sites on tissue (green dots on each frame) using the DRS probe and marker-less tracking.

5.3.4 Usability Questionnaire Analysis

5.3.4.1 Respondent Characteristics

A total of six respondents completed the usability questionnaire, including one consultant surgeon, one surgical registrar/trainee, and one scrub nurse. The results revealed a diverse pool of respondents, including individuals with varying roles within the surgical environment. While the study included participants from surgical and nursing backgrounds, it also incorporated perspectives from engineers and a PhD student, providing a comprehensive range of insights into the usability of the DRS probe across different user groups. The majority of respondents had used the probe more than once, suggesting a level of familiarity and repeated exposure to its functionality, whilst two respondents had only used it once (Figure 61).

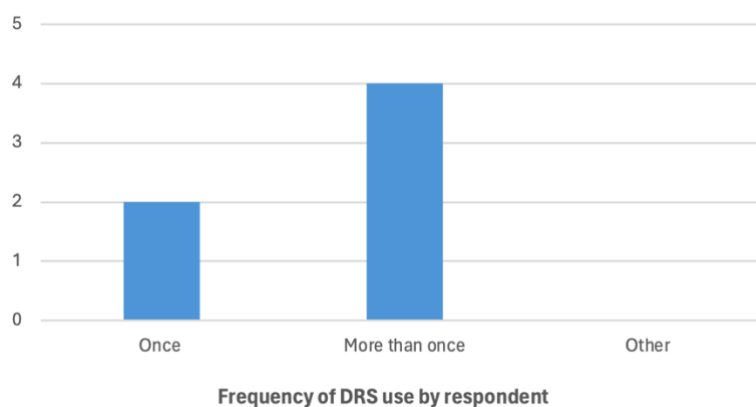


Figure 61 Bar chart representing frequency of DRS probe use by respondents.

5.3.4.2 Ease of Use

All respondents agreed that the DRS probe was easy to use. This high level of agreement underscores the satisfaction with the probe's performance and its ability to meet users' expectations in terms of functionality and utility within the surgical setting.

5.3.4.3 Effectiveness and Ease of Learning

The respondents overwhelmingly expressed positive views regarding the usability and effectiveness of the DRS probe. They strongly agreed, with one somewhat agreeing, that it required minimal steps to accomplish tasks, indicating its efficiency. Additionally, they commonly agreed, with one somewhat disagreeing, on its flexibility and ease of manoeuvrability, highlighting its ergonomic design. The majority strongly agreed that they can operate it without written instructions, notice no inconsistencies in its use, and find it suitable for both occasional and regular users. Furthermore, they agreed that they can recover quickly from mistakes and use it successfully every time, demonstrating its reliability. Regarding ease of learning, all respondents strongly agreed that they quickly learned to use it, easily remember its operation, and rapidly gained proficiency, indicating its user-friendly nature and short learning curve. This is highlighted in Table 24.

Table 24 Usability questionnaire results on ease of use of the DRS probe, ease of learning and future practice.

	Strongly agree	Somewhat agree	Neither agree nor disagree	Somewhat disagree	Strongly disagree	Not applicable
Ease of Use & Effectiveness						
It is easy to use	6	0	0	0	0	0
It requires the fewest steps possible to accomplish what I want to do with it	5	1	0	0	0	0
It is flexible and can be manoeuvred easily	5	0	0	1	0	0
I can use it without written instruction	5	0	0	1	0	0
I don't notice any inconsistencies when I use it	5	1	0	0	0	0
Both occasional and regular users would like it	5	1	0	0	0	0
I can recover from mistakes quickly and easily	4	1	1	0	0	0
I can use it successfully every time	5	1	0	0	0	0

Ease of Learning						
I learned to use it quickly	6	0	0	0	0	0
I easily remember how to use it	6	0	0	0	0	0
I quickly became skilful with it	5	1	0	0	0	0
Future Practice						
It has potential to change clinical practice within surgery (clinician)	3	1	0	0	0	2

5.3.4.4 Satisfaction

As demonstrated in Figure 62, the respondents demonstrated high satisfaction rates with the DRS probe, with five strongly agreeing and one somewhat agreeing that they were satisfied with it. Similarly, they unanimously expressed willingness to recommend it to a colleague, with five strongly agreeing and one somewhat agreeing, suggesting a high level of confidence in its efficacy. Moreover, the majority strongly agreed, with one neither agreeing nor disagreeing, that it functioned as desired, underscoring its alignment with their expectations. Regarding its potential impact on clinical practice within surgery, those that responded strongly agreed or somewhat agreed that the probe had potential to change clinical practice.

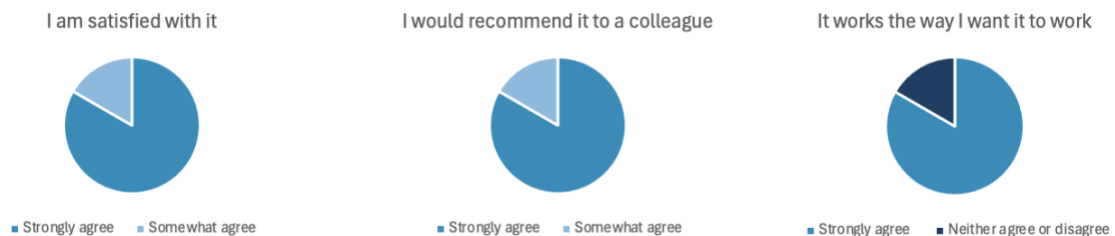


Figure 62 Pie charts showing satisfaction rates with the use of the DRS probe.

5.3.4.5 Positive and Negative Impact

The feedback collected from open questions regarding the DRS probe highlighted several positive aspects. Users found it “easy to use”, “easy to handle”, and stated that it was “easy to integrate into the surgical field”. The provision of real-time results and feedback was also appreciated, enhancing efficiency and decision-making during procedures. Additionally, the probe's size was deemed suitable for the surgical workflow, ensuring practicality. However, concerns were raised regarding its

flexibility, which could potentially lead to “the probe being accidentally dropped”. Furthermore, users noted a lack of information regarding the optimal angle and pressure required for its use, indicating a need for clearer guidance to optimise its utility.

5.4 Discussion

This *in vivo* feasibility study showed a high diagnostic accuracy, of over 90% for stomach tissue and over 83% for oesophageal tissue, for the DRS system to differentiate tumour and non-tumour tissue intra-operatively, with the potential to aid margin assessment. The DRS system was utilised with remarkable efficiency during surgical procedures, seamlessly delivering live on-screen feedback for tissue classification. The findings and the experience of end-users suggest a minimal perceived impact on workflow and high satisfaction rate with its ergonomics, performance and capabilities. Such results strengthen the acceptance and adoption of the DRS probe as a valuable tool in surgical practice.

5.4.1 Strengths of Study

The performance of the DRS system in differentiating tumour and normal tissue *in vivo* demonstrated excellent diagnostic accuracy, closely matching but not exceeding the success observed in *ex vivo* studies. This achievement can be attributed to several key factors that ensured the accuracy and reliability of the results. Firstly, the integration of a sterilisable probe proved invaluable in seamlessly adapting the system to fit within the surgical workflow. This feature ensured that the probe could be safely and effectively used in clinical settings without compromising patient safety or sterility standards. Additionally, the fast processing speeds of data acquisition and spectral measurements facilitated efficient collection of a large volume of data. This contributed to the reliability and accuracy of the findings by providing a more comprehensive dataset for analysis.

Furthermore, the histopathology method employed in this study further bolstered the credibility of the results. By utilising a sterile marker pen to demarcate the suspected tumour area prior to resection, followed by meticulous painting of the area post-resection, the integrity of the tissue samples was preserved, thereby avoiding any disruption that could potentially affect histopathological assessment and analysis. This approach not only ensured the accuracy of

histopathological correlation but also facilitated the creation of ground truth labels for the data analysis process.

The development and use of a marker-less probe tracking system to detect and track the DRS probe ensured effective use of the probe and robust results. The lack of a colour marker, which can be limited by many factors, allowed efficient set-up and use of the DRS probe, without the need to waste time in attaching and detaching it from the probe, and eliminated the risk of the marker being obscured by other instruments, ensuring uninterrupted detection and accurate tracking, regardless of lighting conditions.

Previous *in vivo* studies utilising DRS technology have primarily focused on the diagnostic efficacy and accuracy of the system in differentiating between tumour and normal tissue, often overlooking crucial aspects such as ergonomics and usability in real-life surgical settings. [165,369] However, this study sought to address this gap by conducting a comprehensive analysis of the experiences and views of end-users regarding the usability and functionality of the DRS probe. By soliciting feedback from surgeons, theatre staff, and other healthcare professionals, this study provided valuable insights into the practical implications of integrating DRS technology into surgical practice.

The findings from the usability study revealed several positive outcomes, underscoring the user-friendly nature and effectiveness of the DRS probe. Respondents overwhelmingly expressed satisfaction with the device, citing its ease of use, manoeuvrability, and intuitive design. Moreover, the majority of users reported that the probe met or exceeded their expectations, demonstrating its capability to seamlessly integrate into the surgical workflow without disrupting established protocols.

5.4.2 Limitations of Study

While the single-centre feasibility study provided valuable insights into the application of the DRS system in a controlled environment, its limitations warrant consideration. Firstly, the study's single-centre design restricts the generalisability of its findings, as the patient population and surgical protocols may differ across institutions. To address this limitation, future research will endeavour to conduct multi-centre trials involving a broader spectrum of patients, thus enhancing the study's external validity and allowing for more robust conclusions.

Moreover, the absence of external validation poses another challenge to the study's findings. Without validation from independent sources or external cohorts, there is a risk of bias or

overestimation of the DRS system's efficacy. Therefore, future studies should prioritise external validation to corroborate the system's performance across diverse clinical settings and patient populations.

Lastly, the use of a sterile marker in the *in vivo* setting introduced a potential risk of marker displacement or removal during tissue manipulation, particularly during resection. This could compromise the accuracy of histopathological correlation and the creation of ground truth labels. Addressing this challenge requires the development of alternative techniques or markers that are more resilient to surgical manipulation, ensuring the integrity of tissue marking *in vivo* and subsequent histopathological analysis.

In reflecting on the usability study conducted, it's important to acknowledge its limitations also, notably the relatively small sample size of participants. This constraint stemmed primarily from the practicality of involving the same surgical team in the procedures under study, resulting in a limited pool of individuals who interacted with and handled the probe. However, as we expand the scope of the project to encompass a broader range of surgical procedures and involve multiple centres, we anticipate a significant increase in the participant pool for the usability study. This expansion will afford us a larger and more diverse sample size, allowing for a more comprehensive assessment of the DRS probe's usability across various surgical settings and enhancing the generalisability of our findings.

5.4.3 Further Work

The success of the feasibility study for utilising the DRS system *in vivo* marks a significant milestone, paving the way for further advancements in clinical applications. Moving forward, the focus will be on expanding the use of the sterilisable probe *in vivo*, particularly for colorectal tissue differentiation during CRC resection procedures. By differentiating normal colorectal tissue from cancerous tissue, this approach promises to revolutionise real-time cancer margin assessment in the operating theatre. This advancement holds potential to improve decision-making processes and ultimately improve patient outcomes and care.

Furthermore, ongoing developments include the adaptation of the DRS probe for minimally invasive and laparoscopic surgeries. To facilitate its use in these procedures, a gripper tool is being developed to attach to the ferrule of the DRS probe (Figure 63). The development of an “add-on gripper” tool for the DRS probe has several technical considerations to take into account:

- The DRS probe with add-on gripper must fit through the laparoscopic port (5 mm width).
- A laparoscopic grasper must be able to grasp the probe strongly to maintain grip.
- The grasped probe must have enough range of movement to scan intraperitoneal tissue.
- The DRS probe and gripper tool must be sterilisable.

This innovative solution will enable surgeons to manipulate the probe using laparoscopic instruments, thereby facilitating data collection during minimally invasive procedures in addition to open cases. This integrated approach underscores the commitment to advancing surgical techniques and ensuring the seamless integration of cutting-edge technology into clinical practice to aid margin assessment in real-time.

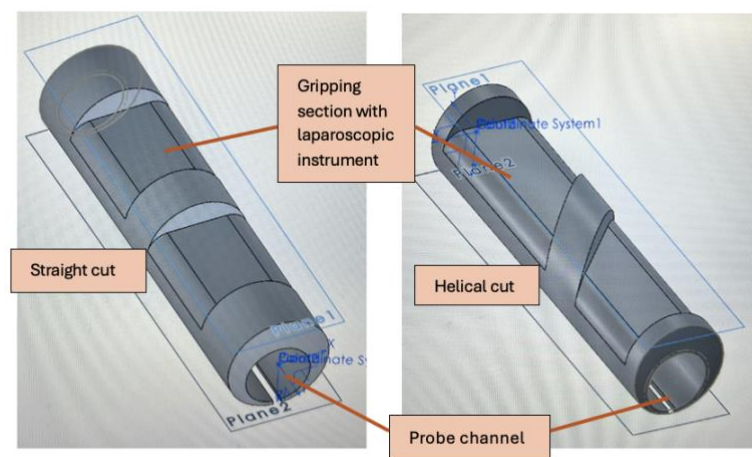


Figure 63 Potential designs for the DRS gripper that will attach onto the DRS ferrule to allow laparoscopic use of the probe *in vivo*.

5.5 Conclusion

In conclusion, this study has highlighted a successful translation of the DRS system into an *in vivo* setting, offering a promising avenue for aiding tissue margin assessment in GI cancer surgeries. By seamlessly integrating the DRS system into real-world surgical settings, we have taken a significant step towards providing surgeons with invaluable real-time tissue differentiation capabilities directly in the operating theatre. This advancement not only has the potential to replace traditional methods, such as frozen sections, but also stands to streamline the surgical workflow, ultimately saving on costs and resources by eliminating the need for an additional workforce and time-consuming procedures.

Moreover, the implementation of the DRS system holds immense promise for improving patient outcomes by facilitating more accurate and efficient surgical interventions. With its ability to provide instant feedback on tissue classification, surgeons can make informed decisions during procedures, leading to more precise resections and potentially reducing the need for repeat surgeries. As we move forward, continued research and development in this field will further refine the capabilities of the DRS system, solidifying its position as a cornerstone technology in modern surgical practice.

Chapter 6 Conclusions & Future Directions

6.1 Conclusions

This thesis aimed to improve pathology detection and resection in the lower GI tract by reviewing novel screening tools for lesion detection and characterisation, and by developing improved surgical techniques for differentiation of tumour and normal tissue to aid margin assessment in CRC resection surgeries. Each chapter has contributed unique insights and conclusions, collectively enhancing our understanding and approach to CRC management.

In Chapter 2, a systematic review and meta-analysis was conducted, revealing the burgeoning interest in AI for polyp detection during colonoscopy. The findings underscore the potential of ML systems to bolster adenoma detection rates and subsequently reduce CRC incidence. However, the field of AI in endoscopy is still evolving, necessitating further high-quality research and prospective validation in clinical settings. Additionally, ethical and legal frameworks must be established to ensure the responsible integration of AI systems into clinical practice.

Chapter 2 also focused on CADe systems in colonoscopy, highlighting their capacity to improve adenoma detection and mitigate the risk of post-colonoscopy colorectal cancer. While CADe systems show promise, further trials comparing different systems and endoscopic expertise levels are warranted before widespread adoption.

The exploration of stakeholder perspectives in Chapter 3 elucidated the optimistic outlook of endoscopists regarding the transformative potential of AI in endoscopy. Despite the perceived benefits, significant challenges and barriers persist, emphasising the need for systematic resolution and evidence-based guidelines. Large RCTs are imperative to fully comprehend the cost-benefit outcomes and establish robust guidelines for AI integration in endoscopic practice.

In Chapter 4, we developed a novel DRS tracking system and introduced the application of DRS for real-time tumour margin assessment in CRC surgeries. Analysis of the data highlighted a high diagnostic accuracy for the DRS probe and tracking system to differentiate tumour and non-tumour colorectal tissue. Real-time tracking of the tissue was shown live on-screen in the operating room with real-time feedback as to the tissue type being scanned. The findings demonstrate the feasibility and potential clinical utility of DRS in aiding surgical decision-making processes. Future endeavours

will focus on expanding the study to include rectal cancer resections and leveraging advanced neural networks for augmentation of probe tracking and spectral data classification.

The culmination of my efforts is presented in Chapter 5, where successful translation of the DRS system was achieved into an *in vivo* setting for GI cancer surgeries. The DRS system was able to differentiate tumour and normal tissue in gastric and oesophageal cancers *in vivo* with high diagnostic accuracy. In addition, the development of a marker-less tracking system for the DRS probe enhanced the robustness and accuracy of probe detection and tracking and was more effective since there was no necessity for reliance on a colour marker. This achievement marks a significant advancement in real-time tissue differentiation capabilities within the operating theatre, with the potential to replace traditional methods and streamline surgical workflows. Moreover, the implementation of the DRS system holds promise for improving patient outcomes and management by facilitating more precise resections and reducing the need for repeat surgeries.

In conclusion, this thesis underscores the transformative potential of existing innovative technologies, such as AI and DRS, in enhancing CRC management. Through rigorous research and development, I have laid the groundwork for a future where advanced screening tools and surgical techniques revolutionise patient care and outcomes in lower gastrointestinal tract pathology. To advance in this field, ongoing collaboration and interdisciplinary efforts will be essential to realising the full potential of these ground-breaking technologies in clinical practice.

6.2 Future Directions

Through ongoing research and development efforts, the DRS system is poised to revolutionise cancer management by providing clinicians with invaluable real-time tissue differentiation capabilities, ultimately leading to more precise and effective oncologic interventions. As advancements in technology continue to unfold, the potential impact of the DRS system on improving patient outcomes and enhancing surgical practice will be transformative.

6.2.1 Tumour Differentiation in Cancer Surgery

By leveraging insights gained from both *ex vivo* and *in vivo* studies in GI cancers, future research aims to evaluate the diagnostic accuracy of the DRS system in differentiating cancer and normal tissue in other surgical disciplines, including endocrine surgery and gynaecological surgery. Through rigorous

evaluation and validation, the DRS system holds the potential to enhance surgical precision in a wide range of surgical procedures.

In thyroid cancer surgery, positive surgical margins are a poor prognostic indicator. The modified 2015 American Thyroid Association risk stratification system identifies the completeness of surgical resection with negative surgical margins as an important determinant of outcome. Positive margins have been associated with an intermediate or high risk of recurrence, and a decrease in overall survival. [370,371] Intraoperatively, surgeons rely on macroscopic visual and tactile properties of the tumour to determine the extent of the lesion and plan the resection. This approach has significant limitations. The development and use of the DRS system as an intra-operative margin assessment (IMA) technique in thyroid cancer surgery can address these shortcomings and reduce positive margin rates, facilitating complete resection at the first attempt. In addition, it can aid in the diagnosis of indeterminate nodules within the thyroid. Results can be provided rapidly, in real-time, to avoid increase operative duration and complexity.

Gynaecological cancers represent a substantial health burden and endometrial cancer accounts for nearly 50% of all new gynaecological cancers diagnosed in the Western world. Endometrial cancer necessitates surgical interventions for tumour removal. However, there is a lack of consensus on the extent of surgical staging in endometrial carcinoma. The ability of surgical staging to accurately identify lymphatic spread and how this information affects prognosis and alters the use of adjuvant therapies are a source of controversy. [372] In endometrial cancer, where hysterectomy is a standard procedure, the DRS system holds the potential to enhance surgical precision by differentiating cancerous areas from healthy tissue. This nuanced discrimination could enable surgeons to tailor their resections, minimising the removal of healthy tissue and optimising patient outcomes.

In addition to expanding its application across various surgical specialties, the aim is to refine the DRS system's capabilities within GI cancer surgery. One such endeavour involves evaluating its diagnostic accuracy in distinguishing fibrotic tissue from tumour tissue. Following neoadjuvant therapy, areas of fibrosis can be very difficult to differentiate from tumour tissue intra-operatively by visual and haptic feedback when making decisions about tumour margin assessment. This may lead to the unnecessary removal of healthy tissue and negatively impact morbidity. By enabling accurate detection of fibrosis tissue, the DRS system can aid in decision-making and optimise surgical outcome. Moreover, expansion of the *ex vivo* and *in vivo* studies to include rectal cancer resections

will be undertaken since a positive resection margin in this cohort of patients has been associated with local recurrence and reduced survival. Assessing spectral measurements for fat or adipose tissue will help to evaluate classification of rectal tissue since the rectum is surrounded by the mesorectum, a perirectal fat layer.

Future application of the DRS system for assessment of tissue infiltration of the peritoneum and other organs, to be able to assess for metastasis of tumour in real-time during cancer procedures, will be invaluable. By providing surgeons with immediate feedback on tissue type, the DRS system has potential to enable early detection of metastatic spread, allowing for prompt decision-making and potentially altering the course of surgical management. It would facilitate more comprehensive tumour staging and guide personalised treatment strategies.

6.2.2 Virtual Reality & Autofluorescence Spectroscopy

The current work in this thesis has introduced augmented reality technology to enhance real-time tracking of the DRS probe and accurate visualisation of the probe tip to aid the decision-making process during surgical procedures. A marker-less machine learning tracking system was developed to recognise the DRS probe's position and orientation within the surgical field in real-time without the need for physical markers.

In future work, the potential transition to virtual reality (VR) technology for the application of the DRS probe will be explored. This type of technology will provide an unparalleled depth of perception, allowing for more precise and intuitive manipulation of the DRS probe. Surgeons will have the ability to interact with 3D representations of patient anatomy in real-time, facilitating a comprehensive understanding of tissue characteristics and spatial relationships, and allowing more accurate resection of tissue. Moreover, those outside of the surgical field will be able to follow the sampling process live with accurate 3D visualisation of the procedure.

In addition, autofluorescence spectroscopy will be incorporated with additional delivery fibres for multimodal sensing. Autofluorescence spectroscopy exploits the natural fluorescence properties of biological tissues, offering complementary information to DRS. By combining these modalities, the DRS system can leverage the distinctive spectral signatures of autofluorescence emitted by different tissue types, augmenting its ability to discriminate between normal and cancerous tissues. The integration of autofluorescence spectroscopy holds promise for refining the diagnostic capabilities of

the DRS probe, particularly in scenarios where subtle differences in tissue composition may impact clinical decision-making.

6.2.3 Minimally Invasive Use of DRS

6.2.3.1 Laparoscopic Use of DRS

In regard to minimally invasive surgery, ongoing efforts are underway to adapt the DRS probe and system for use in laparoscopic procedures. Laparoscopic surgery is limited when compared to open surgery, by its reduced tactile feedback, which is used by surgeons as a crude method of tumour delineation intra-operatively. This makes it more challenging to assess tumour margins accurately. Introducing the DRS system in such scenarios, will significantly improve the accuracy of CRM determination. A significant aspect of the endeavour to transform the DRS system involves the re-design of the probe to broaden its applicability across various surgical operations beyond traditional open procedures. This re-design initiative encompasses the integration of innovative features aimed at enhancing the usability and manoeuvrability of the probe within the constrained environment of minimally invasive surgeries. One such initiative includes the incorporation of a gripper mechanism onto the probe's ferule, facilitating secure attachment and manipulation with laparoscopic instruments. This addition not only ensures better grasp and control of the probe but also enables improved dexterity and movement, thereby facilitating more precise data collection during laparoscopic procedures. By augmenting the probe's design with such functionalities, surgeons can seamlessly integrate DRS technology into minimally invasive surgical workflows, thereby extending the benefits of real-time tissue differentiation and margin assessment to a broader spectrum of surgical interventions.

6.2.3.2 Endoscopic Use of DRS

Conventional colonoscopy is unable to accurately differentiate hyperplastic polyps from adenomatous lesions. As described in the introductory chapter of this thesis, this leads to missed polyps and the risk of post-colonoscopy CRC. As a result, there has been a recent emergence of optical imaging techniques to help tackle this challenge and aid in detection and characterisation of lesions in the lower GI tract. One such technique has been optical spectroscopy. This offers the potential for rapid *in vivo* diagnosis of lesions.

A variety of DRS probes have been explored for tissue spectroscopy, with the probes designed for endoscopic GI tissue characterisation typically comprising of compact optical fibre bundles featuring a single illumination fibre and one or more collection fibres. The optical probe is passed through the working channel of an endoscope and is placed in direct contact with tissue during the endoscopy procedure. This process allows luminal characterisation of tissue type within the GI tract. [Figure 64]

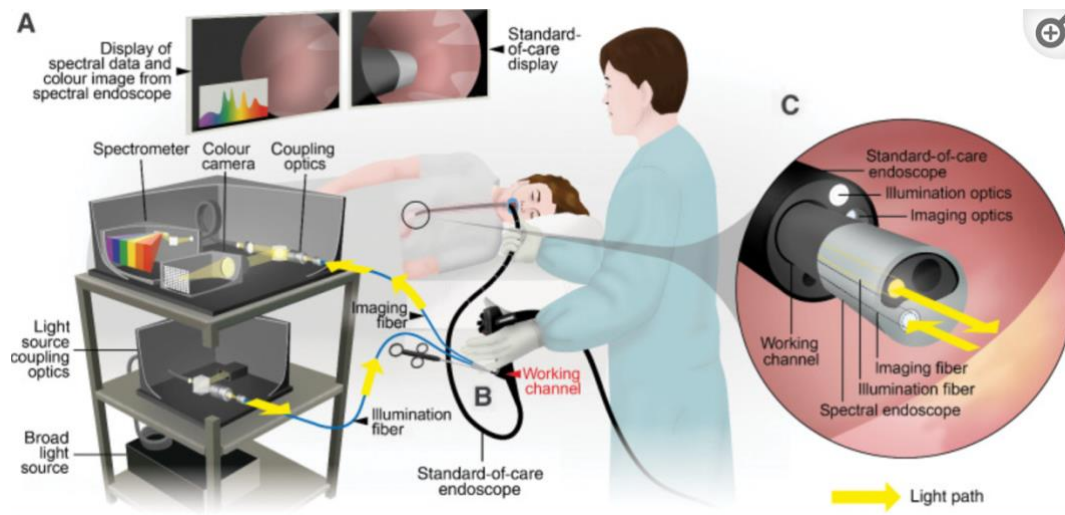


Figure 64 Schematic diagram of endoscopic optical probe in use during gastroscopy, adapted from paper by Waterhouse et al. [373] (A) Light is transferred using the illumination fibre and reflected light is relayed back to the spectrometer. (B) The spectral endoscope is inserted via the working channel of the gastroscopy. (C) Light is delivered to the tissue area of interest to aid tissue characterisation.

There are few studies evaluating the application of optical spectroscopy *in vivo* for endoscopic classification of lesions. Recent studies include that of Dhar et al. who utilised elastic scattering spectroscopy (ESS) to differentiate normal and cancer tissue during colonoscopy with a sensitivity of 80% and specificity of 86%, from a total of 483 spectra across all pathologies. [346] Waterhouse et al. focussed on Barrett's oesophagus and highlighted an accuracy of 84.8%, sensitivity of 83.7% and specificity of 85.5% when comparing neoplasia with non-dysplastic tissue, from 715 spectra. [373] In these studies, small datasets limited the performance of the systems and affected the reliability of the results. However, the potential for optical biopsy to classify diminutive adenomas in order to reduce the risk of missed polyps and increase the ADR was demonstrated. In addition, optical spectroscopy can classify normal tissue with the aim of aiding decision-making within endoscopy in regard to a "resect and discard" or a "diagnose and leave" strategy.

The work from this thesis regarding the use of DRS and its development offer promising implications for improving the delineation of resection margins during advanced endoscopic techniques for complex colorectal lesions. The success of minimally invasive techniques, such as EMR and ESD, for

the removal of abnormal tissue in the GI tract, particularly early-stage cancers, relies heavily on ensuring complete resection of the affected area with clear margins. The integration of DRS into EMR and ESD procedures holds the potential to provide real-time feedback on tissue composition and differentiation, aiding endoscopists in identifying the precise boundaries of lesions. By leveraging the spectral information obtained from DRS, clinicians can discern subtle differences between normal and abnormal tissue, enhancing their ability to achieve clear resection margins. This could significantly reduce the likelihood of incomplete resections and the need for subsequent interventions or surgery due to residual disease.

Furthermore, the development of tailored DRS probes specifically designed for endoscopic use could further facilitate the implementation of this technology in EMR and ESD procedures. These probes could be engineered to be compatible with existing endoscopic instruments, allowing seamless integration into routine clinical practice. Additionally, advancements in data processing algorithms and ML techniques would aid in real-time analysis and interpretation of spectral data, as well as offering real-time tracking of optical biopsy sites, providing immediate feedback to endoscopists during the procedure.

Overall, the incorporation of DRS into EMR and ESD procedures has the potential to revolutionise the management of early-stage disease and reduce the risk of subsequent CRC. Continued research and development in this area will be essential to further refine the technology and validate its clinical utility, ultimately leading to improved patient outcomes and enhanced cancer care in the lower GI tract.

6.2.4 Unmet Clinical Need & Commercial Opportunity

The evidence of an unmet clinical need stems from the challenges of intraoperative margin assessment, for which frozen section analysis remains as the gold standard. Despite advances in surgical techniques and instruments, achieving precise resection margin assessment remains a challenge, often resulting in sub-optimal outcomes and the need for additional interventions. The proposed DRS system in this thesis addresses this critical gap by offering real-time, accurate tissue classification enhancing surgical precision, and potentially minimising the recurrence of disease. Furthermore, the integration of augmented reality improves the surgical experience, providing dynamic visualisation and aiding decision-making intraoperatively. This technology holds significant

commercial promise, given its potential to revolutionise cancer surgery practices, improve patient outcomes, and reduce healthcare costs associated with resource allocation.

While specific market research tailored to our proposed product is pending, existing data on the global market for surgical navigation and imaging systems provides valuable insights. The increasing prevalence of CRC worldwide, coupled with the rising demand for minimally invasive and precision-guided surgical interventions, underscores the market potential for this technology. Additionally, health economics and reimbursement considerations will be addressed through robust health economic evaluations, leveraging real-world data and cost-effectiveness analyses to demonstrate the value proposition of the DRS system to the NHS, regulatory authorities and other stakeholders. This clinical need can be set against the very modest system cost, which gives us confidence that there is a market opportunity for this technology.

The interactions with a PPI focus group, surgical teams, clinicians, and healthcare stakeholders have provided valuable insights into the clinical utility and market demand for the DRS technology. Feedback from key opinion leaders and end-users has informed our development roadmap, ensuring alignment with clinical needs and user preferences. Moving forward, I will continue to engage stakeholders to validate our technology, refine product specifications, and tailor our offering to meet the evolving needs of the healthcare market.

To date, the core technology driving the proposed project has not been subject to disclosure, nor has any patent application been filed. A preliminary internal freedom to operate assessment has been conducted, which indicates a favourable position for the continued development of the DRS technology without infringing on existing intellectual property (IP) rights. An invention disclosure form has also been completed on the DRS device, and extensive expertise and knowledge has been gathered in operating theatre-compatible hardware and real-time processing methods. The next step will be to actively engage with the Industry Partnership and Commercialisation (IPC) office on this project to confirm the necessity and timing of protective measures for the technology. This collaborative effort aims to ensure that, when the time comes to file for patents, we have a robust and strategic approach that aligns with our project goals and commercial potential. Through ongoing discussions, we will evaluate the commercial viability of our technology, identify potential market pathways, and devise strategies for IP protection and commercialisation.

The future landscape for potential collaboration with external industry partners that could scale the technology and maximise its clinical impact is being considered. This would explore interest in further development of the technology for commercialisation. These considerations will be critical as we move closer to the application of our technology in a clinical setting and will be managed in line with our IP strategy to ensure alignment with project objectives.

References

1. Sung H, Ferlay J, Siegel RL, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin*. 2021;71(3):209-249. doi:10.3322/CAAC.21660
2. Morgan E, Arnold M, Gini A, et al. Global burden of colorectal cancer in 2020 and 2040: incidence and mortality estimates from GLOBOCAN. *Gut*. 2023;72(2):338-344. doi:10.1136/GUTJNL-2022-327736
3. Bowel cancer statistics | Cancer Research UK. Accessed September 18, 2023. <https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bowel-cancer>
4. Colorectal Cancer Survival by Stage.
5. Le Clercq CMC, Bouwens MWE, Rondagh EJA, et al. Postcolonoscopy colorectal cancers are preventable: a population-based study. *Gut*. 2014;63(6):957-963. doi:10.1136/GUTJNL-2013-304880
6. Erichsen R, Baron JA, Stoffel EM, Laurberg S, Sandler RS, Sørensen HT. Characteristics and survival of interval and sporadic colorectal cancer patients: a nationwide population-based cohort study. *Am J Gastroenterol*. 2013;108(8):1332-1340. doi:10.1038/AJG.2013.175
7. Morris EJA, Rutter MD, Finan PJ, Thomas JD, Valori R. Post-colonoscopy colorectal cancer (PCCRC) rates vary considerably depending on the method used to calculate them: a retrospective observational population-based study of PCCRC in the English National Health Service. *Gut*. 2015;64(8):1248-1256. doi:10.1136/GUTJNL-2014-308362
8. Govindarajan A, Rabeneck L, Yun L, Tinmouth J, Paszat LF, Baxter NN. Population-based assessment of the outcomes in patients with postcolonoscopy colorectal cancers. *Gut*. 2016;65(6):971-976. doi:10.1136/GUTJNL-2014-308578
9. Forsberg A, Widman L, Bottai M, Ekblom A, Hultcrantz R. Postcolonoscopy Colorectal Cancer in Sweden From 2003 to 2012: Survival, Tumor Characteristics, and Risk Factors. *Clin Gastroenterol Hepatol*. 2020;18(12):2724-2733.e3. doi:10.1016/j.cgh.2020.06.010
10. Bujanda L, Cosme A, Gil I, Arenas-Mirave JI. Malignant colorectal polyps. *World J Gastroenterol*. 2010;16(25):3103-3111. doi:10.3748/wjg.v16.i25.3103
11. Kawasaki K, Kurahara K, Yanai S, et al. Colonoscopic features and malignant potential of sessile serrated adenomas: comparison with other serrated lesions and conventional adenomas. *Color Dis*. 2016;18(8):795-802. doi:10.1111/codi.13276
12. Fearon ER, Vogelstein B. A genetic model for colorectal tumorigenesis. *Cell*. 1990;61(5):759-767. doi:10.1016/0092-8674(90)90186-I
13. Pino MS, Chung DC. The chromosomal instability pathway in colon cancer. *Gastroenterology*. 2010;138(6):2059-2072. doi:10.1053/J.GASTRO.2009.12.065
14. Fearon ER. Molecular Genetics of Colorectal Cancer. <https://doi.org/10.1146/annurev-pathol-011110-130235>. 2011;6:479-507. doi:10.1146/ANNUREV-PATHOL-011110-130235
15. Bateman AC. Pathology of serrated colorectal lesions. *J Clin Pathol*. 2014;67(10):865-874. doi:10.1136/JCLINPATH-2014-202175
16. Kambara T, Simms LA, Whitehall VLJ, et al. BRAF mutation is associated with DNA methylation in serrated polyps and cancers of the colorectum. *Gut*. 2004;53(8):1137-1144. doi:10.1136/GUT.2003.037671
17. Seely KD, Morgan AD, Hagenstein LD, Florey GM, Small JM. Bacterial Involvement in Progression and Metastasis of Colorectal Neoplasia. *Cancers (Basel)*. 2022;14(4). doi:10.3390/CANCERS14041019
18. Hong SN. Genetic and epigenetic alterations of colorectal cancer. *Intest Res*. 2018;16(3):327. doi:10.5217/IR.2018.16.3.327
19. Issa JP. CpG island methylator phenotype in cancer. *Nat Rev Cancer* 2004 412. 2004;4(12):988-993. doi:10.1038/nrc1507

20. Biller LH, Schrag D. Diagnosis and Treatment of Metastatic Colorectal Cancer: A Review. *JAMA*. 2021;325(7):669-685. doi:10.1001/JAMA.2021.0106
21. Bowel cancer screening: programme overview - GOV.UK. Accessed October 31, 2024. <https://www.gov.uk/guidance/bowel-cancer-screening-programme-overview>
22. Chiu HM, Jen GHH, Wang YW, et al. Long-term effectiveness of faecal immunochemical test screening for proximal and distal colorectal cancers. *Gut*. 2021;70(12):2321-2329. doi:10.1136/GUTJNL-2020-322545
23. Rees CJ, Gibson ST, Rutter MD, et al. UK key performance indicators and quality assurance standards for colonoscopy. doi:10.1136/gutjnl
24. Toribara NW, Sleisenger MH. Screening for Colorectal Cancer. *N Engl J Med*. 1995;332(13):861-867. doi:10.1056/NEJM199503303321306
25. Pickhardt PJ, Hassan C, Halligan S, Marmo R. Colorectal cancer: CT colonography and colonoscopy for detection-systematic review and meta-analysis. *Radiology*. 2011;259(2):393-405. doi:10.1148/RADIOL.11101887/-/DC1
26. Winawer SJ, Zauber AG, Ho MN, et al. Prevention of colorectal cancer by colonoscopic polypectomy. The National Polyp Study Workgroup. *N Engl J Med*. 1993;329(27):1977-1981. doi:10.1056/NEJM199312303292701
27. Citarda F, Tomaselli G, Capocaccia R, Barcherini S, Crespi M. Efficacy in standard clinical practice of colonoscopic polypectomy in reducing colorectal cancer incidence. *Gut*. 2001;48(6):812-815. doi:10.1136/GUT.48.6.812
28. Thiis-Evensen E, Hoff GS, Sauar J, Langmark F, Majak BM, Vatn MH. Population-based surveillance by colonoscopy: effect on the incidence of colorectal cancer. Telemark Polyp Study I. *Scand J Gastroenterol*. 1999;34(4):414-420. doi:10.1080/003655299750026443
29. Martín-López JE, Beltrán-Calvo C, Rodríguez-López R, Molina-López T. Comparison of the accuracy of CT colonography and colonoscopy in the diagnosis of colorectal cancer. *Color Dis*. 2014;16(3). doi:10.1111/codi.12506
30. Franco DL, Leighton JA, Gurudu SR. Approach to incomplete colonoscopy: New techniques and technologies. *Gastroenterol Hepatol*. 2017;13(8):476-483. Accessed November 25, 2020. </pmc/articles/PMC5572961/?report=abstract>
31. East JE, Saunders BP, Burling D, Boone D, Halligan S, Taylor SA. Surface visualization at CT colonography simulated colonoscopy: Effect of varying field of view and retrograde view. *Am J Gastroenterol*. 2007;102(11):2529-2535. doi:10.1111/j.1572-0241.2007.01429.x
32. Corley DA, Jensen CD, Marks AR, et al. Adenoma detection rate and risk of colorectal cancer and death. *N Engl J Med*. 2014;370(14):1298-1306. doi:10.1056/NEJMoa1309086
33. Malhotra A, Shaikat A. What is the criterion for high-performing colonoscopists? Being meticulous! *Gastrointest Endosc*. 2015;82(1):130-132. doi:10.1016/j.gie.2015.02.019
34. Wang H, Wang P, Liu X, et al. Factors predicting the colorectal adenoma detection rate in colonoscopic screening of a Chinese population. *Medicine (Baltimore)*. 2019;98(15):e15103. doi:10.1097/MD.00000000000015103
35. Amersi F, Agustin M, Ko CY. Colorectal cancer: Epidemiology, risk factors, and health services. *Clin Colon Rectal Surg*. 2005;18(3):133-140. doi:10.1055/s-2005-916274
36. Roy HK, Bianchi LK. Differences in colon adenomas and carcinomas among women and men: Potential clinical implications. *JAMA - J Am Med Assoc*. 2009;302(15):1696-1697. doi:10.1001/jama.2009.1499
37. Martínez ME, Baron JA, Lieberman DA, et al. A pooled analysis of advanced colorectal neoplasia diagnoses after colonoscopic polypectomy. *Gastroenterology*. 2009;136(3):832-841. doi:10.1053/J.GASTRO.2008.12.007
38. Lee TJW, Rees CJ, Blanks RG, et al. Colonoscopic factors associated with adenoma detection in a national colorectal cancer screening program. *Endoscopy*. 2014;46(3):203-211. doi:10.1055/s-0033-1358831
39. Jover R, Zapater P, Polanía E, et al. Modifiable endoscopic factors that influence the adenoma

- detection rate in colorectal cancer screening colonoscopies. *Gastrointest Endosc*. 2013;77(3):381-389.e1. doi:10.1016/j.gie.2012.09.027
40. Bernstein C, Thorn M, Monsees K, Spell R, O'Connor JB. A prospective study of factors that determine cecal intubation time at colonoscopy. *Gastrointest Endosc*. 2005;61(1):72-75. doi:10.1016/S0016-5107(04)02461-7
 41. Se YO, Chong IS, In KS, et al. Factors affecting the technical difficulty of colonoscopy. *Hepatogastroenterology*. 2007;54(77):1403-1406. Accessed January 20, 2021. <https://europepmc.org/article/med/17708264>
 42. Witte TN, Enns R. The difficult colonoscopy. *Can J Gastroenterol*. 2007;21(8):487-490. doi:10.1155/2007/520431
 43. Tholey DM, Shelton CE, Francis G, et al. Adenoma detection in excellent versus good bowel preparation for colonoscopy. *J Clin Gastroenterol*. 2015;49(4):313-319. doi:10.1097/MCG.0000000000000270
 44. Ness RM, Manam R, Hoen H, Chalasani N. Predictors of inadequate bowel preparation for colonoscopy. *Am J Gastroenterol*. 2001;96(6):1797-1802. doi:10.1111/j.1572-0241.2001.03874.x
 45. Nguyen DL, Wieland M. *Risk Factors Predictive of Poor Quality Preparation during Average Risk Colonoscopy Screening: The Importance of Health Literacy*. Vol 19.; 2010. Accessed January 20, 2021. <https://www.jgld.ro/jgld/index.php/jgld/article/view/2010.4.5>
 46. Köksal AŞ, Kalkan IH, Torun S, et al. A simple method to improve adenoma detection rate during colonoscopy: Altering patient position. *Can J Gastroenterol*. 2013;27(9):509-512. doi:10.1155/2013/276043
 47. Kaminski MF, Regula J, Kraszewska E, et al. Quality Indicators for Colonoscopy and the Risk of Interval Cancer. *N Engl J Med*. 2010;362(19):1795-1803. doi:10.1056/NEJMoa0907667
 48. Bhangu A, Bowley DM, Horner R, Baranowski E, Raman S, Karandikar S. Volume and accreditation, but not specialty, affect quality standards in colonoscopy. *Br J Surg*. 2012;99(10):1436-1444. doi:10.1002/bjs.8866
 49. Rabeneck L, Paszat LF, Saskin R. Endoscopist Specialty Is Associated With Incident Colorectal Cancer After a Negative Colonoscopy. *Clin Gastroenterol Hepatol*. 2010;8(3):275-279. doi:10.1016/j.cgh.2009.10.022
 50. Toyoshima O, Nishizawa T, Yoshida S, et al. Expert endoscopists with high adenoma detection rates frequently detect diminutive adenomas in proximal colon. *Endosc Int Open*. 2020;08(06):E775-E782. doi:10.1055/a-1136-9971
 51. James P, Hegagi M, Hegagi M, et al. Variable Endoscopist performance in proximal and distal adenoma detection during colonoscopy: A retrospective cohort study. *BMC Gastroenterol*. 2018;18(1):73. doi:10.1186/s12876-018-0800-4
 52. Anderson JC, Butterly LF. Colonoscopy: Quality Indicators. *Clin Transl Gastroenterol*. 2015;6(2):e77. doi:10.1038/ctg.2015.5
 53. Kaminski MF, Anderson J, Valori R, et al. Leadership training to improve adenoma detection rate in screening colonoscopy: A randomised trial. *Gut*. 2016;65(4):616-624. doi:10.1136/gutjnl-2014-307503
 54. Coe SG, Crook JE, Diehl NN, Wallace MB. An endoscopic quality improvement program improves detection of colorectal adenomas. *Am J Gastroenterol*. 2013;108(2):219-226. doi:10.1038/ajg.2012.417
 55. Ussui V, Coe S, Rizk C, Crook JE, Diehl NN, Wallace MB. Stability of increased adenoma detection at colonoscopy. Follow-up of an endoscopic quality improvement program-EQUIP-II. *Am J Gastroenterol*. 2015;110(4):489-496. doi:10.1038/ajg.2014.314
 56. Overholt BF, Brooks-Belli L, Grace M, et al. Withdrawal Times and Associated Factors in Colonoscopy: A Quality Assurance Multicenter Assessment. *J Clin Gastroenterol*. 2010;44(4):e80-e86. doi:10.1097/MCG.0b013e3181bf9b02
 57. Barclay RL, Vicari JJ, Doughty AS, Johanson JF, Greenlaw RL. Colonoscopic Withdrawal Times

- and Adenoma Detection during Screening Colonoscopy. *N Engl J Med*. 2006;355(24):2533-2541. doi:10.1056/nejmoa055498
58. Kumar S, Thosani N, Ladabaum U, et al. Adenoma miss rates associated with a 3-minute versus 6-minute colonoscopy withdrawal time: a prospective, randomized trial. *Gastrointest Endosc*. 2017;85(6):1273-1280. doi:10.1016/j.gie.2016.11.030
 59. Hashimoto K, Higaki S, Nishiaki M, Fujiwara K, Gondo T, Sakaida I. Does chromoendoscopy improve the colonoscopic adenoma detection rate? *Hepatogastroenterology*. 2010;57(104):1399-1404. Accessed February 8, 2021. <https://europepmc.org/article/med/21443093>
 60. Leung JW, Yen AW, Jia H, et al. A prospective RCT comparing combined chromoendoscopy with water exchange (CWE) vs water exchange (WE) vs air insufflation (AI) in adenoma detection in screening colonoscopy. *United Eur Gastroenterol J*. 2019;7(4):477-487. doi:10.1177/2050640619832196
 61. Chung SJ, Kim D, Song JH, et al. Comparison of detection and miss rates of narrow band imaging, flexible spectral imaging chromoendoscopy and white light at screening colonoscopy: A randomised controlled back-to-back study. *Gut*. 2014;63(5):785-791. doi:10.1136/gutjnl-2013-304578
 62. van den Broek FJC, Reitsma JB, Curvers WL, Fockens P, Dekker E. Systematic review of narrow-band imaging for the detection and differentiation of neoplastic and nonneoplastic lesions in the colon (with videos). *Gastrointest Endosc*. 2009;69(1):124-135. doi:10.1016/j.gie.2008.09.040
 63. Omata F, Ohde S, Deshpande GA, Kobayashi D, Masuda K, Fukui T. Image-enhanced, chromo, and cap-assisted colonoscopy for improving adenoma/neoplasia detection rate: a systematic review and meta-analysis. *Scand J Gastroenterol*. 2014;49(2):222-237. doi:10.3109/00365521.2013.863964
 64. Li L, Ou Y, Yue H, et al. Comparison of the detection of colorectal lesions in different endoscopic modalities: A network meta-analysis and systematic review. *Exp Ther Med*. 2019;18(1):154. doi:10.3892/etm.2019.7535
 65. Dik VK, Gralnek IM, Segol O, et al. Multicenter, randomized, tandem evaluation of EndoRings colonoscopy - Results of the CLEVER study. *Endoscopy*. 2015;47(12):1151-1158. doi:10.1055/s-0034-1392421
 66. Floer M, Biecker E, Fitzlaff R, et al. Higher adenoma detection rates with endocuff-assisted colonoscopy - A randomized controlled multicenter trial. *PLoS One*. 2014;9(12). doi:10.1371/journal.pone.0114267
 67. Gralnek IM, Siersema PD, Halpern Z, et al. Standard forward-viewing colonoscopy versus full-spectrum endoscopy: An international, multicentre, randomised, tandem colonoscopy trial. *Lancet Oncol*. 2014;15(3):353-360. doi:10.1016/S1470-2045(14)70020-8
 68. Choi J, Shin K, Jung J, et al. Convolutional neural network technology in endoscopic imaging: Artificial intelligence for endoscopy. *Clin Endosc*. 2020;53(2):117-126. doi:10.5946/ce.2020.054
 69. Jaremko JL, Azar M, Bromwich R, et al. Canadian Association of Radiologists White Paper on Ethical and Legal Issues Related to Artificial Intelligence in Radiology. *Can Assoc Radiol J*. 2019;70(2):107-118. doi:10.1016/J.CARJ.2019.03.001
 70. McKinney SM, Sieniek M, Godbole V, et al. International evaluation of an AI system for breast cancer screening. *Nature*. 2020;577(7788):89-94. doi:10.1038/s41586-019-1799-6
 71. Rodríguez-Ruiz A, Krupinski E, Mordang JJ, et al. Detection of breast cancer with mammography: Effect of an artificial intelligence support system. *Radiology*. 2019;290(3):1-10. doi:10.1148/radiol.2018181371
 72. Wu N, Phang J, Park J, et al. Deep Neural Networks Improve Radiologists' Performance in Breast Cancer Screening. *IEEE Trans Med Imaging*. 2020;39(4):1184-1194. doi:10.1109/TMI.2019.2945514

73. Rudnicka AR, Welikala R, Barman S, et al. Artificial intelligence-enabled retinal vasculometry for prediction of circulatory mortality, myocardial infarction and stroke. *Br J Ophthalmol*. 2022;106(12):1722-1729. doi:10.1136/BJO-2022-321842
74. A. E-B, G. G, Suzuki K. AO - El-Baz AO <http://orcid.org/000.-0001-7264-1323>. Machine Learning Applications in Medical Image Analysis. *Comput Math Methods Med*. 2017;2017:2361061. doi:http://dx.doi.org/10.1155/2017/2361061
75. Suzuki K, Zhang J, Xu J. Massive-training artificial neural network coupled with Laplacian-eigenfunction-based dimensionality reduction for computer-aided detection of polyps in CT colonography. *IEEE Trans Med Imaging*. 2010;29(11):1907-1917. doi:https://dx.doi.org/10.1109/TMI.2010.2053213
76. Lecun Y, Bengio Y, Hinton G. Deep learning. *Nature*. 2015;521(7553):436-444. doi:10.1038/nature14539
77. Mori Y, Kudo SE, Berzin TM, Misawa M, Takeda K. Computer-aided diagnosis for colonoscopy. *Endoscopy*. 2017;49(8):813-819. doi:10.1055/s-0043-109430
78. Linguraru MG, Zhao S, Van Uitert RL, et al. CAD of colon cancer on CT colonography cases without cathartic bowel preparation. *Conf Proc . Annu Int Conf IEEE Eng Med Biol Soc IEEE Eng Med Biol Soc Annu Conf*. 2008;2008:2996-2999. doi:https://dx.doi.org/10.1109/IEMBS.2008.4649833
79. G. T, B. B, Osman O. AO - Tulum GO <http://orcid.org/000.-0003-2865-5781>. A CAD of fully automated colonic polyp detection for contrasted and non-contrasted CT scans. *Int J Comput Assist Radiol Surg*. 2017;12(4):627-644. doi:http://dx.doi.org/10.1007/s11548-017-1521-9
80. Alagappan M, Brown JRG, Mori Y, Berzin TM. Artificial intelligence in gastrointestinal endoscopy: The future is almost here. *World J Gastrointest Endosc*. 2018;10(10):239-249. doi:10.4253/wjge.v10.i10.239
81. Namikawa K, Hirasawa T, Yoshio T, et al. Utilizing artificial intelligence in endoscopy: a clinician's guide. *Expert Rev Gastroenterol Hepatol*. 2020;14(8):689-706. doi:https://dx.doi.org/10.1080/17474124.2020.1779058
82. Wang P, Berzin TM, Romek Glissen Brown J, et al. Real-time automatic detection system increases colonoscopic polyp and adenoma detection rates: a prospective randomised controlled study. *Gut*. 2019;68:1813-1819. doi:10.1136/gutjnl-2018-317500
83. Urban G, Tripathi P, Alkayali T, et al. Deep Learning Localizes and Identifies Polyps in Real Time With 96% Accuracy in Screening Colonoscopy. *Gastroenterology*. 2018;155:1069-1078.e8. doi:10.1053/j.gastro.2018.06.037
84. F. D, F.M. B, Wahid K.A. AO - Deeba FO <http://orcid.org/000.-0001-9217-5032>. Computer-aided polyp detection based on image enhancement and saliency-based selection. *Biomed Signal Process Control*. 2020;55:101530. doi:http://dx.doi.org/10.1016/j.bspc.2019.04.007
85. Karkanis SA, Iakovidis DK, Maroulis DE, Karras DA, Tzivras M. Computer-Aided Tumor Detection in Endoscopic Video Using Color Wavelet Features. *IEEE Trans Inf Technol Biomed*. 2003;7(3):141-152. doi:10.1109/TITB.2003.813794
86. Wang Y, Tavanapong W, Wong J, Oh JH, de Groen PC. Polyp-Alert: Near real-time feedback during colonoscopy. *Comput Methods Programs Biomed*. 2015;120(3):164-179. doi:10.1016/j.cmpb.2015.04.002
87. Hwang S, Oh J, Tavanapong W, Wong J, De Groen PC. Polyp detection in colonoscopy video using elliptical shape feature. *Proc - Int Conf Image Process ICIP*. 2007;2:465-468. doi:10.1109/ICIP.2007.4379193
88. Fernández-Esparrach G, Bernal J, López-Cerón M, et al. Exploring the clinical potential of an automatic colonic polyp detection method based on the creation of energy maps. *Endoscopy*. 2016;48(9):837-842. doi:10.1055/s-0042-108434
89. Zhang R, Zheng Y, Mak TWC, et al. Automatic Detection and Classification of Colorectal Polyps by Transferring Low-Level CNN Features From Nonmedical Domain. *IEEE J Biomed Heal informatics*. 2017;21(1):41-47. doi:https://dx.doi.org/10.1109/JBHI.2016.2635662

90. M. M, S.-E. K, Y. M, et al. Artificial Intelligence-Assisted Polyp Detection for Colonoscopy: Initial Experience. *Gastroenterology*. 2018;154(8):2027. doi:http://dx.doi.org/10.1053/j.gastro.2018.04.003
91. C. H, M.B. W, P. S, R. M, V. C, M. S. New artificial intelligence system: First validation study versus experienced endoscopists for colorectal polyp detection. *Gut*. 2020;69(5):799-800. doi:http://dx.doi.org/10.1136/gutjnl-2019-319914
92. Wang P, Berzin TM, Glissen Brown JR, et al. Real-time automatic detection system increases colonoscopic polyp and adenoma detection rates: a prospective randomised controlled study. *Gut*. Published online 2019. doi:10.1136/gutjnl-2018-317500
93. Gong D, Wu L, Zhang J, et al. Detection of colorectal adenomas with a real-time computer-aided system (ENDOANGEL): a randomised controlled study. *Lancet Gastroenterol Hepatol*. 2020;0(0). doi:10.1016/s2468-1253(19)30413-3
94. A. R, M. B, R. M, et al. Efficacy of Real-Time Computer-Aided Detection of Colorectal Neoplasia in a Randomized Trial. *Gastroenterology*. 2020;159(2):512. doi:http://dx.doi.org/10.1053/j.gastro.2020.04.062
95. Attardo S, Chandrasekar VT, Spadaccini M, et al. Artificial intelligence technologies for the detection of colorectal lesions: The future is now. *World J Gastroenterol*. 2020;26(37):5606-5616. doi:10.3748/wjg.v26.i37.5606
96. Nazarian S, Koo H., Carrington E, Darzi A, Patel N. The future of endoscopy – what are the thoughts on artificial intelligence? <https://doi.org/10.1080/0952813X.2023.2178516>. Published online February 20, 2023:1-10. doi:10.1080/0952813X.2023.2178516
97. Jovanovic I. AI in endoscopy and medicolegal issues: the computer is guilty in case of missed cancer? *Endosc Int Open*. 2020;8(10):E1385. doi:10.1055/A-1214-5858
98. Aisu N, Miyake M, Takeshita K, et al. Regulatory-approved deep learning/machine learning-based medical devices in Japan as of 2020: A systematic review. *PLOS Digit Heal*. 2022;1(1):e0000001. doi:10.1371/JOURNAL.PDIG.0000001
99. Lee TJW, Rutter MD, Blanks RG, et al. Colonoscopy quality measures: experience from the NHS Bowel Cancer Screening Programme. *Gut*. 2012;61(7):1050-1057. doi:10.1136/GUTJNL-2011-300651
100. O'brien MJ, Winawer SJ, Zauber AG, et al. The National Polyp Study Patient and Polyp Characteristics Associated With High-Grade Dysplasia in Colorectal Adenomas.
101. Ferlitsch M, Moss A, Hassan C, et al. Colorectal polypectomy and endoscopic mucosal resection (EMR): European Society of Gastrointestinal Endoscopy (ESGE) Clinical Guideline. *Endoscopy*. 2017;49(3):270-297. doi:10.1055/S-0043-102569
102. Axon A, Diebold MD, Fujino M, et al. Update on the Paris classification of superficial neoplastic lesions in the digestive tract. *Endoscopy*. 2005;37(6):570-578. doi:10.1055/s-2005-861352
103. Tanaka S, Sano Y. Aim to unify the narrow band imaging (NBI) magnifying classification for colorectal tumors: current status in Japan from a summary of the consensus symposium in the 79th Annual Meeting of the Japan Gastroenterological Endoscopy Society. *Dig Endosc*. 2011;23 Suppl 1(SUPPL. 1):131-139. doi:10.1111/J.1443-1661.2011.01106.X
104. Hewett DG, Kaltenbach T, Sano Y, et al. Validation of a simple classification system for endoscopic diagnosis of small colorectal polyps using narrow-band imaging. *Gastroenterology*. 2012;143(3). doi:10.1053/J.GASTRO.2012.05.006
105. Mathews AA, Draganov P V, Yang D. Endoscopic management of colorectal polyps: From benign to malignant polyps. *World J Gastrointest Endosc*. 2021;13(9):356. doi:10.4253/WJGE.V13.I9.356
106. Zauber AG, Winawer SJ, O'Brien MJ, et al. Colonoscopic polypectomy and long-term prevention of colorectal-cancer deaths. *N Engl J Med*. 2012;366(8):687-696. doi:10.1056/NEJMoa1100370
107. Mann R, Gajendran M, Umapathy C, et al. Endoscopic Management of Complex Colorectal

- Polyps: Current Insights and Future Trends. *Front Med.* 2022;8:728704. doi:10.3389/FMED.2021.728704/BIBTEX
108. Efthymiou M, Taylor ACF, Desmond P V., Allen PB, Chen RY. Biopsy forceps is inadequate for the resection of diminutive polyps. *Endoscopy.* 2011;43(4):312-316. doi:10.1055/S-0030-1256086
 109. Kim JS, Lee BI, Choi H, et al. Cold snare polypectomy versus cold forceps polypectomy for diminutive and small colorectal polyps: a randomized controlled trial. *Gastrointest Endosc.* 2015;81(3):741-747. doi:10.1016/J.GIE.2014.11.048
 110. Draganov P V., Chang MN, Alkhasawneh A, et al. Randomized, controlled trial of standard, large-capacity versus jumbo biopsy forceps for polypectomy of small, sessile, colorectal polyps. *Gastrointest Endosc.* 2012;75(1):118-126. doi:10.1016/J.GIE.2011.08.019
 111. Ishibashi F, Suzuki S, Nagai M, Mochida K, Morishita T. Colorectal cold snare polypectomy: Current standard technique and future perspectives. *Dig Endosc.* 2023;35(3):278-286. doi:10.1111/DEN.14420
 112. Cold Snare Versus Hot Snare for Removal of Colorectal Polyps - The British Society of Gastroenterology. Accessed September 25, 2023. <https://www.bsg.org.uk/web-education-articles-list/cold-snare-versus-hot-snare-for-removal-of-colorectal-polyps/>
 113. Siau K, Murugananthan A, Dunckley P, Smith G V., Thomas-Gibson S. National survey of UK endoscopists showing variation in diathermy practice for colonic polypectomy: a JAG perspective. *Frontline Gastroenterol.* 2019;10(4):444. doi:10.1136/FLGASTRO-2018-101133
 114. Rutter MD, Chattree A, Barbour JA, et al. British Society of Gastroenterology/Association of Coloproctologists of Great Britain and Ireland guidelines for the management of large non-pedunculated colorectal polyps. *Gut.* 2015;64(12):1847-1873. doi:10.1136/GUTJNL-2015-309576
 115. Castro R, Libânio D, Pita I, Dinis-Ribeiro M. Solutions for submucosal injection: What to choose and how to do it. *World J Gastroenterol.* 2019;25(7):777. doi:10.3748/WJG.V25.I7.777
 116. Horiuchi A, Makino T, Kajiyama M, Tanaka N, Sano K, Graham DY. Comparison between endoscopic mucosal resection and hot snare resection of large nonpedunculated colorectal polyps: a randomized trial. *Endoscopy.* 2016;48(7):646-651. doi:10.1055/S-0042-105557
 117. Kaltenbach T, Anderson JC, Burke CA, et al. Endoscopic Removal of Colorectal Lesions: Recommendations by the US Multi-Society Task Force on Colorectal Cancer. *Am J Gastroenterol.* 2020;115(3):435-464. doi:10.14309/AJG.0000000000000555
 118. Tanaka S, Kashida H, Saito Y, et al. Japan Gastroenterological Endoscopy Society guidelines for colorectal endoscopic submucosal dissection/endoscopic mucosal resection. *Dig Endosc.* 2020;32(2):219-239. doi:10.1111/DEN.13545
 119. Chandan S, Khan SR, Kumar A, et al. Efficacy and histologic accuracy of underwater versus conventional endoscopic mucosal resection for large (>20 mm) colorectal polyps: a comparative review and meta-analysis. *Gastrointest Endosc.* 2021;94(3):471-482.e9. doi:10.1016/J.GIE.2020.12.034
 120. Yandrapu H, Desai M, Siddique S, et al. Normal saline solution versus other viscous solutions for submucosal injection during endoscopic mucosal resection: a systematic review and meta-analysis. *Gastrointest Endosc.* 2017;85(4):693-699. doi:10.1016/J.GIE.2016.12.003
 121. Tim G Belderbos AD, Leenders M, G Moons LM, Siersema PD, G Belderbos TD. Local recurrence after endoscopic mucosal resection of nonpedunculated colorectal lesions: systematic review and meta-analysis. *Endoscopy.* 2014;46:388-400. doi:10.1055/s-0034-1364970
 122. Moss A, Williams SJ, Hourigan LF, et al. Long-term adenoma recurrence following wide-field endoscopic mucosal resection (WF-EMR) for advanced colonic mucosal neoplasia is infrequent: results and risk factors in 1000 cases from the Australian Colonic EMR (ACE) study. *Gut.* 2015;64(1):57-65. doi:10.1136/GUTJNL-2013-305516
 123. Shigita K, Oka S, Tanaka S, et al. Long-term outcomes after endoscopic submucosal dissection

- for superficial colorectal tumors. *Gastrointest Endosc.* 2017;85(3):546-553. doi:10.1016/J.GIE.2016.07.044
124. Ohata K, Kobayashi N, Sakai E, et al. Long-term Outcomes After Endoscopic Submucosal Dissection for Large Colorectal Epithelial Neoplasms: A Prospective, Multicenter, Cohort Trial From Japan. *Gastroenterology.* 2022;163(5):1423-1434.e2. doi:10.1053/j.gastro.2022.07.002
 125. Patel N, Patel K, Ashrafian H, Athanasiou T, Darzi A, Teare J. Colorectal endoscopic submucosal dissection: Systematic review of mid-term clinical outcomes. *Dig Endosc.* 2016;28(4):405-416. doi:10.1111/DEN.12597
 126. Fuccio L, Hassan C, Ponchon T, et al. Clinical outcomes after endoscopic submucosal dissection for colorectal neoplasia: a systematic review and meta-analysis. *Gastrointest Endosc.* 2017;86(1):74-86.e17. doi:10.1016/J.GIE.2017.02.024
 127. Fujiya M, Tanaka K, Dokoshi T, et al. Efficacy and adverse events of EMR and endoscopic submucosal dissection for the treatment of colon neoplasms: a meta-analysis of studies comparing EMR and endoscopic submucosal dissection. *Gastrointest Endosc.* 2015;81(3):583-595. doi:10.1016/J.GIE.2014.07.034
 128. Gao P, Zhou K, Su W, Yu J, Zhou P. Endoscopic management of colorectal polyps. *Gastroenterol Rep.* 2023;11. doi:10.1093/GASTRO/GOAD027
 129. McCarty TR, Aihara H. Current state of education and training for endoscopic submucosal dissection: Translating strategy and success to the USA. *Dig Endosc.* 2020;32(6):851-860. doi:10.1111/DEN.13591
 130. Berr F, Ponchon T, Neureiter D, et al. Experimental endoscopic submucosal dissection training in a porcine model: learning experience of skilled Western endoscopists. *Dig Endosc.* 2011;23(4):281-289. doi:10.1111/J.1443-1661.2011.01129.X
 131. Oda I, Odagaki T, Suzuki H, Nonaka S, Yoshinaga S. LEARNING CURVE FOR ENDOSCOPIC SUBMUCOSAL DISSECTION OF EARLY GASTRIC CANCER BASED ON TRAINEE EXPERIENCE. *Dig Endosc.* 2012;24(SUPPL. 1):129-132. doi:10.1111/J.1443-1661.2012.01265.X
 132. De Ceglie A, Hassan C, Mangiavillano B, et al. Endoscopic mucosal resection and endoscopic submucosal dissection for colorectal lesions: A systematic review. *Crit Rev Oncol Hematol.* 2016;104:138-155. doi:10.1016/J.CRITREVNOC.2016.06.008
 133. Kim ES, Cho KB, Park KS, et al. Factors predictive of perforation during endoscopic submucosal dissection for the treatment of colorectal tumors. *Endoscopy.* 2011;43(7):573-578. doi:10.1055/S-0030-1256339
 134. Rutter MD, Nickerson C, Rees CJ, Patnick J, Blanks RG. Risk factors for adverse events related to polypectomy in the English Bowel Cancer Screening Programme. *Endoscopy.* 2014;46(2):90-97. doi:10.1055/S-0033-1344987
 135. Mönkemüller K, Neumann H, Fry LC, Ivekovic H, Malfertheiner P. Polypectomy techniques for difficult colon polyps. *Dig Dis.* 2008;26(4):342-346. doi:10.1159/000177020
 136. Pattarajierapan S, Takamaru H, Khomvilai S. Difficult colorectal polypectomy: Technical tips and recent advances. *World J Gastroenterol.* 2023;29(17):2600-2615. doi:10.3748/wjg.v29.i17.2600
 137. Jung M. The “difficult” polyp: pitfalls for endoscopic removal. *Dig Dis.* 2012;30 Suppl 2(SUPPL.2):74-80. doi:10.1159/000341898
 138. Rex DK, Bond JH, Feld AD. Medical-legal risks of incident cancers after clearing colonoscopy. *Am J Gastroenterol.* 2001;96(4):952-957. doi:10.1111/J.1572-0241.2001.03677.X
 139. Choi JY, Jung SA, Shim KN, et al. Meta-analysis of Predictive Clinicopathologic Factors for Lymph Node Metastasis in Patients with Early Colorectal Carcinoma. *J Korean Med Sci.* 2015;30(4):398. doi:10.3346/JKMS.2015.30.4.398
 140. Kawaguti FS, Franco MC, Martins BC, et al. Role of Magnification Chromoendoscopy in the Management of Colorectal Neoplastic Lesions Suspicious for Submucosal Invasion. *Dis Colon Rectum.* 2019;62(4):422-428. doi:10.1097/DCR.0000000000001343
 141. Yanai S, Nakamura S, Matsumoto T. Role of magnifying colonoscopy for diagnosis of

- colorectal neoplasms: From the perspective of Japanese colonoscopists. *Dig Endosc.* 2016;28(3):274-280. doi:10.1111/DEN.12568
142. Beets-Tan RGH, Beets GL. Rectal cancer: review with emphasis on MR imaging. *Radiology.* 2004;232(2):335-346. doi:10.1148/RADIOL.2322021326
 143. Rockey DC, Paulson E, Niedzwiecki D, et al. Analysis of air contrast barium enema, computed tomographic colonography, and colonoscopy: Prospective comparison. *Lancet.* 2005;365(9456):305-311. doi:10.1016/S0140-6736(05)70194-X
 144. Piscatelli N, Hyman N, Osler T. Localizing colorectal cancer by colonoscopy. *Arch Surg.* 2005;140(10):932-935. doi:10.1001/ARCHSURG.140.10.932
 145. Cho YB, Lee WY, Yun HR, Lee WS, Yun SH, Chun HK. Tumor localization for laparoscopic colorectal surgery. *World J Surg.* 2007;31(7):1491-1495. doi:10.1007/S00268-007-9082-7
 146. Trakarnsanga A, Akaraviputh T. Endoscopic tattooing of colorectal lesions: Is it a risk-free procedure? *World J Gastrointest Endosc.* 2011;3(12):256. doi:10.4253/WJGE.V3.I12.256
 147. McArthur CS, Roayaie S, Wayne JD. Safety of preoperation endoscopic tattoo with india ink for identification of colonic lesions. *Surg Endosc.* 1999;13(4):397-400. doi:10.1007/S004649900997
 148. EBSCOhost | 16064848 | Colonic Tattooing with India Ink: Benefits, Risks, and Alternatives. Accessed October 18, 2023. <https://web.s.ebscohost.com/abstract?direct=true&profile=ehost&scope=site&authtype=crawler&jrnl=00029270&AN=16064848&h=ogeNte%2Fbiosg3ZIEmRmXy%2BE4FeizVGMEP9vpEt7ZgkxMgxyBbgRU2ejpfabrOGs5gfX5t0e5%2B%2FqEiuyqTTdsMg%3D%3D&crl=f&resultNs=AdminWebAuth&resultLocal=ErrCrlNotAuth&crlhashurl=login.aspx%3Fdirect%3Dtrue%26profile%3Dehost%26scope%3Dsite%26authtype%3Dcrawler%26jrnl%3D00029270%26AN%3D16064848>
 149. Amri R, Bordeianou LG, Sylla P, Berger DL. Association of Radial Margin Positivity With Colon Cancer. *JAMA Surg.* 2015;150(9):890-898. doi:10.1001/JAMASURG.2015.1525
 150. Quirke P, Dixon MF, Durdey P, Williams NS. Local recurrence of rectal adenocarcinoma due to inadequate surgical resection. Histopathological study of lateral tumour spread and surgical excision. *Lancet (London, England).* 1986;2(8514):996-999. doi:10.1016/S0140-6736(86)92612-7
 151. Khan MAS, Hakeem AR, Scott N, Saunders RN. Significance of R1 resection margin in colon cancer resections in the modern era. *Color Dis.* 2015;17(11):943-953. doi:10.1111/CODI.12960
 152. Wibe A, Rendedal PR, Svensson E, et al. Prognostic significance of the circumferential resection margin following total mesorectal excision for rectal cancer. *Br J Surg.* 2002;89(3):327-334. doi:10.1046/J.0007-1323.2001.02024.X
 153. Zeng WG, Liu MJ, Zhou ZX, Wang ZJ. A Distal Resection Margin of ≤ 1 mm and Rectal Cancer Recurrence After Sphincter-Preserving Surgery: The Role of a Positive Distal Margin in Rectal Cancer Surgery. *Dis Colon Rectum.* 2017;60(11):1175-1183. doi:10.1097/DCR.0000000000000900
 154. Rickles AS, Dietz DW, Chang GJ, et al. High Rate of Positive Circumferential Resection Margins Following Rectal Cancer Surgery: A Call to Action. *Ann Surg.* 2015;262(6):891-898. doi:10.1097/SLA.0000000000001391
 155. Nakanishi K, Morita S, Taniguchi H, Otsuki S, Fukagawa T, Katai H. Diagnostic Accuracy and Usefulness of Intraoperative Margin Assessment by Frozen Section in Gastric Cancer. *Ann Surg Oncol.* 2019;26(6). doi:10.1245/S10434-019-07302-1
 156. St John ER, Al-Khudairi R, Ashrafian H, et al. Diagnostic accuracy of intraoperative techniques for margin assessment in breast cancer surgery a meta-analysis. *Ann Surg.* 2017;265(2):300-310. doi:10.1097/SLA.0000000000001897
 157. DiNardo LJ, Lin J, Karageorge LS, Powers CN. Accuracy, utility, and cost of frozen section margins in head and neck cancer surgery. *Laryngoscope.* 2000;110(10 Pt 1):1773-1776.

- doi:10.1097/00005537-200010000-00039
158. Kobashi M, Ishikawa S, Inaba T, et al. Diagnostic accuracy of frozen section biopsy for early gastric cancer extent during endoscopic submucosal dissection: a prospective study. *Surg Endosc.* 2023;37(9):6736-6748. doi:10.1007/S00464-023-10100-2/TABLES/4
 159. Heidkamp J, Scholte M, Rosman C, Manohar S, Fütterer JJ, Rovers MM. Novel imaging techniques for intraoperative margin assessment in surgical oncology: A systematic review. *Int J Cancer.* 2021;149(3):635. doi:10.1002/IJC.33570
 160. Jacques SL. Optical properties of biological tissues: a review. *Phys Med Biol.* 2013;58(11). doi:10.1088/0031-9155/58/11/R37
 161. O'Sullivan TD, Cerussi AE, Cuccia DJ, Tromberg BJ. Diffuse optical imaging using spatially and temporally modulated light. *J Biomed Opt.* 2012;17(7):071311. doi:10.1117/1.JBO.17.7.071311
 162. Dj E, B H, G L, T R. Optical spectroscopy: current advances and future applications in cancer diagnostics and therapy. *Future Oncol.* 2012;8(3):307-320. doi:10.2217/FON.12.15
 163. Evers DJ, Nachabé R, Klomp HM, et al. Diffuse reflectance spectroscopy: a new guidance tool for improvement of biopsy procedures in lung malignancies. *Clin Lung Cancer.* 2012;13(6):424-431. doi:10.1016/J.CLLC.2012.02.001
 164. Zhang Y, Austin J, Moy J, et al. Diffuse reflectance spectroscopy as a potential method for nonmelanoma skin cancer margin assessment. *Transl Biophotonics.* 2020;2(3):e202000001. doi:10.1002/TBIO.202000001
 165. De Boer LL, Bydlon TM, Van Duijnhoven F, et al. Towards the use of diffuse reflectance spectroscopy for real-time in vivo detection of breast cancer during surgery. *J Transl Med.* 2018;16(1):1-14. doi:10.1186/S12967-018-1747-5/FIGURES/5
 166. Shaikh R, Prabitha VG, Dora TK, et al. A comparative evaluation of diffuse reflectance and Raman spectroscopy in the detection of cervical cancer. *J Biophotonics.* 2017;10(2):242-252. doi:10.1002/JBIO.201500248
 167. Skyrman S, Burström G, Lai M, et al. Diffuse reflectance spectroscopy sensor to differentiate between glial tumor and healthy brain tissue: a proof-of-concept study. *Biomed Opt Express.* 2022;13(12):6470. doi:10.1364/BOE.474344
 168. Qiu L, Pleskow DK, Chuttani R, et al. In Vivo Optical Detection of Dysplasia in Barrett's Esophagus with Endoscopic Light Scattering Spectroscopy. *Barrett's Esophagus Emerg Evid Improv Clin Pract.* Published online January 1, 2016:91-104. doi:10.1016/B978-0-12-802511-6.00007-7
 169. Greening G, Mundo A, Rajaram N, Muldoon TJ. Sampling depth of a diffuse reflectance spectroscopy probe for in-vivo physiological quantification of murine subcutaneous tumor allografts. *J Biomed Opt.* 2018;23(8):1. doi:10.1117/1.JBO.23.8.085006
 170. Strattonnikov AA, Loschenov VB. Evaluation of blood oxygen saturation in vivo from diffuse reflectance spectra. *J Biomed Opt.* 2001;6(4):457. doi:10.1117/1.1411979
 171. Nguyen MH, Zhang Y, Wang F, Linan JDLGE, Markey MK, Tunnell JW. Machine learning to extract physiological parameters from multispectral diffuse reflectance spectroscopy. *J Biomed Opt.* 2021;26(5). doi:10.1117/1.JBO.26.5.052912
 172. *Colorectal Cancer Source: Globocan 2018 Number of New Cases in 2018, Both Sexes, All Ages.*; 2018. Accessed November 1, 2019. <http://gco.iarc.fr/today>
 173. Zhao S, Wang S, Pan P, et al. Magnitude, Risk Factors, and Factors Associated With Adenoma Miss Rate of Tandem Colonoscopy: A Systematic Review and Meta-analysis. *Gastroenterology.* 2019;156(6):1661-1674.e11. doi:10.1053/j.gastro.2019.01.260
 174. Pohl H, Bensen SP, Toor A, et al. Cap-assisted colonoscopy and detection of Adenomatous Polyps (CAP) study: A randomized trial. *Endoscopy.* 2015;47(10):891-897. doi:10.1055/s-0034-1392261
 175. Halpern Z, Gross SA, Gralnek IM, et al. Comparison of adenoma detection and miss rates between a novel balloon colonoscope and standard colonoscopy: A randomized tandem

- study. *Endoscopy*. 2015;47(3):238-244. doi:10.1055/s-0034-1391437
176. Papanikolaou IS, Apostolopoulos P, Tziatzios G, et al. Lower adenoma miss rate with FUSE vsconventional colonoscopy with proximal retroflexion: A randomized back-to-back trial. *Endoscopy*. 2017;49(5):468-475. doi:10.1055/s-0042-124415
177. Leufkens AM, Demarco DC, Rastogi A, et al. Effect of a retrograde-viewing device on adenoma detection rate during colonoscopy: The TERRACE study. *Gastrointest Endosc*. 2011;73(3):480-489. doi:10.1016/j.gie.2010.09.004
178. Siersema PD, Rastogi A, Leufkens AM, et al. Retrograde-viewing device improves adenoma detection rate in colonoscopies for surveillance and diagnostic workup. *World J Gastroenterol*. 2012;18(26):3400-3408. doi:10.3748/wjg.v18.i26.3400
179. Chokshi R V., Hovis CE, Hollander T, Early DS, Wang JS. Prevalence of missed adenomas in patients with inadequate bowel preparation on screening colonoscopy. *Gastrointest Endosc*. 2012;75(6):1197-1203. doi:10.1016/J.GIE.2012.01.005
180. Rex DK, Cutler CS, Lemmel GT, et al. Colonoscopic miss rates of adenomas determined by back-to-back colonoscopies. *Gastroenterology*. 1997;112(1):24-28. doi:10.1016/S0016-5085(97)70214-2
181. Coe SG, Crook JE, Diehl NN, Wallace MB. An endoscopic quality improvement program improves detection of colorectal adenomas. *Am J Gastroenterol*. 2013;108(2):219-226. doi:10.1038/ajg.2012.417
182. Lim S, Hammond S, Park J, et al. Training interventions to improve adenoma detection rates during colonoscopy: a systematic review and meta-analysis. *Surg Endosc* 2019 349. 2019;34(9):3870-3882. doi:10.1007/S00464-019-07153-7
183. Killock D. AI outperforms radiologists in mammographic screening. *Nat Rev Clin Oncol*. 2020;17(3):134. doi:10.1038/s41571-020-0329-7
184. ELKarazle K, Raman V, Then P, Chua C. Detection of Colorectal Polyps from Colonoscopy Using Machine Learning: A Survey on Modern Techniques. *Sensors (Basel)*. 2023;23(3). doi:10.3390/S23031225
185. Ponugoti PL, Cummings OW, Rex DK. Risk of cancer in small and diminutive colorectal polyps. *Dig Liver Dis*. 2017;49(1):34-37. doi:10.1016/J.DLD.2016.06.025
186. Hajjar A El, Rey JF. Artificial intelligence in gastrointestinal endoscopy: general overview. *Chin Med J (Engl)*. 2020;133(3):326-334. doi:10.1097/CM9.0000000000000623
187. Rex DK, Kahi C, O'Brien M, et al. The American Society for Gastrointestinal Endoscopy PIVI (Preservation and Incorporation of Valuable Endoscopic Innovations) on real-time endoscopic assessment of the histology of diminutive colorectal polyps. *Gastrointest Endosc*. 2011;73(3):419-422. doi:10.1016/j.gie.2011.01.023
188. Ignjatovic A, East JE, Suzuki N, Vance M, Guenther T, Saunders BP. Optical diagnosis of small colorectal polyps at routine colonoscopy (Detect InSpect ChAracterise Resect and Discard; DISCARD trial): a prospective cohort study. *Lancet Oncol*. 2009;10(12):1171-1178. doi:10.1016/S1470-2045(09)70329-8
189. Kominami Y, Yoshida S, Tanaka S, et al. Computer-aided diagnosis of colorectal polyp histology by using a real-time image recognition system and narrow-band imaging magnifying colonoscopy. *Gastrointest Endosc*. 2016;83(3):643-649. doi:10.1016/j.gie.2015.08.004
190. Rees CJ, Rajasekhar PT, Wilson A, et al. Narrow band imaging optical diagnosis of small colorectal polyps in routine clinical practice: The Detect Inspect Characterise Resect and Discard 2 (DISCARD 2) study. *Gut*. 2016;66(5):887-895. doi:10.1136/gutjnl-2015-310584
191. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Med*. 2009;6(7):e1000097. doi:10.1371/journal.pmed.1000097
192. Whiting PF, Rutjes AWS, Westwood ME, et al. Quadas-2: A revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med*. 2011;155(8):529-536. doi:10.7326/0003-4819-155-8-201110180-00009

193. Jadad AR, Moore RA, Carroll D, et al. Assessing the quality of reports of randomized clinical trials: Is blinding necessary? *Control Clin Trials*. 1996;17(1):1-12. doi:10.1016/0197-2456(95)00134-4
194. Nazarian S, Glover B, Ashrafian H, Darzi A, Teare J. Diagnostic Accuracy of Artificial Intelligence and Computer-Aided Diagnosis for the Detection and Characterization of Colorectal Polyps: Systematic Review and Meta-analysis. *J Med Internet Res*. 2021;23(7). doi:10.2196/27370
195. Wang P, Xiao X, Glissen Brown JR, et al. Development and validation of a deep-learning algorithm for the detection of polyps during colonoscopy. *Nat Biomed Eng*. 2018;2(10):741-748. doi:10.1038/s41551-018-0301-3
196. Fu JJC, Yu Y-W, Lin H-M, Chai J-W, Chen CC-C. Feature extraction and pattern classification of colorectal polyps in colonoscopic imaging. *Comput Med Imaging Graph*. 2014;38(4):267-275. doi:https://dx.doi.org/10.1016/j.compmedimag.2013.12.009
197. Tajbakhsh N, Gurudu SR, Liang J. Automated Polyp Detection in Colonoscopy Videos Using Shape and Context Information. *IEEE Trans Med Imaging*. 2016;35(2):630-644. doi:https://dx.doi.org/10.1109/TMI.2015.2487997
198. Park SY, Sargent D. Colonoscopic polyp detection using convolutional neural networks. In: *Medical Imaging 2016: Computer-Aided Diagnosis*. Vol 9785. SPIE; 2016:978528. doi:10.1117/12.2217148
199. Urban G, Tripathi P, Alkayali T, et al. Deep Learning Localizes and Identifies Polyps in Real Time With 96% Accuracy in Screening Colonoscopy. *Gastroenterology*. 2018;155(4):1069-1078.e8. doi:https://dx.doi.org/10.1053/j.gastro.2018.06.037
200. Misawa M, Kudo S ei, Mori Y, et al. Artificial Intelligence-Assisted Polyp Detection for Colonoscopy: Initial Experience. *Gastroenterology*. 2018;154(8):2027-2029.e3. doi:10.1053/j.gastro.2018.04.003
201. Figueiredo PN, Figueiredo IN, Pinto L, Kumar S, Tsai Y-HR, Mamonov A V. Polyp detection with computer-aided diagnosis in white light colonoscopy: comparison of three different methods. *Endosc Int open*. 2019;7(2):E209-E215. doi:https://dx.doi.org/10.1055/a-0808-4456
202. Yamada M, Saito Y, Imaoka H, et al. Development of a real-time endoscopic image diagnosis support system using deep learning technology in colonoscopy. *Sci Rep*. 2019;9(1). doi:10.1038/s41598-019-50567-5
203. Tischendorf JJW, Gross S, Winograd R, et al. Computer-aided classification of colorectal polyps based on vascular patterns: A pilot study. *Endoscopy*. 2010;42(3):203-207. doi:10.1055/s-0029-1243861
204. Gross S, Trautwein C, Behrens A, et al. Computer-based classification of small colorectal polyps by using narrow-band imaging with optical magnification. *Gastrointest Endosc*. 2011;74(6):1354-1359. doi:10.1016/j.gie.2011.08.001
205. Ganz M, Yang X, Slabaugh G. Automatic segmentation of polyps in colonoscopic narrow-band imaging data. *IEEE Trans Biomed Eng*. 2012;59(8):2144-2151. doi:https://dx.doi.org/10.1109/TBME.2012.2195314
206. Takemura Y, Yoshida S, Tanaka S, et al. Computer-aided system for predicting the histology of colorectal tumors by using narrow-band imaging magnifying colonoscopy (with video). *Gastrointest Endosc*. 2012;75(1):179-185. doi:10.1016/j.gie.2011.08.051
207. Mori Y, Kudo SE, Wakamura K, et al. Novel computer-aided diagnostic system for colorectal lesions by using endocytoscopy (with videos). *Gastrointest Endosc*. 2015;81(3):621-629. doi:10.1016/j.gie.2014.09.008
208. Misawa M, Kudo SE, Mori Y, et al. Characterization of Colorectal Lesions Using a Computer-Aided Diagnostic System for Narrow-Band Imaging Endocytoscopy. *Gastroenterology*. 2016;150(7):1531-1532.e3. doi:10.1053/j.gastro.2016.04.004
209. Mesejo P, Pizarro D, Abergel A, et al. Computer-Aided Classification of Gastrointestinal Lesions in Regular Colonoscopy. *IEEE Trans Med Imaging*. 2016;35(9):2051-2063.

- doi:<https://dx.doi.org/10.1109/TMI.2016.2547947>
210. Mori Y, Kudo S-E, Wai P, et al. Impact of an automated system for endocytoscopic diagnosis of small colorectal lesions: an international web-based study. *Endoscopy*. 2016;48:1110-1118. doi:10.1055/s-0042-113609
 211. Takeda K, Kudo SE, Mori Y, et al. Accuracy of diagnosing invasive colorectal cancer using computer-aided endocytoscopy. *Endoscopy*. 2017;49(8):798-802. doi:10.1055/s-0043-105486
 212. Byrne MF, Chapados N, Soudan F, Oertel C. Real-time differentiation of adenomatous and hyperplastic diminutive colorectal polyps during analysis of unaltered videos of standard colonoscopy using a deep learning model. *Gut*. 2019;68:94-100. doi:10.1136/gutjnl-2017-314547
 213. Komeda Y, Handa H, Watanabe T, et al. Computer-Aided Diagnosis Based on Convolutional Neural Network System for Colorectal Polyp Classification: Preliminary Experience. *Oncology*. 2017;93 Suppl 1:30-34. doi:<https://dx.doi.org/10.1159/000481227>
 214. Misawa M, Kudo S-E, Mori Y, et al. Accuracy of computer-aided diagnosis based on narrow-band imaging endocytoscopy for diagnosing colorectal lesions: comparison with experts. *Int J Comput Assist Radiol Surg*. 2017;12(5):757-766. doi:<https://dx.doi.org/10.1007/s11548-017-1542-4>
 215. Mori Y, Kudo S-E, Mori K. Potential of artificial intelligence-assisted colonoscopy using an endocytoscope (with video). *Dig Endosc*. 2018;30 Suppl 1:52-53. doi:<https://dx.doi.org/10.1111/den.13005>
 216. Chen PJ, Lin MC, Lai MJ, Lin JC, Lu HHS, Tseng VS. Accurate Classification of Diminutive Colorectal Polyps Using Computer-Aided Analysis. *Gastroenterology*. 2018;154(3):568-575. doi:10.1053/j.gastro.2017.10.010
 217. Renner J, Philipsen H, Haller B, et al. Optical classification of neoplastic colorectal polyps—a computer-assisted approach (the COACH study). *Scand J Gastroenterol*. 2018;53(9):1100-1106. doi:10.1080/00365521.2018.1501092
 218. Mori Y, Kudo S-E, Misawa M, et al. Real-Time Use of Artificial Intelligence in Identification of Diminutive Polyps During Colonoscopy: A Prospective Study. *Ann Intern Med*. 2018;169(6):357-366. doi:<https://dx.doi.org/10.7326/M18-0249>
 219. Kudo S-E, Misawa M, Mori Y, et al. Artificial Intelligence-assisted System Improves Endoscopic Identification of Colorectal Neoplasms. *Clin Gastroenterol Hepatol*. Published online September 13, 2019. doi:10.1016/j.cgh.2019.09.009
 220. Figueiredo IN, Pinto L, Figueiredo PN, Tsai R. Unsupervised segmentation of colonic polyps in narrow-band imaging data based on manifold representation of images and Wasserstein distance. *Biomed Signal Process Control*. 2019;53. doi:10.1016/j.bspc.2019.101577
 221. Wang P, Liu X, Berzin TM, et al. Effect of a deep-learning computer-aided detection system on adenoma detection during colonoscopy (CADE-DB trial): a double-blind randomised study. *Lancet Gastroenterol Hepatol*. 2020;0(0). doi:10.1016/s2468-1253(19)30411-x
 222. Su JR, Li Z, Shao XJ, et al. Impact of a real-time automatic quality control system on colorectal polyp and adenoma detection: a prospective randomized controlled study (with videos). *Gastrointest Endosc*. 2020;91(2):415-424.e4. doi:10.1016/j.gie.2019.08.026
 223. W.-N. L, Y.-Y. Z, X.-Q. B, L.-J. W, Q. Y, X.-D. Z. Study on detection rate of polyps and adenomas in artificial-intelligence-aided colonoscopy. *Saudi J Gastroenterol*. 2020;26(1):13-19. doi:http://dx.doi.org/10.4103/sjg.SJG_377_19
 224. Y. L, Y. Z, M. L, et al. Artificial Intelligence-Assisted Colonoscopy for Detection of Colon Polyps: a Prospective, Randomized Cohort Study. *J Gastrointest Surg*. Published online 2020. doi:<http://dx.doi.org/10.1007/s11605-020-04802-4>
 225. Wang P, Liu P, Glissen Brown JR, et al. Lower Adenoma Miss Rate of Computer-Aided Detection-Assisted Colonoscopy vs Routine White-Light Colonoscopy in a Prospective Tandem Study. *Gastroenterology*. Published online 2020. doi:<https://dx.doi.org/10.1053/j.gastro.2020.06.023>

226. Liu W-N, Zhang Y-Y, Bian X-Q, et al. Study on detection rate of polyps and adenomas in artificial-intelligence-aided colonoscopy. *Saudi J Gastroenterol*. 2020;26(1):13-19. doi:https://dx.doi.org/10.4103/sjg.SJG_377_19
227. Kim NH, Jung YS, Jeong WS, et al. Miss rate of colorectal neoplastic polyps and risk factors for missed polyps in consecutive colonoscopies. *Intest Res*. 2017;15(3):411-418. doi:10.5217/ir.2017.15.3.411
228. Leufkens AM, Van Oijen MGH, Vleggaar FP, Siersema PD. Factors influencing the miss rate of polyps in a back-to-back colonoscopy study. *Endoscopy*. 2012;44(5):470-475. doi:10.1055/s-0031-1291666
229. S. G, C. T, A. B, et al. Computer-based classification of small colorectal polyps by using narrow-band imaging with optical magnification. *Gastrointest Endosc*. 2011;74(6):1354-1359. doi:<http://dx.doi.org/10.1016/j.gie.2011.08.001>
230. Clark BT, Rustagi T, Laine L. What level of bowel prep quality requires early repeat colonoscopy: Systematic review and meta-analysis of the impact of preparation quality on adenoma detection rate. *Am J Gastroenterol*. 2014;109(11):1714-1723. doi:10.1038/ajg.2014.232
231. Lai EJ, Calderwood AH, Doros G, Fix OK, Jacobson BC. The Boston bowel preparation scale: a valid and reliable instrument for colonoscopy-oriented research. *Gastrointest Endosc*. 2009;69(3 SUPPL.):620-625. doi:10.1016/j.gie.2008.05.057
232. Shaukat A, Rector TS, Church TR, et al. Longer Withdrawal Time Is Associated With a Reduced Incidence of Interval Cancer After Screening Colonoscopy. *Gastroenterology*. 2015;149(4):952-957. doi:10.1053/j.gastro.2015.06.044
233. Reporting guidelines under development for other study designs | The EQUATOR Network. Accessed March 24, 2020. <https://www.equator-network.org/library/reporting-guidelines-under-development/reporting-guidelines-under-development-for-other-study-designs/#STARDAI>
234. Sounderajah V, Ashrafian H, Aggarwal R, De Fauw J, Denniston AK, Greaves F, Karthikesalingam A, King D, Liu X, Markar SR, McInnes MDF, Panch T, Pearson-Stuttard J, Ting D, Golub R, Moher D, Bossuyt P DA. Developing specific reporting guidelines for diagnostic accuracy studies assessing artificial intelligence interventions: The STARD-AI Steering Group. *Press*. Published online 2020.
235. Hur J, Baek MJ. Limitation and value of using the adenoma detection rate for colonoscopy quality assurance. *Ann Coloproctol*. 2017;33(3):81. doi:10.3393/ac.2017.33.3.81
236. Mori Y, Kudo S ei, East JE, et al. Cost savings in colonoscopy with artificial intelligence-aided polyp diagnosis: an add-on analysis of a clinical trial (with video). *Gastrointest Endosc*. 2020;92(4):905-911.e1. doi:10.1016/j.gie.2020.03.3759
237. Topol E. Preparing the healthcare workforce to deliver the digital future The Topol Review An independent report on behalf of the Secretary of State for Health and Social Care. *Nhs*. Published online 2019.
238. DeCamp M, Lindvall C. Latent bias and the implementation of artificial intelligence in medicine. *J Am Med Informatics Assoc*. 2020;27(12):2020-2023. doi:10.1093/jamia/ocaa094
239. Rigby MJ. Ethical dimensions of using artificial intelligence in health care. *AMA J Ethics*. 2019;21(2):121-124. doi:10.1001/amajethics.2019.121
240. Eide TJ. Risk of colorectal cancer in adenoma-bearing individuals within a defined population. *Int J Cancer*. 1986;38(2):173-176. doi:10.1002/ijc.2910380205
241. Church JM. Clinical Significance of Small Colorectal Polyps. *Dis Colon Rectum*. 2004;47(4):481-485. doi:10.1007/s10350-003-0078-6
242. Heitman SJ, Ronksley PE, Hilsden RJ, Manns BJ, Rostom A, Hemmelgarn BR. Prevalence of Adenomas and Colorectal Cancer in Average Risk Individuals: A Systematic Review and Meta-analysis. *Clin Gastroenterol Hepatol*. 2009;7(12):1272-1278. doi:10.1016/j.cgh.2009.05.032
243. Winawer SJ, Zauber AG, O'Brien MJ, et al. Randomized Comparison of Surveillance Intervals

- after Colonoscopic Removal of Newly Diagnosed Adenomatous Polyps. *N Engl J Med*. 1993;328(13):901-906. doi:10.1056/nejm199304013281301
244. Martínez ME, Sampliner R, Marshall JR, Bhattacharyya AK, Reid ME, Alberts DS. Adenoma characteristics as risk factors for recurrence of advanced adenomas. *Gastroenterology*. 2001;120(5):1077-1083. doi:10.1053/gast.2001.23247
 245. Levine JS, Ahnen DJ. Adenomatous Polyps of the Colon. *N Engl J Med*. 2006;355(24):2551-2557. doi:10.1056/NEJMcp063038
 246. Kahi CJ, Myers LJ, Slaven JE, et al. Lower endoscopy reduces colorectal cancer incidence in older individuals. *Gastroenterology*. 2014;146(3):718-725.e3. doi:10.1053/j.gastro.2013.11.050
 247. Brenner H, Chang-Claude J, Jansen L, Knebel P, Stock C, Hoffmeister M. Reduced risk of colorectal cancer up to 10 years after screening, surveillance, or diagnostic colonoscopy. *Gastroenterology*. 2014;146(3):709-717. doi:10.1053/j.gastro.2013.09.001
 248. Shaikat A, Tuskey A, Rao VL, et al. Interventions to improve adenoma detection rates for colonoscopy. *Gastrointest Endosc*. 2022;96(2):171-183. doi:10.1016/J.GIE.2022.03.026
 249. Cubiella J. Closing the gap for post-colonoscopy colorectal cancer. *Lancet Gastroenterol Hepatol*. 2022;7(8):694-695. doi:10.1016/S2468-1253(22)00128-5
 250. Church JM, Muto T, Appau K. Flat lesions of the colorectal mucosa: differences in recognition between Japanese and American endoscopists. *Dis Colon Rectum*. 2004;47(9):1462-1466. doi:10.1007/S10350-004-0608-X
 251. Soetikno RM, Kaltenbach T, Rouse R V., et al. Prevalence of nonpolypoid (flat and depressed) colorectal neoplasms in asymptomatic and symptomatic adults. *JAMA - J Am Med Assoc*. 2008;299(9):1027-1035. doi:10.1001/jama.299.9.1027
 252. Voorham QJM, Carvalho B, Spiertz AJ, et al. Chromosome 5q loss in colorectal flat adenomas. *Clin Cancer Res*. 2012;18(17):4560-4569. doi:10.1158/1078-0432.CCR-11-2385
 253. Parsa N, Byrne MF. Artificial intelligence for identification and characterization of colonic polyps: <https://doi.org/101177/26317745211014698>. 2021;14. doi:10.1177/26317745211014698
 254. Hassan C, Spadaccini M, Iannone A, et al. Performance of artificial intelligence in colonoscopy for adenoma and polyp detection: a systematic review and meta-analysis. *Gastrointest Endosc*. Published online 2020. doi:10.1016/j.gie.2020.06.059
 255. Barua I, Vinsard DG, Jodal HC, et al. Artificial intelligence for polyp detection during colonoscopy: a systematic review and meta-analysis. *Endoscopy*. Published online June 2020. doi:10.1055/a-1201-7165
 256. Delavari A, Salimzadeh H, Bishehsari F, et al. Mean Polyp per Patient Is an Accurate and Readily Obtainable Surrogate for Adenoma Detection Rate: Results from an Opportunistic Screening Colonoscopy Program. *Middle East J Dig Dis*. 2015;7(4):214-219.
 257. Sivananthan A, Nazarian S, Ayaru L, et al. Does computer-aided diagnostic endoscopy improve the detection of commonly missed polyps? A meta-analysis. *Clin Endosc*. 2022;55(3):355-364. doi:10.5946/CE.2021.228
 258. Anderson R, Burr NE, Valori R. Causes of Post-Colonoscopy Colorectal Cancers Based on World Endoscopy Organization System of Analysis. *Gastroenterology*. 2020;158(5):1287-1299.e2. doi:10.1053/j.gastro.2019.12.031
 259. Zachariah R, Samarasena J, Luba D, et al. Prediction of Polyp Pathology Using Convolutional Neural Networks Achieves "Resect and Discard" Thresholds. *Am J Gastroenterol*. 2020;115(1):138-144. doi:https://dx.doi.org/10.14309/ajg.0000000000000429
 260. E. R-D, G. B. Development of computer-assisted real time histology of colorectal polyps based on the nice criteria using nearfocus narrow-band imaging. *Gastroenterology*. 2017;152(5 Supplement 1):S81. <http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emed18&NEWS=N&AN=618670310>

261. Vleugels JLA, Hassan C, Senore C, et al. Diminutive Polyps With Advanced Histologic Features Do Not Increase Risk for Metachronous Advanced Colon Neoplasia. *Gastroenterology*. 2019;156(3):623-634.e3. doi:10.1053/j.gastro.2018.10.050
262. Kim JY, Kim TJ, Baek SY, et al. Risk of Metachronous Advanced Neoplasia in Patients With Multiple Diminutive Adenomas. *Am J Gastroenterol*. 2018;(June):1855-1861. doi:10.1038/s41395-018-0210-9
263. Lui TKL, Leung WK. Is artificial intelligence the final answer to missed polyps in colonoscopy?. *World J Gastroenterol*. 2020;26(35):5248-5255. doi:https://dx.doi.org/10.3748/wjg.v26.i35.5248
264. He YS, Su JR, Li Z, Zuo XL, Li YQ. Application of artificial intelligence in gastrointestinal endoscopy. *J Dig Dis*. 2019;20(12):623-630. doi:10.1111/1751-2980.12827
265. Antonelli G, Gkolfakis P, Tziatzios G, Papanikolaou IS, Triantafyllou K, Hassan C. Artificial intelligence-aided colonoscopy: Recent developments and future perspectives. *World J Gastroenterol*. 2020;26(47):7436-7443. doi:10.3748/wjg.v26.i47.7436
266. Ammenwerth E. Technology Acceptance Models in ealth nformatics: TAM and UTAUT. *Stud Health Technol Inform*. 2019;263:64-71. doi:10.3233/SHTI190111
267. Togashi K. Applications of artificial intelligence to endoscopy practice: The view from Japan Digestive Disease Week 2018. *Dig Endosc*. 2019;31(3):270-272. doi:10.1111/DEN.13354
268. Cohen J, Bosworth BP, Chak A, et al. Preservation and Incorporation of Valuable Endoscopic Innovations (PIVI) on the use of endoscopy simulators for training and assessing skill. *Gastrointest Endosc*. 2012;76(3):471-475. doi:10.1016/J.GIE.2012.03.248
269. Areia M, Mori Y, Correale L, et al. Cost-effectiveness of artificial intelligence for screening colonoscopy: a modelling study. *Lancet Digit Heal*. 2022;4(6):e436-e444. doi:10.1016/S2589-7500(22)00042-5
270. Hassan C, Antonelli G, Repici A. Artificial intelligence for polyp characterization: Don't save on your competence! *Gastrointest Endosc*. 2020;92(4):912-913. doi:10.1016/J.GIE.2020.04.060
271. V. W, M. A, A. G, et al. Physician sentiment toward artificial intelligence (AI) in colonoscopic practice: A survey of US gastroenterologists. *Endosc Int Open*. 2020;8(10):E1379-E1384. doi:http://dx.doi.org/10.1055/a-1223-1926
272. Moran B, Cunningham C, Singh T, et al. Association of Coloproctology of Great Britain & Ireland (ACPGBI): Guidelines for the Management of Cancer of the Colon, Rectum and Anus (2017) – Surgical Management. *Color Dis*. 2017;19:18-36. doi:10.1111/CODI.13704
273. Park JS, Huh JW, Park YA, et al. A circumferential resection margin of 1 mm is a negative prognostic factor in rectal cancer patients with and without neoadjuvant chemoradiotherapy. *Dis Colon Rectum*. 2014;57(8):933-940. doi:10.1097/DCR.0000000000000171
274. Rocha R, Marinho R, Aparício D, et al. Impact of bowel resection margins in node negative colon cancer. *Springerplus*. 2016;5(1):1959. doi:10.1186/S40064-016-3650-Y
275. Vogel JD, Eskicioglu C, Weiser MR, Feingold DL, Steele SR. The American Society of Colon and Rectal Surgeons Clinical Practice Guidelines for the Treatment of Colon Cancer. *Dis Colon Rectum*. 2017;60(10):999-1017. doi:10.1097/DCR.0000000000000926
276. Stelzner S, Hohenberger W, Weber K, West NP, Witzigmann H, Wedel T. Anatomy of the transverse colon revisited with respect to complete mesocolic excision and possible pathways of aberrant lymphatic tumor spread. *Int J Colorectal Dis*. 2016;31(2):377-384. doi:10.1007/S00384-015-2434-0
277. Dieguez A. Rectal cancer staging: Focus on the prognostic significance of the findings described by high-resolution magnetic resonance imaging. *Cancer Imaging*. 2013;13(2):277–97. doi:10.1102/1470-7330.2013.0028
278. Evans J, Patel U, Brown G. Rectal Cancer: Primary Staging and Assessment After Chemoradiotherapy. *Semin Radiat Oncol*. 2011;21(3):169-177. doi:10.1016/j.semradonc.2011.02.002
279. Spicer J, Benay C, Lee L, et al. Diagnostic accuracy and utility of intraoperative microscopic

- margin analysis of gastric and esophageal adenocarcinoma. *Ann Surg Oncol*. 2014;21(8):2580-2586. doi:10.1245/s10434-014-3669-7
280. Bejarano PA, Berho M. Examination of surgical specimens of the Esophagus. *Arch Pathol Lab Med*. 2015;139(11):1446-1454. doi:10.5858/arpa.2014-0506-RA
 281. Haka AS, Volynskaya Z, Gardecki JA, et al. In vivo margin assessment during partial mastectomy breast surgery using raman spectroscopy. *Cancer Res*. 2006;66(6):3317-3322. doi:10.1158/0008-5472.CAN-05-2815
 282. Yun SH, Kwok SJJ. Light in diagnosis, therapy and surgery. *Nat Biomed Eng* 2017 11. 2017;1(1):1-16. doi:10.1038/s41551-016-0008
 283. Vahrmeijer AL, Hutteman M, Van Der Vorst JR, Van De Velde CJH, Frangioni J V. Image-guided cancer surgery using near-infrared fluorescence. *Nat Rev Clin Oncol* 2013 109. 2013;10(9):507-518. doi:10.1038/nrclinonc.2013.123
 284. Nagaya T, Nakamura YA, Choyke PL, Kobayashi H. Fluorescence-guided surgery. *Front Oncol*. 2017;7(DEC):324990. doi:10.3389/FONC.2017.00314/BIBTEX
 285. Galema HA, Meijer RPJ, Lauwerends LJ, et al. Fluorescence-guided surgery in colorectal cancer; A review on clinical results and future perspectives. *Eur J Surg Oncol*. 2022;48(4):810-821. doi:10.1016/J.EJSO.2021.10.005
 286. Keulen S van, Nishio N, Fakurnejad S, et al. The Clinical Application of Fluorescence-Guided Surgery in Head and Neck Cancer. *J Nucl Med*. 2019;60(6):758. doi:10.2967/JNUMED.118.222810
 287. Balog J, Sasi-Szabó L, Kinross J, et al. Intraoperative tissue identification using rapid evaporative ionization mass spectrometry. *Sci Transl Med*. 2013;5(194). doi:10.1126/SCITRANSLMED.3005623
 288. Jung W, Boppart SA. Optical coherence tomography for rapid tissue screening and directed histological sectioning. *Anal Cell Pathol (Amst)*. 2012;35(3):129. doi:10.3233/ACP-2011-0047
 289. Schmitt JM. Optical Coherence Tomography (OCT): a review. *IEEE J Sel Top Quantum Electron*. 1999;5(4):1205-1215. doi:10.1109/2944.796348
 290. St John ER, Balog J, McKenzie JS, et al. Rapid evaporative ionisation mass spectrometry of electrosurgical vapours for the identification of breast pathology: Towards an intelligent knife for breast cancer surgery. *Breast Cancer Res*. 2017;19(1):1-14. doi:10.1186/s13058-017-0845-2
 291. Yao X, Gan Y, Chang E, Hibshoosh H, Feldman S, Hendon C. Visualization and tissue classification of human breast cancer images using ultrahigh-resolution OCT. *Lasers Surg Med*. 2017;49(3):258-269. doi:10.1002/lsm.22654
 292. Achterberg FB, Sibinga Mulder BG, Meijer RPJ, et al. Real-time surgical margin assessment using ICG-fluorescence during laparoscopic and robot-assisted resections of colorectal liver metastases. *Ann Transl Med*. 2020;8(21):1448-1448. doi:10.21037/atm-20-1999
 293. Mohs AM, Mancini MC, Singhal S, et al. Hand-held Spectroscopic Device for In Vivo and Intraoperative Tumor Detection: Contrast Enhancement, Detection Sensitivity, and Tissue Penetration. *Anal Chem*. 2010;82(21):9058-9065. doi:10.1021/ac102058k
 294. Schnabel F, Boolbol SK, Gittleman M, et al. A randomized prospective study of lumpectomy margin assessment with use of marginprobe in patients with nonpalpable breast malignancies. *Ann Surg Oncol*. 2014;21(5):1589-1595. doi:10.1245/s10434-014-3602-0
 295. Roy HK, Backman V. Spectroscopic Applications in Gastrointestinal Endoscopy. *Clin Gastroenterol Hepatol*. 2012;10(12):1335-1341. doi:10.1016/j.cgh.2012.10.002
 296. Kim JA, Wales DJ, Yang GZ. Optical spectroscopy for in vivo medical diagnosis—a review of the state of the art and future perspectives. *Prog Biomed Eng*. 2020;2(4):042001. doi:10.1088/2516-1091/ABAAA3
 297. Bergholt MS, Lin K, Wang J, et al. Simultaneous fingerprint and high-wavenumber fiber-optic Raman spectroscopy enhances real-time in vivo diagnosis of adenomatous polyps during colonoscopy. *J Biophotonics*. 2016;9(4):333-342. doi:10.1002/jbio.201400141

298. Lui H, Zhao J, McLean D, Zeng H. Real-time Raman spectroscopy for in vivo skin cancer diagnosis. *Cancer Res.* 2012;72(10):2491-2500. doi:10.1158/0008-5472.CAN-11-4061
299. Horgan CC, Bergholt MS, Thin MZ, et al. Image-guided Raman spectroscopy probe-tracking for tumor margin delineation. *J Biomed Opt.* 2021;26(03):036002. doi:10.1117/1.jbo.26.3.036002
300. Avila-Rencoret FB, Elson DS, Mylonas G. Towards a robotic-assisted cartography of the colon: A proof of concept. In: *Proceedings - IEEE International Conference on Robotics and Automation*. Vol 2015-June. Institute of Electrical and Electronics Engineers Inc.; 2015:1757-1763. doi:10.1109/ICRA.2015.7139425
301. de Boer LL, Molenkamp BG, Bydlon TM, et al. Fat/water ratios measured with diffuse reflectance spectroscopy to detect breast tumor boundaries. *Breast Cancer Res Treat.* 2015;152(3):509-518. doi:10.1007/s10549-015-3487-z
302. Spliethoff JW, Evers DJ, Klomp HM, et al. Improved identification of peripheral lung tumors by using diffuse reflectance and fluorescence spectroscopy. *Lung Cancer.* 2013;80(2):165-171. doi:10.1016/j.lungcan.2013.01.016
303. Baltussen EJM, Snaebjornsson P, de Koning SGB, et al. Diffuse reflectance spectroscopy as a tool for real-time tissue assessment during colorectal cancer surgery. *J Biomed Opt.* 2017;22(10):1. doi:10.1117/1.jbo.22.10.106014
304. Baltussen EJM, Brouwer De Koning SG, Sanders J, et al. Tissue diagnosis during colorectal cancer surgery using optical sensing: An in vivo study. *J Transl Med.* 2019;17(1):1-10. doi:10.1186/S12967-019-2083-0/TABLES/3
305. Langhout GC, Spliethoff JW, Schmitz SJ, et al. Differentiation of healthy and malignant tissue in colon cancer patients using optical spectroscopy: A tool for image-guided surgery. *Lasers Surg Med.* 2015;47(7):559-565. doi:10.1002/LSM.22388
306. Schols RM, Dunias P, Wieringa FP, Stassen LPS. Multispectral characterization of tissues encountered during laparoscopic colorectal surgery. *Med Eng Phys.* 2013;35(7):1044-1050. doi:10.1016/J.MEDENGP.2013.01.004
307. Zonios G, Perelman LT, Backman V, et al. Diffuse reflectance spectroscopy of human adenomatous colon polyps in vivo. *Appl Opt.* 1999;38(31):6628. doi:10.1364/ao.38.006628
308. Chen H, Lin Z, Wu H, Wang L, Wu T, Tan C. Diagnosis of colorectal cancer by near-infrared optical fiber spectroscopy and random forest. *Spectrochim Acta Part A Mol Biomol Spectrosc.* 2015;135:185-191. doi:10.1016/J.SAA.2014.07.005
309. Hidović-Rowe D, Claridge E. Modelling and validation of spectral reflectance for the colon. *Phys Med Biol.* 2005;50(6):1071. doi:10.1088/0031-9155/50/6/003
310. Speidel S, Delles M, Gutt C, Dillmann R. Tracking of Instruments in Minimally Invasive Surgery for Surgical Skill Analysis. *Lect Notes Comput Sci (including Subser Lect Notes Artif Intell Lect Notes Bioinformatics)*. 2006;4091 LNCS:148-155. doi:10.1007/11812715_19
311. Elfring R, De La Fuente M, Radermacher K. Assessment of optical localizer accuracy for computer aided surgery systems. *Comput Aided Surg.* 2010;15(1-3):1-12. doi:10.3109/10929081003647239
312. Bouget D, Allan M, Stoyanov D, Jannin P. Vision-based and marker-less surgical tool detection and tracking: a review of the literature. *Med Image Anal.* 2017;35:633-654. doi:10.1016/j.media.2016.09.003
313. Krupa A, Gangloff J, Doignon C, et al. Autonomous 3-d positioning of surgical instruments in robotized laparoscopic surgery using visual servoing. *IEEE Trans Robot Autom.* 2003;19(5):842-853. Accessed January 15, 2024. https://www.academia.edu/33847879/Autonomous_3_d_positioning_of_surgical_instruments_in_robotized_laparoscopic_surgery_using_visual_servoing
314. Huang B, Tsai YY, Cartucho J, et al. Tracking and visualization of the sensing area for a tethered laparoscopic gamma probe. *Int J Comput Assist Radiol Surg.* 2020;15(8):1389-1397. doi:10.1007/s11548-020-02205-z

315. Bouarfa L, Akman O, Schneider A, Jonker PP, Dankelman J. In-vivo real-time tracking of surgical instruments in endoscopic video. *Minim Invasive Ther Allied Technol.* 2012;21(3):129-134. doi:10.3109/13645706.2011.580764
316. Bouget D, Allan M, Stoyanov D, Jannin P. Vision-based and marker-less surgical tool detection and tracking: a review of the literature. *Med Image Anal.* 2017;35:633-654. doi:10.1016/j.media.2016.09.003
317. Gautier B, Tugal H, Tang B, Nabi G, Erden MS. Real-Time 3D Tracking of Laparoscopy Training Instruments for Assessment and Feedback. *Front Robot AI.* 2021;8:751741. doi:10.3389/FROBT.2021.751741/BIBTEX
318. I. G. The combined use of fluorescence, reflectance, and light-scattering spectroscopy for evaluating dysplasia in Barrett's esophagus. *Gastrointest Endosc Clin N Am.* 2004;14(3):519-537. doi:http://dx.doi.org/10.1016/j.giec.2004.03.008
319. Panjehpour M, Overholt BF, Schmidhammer JL, Farris C, Buckley PF, Vo-Dinh T. Spectroscopic diagnosis of esophageal cancer: New classification model, improved measurement system. *Gastrointest Endosc.* 1995;41(6):577-581. doi:10.1016/S0016-5107(95)70194-X
320. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer Statistics, 2021. *CA Cancer J Clin.* 2021;71(1):7-33. doi:10.3322/caac.21654
321. Office for National Statistics. Cancer survival in England - adults diagnosed.
322. Choi N., Gibson MK. Radiation therapy, chemoradiotherapy, neoadjuvant approaches, and postoperative adjuvant therapy for localized cancers of the esophagus. UPTODATE.
323. Cunningham D, Allum WH, Stenning SP. Perioperative Chemotherapy versus Surgery Alone for Resectable Gastroesophageal Cancer. *N Engl J Med.* 2006;335(1):11-20.
324. Ajani JA, Mansfield PF, Janjan N, et al. Multi-institutional trial of preoperative chemoradiotherapy in patients with potentially resectable gastric carcinoma. *J Clin Oncol.* 2004;22(14):2774-2780. doi:10.1200/JCO.2004.01.015
325. Dexter SPL, Sue-Ling H, McMahon MJ, Quirke P, Mapstone N, Martin IG. Circumferential resection margin involvement: An independent predictor of survival following surgery for oesophageal cancer. *Gut.* 2001;48(5):667-670. doi:10.1136/gut.48.5.667
326. Evans R, Bundred JR, Kaur P, Hodson J, Griffiths EA. Meta-analysis of the influence of a positive circumferential resection margin in oesophageal cancer. *BJS open.* 2019;3(5):595-605. doi:10.1002/bjs5.50183
327. Woo JW, Ryu KW, Park JY, et al. Prognostic impact of microscopic tumor involved resection margin in advanced gastric cancer patients after gastric resection. *World J Surg.* 2014;38(2):439-446. doi:10.1007/s00268-013-2301-5
328. Barreto JC, Posner MC. Transhiatal versus transthoracic esophagectomy for esophageal cancer. *World J Gastroenterol.* 2010;16(30):3804-3810. doi:10.3748/wjg.v16.i30.3804
329. Royal College of Pathologists. Standards and datasets for reporting cancers: Dataset for histopathological reporting of oesophageal and gastric carcinoma.
330. Gkouzionis I, Nazarian S, Kawka M, et al. Real-time tracking of a diffuse reflectance spectroscopy probe used to aid histological validation of margin assessment in upper gastrointestinal cancer resection surgery. *J Biomed Opt.* 2022;27(2). doi:10.1117/1.JBO.27.2.025001
331. Zhang Z. *Flexible Camera Calibration By Viewing a Plane From Unknown Orientations.*; 1999. Accessed May 28, 2021. <http://research.microsoft.com/~zhang>
332. Nazarian S, Gkouzionis I, Kawka M, et al. Real-time Tracking and Classification of Tumor and Nontumor Tissue in Upper Gastrointestinal Cancers Using Diffuse Reflectance Spectroscopy for Resection Margin Assessment. *JAMA Surg.* 2022;157(11):e223899-e223899. doi:10.1001/jamasurg.2022.3899
333. Wilson BC, Jermyn M, Leblond F. Challenges and opportunities in clinical translation of biomedical optical spectroscopy and imaging. *J Biomed Opt.* 2018;23(3):1. doi:10.1117/1.JBO.23.3.030901

334. Nazarian S, Gkouzionis I, Murphy J, et al. Real-time classification of tumour and non-tumour tissue in colorectal cancer using diffuse reflectance spectroscopy and neural networks to aid margin assessment. *Int J Surg*. Published online January 22, 2024. doi:10.1097/JS9.0000000000001102
335. Breiman L. Random Forests. *Mach Learn* 2001 451. 2001;45(1):5-32. doi:10.1023/A:1010933404324
336. Guyon I, Weston J, Barnhill S, Vapnik V. Gene Selection for Cancer Classification using Support Vector Machines. *Mach Learn* 2002 461. 2002;46(1):389-422. doi:10.1023/A:1012487302797
337. Kursa MB, Rudnicki WR. Feature Selection with the Boruta Package. *J Stat Softw*. 2010;36(11):1-13. doi:10.18637/JSS.V036.I11
338. Barnes RJ, Dhanoo MS, Lister SJ. Standard Normal Variate Transformation and De-Trending of Near-Infrared Diffuse Reflectance Spectra. <http://dx.doi.org/10.1366/0003702894202201>. 1989;43(5):772-777. doi:10.1366/0003702894202201
339. Kursa MB, Jankowski A, Rudnicki WR. Boruta – A System for Feature Selection. *Fundam Informaticae*. 2010;101(4):271-285. doi:10.3233/FI-2010-288
340. Prusty S, Patnaik S, Dash SK. SKCV: Stratified K-fold cross-validation on ML classifiers for predicting cervical cancer. *Front Nanotechnol*. 2022;4:56. doi:10.3389/FNANO.2022.972421/BIBTEX
341. Ho TK. Random decision forests. *Proc Int Conf Doc Anal Recognition, ICDAR*. 1995;1:278-282. doi:10.1109/ICDAR.1995.598994
342. Friedman JH. Greedy function approximation: A gradient boosting machine. <https://doi.org/10.1214/aos/1013203451>. 2001;29(5):1189-1232. doi:10.1214/AOS/1013203451
343. Hsiung P-L, D.V.M. PRN, Fujimoto JG. Effect of tissue preservation on imaging using ultrahigh resolution optical coherence tomography. <https://doi.org/10.1117/12147155>. 2005;10(6):064033. doi:10.1117/1.2147155
344. Nogueira MS, Maryam S, Amissah M, et al. Evaluation of wavelength ranges and tissue depth probed by diffuse reflectance spectroscopy for colorectal cancer detection. *Sci Rep*. 2021;11(1):798. doi:10.1038/s41598-020-79517-2
345. Baltussen EJM, Sterenborg HJCM, Ruers TJM, Dashtbozorg B. Optimizing algorithm development for tissue classification in colorectal cancer based on diffuse reflectance spectra. *Biomed Opt Express*. 2019;10(12):6096. doi:10.1364/boe.10.006096
346. Dhar A, Johnson KS, Novelli MR, et al. Elastic scattering spectroscopy for the diagnosis of colonic lesions: initial results of a novel optical biopsy technique. *Gastrointest Endosc*. 2006;63(2):257-261. doi:10.1016/J.GIE.2005.07.026
347. Parfitt JR, Driman DK. The total mesorectal excision specimen for rectal cancer: a review of its pathological assessment. *J Clin Pathol*. 2007;60(8):849. doi:10.1136/JCP.2006.043802
348. Marr R, Birbeck K, Garvican J, et al. The Modern Abdominoperineal Excision: The Next Challenge After Total Mesorectal Excision. *Ann Surg*. 2005;242(1):74. doi:10.1097/01.SLA.0000167926.60908.15
349. Xiao A, Shen B, Shi X, et al. Intraoperative Glioma Grading Using Neural Architecture Search and Multi-Modal Imaging. *IEEE Trans Med Imaging*. 2022;41(10):2570-2581. doi:10.1109/TMI.2022.3166129
350. Shen B, Zhang Z, Shi X, et al. Real-time intraoperative glioma diagnosis using fluorescence imaging and deep convolutional neural networks. *Eur J Nucl Med Mol Imaging*. 2021;48(11):3482-3492. doi:10.1007/S00259-021-05326-Y/TABLES/2
351. Fernandes L, Carvalho S, Carneiro I, et al. Diffuse reflectance and machine learning techniques to differentiate colorectal cancer ex vivo. *Chaos An Interdiscip J Nonlinear Sci*. 2021;31(5):053118. doi:10.1063/5.0052088
352. Baltussen EJM, Sterenborg HJCM, Ruers TJM, Dashtbozorg B. Optimizing algorithm development for tissue classification in colorectal cancer based on diffuse reflectance

- spectra. *Biomed Opt Express*. 2019;10(12):6096. doi:10.1364/boe.10.006096
353. Brouwer de Koning SG, Baltussen EJM, Karakullukcu MB, et al. Toward complete oral cavity cancer resection using a handheld diffuse reflectance spectroscopy probe. *J Biomed Opt*. 2018;23(12):1. doi:10.1117/1.JBO.23.12.121611
 354. Singh RP, Hom GL, Abramoff MD, Campbell JP, Chiang MF. Current challenges and barriers to real-world artificial intelligence adoption for the healthcare system, provider, and the patient. *Transl Vis Sci Technol*. 2020;9(2):1-6. doi:10.1167/tvst.9.2.45
 355. Lin C, Lin IC, Roan J. Barriers to physicians' adoption of healthcare information technology: an empirical study on multiple hospitals. *J Med Syst*. 2012;36(3):1965-1977. doi:10.1007/S10916-011-9656-7
 356. Bernhardt S, Nicolau SA, Soler L, Doignon C. The status of augmented reality in laparoscopic surgery as of 2016. *Med Image Anal*. 2017;37:66-90. doi:10.1016/J.MEDIA.2017.01.007
 357. Detering R, Rutgers MLW, Bemelman WA, Hompes R, Tanis PJ. Prognostic importance of circumferential resection margin in the era of evolving surgical and multidisciplinary treatment of rectal cancer: A systematic review and meta-analysis. *Surgery*. 2021;170(2):412-431. doi:10.1016/J.SURG.2021.02.029
 358. Liu Q, Luo D, Cai S, Li Q, Li X. Circumferential resection margin as a prognostic factor after rectal cancer surgery: A large population-based retrospective study. *Cancer Med*. 2018;7(8):3673. doi:10.1002/CAM4.1662
 359. Adam IJ, Martin IG, Finan P, et al. Role of circumferential margin involvement in the local recurrence of rectal cancer. *Lancet*. 1994;344(8924):707-711. doi:10.1016/S0140-6736(94)92206-3
 360. Waterhouse DJ, Fitzpatrick CRM, Pogue BW, O'Connor JPB, Bohndiek SE. A roadmap for the clinical implementation of optical-imaging biomarkers. *Nat Biomed Eng* 2019 35. 2019;3(5):339-353. doi:10.1038/s41551-019-0392-5
 361. van Veen RLP, Sterenborg HJCM, Marinelli AWKS, Menke-Pluymers M. Intraoperatively assessed optical properties of malignant and healthy breast tissue used to determine the optimum wavelength of contrast for optical mammography. *J Biomed Opt*. 2004;9(6):1129. doi:10.1117/1.1803547
 362. Bigio IJ, Bown SG, Briggs G, et al. Diagnosis of breast cancer using elastic-scattering spectroscopy: preliminary clinical results. Published online 2000.
 363. Kennedy S, Caldwell M, Bydlon T, et al. Correlation of breast tissue histology and optical signatures to improve margin assessment techniques. *J Biomed Opt*. 2016;21(6):066014. doi:10.1117/1.JBO.21.6.066014
 364. Baltussen EJM, Brouwer De Koning SG, Benno J, et al. Comparing in vivo and ex vivo fiberoptic diffuse reflectance spectroscopy in colorectal cancer. *Transl Biophotonics*. 2019;1(1-2):e201900008. doi:10.1002/TBIO.201900008
 365. Jocher G, Chaurasia A, Stoken A, et al. ultralytics/yolov5: v7.0 - YOLOv5 SOTA Realtime Instance Segmentation. *zndo*. doi:10.5281/ZENODO.3908559
 366. Gkouzionis I, Zhong Y, Nazarian S, et al. A YOLOv5-based network for the detection of a diffuse reflectance spectroscopy probe to aid surgical guidance in gastrointestinal cancer surgery. *Int J Comput Assist Radiol Surg*. 2023;19(1):11. doi:10.1007/S11548-023-02944-9
 367. Lund A. Measuring Usability with the USE Questionnaire. Accessed February 29, 2024. <https://www.researchgate.net/publication/230786746>
 368. Demers L, Weiss-Lambrou R, Ska B. The Quebec User Evaluation of Satisfaction with Assistive Technology (QUEST 2.0): An overview and recent progress. *Technol Disabil*. 2002;14(3):101-105. doi:10.3233/TAD-2002-14304
 369. De Boer LL, W Hendriks BH, Van Duijnhoven F, et al. Using DRS during breast conserving surgery: identifying robust optical parameters and influence of inter-patient variation. *Biomed Opt Express*, Vol 7, Issue 12, pp 5188-5200. 2016;7(12):5188-5200. doi:10.1364/BOE.7.005188

370. Hartl DM, Zago S, Leboulleux S, et al. Resection margins and prognosis in locally invasive thyroid cancer. *Head Neck*. 2014;36(7):1034-1038. doi:10.1002/HED.23406
371. Hay ID, Johnson TR, Thompson GB, Sebo TJ, Reinalda MS. Minimal extrathyroid extension in papillary thyroid carcinoma does not result in increased rates of either cause-specific mortality or postoperative tumor recurrence. *Surgery*. 2016;159(1):11-21. doi:10.1016/J.SURG.2015.05.046
372. Mehasseb MK, Latimer JA. Controversies in the Management of Endometrial Carcinoma: An Update. *Obstet Gynecol Int*. 2012;2012:1-10. doi:10.1155/2012/676032
373. Waterhouse DJ, Januszewicz W, Ali S, Fitzgerald RC, Di Pietro M, Bohndiek SE. Spectral Endoscopy Enhances Contrast for Neoplasia in Surveillance of Barrett's Esophagus. *Cancer Res*. 2021;81(12):3415. doi:10.1158/0008-5472.CAN-21-0474
374. Phelps DL, Balog J, Gildea LF, et al. The surgical intelligent knife distinguishes normal, borderline and malignant gynaecological tissues using rapid evaporative ionisation mass spectrometry (REIMS). *Br J Cancer* 2018 11810. 2018;118(10):1349-1358. doi:10.1038/s41416-018-0048-3
375. Chuchvara N, Rao B, Liu X. Manually scanned single fiber optical coherence tomography for skin cancer characterization. *Sci Reports* 2021 111. 2021;11(1):1-9. doi:10.1038/s41598-021-95118-z

Appendices

Appendix 1 Search Strategy

EMBASE & Medline

- | | |
|-----|---|
| #1 | Colonic Polyps/ |
| #2 | colorect* or colon or colonoscop* or large bowel or rectal or rectum or anal or anus or large intestine and polyp |
| #3 | 1 or 2 |
| #4 | exp algorithms/ |
| #5 | artificial intelligence or machine learning or neural network* |
| #6 | 4 or 5 |
| #7 | computer aided or computer assisted and diagnos* or detect* |
| #8 | diagnosis, computer-assisted/ or image interpretation, computer-assisted/ or neuronavigation/ |
| #9 | image processing, computer-assisted/ or data compression/ or image enhancement/ or imaging, three-dimensional/ |
| #10 | 7 or 8 or 9 |
| #11 | 3 and 6 and 10 |

The Cochrane Library

- | | |
|-----|--|
| #1 | Colonic Polyps |
| #2 | colorect* or colon or colonoscop* or large bowel and polyp |
| #3 | 1 or 2 |
| #4 | exp artificial intelligence |
| #5 | artificial intelligence or machine learning or neural network* |
| #6 | 4 or 5 |
| #7 | exp diagnosis, computer-assisted |
| #8 | computer aided or computer assisted and diagnos* or detect* |
| #9 | 7 or 8 |
| #10 | 3 or 6 or 9 |

Appendix 2 Quality Assessment

Study	RISK OF BIAS				APPLICABILITY CONCERNS		
	PATIENT SELECTION	INDEX TEST	REFERENCE STANDARD	FLOW AND TIMING	PATIENT SELECTION	INDEX TEST	REFERENCE STANDARD
Detection							
Karkanis et al. 2003							
Fu et al. 2014							
Wang et al. 2015							
Tajbakhsh et al. 2015 (a)							
Tajbakhsh et al. 2015 (b)							
Fernández-Esparrach et al. 2016							
Park & Sargent 2016							
Urban et al. 2018							
Wang et al. 2018							
Misawa et al. 2018							
Figueiredo et al. 2019							
Yamada et al. 2019							
Becq et al. 2020							
Gao et al. 2020							
Guo et al. 2020							
Lee et al. 2020							
Ozawa et al. 2020							
Misawa et al. 2020							
Poon et al. 2020							
Characterisation							
Tischendorf et al. 2010							
Gross et al. 2011							
Ganz et al. 2012							
Takemura et al. 2012							
Mori et al. 2015							
Kominami et al. 2016							
Misawa et al. 2016							
Mesejo et al. 2016							
Mori et al. 2016							
Takeda et al. 2017							
Byrne et al. 2017							
Komeda et al. 2017							
Misawa et al. 2017							

Mori et al. 2018							
Chen et al. 2018							
Renner et al. 2018							
Mori et al. 2018 (a)							
Mori et al. 2018 (b)							
Kudo et al. 2019 (a)							
Kudo et al. 2019 (b)							
Figueiredo et al. 2019							
Rodriguez-Diaz et al. 2020							
Yang et al. 2020							
Zachariah et al. 2020							

Table 25 Assessment of bias for each included study in seven domains according to the QUADAS-2 tool.  = low risk of bias;  = high risk of bias;  = unclear risk of bias.

Study	Randomisation (2)	Blinding (2)	An account of all patients (1)	Total / 5
Wang et al. 2019	2	1	1	4
Wang et al. 2020	2	2	1	5
Su et al. 2020	2	2	1	5
Gong et al. 2020	2	2	1	5
Liu et al. 2020	1	0	1	2
Luo et al. 2020	2	0	1	3
Repici et al. 2020	2	1	1	4
Wang et al. 2020	2	2	1	5

Table 26 Assessment of bias for RCTs using the Jadad scale; score ≥ 3 = good quality.

Appendix 3 Limitations of Study

Factors	Variables
Study design	Retrospective or prospective data collection
Datasets	Image, polyp or patient
Size of datasets	Wide variation in size of datasets for testing and validating
Imaging modality	Different imaging modalities used e.g. white light, NBI, EC
Machine learning approach	Variety in type of algorithm used e.g. CNN, SVM
Population	Eastern vs. Western population
Reporting of outcome	Different diagnostic accuracy metric reported - sensitivity, specificity, accuracy, ADR, PDR
Applicability	Pre-clinical studies vs. clinical trials

Table 27 Limitations of current studies assessing polyp detection or characterisation.

Appendix 4 Database Searches

EMBASE & Medline

#1 Colonic Polyps/

#2 colorect* or colon or colonoscop* or large bowel or rectal or rectum or anal or anus or large intestine and polyp

#3 1or2

#4 exp algorithms/

#5 artificial intelligence or machine learning or neural network*

#6 4or5

#7 computer aided or computer assisted and diagnos* or detect*

#8 diagnosis, computer-assisted/ or image interpretation, computer-assisted/ or neuronavigation/

#9 image processing, computer-assisted/ or data compression/ or image enhancement/ or imaging, three-dimensional/

#10 7 or 8 or 9

#11 3 and 6 and 10

The Cochrane Library

#1 Colonic Polyps

#2 colorect* or colon or colonoscop* or large bowel and polyp

#3 1 or 2

#4 exp artificial intelligence

#5 artificial intelligence or machine learning or neural network*

#6 4 or 5

#7 exp diagnosis, computer-assisted

#8 computer aided or computer assisted and diagnos* or detect*

#9 7 or 8

#10 3 or 6 or 9

Appendix 5 Endoscopy Questionnaire

Views on AI in Endoscopy Questionnaire

1. Which region do you work in?
 - a. North East England
 - b. North West England
 - c. Yorkshire and the Humber
 - d. West Midlands
 - e. East Midlands
 - f. East of England
 - g. London
 - h. South East England
 - i. South West England
2. Which best describes your role?
 - a. Consultant gastroenterologist
 - b. Consultant surgeon
 - c. Gastroenterology registrar
 - d. Surgical registrar
 - e. Clinical fellow in gastroenterology
 - f. Clinical fellow in surgery
 - g. Other – please state
3. What is your level of colonoscopy experience?
 - a. 0-200 procedures
 - b. 200-500 procedures
 - c. 500-1000 procedures
 - d. More than 1000 procedures
4. What is your adenoma detection rate (ADR)?
 - a. 0-15%
 - b. 15-25%
 - c. 25-40%
 - d. 40-60%
 - e. >60%
 - f. Unsure
5. How familiar are you with artificial intelligence (AI) in endoscopy?
 - a. Not at all
 - b. Slightly familiar
 - c. Moderately familiar
 - d. Very familiar
6. In your opinion, will AI improve the quality of endoscopy?
 - a. Yes
 - b. No
 - c. Unsure
7. In your opinion, will AI increase the ADR?
 - a. Yes

- b. No
 - c. Unsure
- 8. In your opinion, will AI increase the PDR?
 - a. Yes
 - b. No
 - c. Unsure
- 9. In your opinion, will AI increase the efficiency of endoscopy?
 - a. Yes
 - b. No
 - c. Unsure
- 10. In your opinion, will AI increase the cost of endoscopy?
 - a. Yes
 - b. No
 - c. Unsure
- 11. Do you think AI will be part of routine practice in endoscopy in the next 5 years?
 - a. Yes
 - b. No
 - c. Unsure
- 12. Do you think AI is necessary to endoscopy?
 - a. Yes
 - b. No
 - c. Unsure
- 13. Do you think AI in endoscopy will aid lesion characterisation?
 - a. Yes
 - b. No
 - c. Unsure
- 14. Do you perceive AI as a threat or a tool to endoscopy practice?
 - a. Threat
 - b. Tool
 - c. Neither
 - d. Both
- 15. In your opinion, what are the main barriers to use of AI in endoscopy? (Select all that apply)
 - a. Lack of research
 - b. Cost
 - c. Accessibility
 - d. Accuracy
 - e. Ethics
 - f. Learning curve
 - g. Medico-legal issues
 - h. Data ownership
 - i. Lack of guidelines
 - j. None of the above

16. In your opinion, what are the main benefits of AI in endoscopy? (Select all that apply)
- a. Increase PDR
 - b. Increase ADR
 - c. Increase confidence
 - d. Cost
 - e. Productivity
 - f. Workflow
 - g. Training tool
 - h. None of the above
17. What area within endoscopy would you like AI to be used?
- a. In-vivo lesion detection/ characterisation
 - b. Quality of endoscopy
 - c. Reporting system
 - d. Prioritisation of procedures
 - e. None of the above
18. How would you rank the priority area for AI use in endoscopy? (Rank most important as 1)
- 1. Upper GI
 - 2. Video capsule endoscopy
 - 3. Lower GI
 - 4. HPB e.g. ERCP
 - 5. Endoscopic ultrasound
 - 6. Other – specify

Appendix 6 Usability Questionnaire

DRS Device Usability Questionnaire

Background

1. I am a...
 - a) Consultant Surgeon
 - b) Surgical registrar/ trainee
 - c) Scrub nurse
 - d) Theatre staff
 - e) Non-surgical doctor
 - f) Other - please state below

2. I have used the DRS system...
 - a) Once
 - b) More than once - state below _____
 - c) Other _____

Usefulness

3. It does not affect the surgical workflow:
 - a) Strongly agree
 - b) Somewhat agree
 - c) Neither agree nor disagree
 - d) Somewhat disagree
 - e) Strongly disagree
 - f) Not applicable

4. It does everything I would expect it to do:
 - a) Strongly agree
 - b) Somewhat agree
 - c) Neither agree nor disagree
 - d) Somewhat disagree
 - e) Strongly disagree
 - f) Not applicable

Ease of Use

5. It is easy to use
 - a) Strongly agree
 - b) Somewhat agree
 - c) Neither agree nor disagree
 - d) Somewhat disagree
 - e) Strongly disagree
 - f) Not applicable

6. It requires the fewest steps possible to accomplish what I want to do with it
 - a) Strongly agree
 - b) Somewhat agree
 - c) Neither agree nor disagree
 - d) Somewhat disagree
 - e) Strongly disagree
 - f) Not applicable
7. It is flexible and can be manoeuvred easily
 - a) Strongly agree
 - b) Somewhat agree
 - c) Neither agree nor disagree
 - d) Somewhat disagree
 - e) Strongly disagree
 - f) Not applicable
8. I can use it without written instruction
 - a) Strongly agree
 - b) Somewhat agree
 - c) Neither agree nor disagree
 - d) Somewhat disagree
 - e) Strongly disagree
 - f) Not applicable
9. I don't notice any inconsistencies when I use it
 - a) Strongly agree
 - b) Somewhat agree
 - c) Neither agree nor disagree
 - d) Somewhat disagree
 - e) Strongly disagree
 - f) Not applicable
10. Both occasional and regular users would like it
 - a) Strongly agree
 - b) Somewhat agree
 - c) Neither agree nor disagree
 - d) Somewhat disagree
 - e) Strongly disagree
 - f) Not applicable
11. I can recover from mistakes quickly and easily
 - a) Strongly agree
 - b) Somewhat agree
 - c) Neither agree nor disagree
 - d) Somewhat disagree
 - e) Strongly disagree
 - f) Not applicable

12. I can use it successfully every time
- a) Strongly agree
 - b) Somewhat agree
 - c) Neither agree nor disagree
 - d) Somewhat disagree
 - e) Strongly disagree
 - f) Not applicable

Ease of Learning

13. I learned to use it quickly
- a) Strongly agree
 - b) Somewhat agree
 - c) Neither agree nor disagree
 - d) Somewhat disagree
 - e) Strongly disagree
 - f) Not applicable
14. I easily remember how to use it
- a) Strongly agree
 - b) Somewhat agree
 - c) Neither agree nor disagree
 - d) Somewhat disagree
 - e) Strongly disagree
 - f) Not applicable
15. I quickly became skillful with it
- a) Strongly agree
 - b) Somewhat agree
 - c) Neither agree nor disagree
 - d) Somewhat disagree
 - e) Strongly disagree
 - f) Not applicable

Satisfaction

16. I am satisfied with it
- a) Strongly agree
 - b) Somewhat agree
 - c) Neither agree nor disagree
 - d) Somewhat disagree
 - e) Strongly disagree
 - f) Not applicable

17. I would recommend it to a colleague

- a) Strongly agree
- b) Somewhat agree
- c) Neither agree nor disagree
- d) Somewhat disagree
- e) Strongly disagree
- f) Not applicable

18. It works the way I want it to work

- a) Strongly agree
- b) Somewhat agree
- c) Neither agree nor disagree
- d) Somewhat disagree
- e) Strongly disagree
- f) Not applicable

19. It has potential to change clinical practice within surgery (surgeon)

- a) Strongly agree
- b) Somewhat agree
- c) Neither agree nor disagree
- d) Somewhat disagree
- e) Strongly disagree
- f) Not applicable

20. List the most positive aspects

21. List the most negative aspects

22. Please feel free to share any other comments
