Using health data to predict individuals at risk of gastrointestinal cancers and to refine delivery of care in light of the COVID-19 pandemic

Kai Man Alexander Ho

Thesis submitted for the degree of

Doctor of Philosophy (PhD)

2025

Division of Surgery and Interventional Science

University College London

Declaration

I, Kai Man Alexander Ho confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Statement about Funders

I would like to thank University College London Hospitals Charity and also Wellcome/EPSRC Centre for Interventional and Surgical Sciences (WEISS) for providing funding through block grants to my primary supervisor, and in particular for the WEISS Continuity Award.

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the thesis, presentations and publications arising from the research conducted.

Abstract

Gastrointestinal tract cancers are some of the highest incidence cancers globally. However, certain subtypes such oesophageal cancer have poor prognosis, as patients often present in advanced stages of disease, limiting curative options. Earlier identification of potentially high-risk individuals would lead to improved outcomes. This has become especially important in the aftermath of the COVID-19 pandemic, which caused widespread disruption of healthcare systems and led to disruption of patient referral pathways.

In this thesis, I firstly examined risk factors which could lead to the development of gastrointestinal cancers. I also performed a systematic review and meta-analysis, focussing the role of gastro-oesophageal reflux in the development of oesophageal cancer.

Next, I used routine health data to assess the impact of the pandemic gastroenterology services both within a single hospital trust and nationally. I collated data on the actual number of endoscopic procedures performed during the first wave of the pandemic in England, demonstrating a precipitous drop before a prolonged recovery. Furthermore, I calculated a projected backlog of endoscopic procedures and use modelling techniques to estimate when the backlog may be cleared under different recovery scenarios.

I then sought to develop solutions to overcome the backlog. Using machine learning, I created a model which predicts the presence of oesophageal and gastric cancer, based on patient questionnaire data. I added epigenetic data derived from paired saliva samples to assess for improvements in model performance. Moreover, using primary care data I evaluated the use of faecal immunochemical testing (FIT) in triaging suspected colorectal cancer referrals, including the role for duplicate measurements.

Finally, I also examined the wider effects of the pandemic, including psychosocial effects within a religious worship context.

In conclusion, this thesis analyses the impact of the COVID-19 pandemic on gastroenterology services, with a focus on endoscopy, and offers solutions to overcome its aftereffects.

Impact Statement

My PhD coincided with the early phase of the COVID-19 pandemic, where I adapted to a rapidly changing research environment. Research questions surrounding COVID-19 and its wider effects suddenly became the priority with early dissemination of results required due to the public health emergency, most exemplified by the use of pre-print servers.[1] Likewise, my thesis title has been adapted as a result of the prioritisation of COVID-19 related research, but I believe this has further increased the impact of my work.

Inside Academia

I have published research arising from my thesis in peer-reviewed journals. The greatest impact was my work in predicting endoscopic backlogs in England as a result of the COVID-19 pandemic. This was published in *The Lancet Gastroenterology & Hepatology*, a leading subject journal and has a current impact factor of 35.7.[2] As of October 2024, the article has had 55 citations. I have also had 3 further first authored/co-first authored journal articles published in peer-reviewed journals related to the thesis. In addition, I have had involvement with 5 further peer-reviewed journal articles as first author or co-author.

I have also been active in presenting my research findings at national and international conferences. In particular, I presented my work on developing machine learning models to triage 2-week wait upper gastrointestinal referrals at the British Society of Gastroenterology Annual Meeting in 2022, as an oral presentation at the oesophageal parallel session. I was also awarded the “Best Oesophageal Oral Presentation” prize at the conference. In addition, I have also published an online calculator which could be openly used by anyone for triaging 2-week wait upper gastrointestinal referrals, giving immediate feedback to the user regarding the likelihood of upper gastrointestinal cancer in percentage terms.

Outside Academia

My work on both predicting endoscopic backlogs was picked up by the UCL media office and ran as an exclusive story in the *i* newspaper, which has an average circulation of 126,000 copies per day. The article was also featured on the *i* newspaper website. In addition, several organisations such as DATA-CAN, part of Health Data Research UK, UCL and Bowel Cancer UK also issued press releases highlighting the size of the backlog and the potential impact on cancer diagnostic pathways. Such publicity could have raised the profile of my findings and led to changes in health policy and plans for recovery of diagnostic services.

Furthermore, I was also involved with public engagement with my work on looking at face mask acceptability in religious worship, which was done as part of the CONFESS (COvid aNd FacE maSkS) study. Participants in the CONFESS study were invited to a webinar where they heard me present some of the preliminary results. In addition, the event also saw the launch of a website where participants had access to study data and were invited to discover insights for themselves. The research team received positive verbal feedback from webinar participants. Similar events in the future could lead to improved satisfaction and engagement with research studies, especially if serial follow up by participants may be required.

Supervisors & Thesis Committee Members

Primary Supervisor	Professor Laurence Lovat Professor of Gastroenterology and Biophotonics Division of Surgery and Interventional Science, UCL Honorary Consultant Gastroenterologist University College London Hospitals NHS Foundation Trust
Secondary Supervisor	Professor Amitava Banerjee Professor of Clinical Data Science Institute of Health Informatics, UCL Honorary Consultant Cardiologist University College London Hospitals NHS Foundation Trust & Barts Health NHS Trust
Thesis Committee Member	Dr Alvina Lai Associate Professor of Health Data Analytics Institute of Health Informatics, UCL

Research Paper Declaration Form #1

UCL Research Paper Declaration Form
referencing the doctoral candidate's own published work(s)

Please use this form to declare if parts of your thesis are already available in another format, e.g. if data, text, or figures:

- *have been uploaded to a preprint server*
- *are in submission to a peer-reviewed publication*
- *have been published in a peer-reviewed publication, e.g. journal, textbook.*

This form should be completed as many times as necessary. For instance, if you have seven thesis chapters, two of which containing material that has already been published, you would complete this form twice.

1. For a research manuscript that has already been published (if not yet published, please skip to section 2)

a) **What is the title of the manuscript?**

Spatiotemporal droplet dispersion measurements demonstrate face masks reduce risks from singing

b) **Please include a link to or doi for the work**

https://www.nature.com/articles/s41598-021-03519-x

c) **Where was the work published?**

Scientific Reports

d) **Who published the work?** (e.g. OUP)

Nature Publishing Group

e) **When was the work published?**

2021

f) **List the manuscript's authors in the order they appear on the publication**

Kai Man Alexander Ho, Hywel Davies, Ruth Epstein, Paul Bassett, Áine Hogan, Yusuf Kabir, John Rubin, Gee Yen Shin, Jonathan P Reid, Ryo Torii, Manish K Tiwari, Ramanarayanan Balachandran, Laurence B Lovat

g) **Was the work peer reviewed?**

Yes

h) **Have you retained the copyright?**

Yes

i) **Was an earlier form of the manuscript uploaded to a preprint server?** (e.g. medRxiv). If 'Yes', please give a link or doi)

Yes
https://www.medrxiv.org/content/10.1101/2021.07.09.21260247v1

If ‘No’, please seek permission from the relevant publisher and check the box next to the below statement:

☐

*I acknowledge permission of the publisher named under **1d** to include in this thesis portions of the publication named as included in **1c**.*

2. For a research manuscript prepared for publication but that has not yet been published (if already published, please skip to section 3)

a) **What is the current title of the manuscript?**

Click or tap here to enter text.

b) **Has the manuscript been uploaded to a preprint server?** (e.g. medRxiv; if ‘Yes’, please give a link or doi)

Click or tap here to enter text.

c) **Where is the work intended to be published?** (e.g. journal names)

Click or tap here to enter text.

d) **List the manuscript’s authors in the intended authorship order**

Click or tap here to enter text.

e) **Stage of publication** (e.g. in submission)

Click or tap here to enter text.

3. For multi-authored work, please give a statement of contribution covering all authors (if single-author, please skip to section 4)

KMAH: data analysis, data interpretation, manuscript preparation, manuscript revision.
HD: study design, data collection, data analysis, data interpretation, manuscript preparation, manuscript revision.
RE: study conceptualisation, study design, data analysis, data interpretation, manuscript revision.
PB: data analysis, data interpretation.
AH, YK: participant recruitment.
JR., GYS., JPR., RT: data interpretation, manuscript revision.
MKT: study conceptualisation, data interpretation, manuscript revision, supervisory oversight.

RB: study conceptualisation, study design, data interpretation, manuscript revision, supervisory oversight.
LBL: study conceptualisation, study design, data interpretation, manuscript revision, supervisory oversight.

4. **In which chapter(s) of your thesis can this material be found?**

Chapter 8: Societal Impacts of COVID-19 Pandemic

5. **e-Signatures confirming that the information above is accurate** (this form should be co-signed by the supervisor/ senior author unless this is not appropriate, e.g. if the paper was a single-author work)

Candidate

KMA Ho

Date:

27/03/2024

Supervisor/ Senior Author (where appropriate)

LB Lovat

Date

27/03/2024

Research Paper Declaration Form #2

UCL Research Paper Declaration Form
referencing the doctoral candidate's own published work(s)

Please use this form to declare if parts of your thesis are already available in another format, e.g. if data, text, or figures:

- *have been uploaded to a preprint server*
- *are in submission to a peer-reviewed publication*
- *have been published in a peer-reviewed publication, e.g. journal, textbook.*

This form should be completed as many times as necessary. For instance, if you have seven thesis chapters, two of which containing material that has already been published, you would complete this form twice.

1. For a research manuscript that has already been published (if not yet published, please skip to section 2)

a) **What is the title of the manuscript?**

Predicting endoscopic activity recovery in England after COVID-19: a national analysis

b) **Please include a link to or doi for the work**

https://www.sciencedirect.com/science/article/pii/S2468125321000583

c) **Where was the work published?**

The Lancet Gastroenterology & Hepatology

d) **Who published the work?** (e.g. OUP)

Elsevier

e) **When was the work published?**

2021

f) **List the manuscript's authors in the order they appear on the publication**

Kai Man Alexander Ho, Amitava Banerjee, Mark Lawler, Matthew D Rutter, Laurence B Lovat

g) **Was the work peer reviewed?**

Yes

h) **Have you retained the copyright?**

Yes

i) **Was an earlier form of the manuscript uploaded to a preprint server?** (e.g. medRxiv). If 'Yes', please give a link or doi)

If 'No', please seek permission from the relevant publisher and check the box next to the below statement:

☒

*I acknowledge permission of the publisher named under **1d** to include in this thesis portions of the publication named as included in **1c**.*

2. For a research manuscript prepared for publication but that has not yet been published (if already published, please skip to section 3)

a) **What is the current title of the manuscript?**

Click or tap here to enter text.

b) **Has the manuscript been uploaded to a preprint server?** (e.g. medRxiv; if 'Yes', please give a link or doi)

Click or tap here to enter text.

c) **Where is the work intended to be published?** (e.g. journal names)

Click or tap here to enter text.

d) **List the manuscript's authors in the intended authorship order**

Click or tap here to enter text.

e) **Stage of publication** (e.g. in submission)

Click or tap here to enter text.

3. For multi-authored work, please give a statement of contribution covering all authors (if single-author, please skip to section 4)

KMAH: study conception, data analysis, and manuscript preparation.
MDR: collection of data, manuscript preparation, and supervisory oversight
AB, ML, and LBL: study conception, data analysis, manuscript preparation, and supervisory oversight.

4. In which chapter(s) of your thesis can this material be found?

Chapter 5: Effect of COVID-19 on Gastrointestinal Services

5. e-Signatures confirming that the information above is accurate (this form should be co-signed by the supervisor/ senior author unless this is not appropriate, e.g. if the paper was a single-author work)

Candidate

KMA Ho

Date:

27/03/2024

Supervisor/ Senior Author (where appropriate)

LB Lovat

Date

27/03/2024

Research Paper Declaration Form #3

UCL Research Paper Declaration Form
referencing the doctoral candidate's own published work(s)

Please use this form to declare if parts of your thesis are already available in another format, e.g. if data, text, or figures:

- *have been uploaded to a preprint server*
- *are in submission to a peer-reviewed publication*
- *have been published in a peer-reviewed publication, e.g. journal, textbook.*

This form should be completed as many times as necessary. For instance, if you have seven thesis chapters, two of which containing material that has already been published, you would complete this form twice.

1. For a research manuscript that has already been published (if not yet published, please skip to section 2)

a) **What is the title of the manuscript?**

Face mask acceptability for communal religious worship during the COVID-19 pandemic in the United Kingdom: results from the CONFESS study

b) **Please include a link to or doi for the work**

https://link.springer.com/article/10.1007/s10943-022-01641-2

c) **Where was the work published?**

Journal of Religion and Health

d) **Who published the work?** (e.g. OUP)

Springer Link

e) **When was the work published?**

2023

f) **List the manuscript's authors in the order they appear on the publication**

Kai Man Alexander Ho, Rebecca F Baggaley, Timothy C Stone, Áine Hogan, Yusuf Kabir Christopher Johnson, Robert Merrifield, Laurence B Lovat

g) **Was the work peer reviewed?**

Yes

h) **Have you retained the copyright?**

Yes

i) **Was an earlier form of the manuscript uploaded to a preprint server?** (e.g. medRxiv). If 'Yes', please give a link or doi)

If ‘No’, please seek permission from the relevant publisher and check the box next to the below statement:

☒

*I acknowledge permission of the publisher named under **1d** to include in this thesis portions of the publication named as included in **1c**.*

2. For a research manuscript prepared for publication but that has not yet been published (if already published, please skip to section 3)

f) **What is the current title of the manuscript?**

Click or tap here to enter text.

g) **Has the manuscript been uploaded to a preprint server?** (e.g. medRxiv; if ‘Yes’, please give a link or doi)

Click or tap here to enter text.

h) **Where is the work intended to be published?** (e.g. journal names)

Click or tap here to enter text.

i) **List the manuscript’s authors in the intended authorship order**

Click or tap here to enter text.

j) **Stage of publication** (e.g. in submission)

Click or tap here to enter text.

3. For multi-authored work, please give a statement of contribution covering all authors (if single-author, please skip to section 4)

KMAH: data curation, data analysis, manuscript preparation—original draft, review & editing.
RFB: data curation, data analysis, manuscript preparation—original draft, review & editing.
TCS: data analysis.
AH, YK: participant recruitment.
CJ, RM: data analysis, manuscript preparation—review & editing.
LBL: conceptualisation, participant recruitment, data analysis, manuscript preparation-review & editing, supervision.

4. **In which chapter(s) of your thesis can this material be found?**

Chapter 8: Societal Impacts of COVID-19 Pandemic

5. **e-Signatures confirming that the information above is accurate** (this form should be co-signed by the supervisor/ senior author unless this is not appropriate, e.g. if the paper was a single-author work)

Candidate

KMA Ho

Date:

27/03/2024

Supervisor/ Senior Author (where appropriate)

LB Lovat

Date

27/03/2024

Research Paper Declaration Form #4

UCL Research Paper Declaration Form
referencing the doctoral candidate's own published work(s)

Please use this form to declare if parts of your thesis are already available in another format, e.g. if data, text, or figures:

- *have been uploaded to a preprint server*
- *are in submission to a peer-reviewed publication*
- *have been published in a peer-reviewed publication, e.g. journal, textbook.*

This form should be completed as many times as necessary. For instance, if you have seven thesis chapters, two of which containing material that has already been published, you would complete this form twice.

1. For a research manuscript that has already been published (if not yet published, please skip to section 2)

a) **What is the title of the manuscript?**

Development and validation of a multivariable risk factor questionnaire to detect oesophageal cancer in 2-week wait patients

b) **Please include a link to or doi for the work**

https://www.sciencedirect.com/science/article/pii/S2210740123000128

c) **Where was the work published?**

Clinics and Research in Hepatology and Gastroenterology

d) **Who published the work?** (e.g. OUP)

Elsevier

e) **When was the work published?**

2023

f) **List the manuscript's authors in the order they appear on the publication**

Kai Man Alexander Ho, Avi Rosenfeld, Áine Hogan, Hazel McBain, Margaret Duku, Paul BD Wolfson, Ashley Wilson, Sharon MY Cheung, Laura Hennelly, Lester Macabodbod, David G Graham, Vinay Sehgal, Amitava Banerjee, Laurence B Lovat; SPIT Study Group Collaborators

g) **Was the work peer reviewed?**

Yes

h) **Have you retained the copyright?**

Yes

i) **Was an earlier form of the manuscript uploaded to a preprint server?** (e.g. medRxiv). If 'Yes', please give a link or doi)

If ‘No’, please seek permission from the relevant publisher and check the box next to the below statement:

☒

*I acknowledge permission of the publisher named under **1d** to include in this thesis portions of the publication named as included in **1c**.*

2. For a research manuscript prepared for publication but that has not yet been published (if already published, please skip to section 3)

a) **What is the current title of the manuscript?**

Click or tap here to enter text.

b) **Has the manuscript been uploaded to a preprint server?** (e.g. medRxiv; if ‘Yes’, please give a link or doi)

Click or tap here to enter text.

c) **Where is the work intended to be published?** (e.g. journal names)

Click or tap here to enter text.

d) **List the manuscript’s authors in the intended authorship order**

Click or tap here to enter text.

e) **Stage of publication** (e.g. in submission)

Click or tap here to enter text.

3. For multi-authored work, please give a statement of contribution covering all authors (if single-author, please skip to section 4)

KMAH: conceptualisation, data curation, formal analysis, investigation, methodology, software, validation, visualisation, writing -original draft, writing - review & editing
AR: conceptulisation, formal analysis, methology, supervision, writing - review & editing
AH, HM, MD, PW, AW, SC, LH, LM, DGG, VS: investigation, project administration,
SPIT Study Group Collaborators: investigation
AB: conceptualisation, formal analysis, methodology, supervision, writing - review & editing
LBL: conceptualisation, formal analysis, funding acquisition, investigation, methodology, supervision, validation, writing - review & editing

4. **In which chapter(s) of your thesis can this material be found?**

   Chapter 6: Creating a Risk Prediction Model

5. **e-Signatures confirming that the information above is accurate** (this form should be co-signed by the supervisor/ senior author unless this is not appropriate, e.g. if the paper was a single-author work)

   Candidate

   KMA Ho

   Date:

   27/03/2024

   Supervisor/ Senior Author (where appropriate)

   LB Lovat

   Date

   27/03/2024

Acknowledgements

As I have discovered over the past 4 or so years, embarking on a PhD is a large challenge, and I am grateful to a huge number of people who have supported and believed in me.

First, I thank Professor Laurence Lovat (primary supervisor) for his support throughout my PhD. I admire his dedication to both research and his specialty and his continuous push and enthusiasm for me to excel in what I do. I thank him for believing in my abilities and securing funding for me so I can focus on my research.

Second, I also thank Professor Ami Banerjee (secondary supervisor) for his support. He has offered an alternative insight into my PhD and introduced me to new collaborators. His input has very much enriched my PhD.

Third, I thank Dr Alvina Lai (Thesis Committee Member) for her time, especially in pointing out resources when I initially started on my PhD journey when I was a complete novice to data science.

I also thank members of the Lovat Group, past and present: Abi, Áine, Archana, Ash, Daniele, Hazel, Laura, Lester, Margaret, Paul, Rawen, Nas, Sharon, Tim, Vanessa and Yusuf.

Finally, I would like to thank the unwavering support of my family who have carried me through my PhD and life in general.

Contents

List of Tables

List of Figures

List of Abbreviations

2ww	2-week wait
BBC	British Broadcasting Corporation
BCSP	Bowel Cancer Screening Programme
BMI	Body mass index
BO	Barrett's oesophagus
BSG	British Society of Gastroenterology
CI	Confidence interval
COVID-19	Coronavirus disease 2019 (caused by SARS-CoV-2)
CpG	Cytosine-phosphate-Guanine
CT	Computerised tomography
CTC	Computerised tomography colonography (also known as virtual colonoscopy)
EDS	Edinburgh Dysphagia Score
EMR	Endoscopic mucosal resection
ERCP	Endoscopic retrograde cholangio pancreatography
ESEM	Endoscopically suspected (o)esophageal metaplasia
ESD	Endoscopic submucosal dissection

EUS	Endoscopic ultrasound
FAP	Familial adenomatous polyposis
FIT	Faecal immunochemical test
GI	Gastrointestinal
GIST	Gastrointestinal stromal tumour
GOJ	Gastro-oesophageal junction
GORD	Gastro-oesophageal reflux disease
GP	General Practitioner
HADS	Hospital Anxiety and Depression Scale
HNPCC	Hereditary non-polyposis colorectal cancer (also known as Lynch Syndrome)
HPB	Hepato-pancreato-biliary
IBD	Inflammatory bowel disease
ICD-10	International Classification of Diseases, 10th revision
IM	Intestinal metaplasia
LGI	Lower gastrointestinal
NSAID	Non-steroidal anti-inflammatory drugs
NHS	National Health Service
NPV	Negative predictive value

OAC	Oesophageal adenocarcinoma
OGD	Oesophagogastroduodenoscopy
OSCC	Oesophageal squamous cell carcinoma
OR	Odds ratio
PCCRC	Post-colonoscopy colorectal cancer
PET	Position emission tomography
PPE	Personal protective equipment
PPI	Proton pump inhibitors
PPV	Positive predictive value
RFA	Radiofrequency ablation
SARS-CoV-2	Severe acute respiratory syndrome coronavirus 2
SNP	Single nucleotide polymorphism
TNM	Tumour, Node, Metastasis (cancer staging system), maintained by Union for International Cancer Control (UICC)
UGI	Upper gastrointestinal
UK	United Kingdom
USA	United States of America
UKRI	United Kingdom Research and Innovation

VOC	Volatile organic compounds
WHO	World Health Organization

List of Publications and Presentations

Journal Publications

Related to Thesis

Ho KMA, Banerjee A, Lawler M, Rutter MD, Lovat LB. Predicting endoscopic activity recovery in England after COVID-19: a national analysis. *Lancet Gastroenterol Hepatol* 2021; **6**: 381–90.[3]

Ho KMA, Davies H, Epstein R, et al. Spatiotemporal droplet dispersion measurements demonstrate face masks reduce risks from singing. *Sci Rep* 2021; **11**: 1–11.[4]

Ho KMA, Baggaley RF, Stone TC, et al. Face Mask Acceptability for Communal Religious Worship During the COVID-19 Pandemic in the United Kingdom: Results from the CONFESS Study. *J Relig Health* 2023; **62**: 608–26.[5]

Ho KMA, Rosenfeld A, Hogan Á, et al. Development and validation of a multivariable risk factor questionnaire to detect oesophageal cancer in 2-week wait patients. *Clin Res Hepatol Gastroenterol* 2023; **47**: 102087.[6]

Unrelated to Thesis

Ho KMA, Anandhakrishnan A, Mahay A, Soo Y, Lovat LB, Rochford AP. How COVID-19 has changed the unselected medical take: an observational study. *Clin Med* 2020; **20**: e229–33.[7]

Wolfson P, Ho KMA, Bassett P, et al. Accuracy of clinical staging for T2N0 oesophageal cancer: Systematic review and meta-analysis. *Dis Esophagus* 2021; **34**: 1–12.[8]

Wolfson P, Ho KMA, Wilson A, et al. Endoscopic eradication therapy for Barrett's esophagus–related neoplasia: a final 10-year report from the UK National HALO Radiofrequency Ablation Registry. *Gastrointest Endosc* 2022; **96**: 223–33.[9]

Maity AK, Stone TC, Ward V, et al. Novel epigenetic network biomarkers for early detection of esophageal cancer. *Clin Epigenetics* 2022; **14**: 1–14.[10]

Stone TC, Ward V, Hogan A, *et al.* Using saliva epigenetic data to develop and validate a multivariable predictor of esophageal cancer status. *Epigenomics* 2024; **16**: 109–25.[11]

National and International Presentations

Related to Thesis

Ho KMA, Rosenfeld A, Hogan Á, et al. O5 Using machine learning to develop models for the prediction of upper gastrointestinal cancers. *Gut* 2022; **71(S1)**: A3.[12]

- Presented at British Society of Gastroenterology (BSG) Annual Meeting, Birmingham, UK (oral presentation). Awarded "Best Oesophageal Oral Presentation".

Ho KMA, Rosenfeld A, Hogan Á, et al. P105 Validation of a machine learning model in a prospective cohort: the RISQ study. *Gut* 2022; **71(S1)**: A92–A92.[13]

- Presented at BSG Annual Meeting, Birmingham, UK (poster presentation).

Ho KMA, Rosenfeld A, Hogan Á, et al. PTH-72 A symptom and risk factor questionnaire accurately predicts upper gastrointestinal cancer. *Gut* 2021; **70(S4**): A136.[14]

- Presented at BSG Annual Meeting 2021, virtual (poster presentation).

Ho KMA, Rosenfeld A, Hogan, Á, McBain H, et al. P0074 A symptom and risk factor questionnaire accurately predicts upper gastrointestinal cancer. *United Eur Gastroenterol J* 2021; **9(S8)**: 302.[15]

- Presented at United European Gastroenterology Week 2021, virtual (poster presentation).

Ho A, Wolfson P, Wilson A, et al. P28 Comparison of anxiety and depression scores between 2-week wait and Barrett's surveillance endoscopy referrals. *Gut* 2021; **70(S1)**: A56.[16]

- Presented at BSG Campus 2021, virtual (poster presentation).

Chapter 1: Introduction

In this chapter I will highlight the global burden of gastrointestinal cancers, and also provide an overview of how COVID-19 has disrupted healthcare services.

1.1 The Global Burden of Gastrointestinal Cancers

Cancers affecting the gastrointestinal tract represent a significant disease burden globally. In 2018, there were an estimated 4.8 million cases of gastrointestinal cancers, leading to 3.4 million deaths.[17] This accounted for 26% of all global cancers and 35% of all cancer related deaths.[17] It has been predicted by the year 2040 new cases and deaths from gastrointestinal cancers are projected to increase by 58% and 73% respectively.[17]

In the United Kingdom, they feature prominently in the 20 most common cancers. These include colorectal cancer (4th commonest), oesophageal cancer (14th commonest) and gastric (17th commonest).[18] These cancers are further associated with significant morbidity and mortality; over half of all colorectal cancers are diagnosed at a late stage, while oesophageal and gastric cancers are associated with a 5-year survival rate of 12% and 17% respectively.[18] While over the last decade annual incidence rates of oesophageal and colorectal cancers have remained stable, gastric cancers have risen by 29%.[18] These headline statistics suggest that there can still be improvements in both the early detection and diagnosis of such cancers, which could lead to improvements in patient survival and quality of life. Importantly, more than half of all gastrointestinal cancers are caused by modifiable risk factors such as alcohol and tobacco use.[17] This therefore means primary prevention measures, such as introducing governmental policy to reduce alcohol and tobacco use, must be part of any strategy to reduce the burden of gastrointestinal cancers.[17]

Finally, although the definition for gastrointestinal cancers can also include solid organ cancers such as liver, gall bladder and pancreas, I will predominantly focus on the major luminal cancers, i.e. oesophageal, gastric and colorectal in this thesis.

1.2 Oesophageal Cancer

Globally, there are approximately 473,000 new cases and 436,000 deaths related to oesophageal cancer per year.[19] There are two major histological subtypes of oesophageal cancer: adenocarcinoma (OAC), which is most common in the Western hemisphere, and squamous cell carcinoma (OSCC), which is prevalent in Africa, Central and South-East Asia.[19,20] Worldwide, 90-95% of oesophageal cancer cases are squamous cell.[17,20] Although both subtypes show a male predominance, this is greater for OAC, with a 4.4:1 male: female ratio, compared to 2.7:1 for OSCC.[20] While most Asian countries have noted decreasing incidence trends, there have been increasing trends in historically low-risk populations such as in the United Kingdom and in people of white ethnicity in the United States of America.[17]

1.2.1 Oesophageal Adenocarcinoma (OAC)

OAC is thought to arise from chronic gastro-oesophageal reflux disease (GORD), where acid from stomach travels up the oesophagus.[21] This leads to the development of Barrett's Oesophagus (BO), where the oesophageal lining undergoes an intestinal metaplastic (IM) process and squamous cells are replaced by mucus secreting columnar cells.[22,23] The presence of BO leads to a 30-40 fold increase in the risk of OAC, with an annual incidence rate of 0.1% to 2.9% in the development of oesophageal adenocarcinoma.[22,24] It is estimated that the prevalence of BO is 1.6-6.8% of people worldwide.[24] The risk of progression of BO to oesophageal adenocarcinoma increases with both length of segment as well as the presence of dysplasia.[25] Thus, the diagnosis of BO requires both endoscopic evaluation and histological confirmation.

There are many risk factors which predispose both to the development of BO and OAC. These include male sex, older age, White Caucasian ethnicity, obesity, GORD and smoking which have been found to be most consistent in showing causality.[22,24] In addition, there is a familial form of BO which accounts for approximately 10% of cases.[26]

Other risk factors include *Helicobacter pylori* infection, which has an inverse correlation with BO and OAC. A meta-analysis demonstrated that *H. pylori* infection was protective in the development of BO, with an odds ratio (OR) of 0.46 (95% CI: 0.35-0.60), based primarily on four studies which had minimal selection and information biases.[27] In addition, *H. pylori* infection was inversely associated with a GORD diagnosis.[28] This inverse association was also confirmed for OAC but not for OSCC (OR=0.57, 95% CI: 0.44–0.73).[29] Other factors including non-steroidal anti-inflammatory drugs (NSAID) use and a diet rich in fruits and vegetables have been found to be protective in the development of BO and OAC.[24] In contrast, a diet rich in trans-fats and red meat has been found to be associated with the development of BO and OAC.[24] Table 1-1 summarises risk factors for BO and OAC.

Risk Factor	**Barrett's Oesophagus**	**Oesophageal Adenocarcinoma**
Age	+	+
Sex	+ (Male)	+ (Male)
Ethnicity	White Caucasian	White Caucasian
GORD	+	+
H. pylori infection	-	-
Smoking	+	+
Alcohol Use	None	None
Obesity	+	+
Diet	+ (Red Meat/Trans-Fats) - (Fruits and Vegetables) - Omega 3 Fatty Acids, Vitamin E	+ (Red Meat/Trans-Fats) - (Fruits and Vegetables) - Vitamin A, Beta-Carotene
Drugs	+ (NSAIDS)	+ (NSAIDS)

Table 1-1: Summary of risk factors associated with Barrett's oesophagus (BO) and oesophageal adenocarcinoma (OAC).

+: positive association, -: negative association. Table adapted from Schneider et al.[24]

Psychosocial factors have also been associated with the development of OAC. A Swedish study found lower socio-economic status, as defined by occupation, was associated with a

3.7-fold increase (95% CI: 1.7-7.7) in developing OAC when adjusted for age and sex respectively when compared to professionals.[30] However, when adjusted for smoking, body mass index (BMI) and GORD symptoms this was no longer statistically significant.[30] In addition, a shorter duration of living with a partner (<1 year compared to >31 years) was associated with a 2.3-fold increase (95% CI: 1.2-4.5) in developing OAC, even when adjusted for smoking, BMI and GORD symptoms.[30]

1.2.2 Oesophageal Squamous Cell Carcinoma (OSCC)

Oesophageal squamous cell carcinoma (OSCC) is the dominant histological subtype of oesophageal cancer found globally, representing 90% of incident cases.[31] In contrast to OAC, the highest incidence is found in Asia and Africa in two distinct geographical regions: the first is the "oesophagea" which stretches from Turkey, through Iran, Central Asia and into northern and central China. The second is the "oesophageal cancer corridor" which stretches from Ethiopia to South Africa.[17,32] A further hotspot is found in South America near Uruguay.[31] Notably however, there can be huge variations in OSCC incidence within the same country, notably in China and Iran.[31]

OSCC is thought to develop from the oesophageal mucosa coming in direct contact with carcinogenic compounds.[21] Mechanical injury can occur with known risk factors such as hot beverages and radiotherapy.[21] Other risk factors attributable to OSCC development include increasing age, alcohol use, tobacco use, high temperature beverages and also family history.[19,31,33,34] The risk associated with tobacco use appears to be from all forms of tobacco including cigarette, pipe, cigar smoking and chewing tobacco.[31] The risk associated with alcohol appears to vary depending on population but may vary between 1.6-fold to 5.3-fold between studies in high prevalence countries.[31,35] However, in lower incidence countries such as the USA, there appeared to be a pronounced risk for alcohol, with one study

demonstrating a 16.9-fold increase (95% CI: 10.1-28.1) in high compared to low alcohol consumers.[31,36] Table 1-2 summarises risk factors for OSCC.

Risk Factor	Oesophageal Squamous Cell Carcinoma
Age	+
Smoking	+
Alcohol Use	+
Obesity	-
Diet	+ (Hot Beverages) + (Red Meat) - (Fruits and Vegetables)
Achalasia	+

Table 1-2: Summary of risk factors associated with oesophageal squamous cell carcinoma (OSCC).

+: positive association, -: negative association. Table adapted from Abnet et al. and Lagergren et al. [21,31]

There is limited evidence of the role of GORD in the development of OSCC: two small case-control studies have linked squamous cell carcinoma with non-acid gastro-oesophageal reflux disease diagnosed during multichannel impedance, although these were single centre studies, conducted in South Africa and Japan respectively.[37,38] Gastric atrophy has also been associated with OSCC; It has been postulated that non-acid gastro-oesophageal reflux could be the missing link in this causation. [39,40] This link will be assessed in further detail in Chapter 3: Literature Review of Gastro-Oesophageal Reflux as a Risk Factor for Oesophageal Cancer.

Wider sociodemographic factors may also play a role in the development of OSCC. A study based in Kashmir, India found that people working in farming had a 3.70-fold increase (95%CI: 1.56-9.09) in developing OSCC. However, once this was adjusted for a number of variables including age, ethnicity and education this was no longer statistically significant.[41] A case-control study based in Iran demonstrated even after adjustment for factors including tobacco and alcohol use, risks of OSCC were increased in people with no formal education compared to middle school or higher (OR=5.00, 95% CI: 1.54-16.7) and also in people who were widowed/divorced compared to being married (OR=2.44, 95% CI: 1.59-3.70).[42]

1.3 Gastric Cancer

In 2018 there were more than 1 million new incident cases of gastric cancer globally with nearly 800,000 deaths.[17] It is the fifth most commonly diagnosed cancer worldwide.[43] Countries with the highest incidence rates include South Korea, Mongolia and Japan, with up to 22 cases per 100,000 person years.[17] It is twice as common in males compared to females.[17,43]

Established risk factors include chronic *Helicobacter pylori* infection, which it is the strongest risk factor known for the development of gastric cancer.[44] This contrasts with OAC where it is instead a protective factor. *H pylori* infection leads to chronic inflammation, with the gastric mucosa progressing into atrophic gastritis and intestinal metaplasia, before eventually developing into cancer.[43] Other known risk factors include increasing age, smoking, alcohol use, gastric surgery and pernicious anaemia.[43] 10% of gastric cancer cases also occur with known family history including hereditary syndromes.[43] Risk factors are summarised in Table 1-3.

Risk Factor	Gastric Cancer
Age	+
Sex	+ (Male)
Smoking	+
Alcohol Use	+
Pernicious Anaemia	+
Previous Gastric Surgery	+
Diet	+ Salty Foods - Fruits and Vegetables
Family History	+

Table 1-3: Summary of risk factors associated with gastric cancer.

+: positive association, -: negative association. Table adapted from Smyth et al.[43] NSAID = non-steroidal anti-inflammatory drugs

1.4 Colorectal Cancer

Colorectal cancer represented 1 in 10 cancer cases and deaths worldwide in 2018, with 1.8 million cases and 881,000 deaths.[17] It is the second most common malignancy in females and

third most common in males.[45] The highest incidence globally was found in Australasia, whereas South Central Asia had the lowest.[17] Colorectal cancer remains more common in developed countries, with a higher incidence compared to less developed countries, and the transition to increasing incidence over time can be attributed to societal and economic development.[17]

Most colorectal cancers arise from a polyp, where over a course of 10-15 years an aberrant crypt develops into a polyp and progresses into cancer.[45] 70-90% of colorectal cancers develop from the adenoma-carcinoma pathway, while 10-20% develop from the serrated neoplasia pathway and 2-7% from microsatellite instability.[45] Each pathway is defined by specific genetic changes, such as *APC* gene mutation (adenoma-carcinoma pathway), *KRAS* and *BRAF* gene mutations (serrated neoplasia pathway) and mismatch repair genes (microsatellite instability).[45]

Similar to oesophageal and gastric cancers, modifiable and non-modifiable risk factors play a role in the development of colorectal cancer.[45] Established non-modifiable risk factors include increasing age, male sex, family history and hereditary colorectal cancer syndromes, such as familial adenomatous polyposis (FAP) and hereditary non-polyposis colorectal cancer (HNPCC), also known as Lynch Syndrome.[45]

Modifiable risk factors include a diet high in processed and red meats, smoking alcohol use and also obesity and a sedentary lifestyle, which are all associated with an increased risk of colorectal cancer.[17] In addition, type 2 diabetes mellitus has also been independently associated with an increase in risk, even after adjusting for potential confounding risk factors such as age and obesity.[46] In contrast, a diet with whole grains, fibre, dairy and tree nuts reduces the risk of developing colorectal cancer, as does calcium and vitamin D supplementation.[47] In addition, low dose aspirin and non-steroidal anti-inflammatory drugs

(NSAIDs) have also been found to reduce the risk of colorectal cancer.[48] These risk factors are summarised in Table 1-4.

Risk Factor	Colorectal Cancer
Age	+
Sex	+ (Male)
Smoking	+
Alcohol Use	+
Diet	+ Processed Meat, Red Meat - Whole Grains, Dairy, Fibre, Fish, Tree Nuts, Calcium, Vitamin D, Folate, Fruits and Vegetables
Family History	+ (including Hereditary Cancer Syndromes)
HPV Infection	+ (Anal Cancer)
Obesity	+
Lifestyle	+ Physical Activity - Sedentary Lifestyle
Medical conditions	+ Type 2 Diabetes, Inflammatory Bowel Disease
Medication	- NSAID and Aspirin Use

Table 1-4: Summary of risk factors associated with colorectal cancer.

+: positive association, -: negative association. Table adapted from Dekker et al and Song et al.[45,47] HPV = human papilloma virus. NSAID= non-steroidal anti-inflammatory drugs

1.5 Diagnosis and Treatment of Gastrointestinal Cancers

The mainstay for investigation for gastrointestinal cancers is through endoscopy. Patients undergo either oesophagogastroduodenoscopy (OGD) or colonoscopy for optical and histological diagnosis of upper and lower gastrointestinal cancers respectively.[21,45] Endoscopic ultrasound (EUS) may be helpful in determining depth of invasion.[43] Computed tomography (CT) of the neck, chest, abdomen and pelvis or a whole body positron emission tomography CT (PET-CT) scan are subsequently used to determine the presence of metastasis.[21]

All patients have their cancer staged, often using the Union for International Cancer Control's TNM (tumour, node, metastasis) system.[21] Furthermore, gastro-oesophageal junction (GOJ) cancers and colorectal cancers are additionally staged using the Siewert classification and the Dukes' classification respectively.[21,45]

Management of patients with suspected or confirmed gastrointestinal cancer is usually done through a multidisciplinary approach.[21] Potential treatment options for early oesophageal and gastric lesions include endoscopic based techniques such as radiofrequency ablation (RFA), endoscopic musical resection (EMR) and endoscopic submucosal dissection (ESD).[21,43,45] More advanced lesions usually either chemotherapy, biologics, chemoradiotherapy or surgery, whether singly or in combination to achieve curative intent.[21,45] Unfortunately most patients diagnosed with either oesophageal cancer or gastric cancer have overall poor prognosis due to late diagnosis or recurrence. Such patients may benefit from palliative treatment, which may be for managing local complications such as oesophageal invasion with oesophageal stenting, or managing overall symptoms such as pain and nausea.[21,43] A summary of possible treatment options is shown in Table 1-5.

Cancer	Curative Treatment Options	Palliative Treatment Options
Oesophageal Adenocarcinoma	Endoscopic therapy (RFA, EMR, ESD) Neoadjuvant chemotherapy or chemoradiotherapy + surgery Perioperative chemotherapy + surgery Immunotherapy	Endoscopic therapy (stenting) Chemotherapy Chemoradiotherapy Immunotherapy
Oesophageal Squamous Cell Carcinoma	Endoscopic therapy (RFA, EMR, ESD) Chemoradiotherapy +/- surgery Neoadjuvant chemotherapy + surgery Immunotherapy	Endoscopic therapy (stenting) Chemotherapy Chemoradiotherapy Immunotherapy
Gastric Cancer	Endoscopic therapy (EMR, ESD) Neoadjuvant chemotherapy or chemoradiotherapy + surgery Perioperative chemotherapy + surgery Immunotherapy	Endoscopic therapy (stenting) Chemotherapy Chemoradiotherapy Immunotherapy
Colorectal Cancer	Endoscopic therapy (EMR, ESD) Surgery Neoadjuvant chemotherapy or chemoradiotherapy + surgery Chemotherapy +/- biologics Immunotherapy	Endoscopic therapy (stenting) Chemotherapy+/- biologics Chemoradiotherapy Immunotherapy

Table 1-5: Summary of treatment options available for different types of gastrointestinal cancers.

RFA= radiofrequency ablation, EMR = endoscopic mucosal dissection, ESD = endoscopic submucosal dissection. Adapted from Lagergren et al., Smyth et al. and Dekker et al.[21,43,45]

1.6 Early Detection of Gastrointestinal Cancers

There are two major approaches for early detection of gastrointestinal cancers. Firstly, asymptomatic individuals or high-risk individuals could be invited for screening programmes, such as the national Bowel Cancer Screening Programme (BCSP) in the UK.[49] Secondly, a test could be applied to symptomatic individuals to reduce the time taken between presentation and diagnosis of cancer.[49] There has been increasing interest in early detection programmes,

including funding and initiatives through UK Research and Innovation (UKRI), although it has been acknowledged the current research within the field of early detection of cancer is fragmented and faces multiple barriers in translation to clinical use.[50]

1.6.1 National Screening Programmes

Bowel cancer screening was first introduced in the UK in 2006 and this has led to a fall in the incidence of colorectal cancer in the eligible population, with decreases most pronounced in distal tumour incidence between 2011 and 2014.[51] Currently bowel cancer screening is offered every 2 years in England to adults aged between 50 and 74.[52] There are some differences in age thresholds in Scotland, Wales and Northern Ireland.[53] The first line test is the faecal immunochemical test (FIT), which is completed by patients at home. Patients testing above the threshold of 120 μg Hb/g are offered a screening colonoscopy on a dedicated screening list.[54,55] Historically, there was a second national screening programme called BowelScope, which ran between 2013-2021 in England. A one off flexible sigmoidoscopy was offered for people aged 55.[53,56] Both programmes have been attributed to a stage shift in colorectal cancer diagnosis, with cancers diagnosed earlier.[53] However, there remains suboptimal uptake of the national BCSP, particularly in areas with increased deprivation or ethnic diversity.[57] Overall eligible uptake of colonoscopy was recorded at 79% and uptake for the guaiac faecal occult blood test (gFOBT), the predecessor for FIT, was recorded at around 58%.[53]

In contrast, there is no screening programme at present in the UK for upper gastrointestinal cancers. Current guidelines from the British Society of Gastroenterology (BSG) state however that it is not 'feasible' or 'justified' to screen for BO or oesophageal cancer in unselected patients with GORD.[58]

1.6.2 Targeted Surveillance Programmes

Latest BSG guidelines recommend upper GI endoscopy surveillance for patients with a greater than 3cm of BO every 2-3 years.[58] In addition, the BSG also recommends patients with extensive gastric atrophy or gastric intestinal metaplasia, defined as being present in the body and the antrum, to undergo endoscopic surveillance every 3 years.[59] Patients with inflammatory bowel disease (IBD) or inherited cancer syndromes likewise are recommended to undergo regular surveillance colonoscopy. Surveillance intervals can be as frequent as annually in patients with primary sclerosing cholangitis or polyposis syndromes.[60,61]

1.7 The Clinical Problem

As discussed earlier gastrointestinal cancers represent a significant healthcare burden worldwide. In particular, the poor survival for gastric and oesophageal cancers suggest that a significant majority of cancers are diagnosed late. In addition, the current gold standard test for both investigation and surveillance is an OGD. Not only is this invasive for patients, but also requires them to take time off to attend hospital. Moreover, endoscopy is expensive to deliver, requiring numerous healthcare professionals and specialised equipment. A previous analysis in 2009 suggested that it would cost over $80,000 US dollars to detect a case of cancer in a population with dyspeptic but non-alarm symptoms.[62] Furthermore, the vast majority of endoscopies performed ultimately do not find significant pathology. As an example, in a series of over 2,700 primary care outpatients who met diagnostic criteria for dyspepsia but lacked alarm features such as dysphagia, only 23% of patients had what was defined as significant benign endoscopic abnormality, which included oesophageal, gastric or duodenal lesions or ulcers. Only 6 patients (0.2%) were found to have cancer.[62] Although there is a national bowel cancer screening programme, uptake among the general public remains an issue, so there also needs to be a robust referral pathway to ensure patients who do not engage with screening or fall outside of eligibility are also investigated expeditiously.

In addition, even within symptomatic individuals there are additional sources of delays in patient care. This includes delays from symptoms to primary care consultation (patient related delays), delays from primary care consultation to referral (primary care related delays) and also delays from referral to consultation or consultation to diagnostic test (secondary care related delays).[63] Patient related delays appeared to have the largest effect, suggesting the need for improving patient awareness of symptoms. However, novel diagnostic strategies could have a role especially for patients with less specific symptoms and improving selection of patients for further investigation.[63] Furthermore, oesophageal cancer often only is symptomatic when it is locally advanced, with dysphagia only becoming apparent when it has invaded around two-thirds of the oesophageal circumference.[64] Similarly, gastric cancer often presents with non-specific symptoms such as dyspepsia, anorexia or early satiety. However, if symptoms are present at the time of diagnosis, often this has become locally advanced and is incurable.[43] In addition, pre-cancerous lesions such as Barrett's oesophagus and gastric intestinal metaplasia are asymptomatic hence there are challenges with early detection of patients based purely on symptomatology.[22,65]

The ideal scenario would be for individuals who present with symptoms to firstly be identified using non-invasive means. This could be in completing a questionnaire or obtaining a body tissue or fluid sample from a site which may be more tolerable for patients. Patients could then be referred to further investigations, which should yield an overall higher percentage of finding significant pathology. However, a reasonable balance must be struck. Applying methods for early detection of cancer too liberally will lead to overdiagnosis, defined as when a cancer is detected which would otherwise not have been clinically apparent in a person's lifetime, and thus create unnecessary anxiety.[49] Treatment of overdiagnosed cancers also would not offer any overall survival benefit, despite the burdens of undergoing oncological treatments.[49]

Furthermore, the COVID-19 pandemic has led to significant problems in the delivery of routine healthcare. Backlogs of cases have built up across many medical specialties, leading to rising waiting lists.[66,67] Improved management of referrals with prioritisation of more high risk could form part of a strategy to overcome backlogs while minimising adverse patient outcomes.

1.8 Potential Strategies to Reduce GI Cancer Burden

A number of interventions could be employed to potentially reduce GI cancer burden further, either in a general population or targeted in susceptible individuals.

1.8.1 Public Health Measures

Modifiable Risk Factors

Up to 45% of cancer related deaths are related to modifiable risk factors.[68] These include tobacco use, obesity, alcohol and oncogenic infections, such as *Helicobacter pylori*. A number of measures have been implemented to target tobacco use, such as taxation, use of plain packaging and addition of graphic health warnings.[68] Minimum pricing for alcohol has also been introduced in Scotland and Wales; there is evidence to suggest that this is effective in improving health outcomes such as reduction in deaths and hospitalisations related to alcohol consumption.[69] It has been proposed that a screen and treat strategy for *Helicobacter pylori* could be applied in Europe to reduce the incidence of gastric cancer.[70]

Raising Awareness in Patients

Public health information campaigns have been used to raise the profile of cancer with the UK. These include the “Be Clear on Cancer” campaign which ran between 2011 to 2018.[71] This encouraged patients to approach their GP with cancer red flag symptoms such as dysphagia and per-rectal bleeding.

Population Screening

Apart from the established bowel cancer screening programme for lower GI cancers, population screening for upper GI cancers could also be introduced. Currently, there is no population screening programme in existence in Europe.[70] Worldwide, Japan and South Korea both offer national screening programmes for gastric cancer.[72] Most gastric cancers develop from pre-cancerous lesions. These include gastric mucosal atrophy, intestinal metaplasia and dysplasia.[70] Consequently, upper GI endoscopy would be the only modality which would ensure appropriate risk stratification of pre-cancerous lesions and subsequent earlier detection of dysplastic lesions which would have more curative options available.[70]

1.8.2 Healthcare Measures

Measures to improve delivery of healthcare for patients diagnosed with gastrointestinal cancers can also be introduced to improve overall outcomes for patients. This includes raising awareness in healthcare professionals, especially within primary care, as less than half may be aware of new changes in guidance, and even fewer may be using it to guide referrals.[73] Furthermore, fast track pathways, such as the 2-week wait pathway (2ww) ensure timely referral, prompt investigations within secondary care and help to reduce diagnostic delays.[74] Finally, although beyond the scope of this thesis, cancer treatment advances such as in immunotherapy, radiotherapy and chemotherapy show potenial in improving patient outcomes.[75,76]

1.9 Chapter Summary

In this chapter I have presented the risk factors for the development of oesophageal, gastric and colorectal cancers. I have also discussed about the diagnosis and treatment of these cancers. In addition, I have highlighted some of issues around early detection of cancers, which has become an acute issue due to COVID-19 pandemic related backlogs, and other potential strategies to improve detection and treatment of cancers. Over the coming

chapters, I will define the effect of the pandemic within gastroenterology and assess how backlogs could be tackled.

Chapter 2: Thesis Aims and Objectives

This chapter will outline the aims and objectives of my thesis.

2.1 Chapter Introduction

When I embarked on my PhD back in September 2019 the initial aim was to use large scale datasets to identify patients at risk of development of upper gastrointestinal (UGI) cancers, and to develop a clinical tool which could be used to identify such patients. The emergence of COVID-19 in early 2020 led to significant disruption in all facets of life, but it also presented new opportunities and directions for research. As a result, COVID-19 related themes feature prominently within this thesis.

2.2 Aims and Objectives

In my PhD I aim to

- Identify known risk factors in the development of gastrointestinal cancers, focussing on oesophageal cancer and incorporating a meta-analysis (Chapter 1: Introduction and Chapter 3: Literature Review of Gastro-Oesophageal Reflux as a Risk Factor for Oesophageal Cancer). In particular, I investigate the association between gastro-oesophageal reflux disease (GORD) and the development of oesophageal squamous cell carcinoma (OSCC), which is the predominant subtype worldwide.
- Introduce the datasets that are used in this thesis and how they were prepared for analysis (Chapter 4:Datasets and Data Curation). I introduce both routine and bespoke datasets used in my thesis and discuss some of the drawbacks of each dataset.
- Highlight the effect of the COVID-19 pandemic on gastroenterology services and discuss mitigation measures to help overcome backlogs (Chapter 5: Effect of COVID-19 on Gastrointestinal Services). I use both local and national data to demonstrate

its effects and perform predictive modelling to estimate a national backlog of endoscopic procedures.

- Use machine learning to create both a questionnaire and salivary epigenetic based tool which could be used to triage patients at high risk for upper gastrointestinal (UGI) cancer for endoscopy (Chapter 6: Creating a Risk Prediction Model). I compare different machine learning methods to select the best performing model, then create an online calculator which can estimate an individual's risk for UGI cancer.
- Discuss the use of faecal immunochemical testing (FIT) in the triaging of lower GI referrals (Chapter 7: Faecal Immunochemical Testing for Triaging Lower Gastrointestinal Referrals). In particular, I focus on the use of a duplicate FIT testing strategy using primary care data to assess if it would outperform using a single FIT result and improve the overall sensitivity of the test.
- Highlight some of the wider societal impacts which the COVID-19 pandemic has had (Chapter 8: Societal Impacts of COVID-19 Pandemic). I focus especially on worship within a religious setting and how government rules enforced at the time impacted on the experience of worshippers.
- Discuss the implications of evidence and future work (Chapter 9: Discussion). This is especially relevant given the increasing role of artificial intelligence in healthcare and specifically raises ethical questions on regulation and clinical responsibility.

2.3 Chapter Summary

In this chapter I have outlined the main aims and objectives of my thesis. Principally this includes describing the impact of the COVID-19 pandemic on gastrointestinal services and possible mitigation measures to minimise its effects on patient care.

Chapter 3: Literature Review of Gastro-Oesophageal Reflux as a Risk Factor for Oesophageal Cancer

In this chapter I will outline established risk factors for the development of gastrointestinal cancers. In particular, I will focus on the association between gastro-oesophageal reflux disease and the development of oesophageal cancer, for which I performed a systematic review and meta-analysis of published literature up to July 2022.

3.1 Chapter Introduction

Gastrointestinal cancers can be split into solid organ cancers affecting the liver and pancreas as well as those affecting the digestive tract, termed luminal cancers. The ligament of Treitz, marking the junction between the duodenum and the jejunum has traditionally been used to separate between the upper and lower gastrointestinal tract.[77] This chapter will focus on luminal cancers and so will primarily discuss risk factors for these cancers, with a more in-depth look at the relationship between gastro-oesophageal reflux disease (GORD) and the development of oesophageal cancer.

3.2 Systematic Review and Meta-Analysis of Gastro-Oesophageal Reflux and Development of Oesophageal Cancer

3.2.1 Introduction

Gastro-oesophageal reflux occurs when there is retrograde movement of stomach contents back into the oesophagus. This is defined as GORD when this is so frequent as to cause either "troublesome symptoms and/or complications."[78]. GORD is one of the most significant risk factors for the development of oesophageal adenocarcinoma (OAC). This link was identified initially in the 1990s in two case-control studies.[79,80] A meta-analysis published in 2010 has further strengthened this association.[81] Rubenstein et al. demonstrated that at least weekly

symptoms of GORD increased the odds of OAC by fivefold, while daily symptoms increased the odds by sevenfold.[81] Conversely, *Helicobacter pylori* colonisation was found to be protective for the development of OAC (odds ratio (OR) = 0.56, 95% confidence interval (95% CI) = 0.46-0.68).[82] Broader population level studies using routine primary care data in over 40,000 patients further demonstrated that reflux and dyspepsia was associated with an odds ratio (OR) of 5.7 and 6.0 respectively for the development of oesophago-gastric cancer.[83] However, this study did not stratify both in terms of site and histology of cancer; indeed given it is based on UK data it is likely to be weighted more towards OAC which has a stronger association with GORD.

Further to this, GORD has also been linked to the development of Barrett's oesophagus (BO).[84] However, not all cases of OAC are derived initially from BO. A study on pre-operative patients with oesophageal cancer suggested less than 5% had pre-existing BO.[85] A more recent systematic review and meta-analysis demonstrated a pooled prevalence of 7.2% and 12.0% for histologically confirmed and endoscopically suspected BO in adults with GORD.[86] However, GORD in combination with other risk factors such as obesity or age >50 years acted in tandem to increase the risk of BO: one meta-analysis demonstrated that the prevalence of GORD in patients with BO was 3%, increasing to 12.2% in the presence of another risk factor.[87] It therefore seems likely that there may be distinct pathways as to how BO and GORD both lead to the development of OAC.

Although there has been published evidence of the link between GORD and OAC, very few studies have been published with regards to GORD and oesophageal squamous cell carcinoma (OSCC), despite the fact the OSCC represents over 90% of the world's cases of oesophageal cancers.[17,20] However, there has been increasing evidence that non-acid reflux may play a role, and may be the link between the presence of gastric atrophy and the development of OSCC.[88] In particular, Wang et al., using a cohort of nearly half a million

patients, demonstrated that 17% of OSCC cases were associated with GORD, with an adjusted hazard ratio of 1.99.[89] Within the study GORD prevalence was estimated to be 24%, which also corroborates with range of estimates between 18.1% to 27.8% within North America.[90]

At the time of writing, no meta-analysis had been completed to investigate the association between GORD and OSCC. I therefore aimed to assess the association of both the presence and frequency of GORD symptoms with both OSCC and OAC.

3.2.2 Material and Methods

Search Strategy

Prior to the study I searched the PROSPERO database for registration of systematic reviews and meta-analysis in July 2022.[91] I identified 26 studies which investigated oesophageal cancer and a further study which investigated reflux. However, none of the studies specifically investigated the role of GORD and oesophageal cancer.

I searched for eligible studies using MEDLINE (PubMed), EMBASE (Elsevier) and The Cochrane Library databases up to 27 July 2022 for relevant articles. I used the following search terms, with no language or geographical restrictions:

(oesophageal cancer OR esophageal cancer) AND (reflux OR gastro-oesophageal reflux disease OR gastro oesophageal reflux disease OR GORD OR gastro-esophageal reflux disease OR GERD OR gastro esophageal reflux disease OR dyspepsia OR heartburn OR regurgitation OR reflux)

Study Selection and Outcome of Interest

I used the following inclusion criteria:

- Performed in adults aged 18 years or older

- Provided data on OAC and/or OSCC
- Provided data on diagnosis of GORD or typical symptoms of GORD with no time restrictions
- Provided data on frequency of GORD or typical symptoms of GORD with no time restrictions

I defined typical symptoms of GORD as those derived from the Montreal Definition which included heartburn and regurgitation.[78] I further defined frequent symptoms of GORD as those occurring on at least a weekly interval.

Data Collection

Titles and associated abstracts from the initial searches were independently screened by two reviewers (me and another researcher (PW)). Studies of potential interest were then agreed on and the full text article independently read to assess if it met inclusion criteria. Data was extracted independently and entered into a Microsoft Excel 365 worksheet (Redmond WA, USA). Any disagreement was resolved by discussion, with the thesis primary supervisor making the final decision over any disagreements. I attempted to contact corresponding authors for any missing data from highlighted studies. Two reviewers (me and another researcher (PW)) used the Newcastle-Ottawa Scale (NOS) to independently assess for study quality and assess for bias in included studies (Table 3-1).[92]

For each included study I collected data on year of publication, journal, years of study, country of origin, demographic information such as age and sex and study design. In addition, I collected data on the number of patients who were diagnosed with and without GORD and those who experienced this at least weekly, stratified by their cancer status.

	Selection				Comparability	Exposure			Score
	Adequacy of case definition	Representativeness of the cases	Selection of controls	Definition of controls		Ascertainment of exposure	Same method of ascertainment	Non-response rate	
Anderson	✵	✵	✵	✵	✵	✵	✵		7
Crane	✵	✵	✵	✵	✵		✵		6
Farrow	✵	✵	✵	✵	✵	✵	✵	✵	8
Lagergren	✵	✵	✵	✵	✵	✵	✵		7
Pandeya	✵	✵	✵	✵	✵	✵	✵	✵	8
Wu	✵	✵	✵	✵	✵	✵	✵	✵	8

Table 3-1: Newcastle Ottawa Scale for assessment of included studies

Case definition = oesophageal adenocarcinoma or oesophageal squamous cell carcinoma. Exposure = presence of gastro-oesophageal reflux

Statistical Analysis

I used the Cochran-Mantel-Haenszel Method to pool together incidence rates from different studies. I used the DerSimonian-Laird random-effects method for analysis, regardless of the degree of heterogeneity between the study results. I used a binary measure for the outcome of interest, which was the presence or absence of any GORD and the presence and absence of at least weekly symptoms of GORD. The difference between groups was reported as odds ratios (OR) with corresponding 95% confidence intervals (95% CI). Heterogeneity was quantified assessed using the I^2 statistic. For all statistical calculations, a p value of ≤0.05 was taken as significant. I aimed to assess for possible publication bias using funnel plots. Statistical analyses and graphical plots were created using R software version 4.1.2 (R Core Team, Vienna, Austria) using the "meta" package.[93,94] I conformed to the PRISMA guidelines for the reporting of this meta-analysis.[95]

3.2.3 Results

Characteristics of Included Studies

Figure 3-1 demonstrates the flow diagram for the meta-analysis. 51 studies were assessed for inclusion after screening. After eligibility assessment I identified 10 studies with potential data available for extraction and analysis.[79,96–104]. However, I subsequently excluded four studies due to use of overlapping databases; three studies included the Australian Cancer Study database hence I selected a single representative one, Pandeya et al. for my analysis.[101] Two studies collated data from the Barrett's and Esophageal Adenocarcinoma Consortium (BEACON), an international collaboration for research into Barrett's oesophagus and OAC.[103,104] As these studies included a number of existing data sources such as the Australian Cancer Study and Los Angeles county data,[97,100,101] I also excluded them to avoid the risk of duplication. Moreover, for the study by Wu et al., I extracted data from Rubenstein et al.'s meta-analysis.[81,97] This was because in the original paper, odds ratios rather than case

numbers of patients were presented. Rubenstein et al. subsequently requested and published the raw original data which was preferable for my analysis. Notably, all included studies were conducted in higher income countries. Studies from lower- and middle-income countries were underrepresented, despite OSCC being more prevalent in these areas.

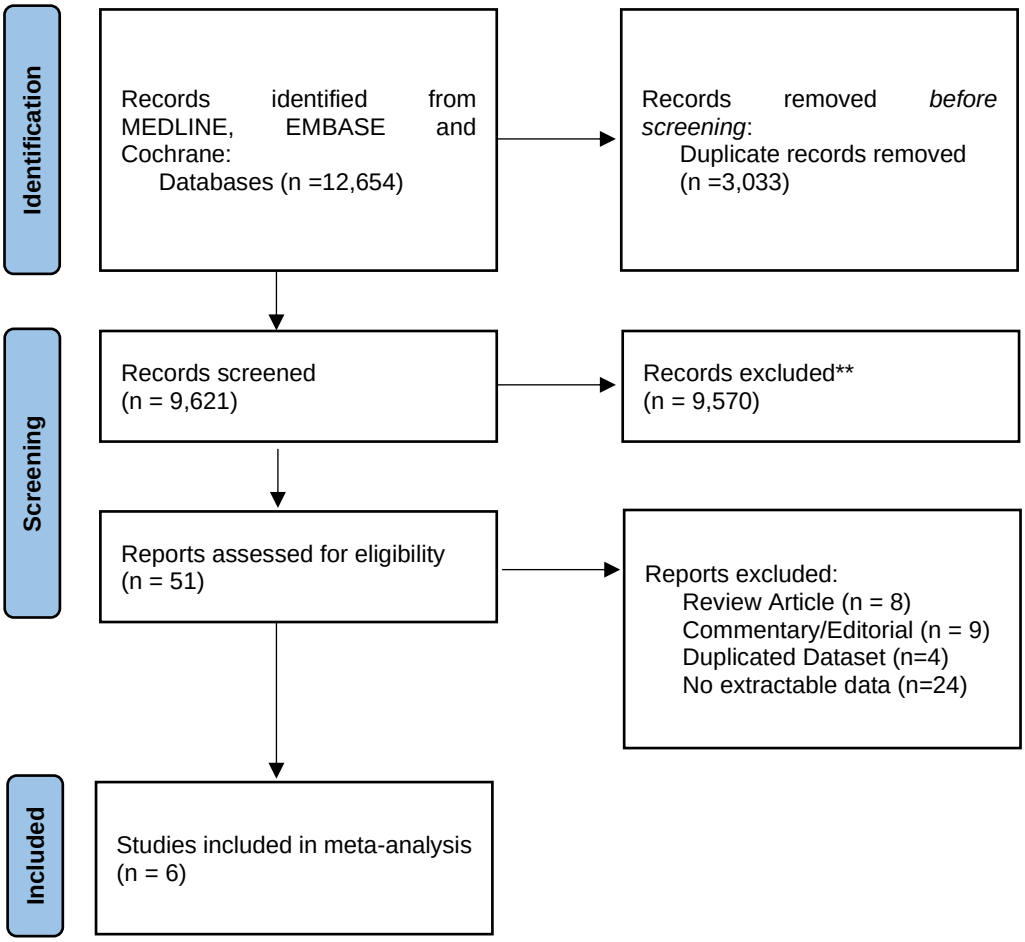


Figure 3-1: Flow diagram of search strategy and selection of studies

Lead Author	Year	Country	Study Design	Recruitment Years	Method of GORD Diagnosis	Controls		OAC		OSCC	
						Mean Age	% Male	Mean Age	% Male	Mean Age	% Male
Anderson LA[98]	2007	NI and ROI	Case Control	2002-2004	In-Person Interview	63	85	64	85		
Crane SJ[99]	2007	USA	Case Control	1971-2000	Medical Records	#	#	70	83		
Farrow DC[96]	2000	USA	Case Control	1993-1995	In-Person Interview	61^	80	63^	84	64^	80
Lagergren J[79]	1999	Sweden	Case Control	1994-1997	In-Person Interview	68*	83	69*	87	67*	72
Pandeya N[101]	2010	Australia	Case Control	2001-2005	Self-completed Questionnaire	61	66	64	91	65	57
Wu AH[97]	2003	USA	Case Control	1992-1994	In-Person Interview	60	74	61	91		

Table 3-2: Demographics of participants in included studies

Data for Farrow et al. and Wu et al. (2003) was taken from Gammon MD et al. and Wu et al. (2001) as referenced in the journal articles respectively.[96,97,105,106]

OAC = Oesophageal Adenocarcinoma; OSCC = Oesophageal Squamous Cell Carcinoma; NI= Northern Ireland, ROI = Republic of Ireland, USA= United States of America

*=median age presented as mean age not available

^=estimated from the proportions of patients within each age band

#=no formal age and % of males were presented but the authors state controls were age matched within a year and sex matched

Lead Author	Country	Total Number			Normal (No GORD)			Any GORD			At Least Weekly GORD		
		OSCC (n=588)	OAC (n=1199)	Controls (n=4672)	OSCC (n=378)	OAC (n=463)	Controls (n=3059)	OSCC (n=210)	OAC (n=736)	Controls (n=1613)	OSCC (n=108)	OAC (n=539)	Controls (n=690)
Anderson LA[98]	All Ireland		227	260		117	211		110	49		110	49
Crane SJ[99]	USA		29	29		16	25		13	4			
Farrow DC[96]	USA	144	198	671	107	72	355	37	126	316	8	67	78
Lagergren J[79]	Sweden	167	189	820	142	76	685	25	113	135	25	113	135
Pandeya N[101]	Australia	277	344	1545	129	74	686	148	270	859	75	145	178
Wu AH[97]	USA		212	1347		108	1097		104	250		104	250

Table 3-3: Distribution of oesophageal cancer subtypes and presence and frequency of GORD

GORD = Gastro-Oesophageal Reflux Disease, OAC = Oesophageal Adenocarcinoma; OSCC = Oesophageal Squamous Cell Carcinoma

Table 3-2 shows the demographic information for the six included studies.[79,96–99,101] All studies had a male predominance with a mean age of over 60 for both cancer groups. The six studies included a total of 588 OSCC, 1,199 OAC and 4,672 control patients; six studies included OAC patients while four studies recruited OSCC patients (Table 3-3).While some studies stratified by frequency of GORD symptoms,[79,96–98,101] two studies did not provide this information.[99,107] In one study, GORD symptom frequency was stratified as follows:[96]

- Never
- 1 or 2 times/year
- 3 to 12 times/year
- 13 to 104 times/year
- 105 to 364 times/year
- 365 or more times/year

I defined patients who experienced reflux at least 105 times a year as equivalent to at least weekly symptoms of GORD, as this would roughly equate to once every 3 to 4 days[96]. I was unable to expand further as the next grouping was between 13 to 104 times a year.[96] Finally two studies provided data only on participants experiencing least weekly symptoms of GORD, so the same figures were entered for weekly symptoms of GORD and the presence of GORD.[97,98]

Oesophageal Squamous Cell Carcinoma (OSCC)

Three studies provided both data on GORD and OSCC and at least weekly GORD symptoms and OSCC (Table 3-4). There was no statistically significant association seen with the presence of GORD and OSCC (Figure 3-2). In addition, there was no association seen with at least weekly symptoms of GORD and OSCC (Figure 3-3). However, importantly there was significant heterogeneity for both analyses, with the I^2 value above % for both instances.

Outcome	Number of Studies	Heterogeneity		Group Comparisons	
		p-value	I^2	Odds Ratio (95% CI) (*)	p-value
OSCC vs. Controls					
Any GORD	3	0.001	85%	0.68 (0.40-1.19)	0.179
At Least Weekly GORD	3	<0.001	94%	1.09(0.37-3.22)	0.883
OAC vs. Controls					
Any GORD	6	<0.001	86%	3.81 (2.54-5.72)	<0.001
At Least Weekly GORD	5	<0.001	61%	4.95 (3.90-6.27)	<0.001

Table 3-4: Heterogeneity and relative risks for comparison of OSCC vs. controls and OAC vs. controls

OSCC = oesophageal squamous cell carcinoma, OAC = oesophageal squamous cell carcinoma

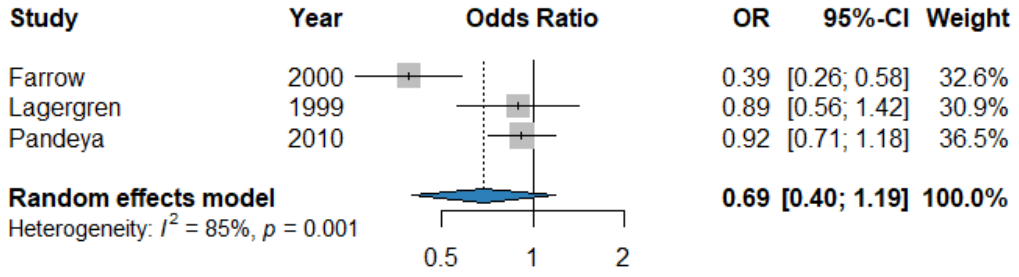


Figure 3-2: Forest plot of presence of GORD and odds ratio (OR) and 95% confidence intervals (95% CI) of OSCC

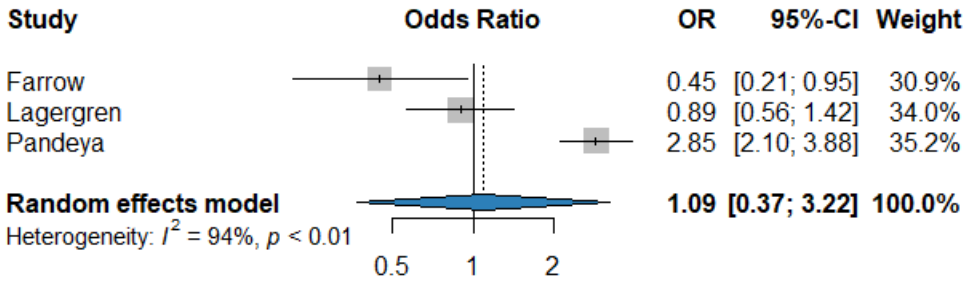


Figure 3-3 Forest plot of at least weekly symptoms of GORD and odds ratio (OR) and 95% confidence intervals (95% CI) of OSCC

GORD = Gastro oesophageal reflux disease, OSCC = oesophageal squamous cell carcinoma

Due to the low number of studies included it was difficult to draw meaningful conclusions about publication bias using funnel plots.

Oesophageal Adenocarcinoma (OAC)

There were 6 studies which provided data on GORD and OAC, while 5 studies provided data on the presence of at least weekly GORD symptoms and OAC (Table 3-4). There was a statistically significant association between the presence of GORD and OAC (p<0.001); where the OR was 3.81 (95% CI: 2.54-5.72) (Figure 3-4). The association was even stronger for at least weekly GORD and OAC (OR=4.95, 95% CI: 3.90-6.27, p<0.001) (Figure 3-5). Once again, there was significant heterogeneity between studies, with the I^2 value above 60% for both instances (p<0.001).

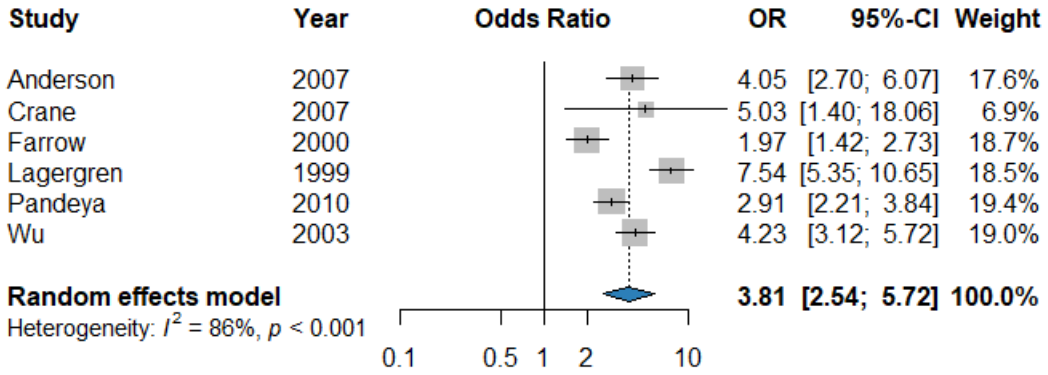


Figure 3-4: Forest plot of presence of GORD symptoms and odds ratio (OR) and 95% confidence intervals (95% CI) of OAC

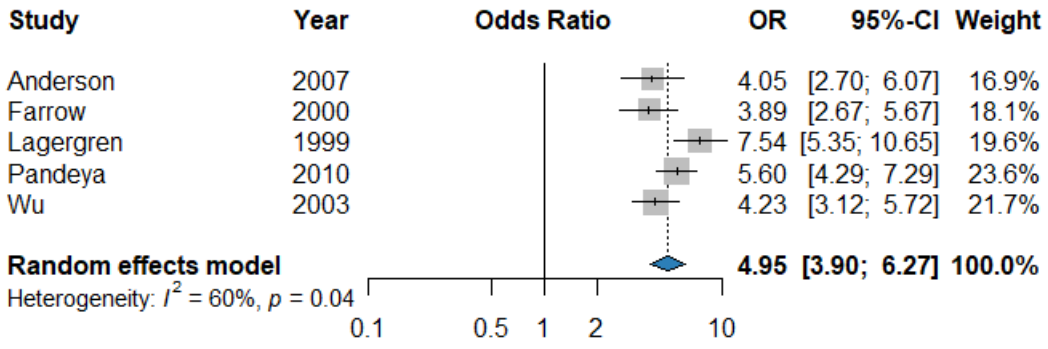


Figure 3-5: Forest plot of at least weekly symptoms of GORD symptoms and odds ratio (OR) and 95% confidence intervals (95% CI) of OAC

GORD = Gastro oesophageal reflux disease, OAC = oesophageal adenocarcinoma

Similarly, due to the low number of studies it was difficult to draw meaningful conclusions on publication bias.

3.2.4 Discussion

My meta-analysis suggests there is no association between GORD symptoms and the development of OSCC. Other evidence using multichannel impedance and pH studies has suggested that there is an association between non-acid gastro-oesophageal reflux and OSCC.[37,38] Non-acid reflux is defined as any refluxate being pH>4, hence it would impossible to differentiate between acid or non-acid reflux based on patient symptoms alone.[108] It has been noted that in GORD patients non-acid reflux can account for 50% of reflux episodes, rising to 75% for healthy controls.[109] Uno et al. in a cohort of 14 OSCC patients and 14 age and sex matched controls diagnosed with superficial differentiated early gastric cancer demonstrated that the total number of reflux and non-acid reflux episodes was higher in the OSCC group compared to the control group, although for acid reflux episode numbers were similar.[38] Notably however none of the participants in this study had any GORD symptoms.[38] Furthermore using similar methods, Kgomo et al. demonstrated an OR of 8.8 (95% CI: 3.2-24.5) in the development of OSCC in patients with non-acid gastro-oesophageal reflux. Interestingly however for acid gastro-oesophageal reflux OR was 0.04 (95% CI: 0.009-0.189), suggesting there may be a protective effect.[37] Importantly however, it was not clear how many of the OSCC patients may have been symptomatic, although some control group patients could have been symptomatic as they were initially referred for investigation due to “dyspeptic or other benign symptoms.”[37] I was unfortunately unable to establish contact with the authorship team to confirm this. Furthermore, in my meta-analysis there was no association between at least weekly symptoms of GORD and the development of OSCC.

Notably however, my findings contradict the findings of a large-scale prospective cohort study of nearly half a million patients. The study estimated that GORD was associated with a near twofold risk of OSCC after correcting for sex, smoking status, alcohol intake and follow up period.[89] There are however a few caveats; the study extracted the diagnosis of GORD

based on medical claims and hence may lead to a reduced prevalence compared to symptomatology surveys.[89] In addition, Medicare diagnosis data was available for around 20% of participants so the authors performed multiple imputation on the remaining data. As highlighted by an independent authorship team, the quality of evidence between GORD and OSCC is moderate at best,[88] and larger prospective studies would be needed to fully elucidate if there is any association.

In addition, I have further confirmed the association between GORD and OAC and have updated the meta-analysis first performed in 2010 by Rubenstein et al.[81] with the addition of a further study.[99] I demonstrated an OR of 4.95 (95% CI: 3.90-6.27) for at least weekly symptoms of GORD compared to OR of 4.92 (95% CI: 3.90-6.22) for a similar comparison in the study by Rubenstein et al.[81] It is however worth pointing out some key differences between ours and Rubenstein et al.'s analysis. Firstly, the authors looked at two symptom frequencies: at least weekly and daily GORD symptoms, whereas I assessed for the presence and absence of GORD symptoms plus at least weekly GORD symptoms.[81] Furthermore, while my control group in both instances only included patients without GORD, the control group in Rubenstein et al.'s was matched with patients without GORD or experienced it less than weekly.[81]

A point to note is that certain secular trends may have influenced the studies, which were from 1999 to 2010. This includes the rising obesity epidemic, which is a known risk factor for OAC.[110] In England, prevalence of obesity has increased from 21% to 26% between 1999 and 2010.[111] Conversely, there has been consistent reductions in tobacco smoking, particularly in Latin America and high sociodemographic countries, which would have had an overall protective effect In the development of OSCC and OAC.[112]

3.2.5 Limitations

My meta-analysis has several limitations. First, there is significant heterogeneity between studies, with only six studies included in the final analyses. This was also noted by Rubenstein et al., where moderate heterogeneity was somewhat resolved by the removal of a study in certain analyses.[81] The small number of overall studies mean it was difficult comment on publication bias, as funnel plots can only be used for 10 studies or more. In addition, the small number of studies identified meant that I was unable to provide a definitive conclusion as to whether GORD is associated with OSCC.

Second, there were some differences in the diagnosis of GORD. While the majority of studies relied on GORD being diagnosed through structured questions,[79,96–98,101,107] one study depended on GORD being diagnosed by a physician during medical chart review.[99] In addition, to reduce the risk of reverse causality, most studies excluded symptoms for a period prior to the diagnosis of cancer, although this was not explicitly stated in two studies.[99,107] Most studies also used a questionnaire or interview approach to diagnose GORD which could also introduce recall bias.

Thirdly, there are several confounders which have not been taken into account in this meta-analysis. These include age, body mass index, tobacco smoking, alcohol use and socioeconomic status. After correction for confounding variables including age and sex, only one study found an overall association between GORD and OSCC.[107] The remaining studies found either an association in certain subgroups of GORD symptom frequency[96,101] or no association at all.[79] Furthermore, Pandeya et al. demonstrated that there may be a synergistic effect with smoking, GORD and the development of OSCC, with a relative risk of 140%.[88,101] This contrasts with the association with OAC where the majority of constituent studies in the meta-analysis found associations with GORD and OAC even after adjustment for confounders in multivariate analyses.[79,96–98,101]

3.2.6 Future Work

Future work should focus on assessing if there is an association between OSCC and GORD. This could be a large scale cohort study in high prevalence areas such as in the "oesophageal cancer belt." Patients with known GORD could be screened for using a validated tool such as the GerdQ.[113] Enrolled patients could be further assessed using other validated tools such as the GORD Impact Scale to evaluate for symptom severity.[114] In addition, comprehensive data on alcohol and tobacco use could be collected to assess for possible confounding. Pragmatically, such a study may be more difficult to conduct given OSCC is highest in prevalence in lower income countries. It may be more realistic to perform a case control study, matching individuals with confirmed OSCC with controls, albeit accepting that it would not be possible to prove causality.

3.2.7 Conclusions

In this study, I have confirmed the association that GORD is a risk factor for the development of OAC. In addition, I have shown there is no association between GORD and OSCC. However, there was very limited evidence, with significant heterogenicity in only a small number of studies. Given that OSCC comprises the majority of the world's oesophageal cancer cases, large scale prospective studies in countries where OSCC is prevalent is needed to fully assess if there is a link between GORD and OSCC.

Chapter 4: Datasets and Data Curation

This chapter will outline the data sources used in my thesis. In addition, I will discuss about access to these databases and the advantages and disadvantages of each dataset.

4.1 Chapter Introduction

Over the past few years there has been increasing interest in the use of health data and how it can be used to drive improvements in patient care. Indeed, many other sectors such as travel and retail have used technology and routine data to improve the experience for the consumer.[115] This has been recognised as a priority by the UK government in a draft policy paper entitled "data saves lives: reshaping health and social care with data."[116] In the document, future strategies are laid out to harness the power of health data research. These could lead to positive outcomes such as better treatment for patients, better health results and better decision making. Priorities outlined in the document include improving ease of access and making data sharing the norm rather than the exception, improving transparency as to how data is being used and also ensuring the correct legal and regulatory framework is in place.[116] There is perhaps no better example how impactful health data research can be than the COVID-19 pandemic.

4.1.1 Data During the COVID-19 Pandemic

The role of health data was particularly amplified during the early phase of COVID-19. The UK government acknowledged the "power of data", noting that it "identified those who are most vulnerable to coronavirus" as well as "powered vital research." One of the most renowned examples was the modelling paper by Ferguson et al. which predicted over half a million deaths related to COVID-19 if no action was taken, with intensive care and inpatient hospital capacity being breached.[117] This paper was widely reported by the media to be

instrumental in the UK government introducing a lockdown on the 23rd March 2020 to curb rising COVID-19 cases and to reduce transmission of the virus.

However at the time, many other groups also sought to assess the collateral impact of the pandemic. Banerjee et al. demonstrated using linked primary and secondary electronic health record data that in a do-nothing scenario there could be more than half a million excess deaths in the UK.[118] In addition, Lai et al. also demonstrated the collateral effects of the pandemic, which could include up to 17,910 excess deaths in patients with cancer over one year as a result of the impact on both demand for and supply of cancer services.[119] Similarly, diagnostic delays for patients with suspected cancers could lead to hundreds of additional deaths depending on cancer type.[120,121] Such examples highlight how near real time data could be especially impactful in a rapidly evolving situation, with implications on national health policy.

4.1.2 Advantages and Disadvantages of Routine Health Data Usage

One of the strengths of the UK is the volume of routine patient care data held by the National Health Service (NHS). For the English NHS, researchers can request data on hospital attendances and admissions, termed Hospital Episode Statistics. Furthermore, the NHS number, a unique identifier, can link a single patient through different domains of care or life events, such as primary and secondary care records as well as registry records such as for cancer and death. This is especially powerful as it would allow researchers access to large numbers of patients and their outcomes, without the necessary expense of running large scale trials.[122] Using real world populations would also increase generalisability of research.[122] However, as routine health data is collected before research questions are posed, there can be issues related to data quality.[123,124] Indeed, it may be impossible to answer certain research questions. Professional clinical coders are often responsible for assigning diagnosis codes such as International Classification of Diseases (ICD-10) codes to

patient records, although there are often barriers which restrict accuracy of coding. These include limitations in how clinical coders are able to interpret or modify physician notes or their ability to communicate with the original physician in the case of a discrepancy.[125] Table 4-1 outlines some further advantages and disadvantage of routine health data for research:

Theme	Advantages	Disadvantages
Population	Normally large sample, generalisable populations	May lack precision in defining population of interest
Research Outcome	Able to define large numbers of research questions and outcome of interest	May be hard to define niche questions and outcomes outside standard of care
Cost	Does not require additional cost for data collection	Costs associated with data access May be some cost associated with infrastructure set up
Data Quality	Continuously updated Large number of data points	Data collection may be incomplete, have issues with coding or be inaccessible for research purposes Differences in clinical practice may not be apparent between institutions/countries
Relevance to Real World	Based on real world data so high external validity	Some health systems may be targeted to particular populations (e.g. veterans)
Regulation	Does not routinely need ethical approval	Less oversight may lead to unnoticed errors or biases

Table 4-1 Some advantages and disadvantages of routine health care data for research use

Furthermore, there have been concerns as to how personalised data could be used, sometimes for more sinister means. A Canadian cross-sectional survey found that although over 93% of respondents felt positive about the use of routine health data for research, only 58% felt confident over data security and privacy.[126] As a result, stronger safeguards have been introduced, such as the Data Protection Act 2018 which places legal obligations on organisations who process data and allows the individual rights to access personal data.[127]

4.2 Datasets in Thesis

For my thesis I have used a range of both bespoke and also openly available datasets.

4.3 Bespoke Datasets

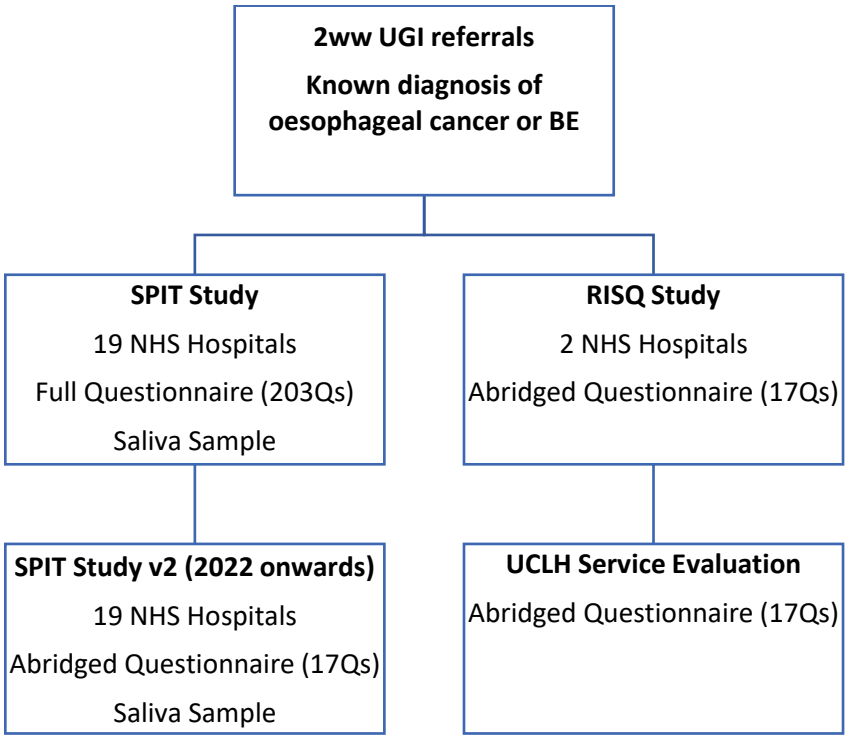


Figure 4-1: Schematic showing the inter-relationship between SPIT and RISQ Studies and the UCLH Service Evaluation which was incorporated into the RISQ Study during analysis

4.3.1 Saliva to Predict Risk of Disease Using Transcriptomics and Epigenetics (SPIT) Study

The Saliva to Predict Risk of Disease Using Transcriptomics and Epigenetics (SPIT) study was set up in 2017. This is an ongoing multi-centre cross-sectional study of 19 NHS hospitals across the UK. The overall aim of the study is to identify a cheap and non-invasive bedside test which could aid in predicting risk of serious oesophageal disease, namely oesophageal cancer.

The main inclusion criteria are:

- Patients who are referred under the 2-week wait (2ww) upper gastrointestinal pathway for investigation using gastroscopy
- Patients with known Barrett's oesophagus or oesophageal cancer
- Able to provide a saliva sample

The main exclusion criteria are:

- Patients who were unable to undergo definitive diagnostic tests such as gastroscopy
- Pregnancy
- Patients aged under 21

All patients provided informed consent prior to enrolment. Patients who were enrolled completed a symptom risk factor questionnaire and also provided a saliva sample for epigenetic analysis. The original questionnaire consisted of 209 questions across several sections as shown in Table 4-2.

Domain	**Maximum Number of Questions**
Personal Questions	8
Dental Health	9
Heartburn/Acid Reflux Symptoms	55
Swallowing Difficulties	8
Unexplained Weight Loss	4
Nausea and Vomiting	6
Smoking History	14
Alcohol Consumption	9
Obesity	4
Diet	19
Physical Exercise	6
Family History	12
Mental Wellbeing (based on the Hospital Anxiety and Depression Scale (HADS) questionnaire)[128]	14
Medical History	26
General Questions	15

Table 4-2: Question domains in the Saliva to Predict Risk of Disease Using Transcriptomics and Epigenetics (SPIT) Study

As a result of the work performed in Chapter 6: Creating a Risk Prediction Model, 6.3.3 Results: Selection of Features, the questionnaire was shortened to 17 questions in 2022 (SPIT Study v2).

Ethical approval for the study was obtained from Coventry and Warwickshire Research Ethics Committee (17/WM/0079).

4.3.2 Predicting Risk of Disease Using Detailed Questionnaires (RISQ) Study

The Predicting Risk of Disease Using Detailed Questionnaires (RISQ) study was established in 2019. It was initially designed to complement the SPIT study. This involved the participant completing a questionnaire but without collection of a saliva sample. This is an ongoing multi-centre cross-sectional study consisting of 2 NHS hospitals in the UK but with ongoing expansion to other sites. The overall aim of the study is to assess whether by using a brief questionnaire it would be possible to identify patients who would be at high risk of developing oesophageal cancer.

The main inclusion criteria are:

- Patients who are referred under the 2-week wait (2ww) upper gastrointestinal pathway for investigation using gastroscopy
- Patients with known Barrett's oesophagus or oesophageal cancer

The main exclusion criteria are:

- Pregnancy
- Patients aged under 18

Patients provided informed consent prior to enrolment. After enrolment, patients completed a symptom risk factor questionnaire of 17 questions. This is also the same questionnaire as version 2 of the SPIT questionnaire (SPIT Study v2).

Due to the onset of the COVID-19 pandemic, recruitment for the study was paused. As a direct consequence of the pandemic, many NHS trusts accumulated large backlogs of patients waiting for endoscopic procedures. To improve prioritisation of referrals, our team incorporated this questionnaire as a service evaluation pilot for 2ww suspected upper gastrointestinal cancer pathway patients at University College London Hospitals NHS Foundation Trust. Patients who were referred under this pathway were invited to complete the symptom risk factor questionnaire in advance of their appointment at the hospital. Completed questionnaires generated a risk score which was under evaluation to assess if it could be used for resource allocation purposes, with higher risk patients upgraded to undergo endoscopy sooner. For ease of analysis the two cohorts of directly recruited RISQ patients and service evaluation pilot patients were combined together. Table 4-3 shows the number of questions in each domain of the questionnaire.

Domain	Maximum Number of Questions
Personal Questions	2
Dental Health	
Heartburn/Acid Reflux Symptoms	4
Swallowing Difficulties	3
Unexplained Weight Loss	2
Nausea and Vomiting	
Smoking History	1
Diet	
Mental Wellbeing (based on the Hospital Anxiety and Depression Scale (HADS) questionnaire)[128]	
Medical History	1
General Questions	4

Table 4-3: Question domains in the Predicting Risk of Disease Using Detailed Questionnaires (RISQ) Study

Ethical approval for the study was obtained from the South Central – Oxford B Research Ethics Committee (19/SC/0382). Approval for the service evaluation pilot was given from the local hospital audit committee and the Caldicott Guardian, including for secondary research use.

Figure 4-1: shows the interrelationship between the SPIT and RISQ studies.

4.3.3 COvid aNd FacE maSkS (CONFESS) Study

CONFESS (COvid aNd FacE maSkS) was a cross-sectional study comprising of an online questionnaire to assess the effect of the COVID-19 pandemic on religious practice, both at an individual level and also to assess for changes within a communal worship setting. This was previously available at https://confess-study.co.uk, but the website since has been deactivated. Recruitment and questionnaire completion occurred between August and November 2020, with the last questionnaire completed on 5 November 2020. Participants were required to be at least 18 years old and be able to understand English. Participation was voluntary and respondents were recruited via a convenience sampling technique using word of mouth, targeted advertising through religious institutions, social media such as Facebook and WhatsApp, and publicity with the British Broadcasting Corporation (BBC),

which featured the study in the national news bulletin and a religious affairs programme.[129] In addition to the questionnaire, participants could also voluntarily consent to undergo further experiments at UCL Engineering to assess the effectiveness of face masks in reducing aerosol and droplet spread while singing. Questions were asked from the domains as shown in Table 4-4.

Domain	Maximum Number of Questions
Personal Details	14
COVID-19 Status	3
Personal Demographics	24
Religious and Spiritual Beliefs	4
Prayer and Community	13
Musical Styles in Worship	10
General Health	17
Facial Characteristics and Face Mask Use	17
Communal Worship Practices During COVID-19	38
Mental Health	21
Extra Personal Information	24
Mental Wellbeing (based on the Hospital Anxiety and Depression Scale (HADS) questionnaire)[128]	8
Additional Questions on Face Mask Use	18

Table 4-4: Question domains in the COvid aNd FacE maSkS (CONFESS) Study

The majority of questionnaire items were closed questions. However, open-ended questions regarding face mask acceptability and comfort as well as the impact of COVID-19-related restrictions on communal worship were also included in order to gain deeper insights. Due to an initial technical error, it was also possible to skip questions so not all participants answered every question, leading to different number of responses for questions within the questionnaire.

Ethical approval was obtained from University College London research ethics committee (14223/002). Participants completed informed consent online as the start of the questionnaire.

4.3.4 Data Curation Prior to Analysis

Both the SPIT and RISQ datasets were both bespoke datasets so there was minimal data curation required prior to analyses. Results could be downloaded in real time from https://oes.spitstudy.co.uk/sign_in for the SPIT Study and https://endoclassify.myrisq.co.uk/sign_in for the RISQ study. Some data cleaning was performed to correct data entry errors, such as units (e.g.: kg) included in numeric only fields.

4.4 Routine Health Datasets

4.4.1 University College London Hospitals NHS Foundation Trust (UCLH) Electronic Health Records

UCLH uses the Epic electronic healthcare record system, which was introduced in March 2019. I wanted to assess how gastroenterology services had been affected during the pandemic period. In particular, I was interested in how cancer diagnosis and treatment was affected. To do this, I decided to use several metrics in measuring hospital activity. These included:

1. Number of referrals for suspected upper and lower gastrointestinal cancers
2. Number of GI operations which involved removal of a portion of the digestive tract (oesophagectomy, gastrectomy, colectomy and small bowel resection)
3. Number of patients who had digestive cancers in their problem list, which was noted on or after January 2023. This was denoted by ICD-10 codes C15, C16, C17, C18, C19, C20, C21, C22, C23, C24, C25 and C26, which correspond to malignant neoplasms of digestive organs.[130]
4. Number of patients who underwent oesophagogastroduodenoscopy (OGD), colonoscopy, flexible sigmoidoscopy and endoscopic retrograde cholangio pancreatography (ERCP).
5. Number of patients who underwent CT colonography.

Data was provided between 1 January 2019 and 7 December 2020, which meant it covered the start of the 1st wave of the pandemic and the initial recovery phase.

Data Curation Prior to Analysis

Much of the data had already been curated by the inhouse data analytics team at UCLH before it was transferred to me, hence most of the data was provided in summary form, rather than in a raw format with patient details attached. While this meant less time was spent on data curation, it also meant patient level data was unavailable hence there were limitations in what could be analysed. In addition, during the study period, there was a major change in hospital record system in March 2019. It is possible that there could be some inconsistencies between how data was encoded before and after the changeover. As an example, colonoscopy was coded as "colonoscopy" under the old hospital system (Infoflex) but "fibreoptic colonoscopy" under the new system (Epic).

4.4.2 NHS Diagnostics Waiting Times and Activity Dataset (DM01)

The NHS Diagnostics Waiting Times and Activity Dataset (DM01) is openly published by NHS England on a monthly basis, with data available from January 2006 onwards.[131] The data also forms part of National Statistics. This collects monthly data on number of procedures performed and also the number of patients on the waiting list for 15 key diagnostic tests. This includes OGD, flexible sigmoidoscopy and colonoscopy. Patients on the waiting list are grouped by the number of weeks they have been waiting, in weekly increments, with the longest waiters grouped as '13+ weeks'. The number of procedures performed are further stratified by planned tests (such as a repeat endoscopy to check for ulcer healing or polypectomy site), unscheduled tests (such as inpatient procedures for acute upper gastrointestinal bleeding) and waiting lists tests (such as those who are referred to endoscopy after being seen in outpatient clinic). The data is subdivided by both local commissioning region (such as London Commissioning Region) and also the hospital trust

performing the procedure (such as University College London Hospitals NHS Foundation Trust).

Data Curation Prior to Analysis

Procedures from 7 English NHS Trusts were excluded due to incomplete data. These were:

- Bradford Teaching Hospitals NHS Foundation Trust
- East and North Hertfordshire NHS Trust
- North West Anglia NHS Foundation Trust
- Royal Berkshire NHS Foundation Trust
- St Helens and Knowsley Teaching Hospitals NHS Trust
- The Princess Alexandra Hospital NHS Trust
- University College London Hospitals NHS Foundation Trust

In addition, there were some NHS trusts that were merged or were renamed during the study period. For example, there were mergers of hospitals in Bristol and Weston-super-Mare so pre-merger and post-merger trusts were combined together in the final analysis to ensure consistency. Moreover, as I was most interested in the effect of endoscopy within acute NHS trusts, data available for private hospitals, public-private partnership diagnostic centres or community healthcare providers were excluded in analyses.

4.4.3 Barnet General Practice Dataset

In the UK, general practitioners (GP) are often the first port of call when a patient requires medical care, and they have a responsibility for looking after patients within the community. GP data thus contains a large body of information including diagnoses and test results, which could potentially cover a large number of patients. Through a collaboration with the North Central London Cancer Alliance, I was able to obtain curated General Practice (GP) data in Barnet. Barnet is an outer London borough with 395,000 residents. The English Indices of

Deprivation ranks Barnet at 184/317 in terms of deprivation nationally.[132] The borough is ethnically diverse, with 47% of residents from non-White backgrounds. 62% of residents are working age (18-64 years).[133]

The dataset was provided as part of a service evaluation project where I evaluated the use of the faecal immunochemical test (FIT) across GP practices within Barnet. This dataset contained anonymised data on 23,383 individual patients, with FIT results available between January 2019 and March 2022. Available data fields included:

- Sex
- Age
- FIT results
    - Numeric value
    - Date of numeric value
    - Associated free text comments
- Lower GI 2-week wait (2ww) referral and associated date
- Cancer diagnoses
    - All cancer diagnoses and associated date
    - Lower GI diagnoses and associated date
- Significant GI diagnoses and associated date; these included:
    - Crohn's disease
    - Ulcerative colitis
    - Diverticulosis
    - Polyp
    - Other GI specific investigations e.g. CT colonography or colonoscopy
- Blood test results – 1 year pre or post FIT result
    - Haemoglobin (Hb)

- Mean cell volume (MCV)
- C-reactive protein (CRP)
- Ferritin
- Platelet count

Data Curation Prior to Analysis

This dataset needed significant curation prior to use in statistical analysis. Firstly, blood test and FIT results were sometimes listed as part of a numeric field, and sometimes written as part of free text prose. As a result, I wrote code which could detect either of these to maximise the amount of data available. In addition, FIT results were sometime written qualitatively (e.g. positive/negative); these had to be disregarded. Furthermore, some results were listed as below the lower limit of detection (e.g.<4µg Hb/g). To allow for quantitative analyses there were recoded without the 'less than' or 'greater than' operators. Secondly, there was often duplication of dates of events, such as for dates of diagnosis of medical conditions or dates of referral to secondary care. This occurred when the same event was encoded twice by separate individuals, such as by the GP at the time of referral and also by administrators when the referral was actually sent onward to secondary care. To avoid double counting of the same event or result, if the same event or result was found to be within 30 days of each other this was treated as a duplicate and the later result was removed from final analyses.

4.5 Chapter Conclusions

In this chapter I have introduced the datasets that will be used in my thesis, which includes both bespoke and routine datasets. I have also outlined what data curation needed to be completed prior to analysis. In addition, I have given an overview of both the advantages and disadvantages of using both bespoke and routine datasets.

Chapter 5: Effect of COVID-19 on Gastrointestinal Services

Contents of this chapter have previously been presented and published as detailed below:

Publications

Ho KMA, Banerjee A, Lawler M et al. Predicting endoscopic activity recovery in England after COVID-19: a national analysis. *Lancet Gastroenterology and Hepatology*. 2021; 6: 381–90

N.B. the bulk of this work was completed during winter 2020 and spring 2021 hence some of the data at time of thesis submission may appear to be outdated

5.1 Chapter Introduction

This chapter will explore the COVID-19 pandemic and its effect on gastrointestinal services, both locally with University College London Hospitals NHS Foundation Trust and also nationally through openly available data, with a particular focus on endoscopy provision. In particular I will highlight the scale of the backlog caused by the COVID-19 pandemic and the challenge this poses.

5.2 Background

5.2.1 The COVID-19 Pandemic

The defining event of the 2020s was the emergence of a novel coronavirus from Wuhan, China in December 2019 and the subsequent rapid transmission of this virus around the world. The virus, named SARS-CoV-2, led to the development of multisystem disease known as coronavirus disease 2019 (COVID-19). March 2020 was a pivotal moment in the UK's response to the virus. Within a single month, the spread of COVID-19 meant life in the UK changed dramatically with ramifications on public health, society and the economy. At the beginning of March 2020, life in the UK remained more or less normal. However, by 13 March 2020 the World Health Organization (WHO) declared a global pandemic and Europe to be the epicentre.[134,135] By 23 March 2020, strict curbs on life were introduced as the UK entered a nationwide lockdown, which would only be relaxed some 3 months later.[136,137]

Mild symptoms of COVID-19 include fever, cough, fatigue, myalgia, anosmia and gastrointestinal symptoms.[138,139] However, severe COVID-19 typically occurred one week post infection, with rapid progression to respiratory failure and possible endotracheal intubation with ventilatory support.[139] In addition, multisystem organ dysfunction affecting the heart, kidneys, liver, coagulation and circulation may also occur.[139] Initial data from China suggested that COVID-19 led to mild infection in 81% of cases. However, 14% had severe disease with evidence of respiratory dysfunction and 5% were classed as critical, with

evidence of single or multiple organ failure.[140] Although there was an overall case fatality rate of 2.3%, this rose to 49% for those who were critically unwell.[140] The appearance of large numbers of patients who were simultaneously unwell within a limited geographic area and with high healthcare needs such as requiring admission to intensive care units and invasive ventilation meant standard capacity was quickly exceeded.[141] Intensive care units as well as medical inpatient capacity had to be expanded rapidly, with redeployment of healthcare staff and repurposing of clinical areas.[141] A further issue arose as healthcare personnel became infected. This was as high as 63% in Wuhan, with 14.8% of cases classified as severe or critical.[140] This meant stretched healthcare personnel became even more depleted, adding additional pressures in the delivery of patient care.

Although the WHO declared an end to the “global health emergency" on 5 May 2023, the COVID-19 pandemic has claimed 7 million lives (and counting), and estimates that the true death toll could be at least 20 million.[142] The WHO however still regard COVID-19 as an established and ongoing health matter, and its ramifications will continue for some years to come.[142]

5.3 Effect on Gastrointestinal Services Locally

5.3.1 Introduction

The rapid influx of patients with high medical needs led to wholesale reconfiguration of healthcare services around the world. Writing in the *Lancet Oncology*, Hamilton noted that as a result of the pandemic the UK’s National Health Service (NHS) reorganised from a comprehensive healthcare service to one that was solely focussed on treating COVID-19 patients.[143] This led to huge disruption in the delivery of routine care, with outpatient activity such as appointments and procedures cancelled so that clinical areas could be reconfigured to accommodate more inpatients, and medical and nursing staff redeployed to other clinical areas. Guidance issued by the British Society of Gastroenterology (BSG) in March 2020

supported such efforts and recommended a 6-week pause in endoscopic activity, such that all but emergency and essential endoscopy stopped, including bowel cancer screening, allowing time to refine triage systems.[144] Early data from the UK National Endoscopy Database (NED), encompassing 79% of endoscopic units in the UK, demonstrated that activity fell to as low as 5% of normal levels in the COVID-19 impacted period from 23 March to 31 May 2020.[145] Moreover, one important observation was that during the COVID-19 impacted period, while the per-procedure cancer detection rate increased from a baseline of 1.9% to 6.6% ($p<0.001$), the weekly number of cancers detected decreased by 58%.[145] This therefore suggests that while there was increased vetting of procedures, which increased the per procedure cancer detection rate, a large number of procedures were simply not being performed, leading to significant concerns over missed or delayed diagnoses, and consequently potentially worse outcomes for patients. Furthermore, early in the COVID-19 pandemic, the general population faced lockdown restrictions and had restrictions placed on their livelihoods.[137] This led to altered health behaviour, such as reduced emergency department attendances and delayed presentation of patients with symptoms.[7] Finally, the transmissibility of the COVID-19 virus also meant that staff became sick with the virus, requiring self-isolation and hence impacting on the delivery of clinical services.

5.3.2 Effect at a Local Level - University College London Hospitals NHS Foundation Trust (UCLH)

5.3.3 Introduction

I was interested to see how the COVID-19 pandemic affected gastrointestinal services at our affiliated hospital trust, University College London Hospitals NHS Foundation Trust (UCLH). UCLH is an acute NHS trust providing a range of emergency and scheduled care services located in Central London. Services are provided across 10 sites, with University College Hospital (UCH) being the main acute hub and provider of gastrointestinal services. It is also

a major hub for cancer treatment and includes services for tertiary care such as upper gastrointestinal surgery within the North Central London region. Like many hospitals nationally and internationally, there were significant changes including workforce redeployment and recommissioning of clinical areas to tackle the challenges of the pandemic. While in the early phase of the pandemic the clinical priority was to care for patients who were infected with COVID-19, there were significant concerns on the collateral damage the pandemic would cause, in particular relating to disruptions to cancer pathways and delayed cancer diagnoses.[119–121]

5.3.4 Aims

My primary aim was to analyse how the COVID-19 pandemic affected gastrointestinal clinical services at UCLH. I would achieve this by looking at different aspects of the service, such as referrals to the service, number of diagnostic procedures performed and the number of curative surgical operations performed. In addition, I aimed to perform subgroup analyses with only suspected or confirmed cancer patients included to assess the impact of the pandemic on this cohort.

5.3.5 Methods

Data was provided by the UCLH data analytics department for the purpose of internal audit and service evaluation and was available from January 2019 to December 2020. However, as data was only reliable from March 2019 onwards due to IT system migration, I excluded data from January to February 2019 for certain analyses. The database has been introduced previously in Chapter 4:Datasets and Data Curation. In brief, the database contains patient level data on several key metrics for gastrointestinal services activity. The database included:

- Number of referrals under suspected cancer pathways for both upper gastrointestinal and lower gastrointestinal cancer

- Number of new gastrointestinal cancer diagnoses made (e.g. gastric cancer and colorectal cancer).
- Number of gastrointestinal surgical resections performed (e.g. oesophagectomy and colectomy)
- Number of gastrointestinal endoscopic procedures performed (e.g. oesophagogastroduodenoscopy (OGD) and colonoscopy)
- Number of computerised tomography colonography (CTC) scans performed (N.B. CTC is also known as virtual colonoscopy).

Data cleaning was kindly performed by the data analytics team hence potential duplicated records had already been removed. The provided dataset contained patient level data, with information on which procedure or operation was performed, indication and also the date of procedure or operation, except for CTC data which was provided on an aggregated basis and stratified by referrer.

Data was provided in a Microsoft Excel spreadsheet. Data analysis and generation of statistical plots were completed in R version 4.2.2.[93]

5.3.6 Results

Number of Cancer Referrals

During the study period there were 6,740 upper and lower GI referrals made. However, 278 were duplicated and hence were excluded in further analyses, meaning there were 6,462 distinct referrals. This was split between 1,829 upper GI referrals (28%) and 4,633 lower GI referrals (72%) (Figure 5-1 and Table 5-1).

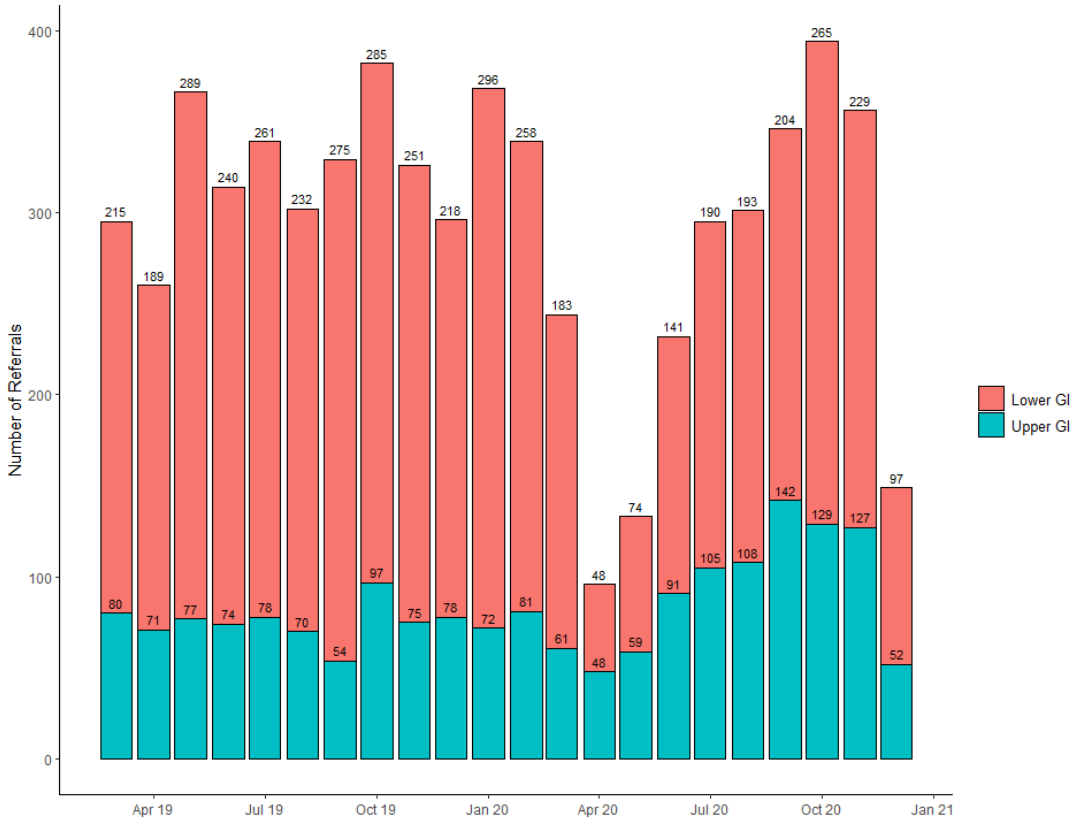


Figure 5-1: Number of referrals for suspected upper and lower gastrointestinal cancers between March 2019 and December 2020.

Month	Lower GI		Upper GI		Total	
	No. of Referrals	% Compared to 2019	No. of Referrals	% Compared to 2019	No. of Referrals	% Compared to 2019
Mar-19	215		80		295	
Apr-19	189		71		260	
May-19	289		77		366	
Jun-19	240		74		314	
Jul-19	261		78		339	
Aug-19	232		70		302	
Sep-19	275		54		329	
Oct-19	285		97		382	
Nov-19	251		75		326	
Dec-19	218		78		296	
Jan-20	296		72		368	
Feb-20	258		81		339	
Mar-20	183	85	61	76	244	83
Apr-20	48	25	48	68	96	37
May-20	74	26	59	77	133	36
Jun-20	141	59	91	123	232	74
Jul-20	190	73	105	135	295	87
Aug-20	193	83	108	154	301	100
Sep-20	204	74	142	263	346	105
Oct-20	265	93	129	133	394	103
Nov-20	229	91	127	169	356	109
Dec-20	97	44	52	67	149	50

Table 5-1: Number of referrals for suspected upper and lower gastrointestinal cancers between March 2019 and December 2020, with comparison to same month in 2019.

There was a decrease in the number of GI cancer referrals during the first wave of the pandemic, most evident in April and May 2020. In April 2020 there were 96 referrals in total, compared to 260 or 37% of the level in April 2019. However, lower GI referrals were more affected compared to upper GI referrals. For lower GI this was at 25% (48/189) of April 2019 levels, compared to 66% (48/71) of April 2019 levels for upper GI. This was statistically significant (X^2=16.2, df=1, p<0.001). A similar trend was also observed for May 2020 compared to May 2019. For upper GI this was at 77% (59/77) compared to 25% for lower GI (74/289). Similarly, this was statistically significant (X^2=26.8, df=1, p<0.001). Interestingly, by June 2020 upper GI referrals had exceeded the numbers seen in June 2019 (91 referrals

compared to 74 referrals, 123%). However, it would take until October 2020 for lower GI referrals to recover to >90% of 2019 levels (265 referrals compared to 285 referrals, 93%).

5.3.7 Number of Cancer Diagnoses

For this analysis I focussed purely on digestive system cancers. I grouped cancers into 4 main groups: colorectal and anal, hepato-pancreato-biliary (HPB) cancers, including hepatocellular carcinoma and cholangiocarcinoma, oesophago-gastric cancers and other gastrointestinal cancers, such as gastrointestinal stromal tumours (GIST) and where the actual site of the cancer was unclear, such as the term “bowel cancer”.

During the study period I identified 1,908 cancers diagnosed in 1,774 patients. However, I identified 69 duplicate diagnoses, hence there were 1,839 cancer diagnoses in 1,774 patients during analysis (Figure 5-2 and Table 5-2).

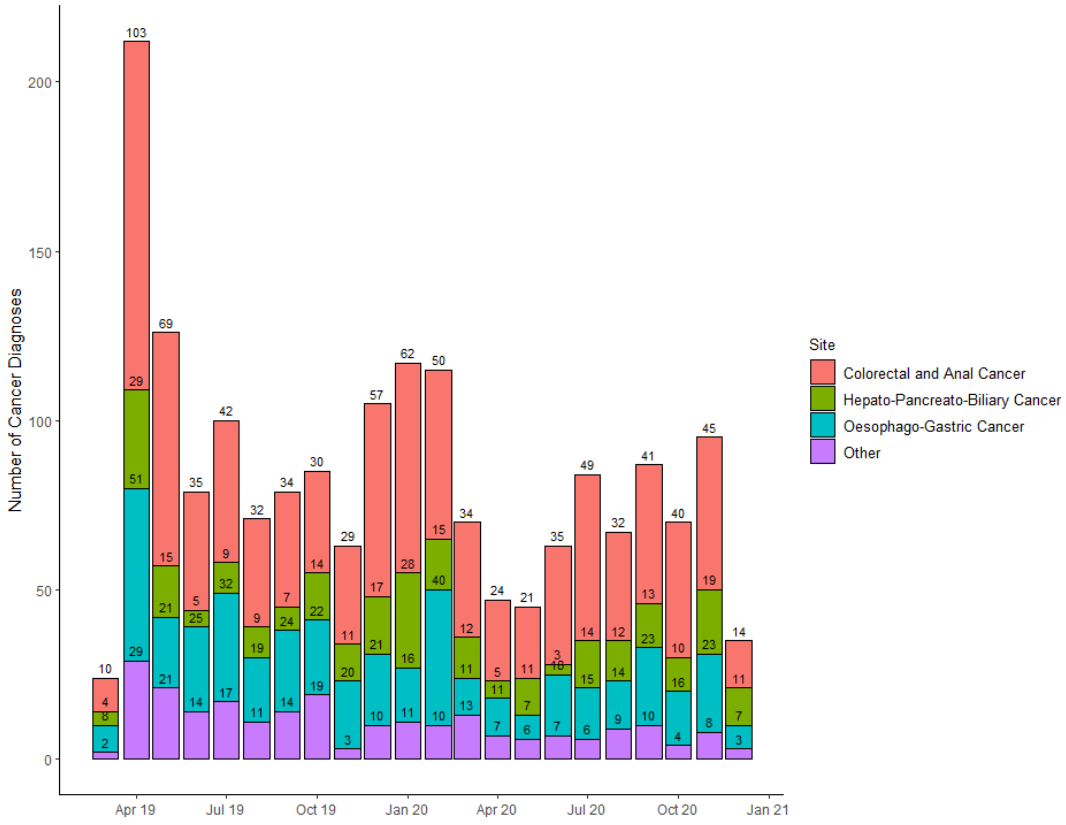


Figure 5-2: Number of cancer diagnoses and their anatomical site between March 2019 and December 2020.

Month	Colorectal		Hepato-Pancreato-Biliary		Oesophageal		Other		Total	
	No. of Referrals	% Compared to 2019	No. of Referrals	% Compared to 2019	No. of Referrals	% Compared to 2019	No. of Referrals	% Compared to 2019	No. of Referrals	% Compared to 2019
Mar-19	10		4		8		2		24	
Apr-19	103		29		51		29		212	
May-19	69		15		21		21		126	
Jun-19	35		5		25		14		79	
Jul-19	42		9		32		17		100	
Aug-19	32		9		19		11		71	
Sep-19	34		7		24		14		79	
Oct-19	30		14		22		19		85	
Nov-19	29		11		20		3		63	
Dec-19	57		17		21		10		105	
Jan-20	62		28		16		11		117	
Feb-20	50		15		40		10		115	
Mar-20	34	340	12	300	11	138	13	650	70	292
Apr-20	24	23	5	17	11	22	7	24	47	22
May-20	21	30	11	73	7	33	6	29	45	36
Jun-20	35	100	3	60	18	72	7	50	63	80
Jul-20	49	117	14	156	15	47	6	35	84	84
Aug-20	32	100	12	133	14	74	9	82	67	94
Sep-20	41	121	13	186	23	96	10	71	87	110
Oct-20	40	133	10	71	16	73	4	21	70	82
Nov-20	45	155	19	173	23	115	8	267	95	151
Dec-20	14	25	11	65	7	33	3	30	35	33

Table 5-2: Number of cancer diagnoses and their anatomical site between March 2019 and December 2020, with percentage compared to 2019. N/A = not available.

The highest overall number of cancers diagnosed was in April 2019, with a total of 212 cancers. However, this figure may be erroneous as this was the first complete month after migration to Epic electronic health record software. This is likely to have led to some patients being coded for cancer as if they were new diagnoses, when the condition may have already been pre-existing in the patient. However, there was still a decrease in the number of cancer diagnoses in April 2020 (total of 47 new diagnoses) and May 2020 (total of 45 new diagnoses), compared with the preceding months of February 2020 (total of 115 new diagnoses) and March 2020 (total of 70 new diagnoses). By July 2020, cancer diagnoses had reached 84% (84/100) of July 2019 levels. There was a statistically significant difference between the proportions of the subtypes of gastrointestinal cancer cases being diagnosed (X^2=11.7, df=3, p-value=0.008), with primarily an increase of colorectal cancers (58% (49/84) of cases in July 2020 compared to 42% (42/100) in July 2019) and a decrease of oesophago-gastric cancers (18% (15/84) of cases in July 2020 compared to 32% (32/100) in July 2019). By August 2020 overall cancer diagnoses had reached 94% (67/71) of August 2019 levels.

5.3.8 Gastrointestinal Surgical Procedures

To understand the impact of the pandemic on surgical procedures for cancer cases I have split the analysis into looking at gastrointestinal surgical procedures for all indications and for cancer only. I only included procedures involving the luminal digestive tract such as excision of bowel or manipulation of stomas. Procedures involving the hepato-pancreato-biliary tract such as cholecystectomy were excluded. Other abdominal procedures such as inguinal hernia repair which do not involvement operating on the luminal digestive tract were also excluded.

Procedures were grouped by indication such as for acute abdomen (i.e. non-elective, urgent cases such as acute bowel obstruction), cancer, inflammatory bowel disease (IBD), or other benign indications such as diverticular disease or fistulae.

Finally, as the raw numbers for certain indications per month were low, I elected to analyse in 3-month blocks (i.e. quarters of the year).

5.3.9 All Indications

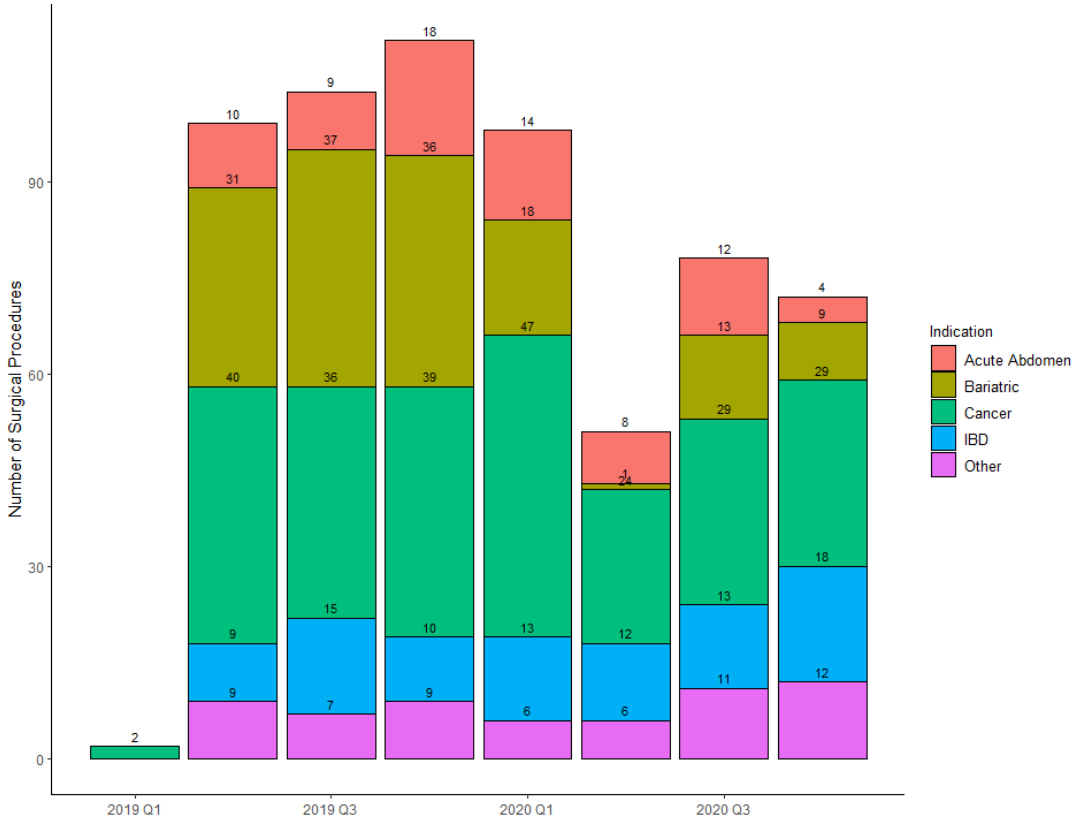


Figure 5-3: Number of surgical procedures and their indications between March 2019 and December 2020, stratified by quarter and year.

IBD = inflammatory bowel disease

616 surgical procedures were performed on 606 individual patients during the study period (Figure 5-3 and Table 5-3). The highest number of surgical procedures was performed in the 4th quarter of 2019 (112 cases). Similar to previous trends, there was a decrease in the overall number of procedures in the 2nd quarter of 2020, at 52% of 2019 levels. Between April 2020 to June 2020 only a single bariatric case was carried out, compared with 31 cases between April 2019 to June 2019. Recovery continued such that in the 3rd quarter of 2020 there were 78 surgical procedures carried out, at 75% of the level in 2019 (104 surgical procedures).

Statistical testing using Wilcoxon Signed Rank test with Bonferroni correction for each indication was not significant.

	Acute Abdomen		Bariatric		Cancer		IBD		Other		Total	
Quarter	No. of Referrals	% Comp to 2019	No. of Referrals	% Comp to 2019	No. of Referrals	% Comp to 2019	No. of Referrals	% Comp to 2019	No. of Referrals	% Comp to 2019	No. of Referrals	% Comp to 2019
Q1-2019	0		0		2		0		0		2	
Q2-2019	10		31		40		9		9		99	
Q3-2019	9		37		36		15		7		104	
Q4-2019	18		36		39		10		9		112	
Q1-2020	14		18		47		13		6		98	
Q2-2020	8	80	1	3	24	60	12	133	6	67	51	52
Q3-2020	12	133	13	35	29	81	13	87	11	157	78	75
Q4-2020	4	22	9	25	29	74	18	180	12	133	72	64
Wilcoxon p-value	Q2-Q4 2019 v Q2-Q4 2020 0.400		Q2-Q4 2019 v Q2-Q4 2020 0.100		Q2-Q4 2019 v Q2-Q4 2020 0.077		Q2-Q4 2019 v Q2-Q4 2020 0.400		Q2-Q4 2019 v Q2-Q4 2020 0.658			

Table 5-3: Number of surgical procedures and their indications between March 2019 and December 2020, stratified by quarter and year.

IBD = inflammatory bowel disease

5.3.10 Cancer Only

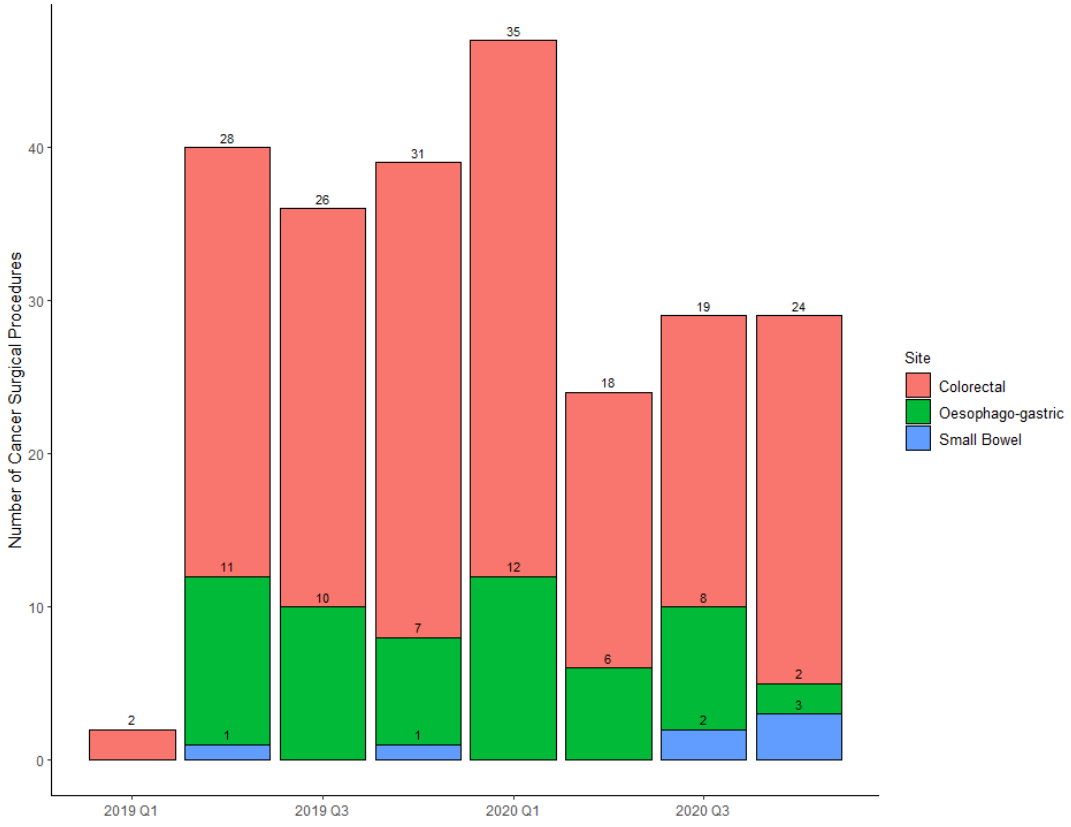


Figure 5-4: Number of cancer related surgical procedures performed and the related anatomical site between March 2019 and December 2020, stratified by quarter and year.

Figure 5-4 and Table 5-4 demonstrate the number of cancer surgical procedures that were performed between March 2019 and December 2020. The highest number of procedures performed was in the 1st quarter of 2020 with 47 procedures. In the 2nd quarter of 2020, 24 procedures were performed, 60% (24/40) of the level in the same quarter in 2019. By the 3rd and 4th quarters of 2020, procedure numbers had increased to 29 in both quarters, although this was still below 2019 levels (81% (29/36) in the 3rd quarter, 74% (29/39) in the 4th quarter). Statistical testing using Wilcoxon Signed Rank test with Bonferroni correction was not significant for each site of gastrointestinal surgery.

	Colorectal		Oesophago-Gastric		Small Bowel		Total	
Quarter	No. of Referrals	% Compared to 2019	No. of Referrals	% Compared to 2019	No. of Referrals	% Compared to 2019	No. of Referrals	% Compared to 2019
Q1-2019	2		0		0		2	
Q2-2019	28		11		1		40	
Q3-2019	26		10		0		36	
Q4-2019	31		7		1		39	
Q1-2020	35		12		0		47	
Q2-2020	18	64	6	80	0	0	24	67
Q3-2020	19	73	8	29	2	NA	29	70
Q4-2020	24	77	2	33	3	300	29	70
Wilcoxon p-value	Q2-Q4 2019 v Q2-Q4 2020 0.100		Q2-Q4 2019 v Q2-Q4 2020 0.200		Q2-Q4 2019 v Q2-Q4 2020 0.221			

Table 5-4: Number of cancer related surgical procedures performed and the related anatomical site between March 2019 and December 2020, stratified by quarter and year.

Endoscopy Procedures

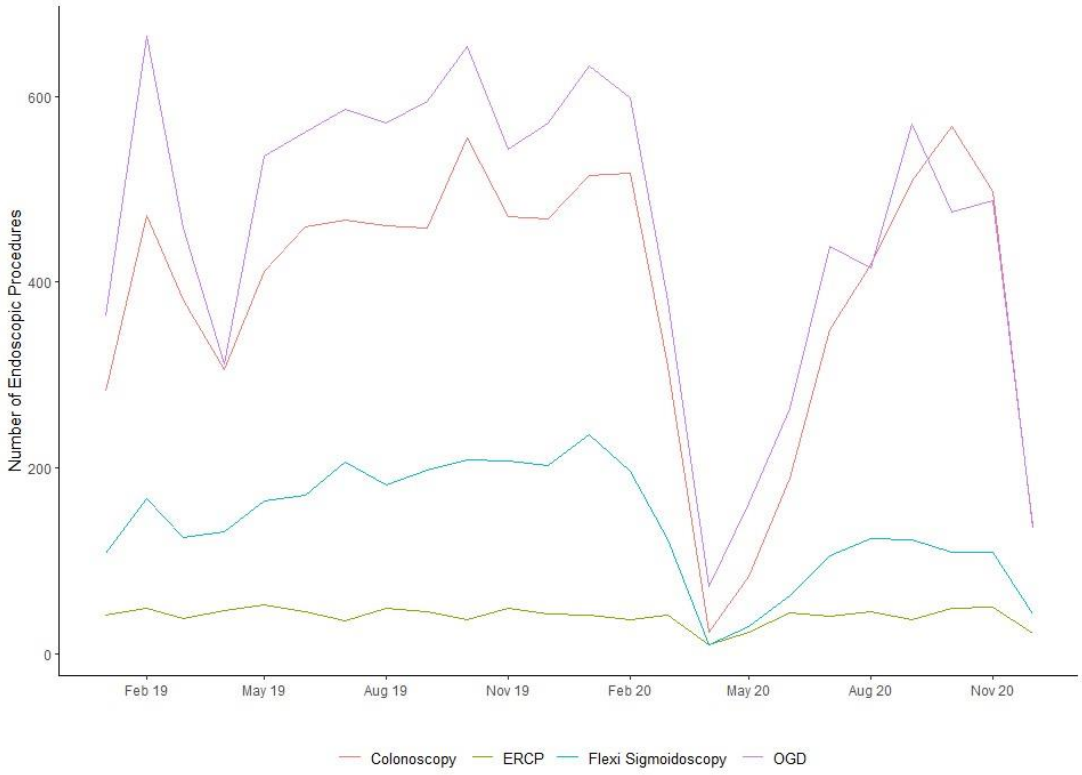


Figure 5-5: Number of endoscopic procedures performed for colonoscopy, endoscopic retrograde cholangio pancreatography (ERCP), flexible sigmoidoscopy and oesophagogastroduodenoscopy (OGD) between January 2019 and December 2020.

There were notable decreases in the number of endoscopic procedures with the arrival of the COVID-19 pandemic (Table 5-5 and Figure 5-5). The most impacted month was April 2020, where compared to the levels of April 2019, only 8% (23/306) of colonoscopies, 22% (10/46) of ERCPs, 9% (9/131) of flexible sigmoidoscopies and 23% (72/312) of OGDs were performed. However, recovery to at least 90% of the previous year's capacity was fastest for ERCPs, which was reached in June 2020 (98%, 44/45), followed by colonoscopy in August 2020 (91%, 419/460) and OGD in September 2020 (96%, 570/594). Flexible sigmoidoscopy capacity never reached this threshold in the study period, suggesting of the endoscopic modalities studied it was the slowest to recover.

	Colonoscopy		ERCP		Flexible Sigmoidoscopy		OGD		Total	
Month	No. of Procedures	% Compared to 2019	No. of Procedures	% Compared to 2019	No. of Procedures	% Compared to 2019	No. of Procedures	% Compared to 2019	No. of Procedures	% Compared to 2019
Jan-19	282		42		108		363		795	
Feb-19	472		49		167		664		1352	
Mar-19	381		38		125		458		1002	
Apr-19	306		46		131		312		795	
May-19	412		53		164		536		1165	
Jun-19	459		45		170		561		1235	
Jul-19	467		35		206		586		1294	
Aug-19	460		49		182		571		1262	
Sep-19	458		45		197		594		1294	
Oct-19	555		36		209		653		1453	
Nov-19	471		49		208		543		1271	
Dec-19	468		43		203		571		1285	
Jan-20	515	183	41	98	236	219	632	174	1424	179
Feb-20	517	110	37	76	196	117	598	90	1348	100
Mar-20	309	81	41	108	123	98	379	83	852	85
Apr-20	23	8	10	22	9	7	72	23	114	14
May-20	83	20	23	43	29	18	162	30	297	25
Jun-20	189	41	44	98	62	36	264	47	559	45
Jul-20	349	75	40	114	105	51	438	75	932	72
Aug-20	419	91	45	92	124	68	415	73	1003	79
Sep-20	509	111	36	80	123	62	570	96	1238	96
Oct-20	568	102	49	136	109	52	475	73	1201	83
Nov-20	497	106	50	102	109	52	488	90	1144	90
Dec-20	135	29	22	51	43	21	137	24	337	26

Table 5-5: Number of endoscopic procedures performed for colonoscopy, endoscopic retrograde cholangio pancreatography (ERCP), flexible sigmoidoscopy and oesophagogastroduodenoscopy (OGD) between January 2019 and December 2020.

5.3.11 Bowel Cancer Screening Colonoscopy and Flexible Sigmoidoscopy

Some endoscopic procedures in the study were performed as part of population level screening for colorectal cancer. Flexible sigmoidoscopy was offered as part of Bowel Scope, a now discontinued programme where patients aged 55 were offered a one off flexible sigmoidoscopy.[56] As a result of the pandemic, introduction of the faecal immunochemical test (FIT) and also challenges with endoscopic workforce, this was discontinued in early 2021. In addition, the UK also has a national Bowel Cancer Screening Programme (BCSP) which invites individuals aged between 60-74 years to have a FIT every 2 years.[54] Individuals who have a FIT of ≥120 µg Hb/g are invited to undergo a colonoscopy on a specific bowel cancer screening list.[146]

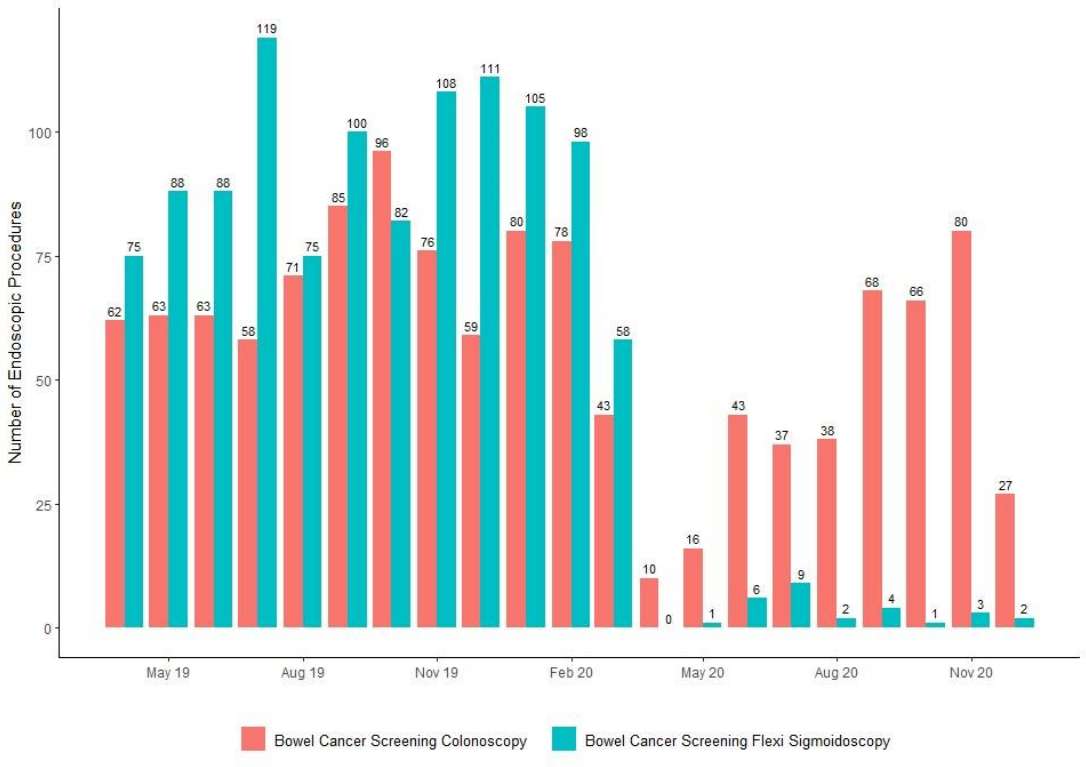


Figure 5-6: Number of Bowel Cancer Screening Colonoscopies and Flexible Sigmoidoscopies performed between April 2019 and December 2020.

Figure 5-6 demonstrates the number of endoscopic screening procedures which were performed between April 2019 to December 2020. Data from January to March 2019 was

not available. Notably, there was a significant decrease in the number of bowel cancer screening flexible sigmoidoscopies being performed after April 2020, with at most 9 procedures performed in July 2020, compared with 119 (7.6%) of July 2019 levels. In comparison, recovery of bowel cancer screening colonoscopy was quicker; in April 2020 only 10 procedures were performed, compared with 62 (16%) in April 2019. By September 2020, number of procedures was at 80% when compared to a year ago (68 in September 2020 compared to 85 in September 2019). By November 2020, cases performed exceeded numbers from the year previously (80 in November 2020 compared with 76 in November 2019; 105%). Fisher's Exact Test demonstrated a statistically significant difference in the distribution of colonoscopies and flexible sigmoidoscopies during the study period (p<0.001).

5.3.12 CT Colonography

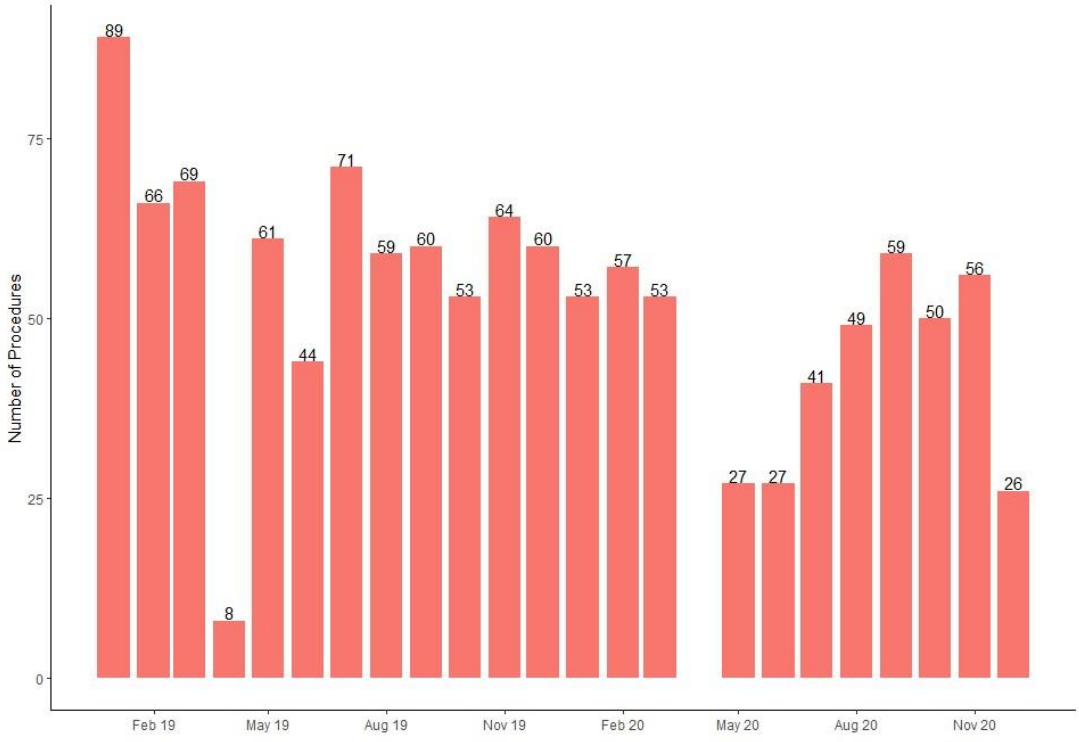


Figure 5-7: Number of CT colonographies (CTC) performed between January 2019 and December 2020.

Figure 5-7 demonstrates the number of CT colonographies (CTC) performed between January 2019 and December 2020. While not even a single CTC was performed in April 2020, 27 procedures were completed in May 2020, compared to 61 in May 2019 (44%). By September 2020 this had almost reached pre-pandemic levels (59 in September 2020 compared to 60 in September 2019; 98%).

5.3.13 Discussion

In this local analysis I have used five different metrics spanning several different parts of the hospital operation as a snapshot to assess how the COVID-19 pandemic has affected gastroenterology services at University College London Hospitals NHS Foundation Trust. COVID-19 led to decreases in number of referrals received, number of gastrointestinal cancer diagnoses made, number of surgical operations performed, number of endoscopies performed and number of CT colonographies performed. The decrease was most apparent in the month of April 2020 in all analyses performed, with some degree of recovery experienced in the months following this.

While I noted a drop in GI referrals during the early phase of the first wave of the pandemic, I demonstrated that the effect on upper GI referrals was less than for lower GI referrals. By June 2020, upper GI referrals had exceeded pre-pandemic levels, whereas it took until October 2020 for lower GI referrals to reach >90% of pre-pandemic levels. Similar trends have also been seen nationally in another NHS trust.[147] One of the difficulties in the diagnosis of lower GI cancer is that alarm symptoms are very non-specific.[148] In addition, there are other factors, including socioeconomic ones which lead to delayed presentation to health services for lower GI cancer. To improve awareness, public health campaigns, such as “Be Clear on Cancer” have run in England, targeting higher risk groups with some success.[149]

Perhaps the most worrying trend was that there were decreases noted in the number of cancers diagnoses and also in diagnostic procedure numbers, such as for endoscopy and CT

colonography, which are key investigations for gastrointestinal cancers. Overall cancer diagnoses reached >90% of the previous year's level in August 2020, while by September 2020 OGD, colonoscopy and CTC had reached >90% of the previous year's procedure numbers. This means that a backlog of procedures and delayed diagnoses is likely to have built up, potentially leading to worse patient outcomes.[120,121] While CTC was touted as an alternative to colonoscopy in the recovery phase of the pandemic, approximately 8 times as many colonoscopies are performed per month compared to CTC, and it is unlikely there is adequate capacity to significantly increase the number the CTCs currently being performed.[150] Indeed, the number of CTC procedures performed after March 2020 never exceeded the number before March 2020, suggesting there may have already been capacity constraints.

The pandemic also had an effect on bowel cancer screening programmes. Most notably, bowel cancer screening flexible sigmoidoscopy (BowelScope) did not recover to pre-pandemic levels by the end of the study period, coinciding with the national decision to withdrawal of this service. However, for bowel cancer screening colonoscopy the numbers of procedures performed had recovered to pre-pandemic levels by November 2020.

5.3.14 Limitations of Study

The major limitation of the study was related to data quality. There was a significant change in the data collection method in the month of March 2019, when UCLH NHS Foundation Trust changed from one electronic healthcare record system to another. Subsequently, this led to changes in coding of procedures and diagnoses such as cancer and this may have led to procedures or diagnoses being miscounted or not linked up appropriately. In addition, it was unclear how the migration process of data worked from moving between the old and the new system. As an example, the data for cancer diagnoses peaked in April 2019, at the time of the system migration. This could suggest data was either migrated en bloc, leading to a

spike of diagnoses or there was potential double counting where diagnoses before and after healthcare system migration may have been counted twice. Furthermore, there were possible issues with data quality and coding: as an example, 8.4% (156/1847) of cases were coded as “bowel cancer” and hence it was not possible to be more precise with the anatomical location of the patient’s cancer. Similarly, there may have been differences in coding before and after system migration, and as a consequence some data from January to February 2019 had to be deliberately excluded from my analyses due to concerns of reliability.

In addition, during the pandemic, NHS England made a national contract with private healthcare providers to provide care to patients requiring COVID-19 care, as well as for urgent diagnostic tests and surgical operations.[151] This means that a number of diagnostic tests and operations may be missing in the data presented as it would have been contracted out to external services. It is unlikely that these contracted out procedures would have been retrospectively added back to hospital record systems and so be included in overall counts.

Finally, the raw data used within these analyses lacked granularity. As an example for cancer, no data was available for stage of cancer at diagnosis. This data would have been especially useful to assess if the pandemic led to later presentations of cancers. Moreover, data on key events along the cancer pathway such as date of referral, date of investigation, date of diagnosis and date of treatment would have helped to build a more complete picture of potential delays affecting cancer diagnosis and care.

5.3.15 Conclusion

I have been able to use gastroenterology as a case example to demonstrate how severely COVID-19 affected clinical services, especially within the early phase of the pandemic. Similar trends are likely to have been replicated across different specialities within the hospital trust and indeed other healthcare providers both in the UK and worldwide.

5.4 Effect on Gastrointestinal Services Nationally

5.4.1 Introduction

Although in the previous section I was able to demonstrate the effect of the COVID-19 pandemic on local gastroenterology services, this effect would have been amplified nationally and internationally. Perhaps the most significant impact was on endoscopic services. This is because endoscopy has a prominent role in the diagnosis of gastrointestinal cancers and are often the first line investigation. Delayed diagnosis for cancer is a particular problem in the UK, where it lags behind other developed nations such as Australia and Canada in survival.[152] The key role for endoscopy was recognised by the BSG in late April as it recommended restarting endoscopy services safely.[153] The disruption in endoscopic services led to a reduction in the number of patients on colorectal cancer pathways and significant decreases in the number of cancers detected at endoscopy.[145,154] Extrapolating from a single centre, it was postulated that a 1 year restriction on endoscopic services could lead to 28,800 undiagnosed gastrointestinal cancers and a 1.46 million endoscopic procedural backlog.[155]

5.4.2 Aims

The aims were to:

- Investigate the impact of the COVID-19 pandemic on endoscopy services in England and calculate an estimate for the nationwide backlog of procedures.
- Present strategies to clear the potential endoscopy backlog related to the COVID-19 pandemic, including a temporary increase in capacity.
- Estimate the effect of a further reduction in endoscopy capacity on the overall backlog secondary to a further pandemic.

5.4.3 Methods

I analysed data from January 2018 to October 2020 from NHS England's Monthly Diagnostic Waiting Times and Activity data.[131] Further information on the dataset could be found in chapter 4.4.2: NHS Diagnostics Waiting Times and Activity Dataset (DM01). In brief, data on the number of colonoscopies, flexible sigmoidoscopies, gastroscopies and total endoscopic procedures performed in 126 English NHS trusts was available. This was further stratified by urgency: planned/surveillance, unplanned (which includes inpatient and emergencies) and waiting list (either GP or hospital referral). Notably, this dataset excludes colonoscopies performed under the Bowel Cancer Screening Programme (BCSP) and flexible sigmoidoscopies performed under BowelScope. Figure 5-8 shows a schematic of the analyses performed.

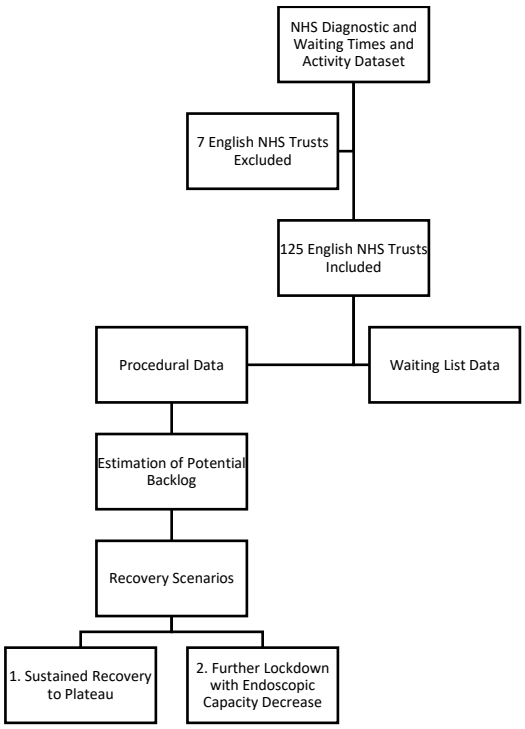


Figure 5-8: Schematic of data analysis performed.

Procedures performed in seven NHS trusts were excluded due to incomplete monthly data (see chapter 4.4.2: NHS Diagnostics Waiting Times and Activity Dataset (DM01)). Several NHS trusts and NHS commissioning regions merged during the study period; pre-merger trusts

and regions were aligned with their post-merger counterparts in the final analysis to ensure consistency. The primary outcomes captured were the change in number of endoscopic procedures compared to the same month in 2019 and an estimate of the backlog of procedures associated with the pandemic. Secondary outcomes were to estimate the effect of increasing capacity and a temporary reduction of capacity due to a further lockdown.

Note that the bulk of this work was performed at the end of 2020, hence data presented here may appear outdated.

Statistical Analyses

Statistical analyses were performed using R software version 4.0.2.[93] Data was analysed on a per month basis and I used the chi squared test to compare between the same month in 2019 and 2020 and Kruskal-Wallis test to compare the overall number of procedures performed, the case mix for each individual procedure and also the number of procedures performed in each region. A p-value of ≤0.05 was taken as significant, with Dunn's test for multiple comparisons.

For predicting the potential endoscopic backlog, I also created two simple linear regression models; the first model employed data from January 2018 to January 2020 to calculate the number of expected procedures in the absence of the pandemic, based on historical demand (Table 5-6). The second model used data from April 2020 to October 2020 to estimate the recovery of endoscopic capacity (Table 5-7). I calculated the potential backlog of procedures by cumulating the difference between the two models I also altered parameters in the second model to simulate different scenarios, including a sustained recovery to plateau at different levels of capacity and a further reduction in endoscopic activity, I assumed these scenarios occurred in December 2020 to ensure consistency for a single timepoint.

	Estimated Effect	Standard Error of Estimated Effect	Test statistic (t-test)	p-value
Intercept	38240.3	2071.3	18.46	$<2.00 \times 10^{-16}$
Number of months from January 2018	103.7	58.1	1.79	0.076

Table 5-6: Linear regression model 1, estimating endoscopic demand in absence of COVID-19.

	Estimated Effect	Standard Error of Estimated Effect	Test statistic (t-test)	p-value
Intercept	403082	84776	-4.76	3.76×10^{-4}
Number of months from April 2020	6340	1284	4.94	2.72×10^{-4}

Table 5-7: Linear regression mode 2, estimating endoscopic capacity recovery.

5.4.4 Results

Endoscopy Procedures

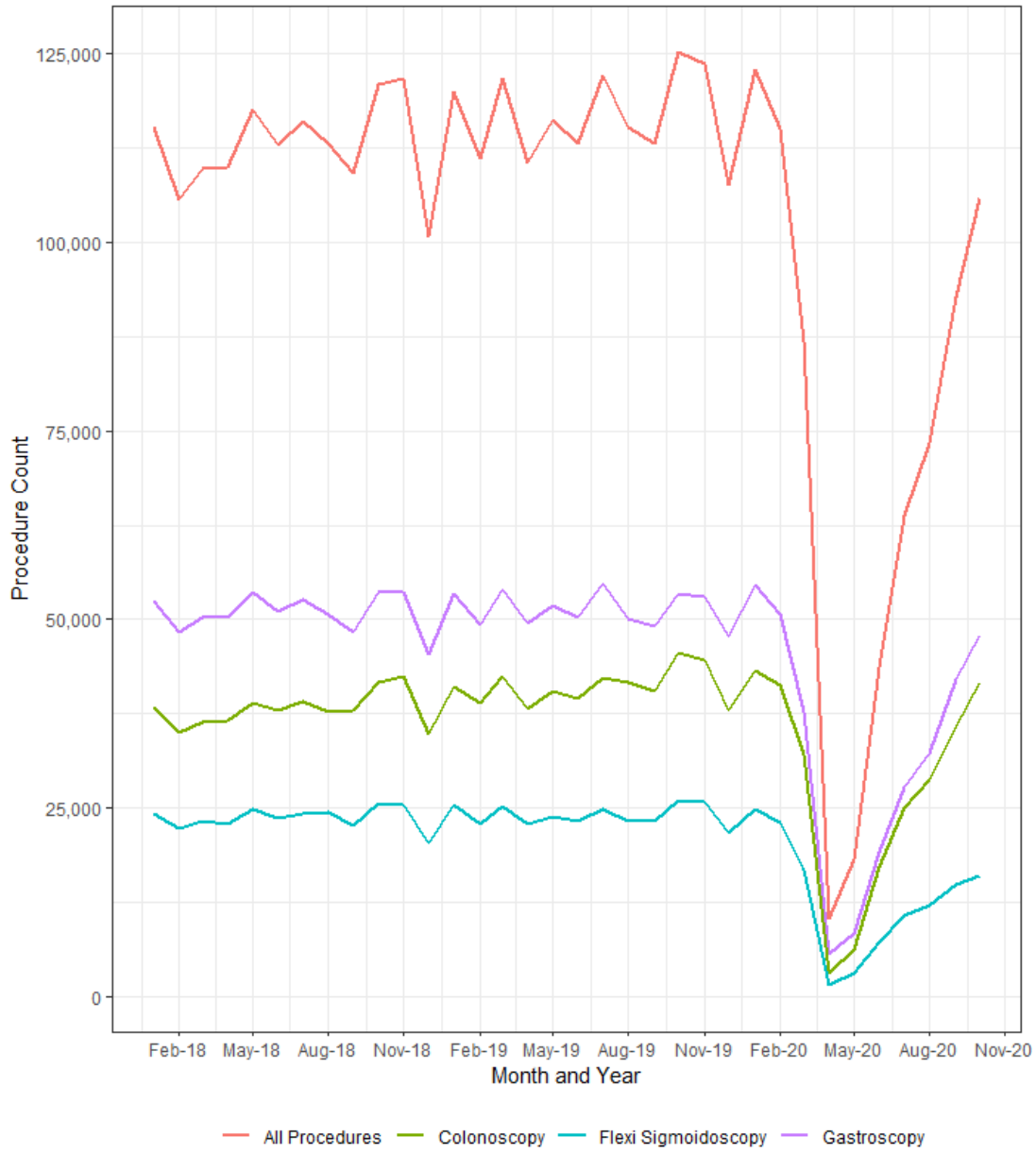


Figure 5-9: Trends in number of endoscopy procedures from January 2018 to October 2020.

Colonoscopy								
2020	**Planned/Surveillance Procedures**		**Unscheduled Procedures**		**Waiting List Procedures**		**Total Procedures**	
	n(%)	**As % of 2019**	**n(%)**	**As % of 2019**	**n(%)**	**As % of 2019**	**n**	**As % of 2019**
Jan	5,868 (13.5)	93.5	364 (0.8)	57.6	37,082 (85.6)	108.2	43,314	105.2
Feb	5,457 (13.2)	91.4	389 (0.9)	74.8	35,381 (85.8)	109.3	41,227	106.1
Mar	4,030 (12.5)	62.3	274 (0.9)	44.6	27,852 (86.6)	78.7	32,156	75.7
Apr	488 (15.1)	9.2	108 (3.3)	26.9	2,639 (81.6)	8.1	3,235	8.5
May	393 (6.2)	6.7	277 (4.3)	60.7	5,701 (89.5)	16.7	6,371	15.8
Jun	1,190 (6.9)	20.1	384 (2.2)	86.7	15,644 (90.9)	47.2	17,218	43.6
Jul	1,897 (7.6)	31.8	546 (2.2)	130.0	22,627 (90.3)	63.0	25,070	59.3
Aug	2,443 (8.5)	42.7	493 (1.7)	70.4	25,933 (89.8)	73.7	28,869	69.4
Sep	3,603 (10.1)	61.4	570 (1.6)	137.7	31,590 (88.3)	92.0	35,763	88.1
Oct	4,467 (10.7)	69.8	616 (1.5)	132.2	36,579 (87.8)	94.3	41,662	91.3
Overall p<0.001								
Flexible Sigmoidoscopy								
Jan	2,878 (11.6)	91.0	1,222 (4.9)	80.1	20,747 (83.5)	100.1	24,847	97.8
Feb	2,610 (11.2)	91.7	1,098 (4.7)	92.0	19,501 (84.0)	103.7	23,209	101.6
Mar	1,829 (10.8)	57.8	939 (5.6)	66.8	14,143 (83.6)	68.2	16,911	66.8
Apr	140 (8.8)	5.5	361 (22.6)	30.4	1,098 (68.7)	5.7	1,599	7.0
May	149 (4.6)	5.5	728 (22.4)	55.3	2,376 (73.0)	12.0	3,253	13.6
Jun	394 (5.4)	14.1	965 (13.3)	84.9	5,875 (81.2)	30.4	7,234	31.1
Jul	732 (6.8)	26.5	1,204 (11.1)	96.6	8,866 (82.1)	42.4	10,802	43.4
Aug	810 (6.7)	32.4	1,108 (9.1)	78.4	10,260 (84.3)	52.7	12,178	52.1
Sep	1,054 (7.1)	39.8	1,242 (8.4)	107.5	12,568 (84.6)	64.5	14,864	63.8
Oct	1,233 (7.6)	39.1	1,170 (7.2)	90.1	13,772 (85.1)	64.0	16,175	62.2
Overall p<0.001								
Gastroscopy								
Jan	6,111 (11.2)	95.9	3,711 (6.8)	94.8	44,797 (82.0)	104.0	54,619	102.4
Feb	5,709 (11.3)	92.1	3,270 (6.5)	94.8	41,622 (82.3)	105.1	50,601	102.7
Mar	3,887 (10.3)	59.2	3,038 (8.1)	77.6	30,800 (81.6)	70.9	37,725	70.0
Apr	507 (9.0)	9.3	1,501 (26.6)	42.5	3,634 (64.4)	9.0	5,642	11.4
May	314 (3.7)	4.9	2,464 (28.8)	63.7	5,768 (67.5)	13.9	8,546	16.5
Jun	999 (5.2)	15.8	3,055 (15.9)	87.3	15,155 (78.9)	37.5	19,209	38.3
Jul	1,676 (6.0)	26.4	3,531 (12.7)	91.0	22,703 (81.3)	51.0	27,910	51.0
Aug	2,100 (6.5)	38.1	3,269 (10.1)	86.8	26,882 (83.4)	65.7	32,251	64.3
Sep	3,222 (7.6)	58.8	3,817 (9.1)	112.6	35,079 (83.3)	87.3	42,118	85.9
Oct	3,966 (8.3)	67.8	3,636 (7.6)	97.0	40,277 (84.1)	91.9	47,879	89.6
Overall p<0.001								
All Procedures								
Jan	14,857 (12.1)	93.9	5,297 (4.3)	87.3	102,626 (83.6)	104.7	122,780	102.4
Feb	13,776 (12.0)	91.8	4,757 (4.1)	92.1	96,504 (83.9)	106.3	115,037	103.7
Mar	9,746 (11.2)	60.2	4,251 (4.9)	71.7	72,795 (83.9)	73.1	86,792	71.3
Apr	1,135 (10.8)	8.5	1,970 (18.8)	38.5	7,371 (70.4)	8.0	10,476	9.5
May	856 (4.7)	5.7	3,469 (19.1)	61.5	13,845 (76.2)	14.5	18,170	15.6
Jun	2,583 (5.9)	17.2	4,404 (10.1)	86.7	36,674 (84.0)	39.5	43,661	38.7
Jul	4,305 (6.7)	28.6	5,281 (8.3)	95.2	54,196 (85.0)	53.5	63,782	52.3
Aug	5,353 (7.3)	39.0	4,870 (6.6)	82.8	63,075 (86.1)	66.0	73,298	63.6
Sep	7,879 (8.5)	56.3	5,629 (6.1)	113.5	79,237 (85.4)	84.3	92,745	82.1
Oct	9,273 (10.0)	60.2	5,422 (5.8)	98.3	90,628 (97.7)	87.0	105,716	84.5
Overall p<0.001								

Table 5-8: Number of endoscopy procedures, with percentages, compared to same month in 2019, split by procedure type and scheduling type from January to October 2020.

In 2018, a mean of 112,680 procedures were performed per month (Figure 5-9). This increased to 116,538 in 2019, a 3.5% rise. The pandemic led to a sudden decrease, to a nadir of 10,476 procedures performed in April 2020, 9.5% of the procedures performed in April 2019 (Table 5-8). There was a partial recovery, with 105,716 procedures in October 2020, representing 84.5% of the procedures performed in October 2019. While individual endoscopic procedures have all followed the same general trend of a sudden decrease followed by a slow recovery, gastroscopies appeared least affected, with 5,642 procedures completed in April 2020 (11.4% compared to April 2019), whereas only 1,599 flexible sigmoidoscopies were completed (7.0% compared to April 2019). Individual procedures have also recovered at different rates: in October 2020, flexible sigmoidoscopy was at 62.2% of October 2019 levels, compared with 89.6% for gastroscopy and 91.3% for colonoscopy. The trend was significant for all procedures (overall p=0.015).

Regional Effect

I also investigated whether there was any regional effect on the total number of procedures performed (Table 5-9). East of England was the most affected region in April 2020, performing 4.5% of the number of overall procedures compared to April 2019. In contrast, the corresponding percentage for the Midlands, the least affected region, was 17.6% (X^2 =0.813, p<0.001). Recovery has also differed between regions: by October 2020, North East and Yorkshire had recovered least well (75.3% compared to October 2019), whereas East of England recovered best, to 106.2% (X^2 =0.411, p<0.001). However, when analysing numbers of procedures performed in each month from January to October 2020, there was no statistically significant effect between different regions. This therefore suggests that the COVID-19 pandemic led to similar patterns of decrease and recovery across regions.

2020	East of England		London		Midlands		North East & Yorkshire		North West		South East		South West	
	n	As % of 2019 procedures	n	As % of 2019 procedures	n	As % of 2019 procedures	n	As % of 2019 procedures	n	As % of 2019 procedures	n	As % of 2019 procedures	n	As % of 2019 procedures
Jan	10,359	99.8	18,953	110.0	23,278	102.8	22,249	107.6	17,627	99.9	17,000	94.9	13,314	99.0
Feb	10,366	102.5	18,465	111.1	21,680	104.8	20,722	108.3	16,139	99.9	15,754	95.7	11,911	101.0
Mar	7,448	70.9	13,175	68.9	16,097	73.7	16,458	80.5	12,356	67.5	11,480	63.8	9,778	72.3
Apr	456	4.5	882	5.4	3,844	17.6	1,649	8.8	745	4.7	1,941	12.6	959	7.8
May	1,097	10.8	2,369	12.7	4,490	20.8	2,505	12.8	2,101	12.3	2,935	17.5	2,673	21.6
Jun	3,839	38.1	5,606	30.8	9,684	44.5	7,408	38.9	4,570	27.1	6,717	43.3	5,837	50.8
Jul	6,837	61.1	9,860	51.3	12,510	54.4	10,255	47.2	7,586	43.4	8,793	54.8	7,941	60.0
Aug	6,405	62.9	14,064	75.9	14,660	69.0	12,054	58.1	9,179	55.3	9,086	58.0	7,850	64.2
Sep	9,328	96.2	16,753	90.6	17,425	84.4	14,876	72.4	11,736	72.8	12,060	79.7	10,567	85.8
Oct	11,240	106.2	19,119	91.6	19,093	84.5	16,852	75.3	13,547	75.9	14,631	83.4	11,234	84.8
Overall p=0.0853 (ns)														

Table 5-9: Endoscopic procedures from January to October 2020 by region.

Case Mix

The pandemic also led to a different case mix of procedures being performed (Table 5-8). There was an increase in the proportion of unscheduled procedures and a decrease in planned/surveillance and waiting list procedures. This was most stark in the month of April 2020, where 3.3% of colonoscopies, 22.6% of flexible sigmoidoscopies and 26.6% of gastroscopies were unscheduled, compared to 1.1%, 5.2% and 7.1% respectively in April 2019 (X^2 =0.71, p<0.001). As endoscopy services started to recover, there has been a reversal of this trend (Table 5-8). A further observation is that recovery for planned/surveillance procedures is slower than waiting list procedures. For all procedures in April 2020, planned/surveillance and waiting list procedures were at 8.5% and 8.0% of April 2019 levels. However, in October 2020 planned/surveillance procedures were at 60.2% of October 2019 levels, compared with 87.0% for waiting list procedures. The differences between the monthly proportions of unscheduled, planned/surveillance and waiting list procedures were statistically significant for each individual endoscopic procedure (all p<0.001).

Waiting List

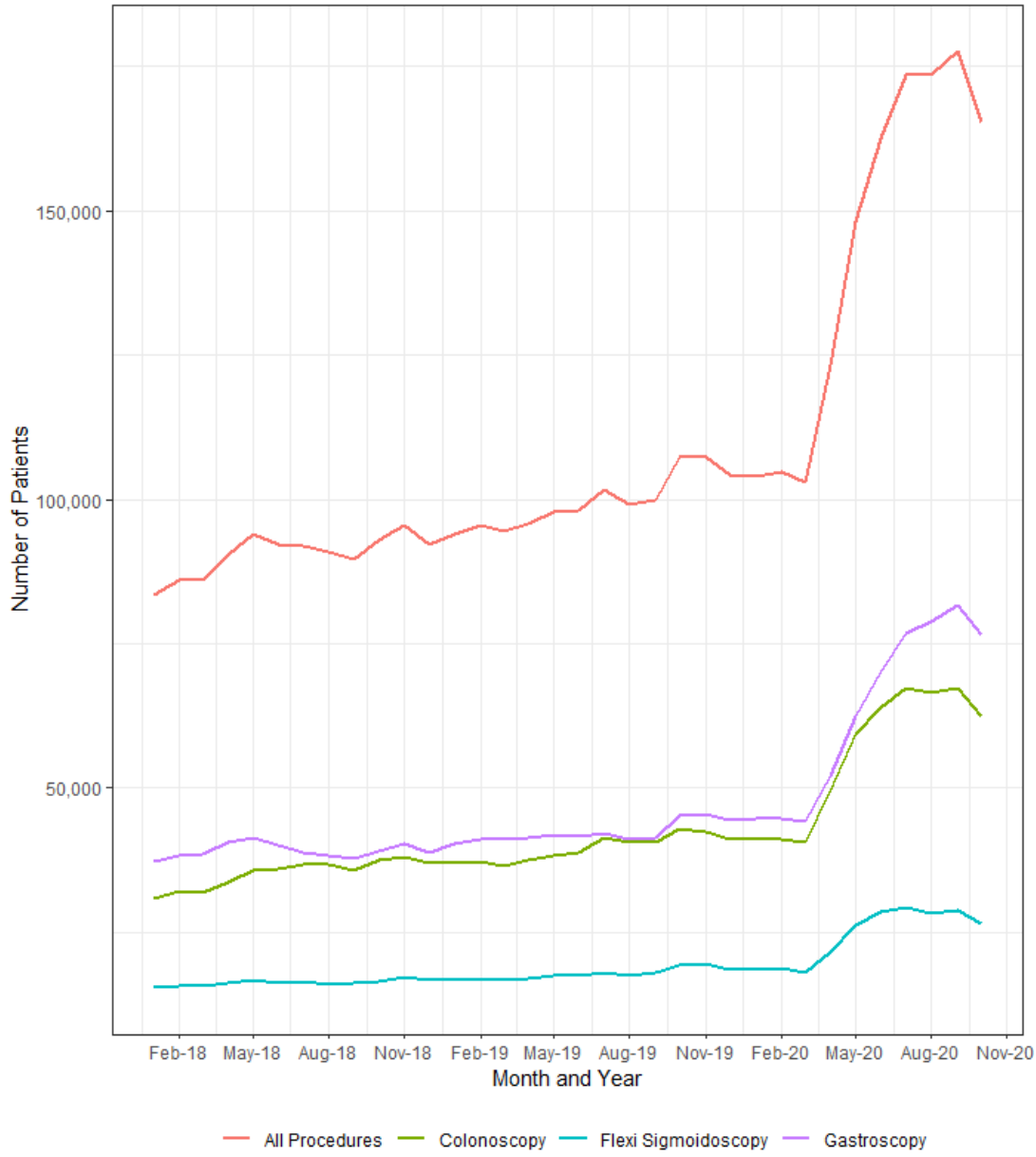


Figure 5-10: Number of patients on procedure waiting list at the end of the month from January 2018 to October 2020.

Figure 5-10 demonstrates the number of patients on the waiting list at the end of each month for each endoscopic procedure. There was a marked increase in the patients on the overall waiting list since March 2020, increasing from 102,891 patients to a peak of 177,557 patients in September 2020 (72.6% increase). However, it appears that this had started to fall in October 2020.

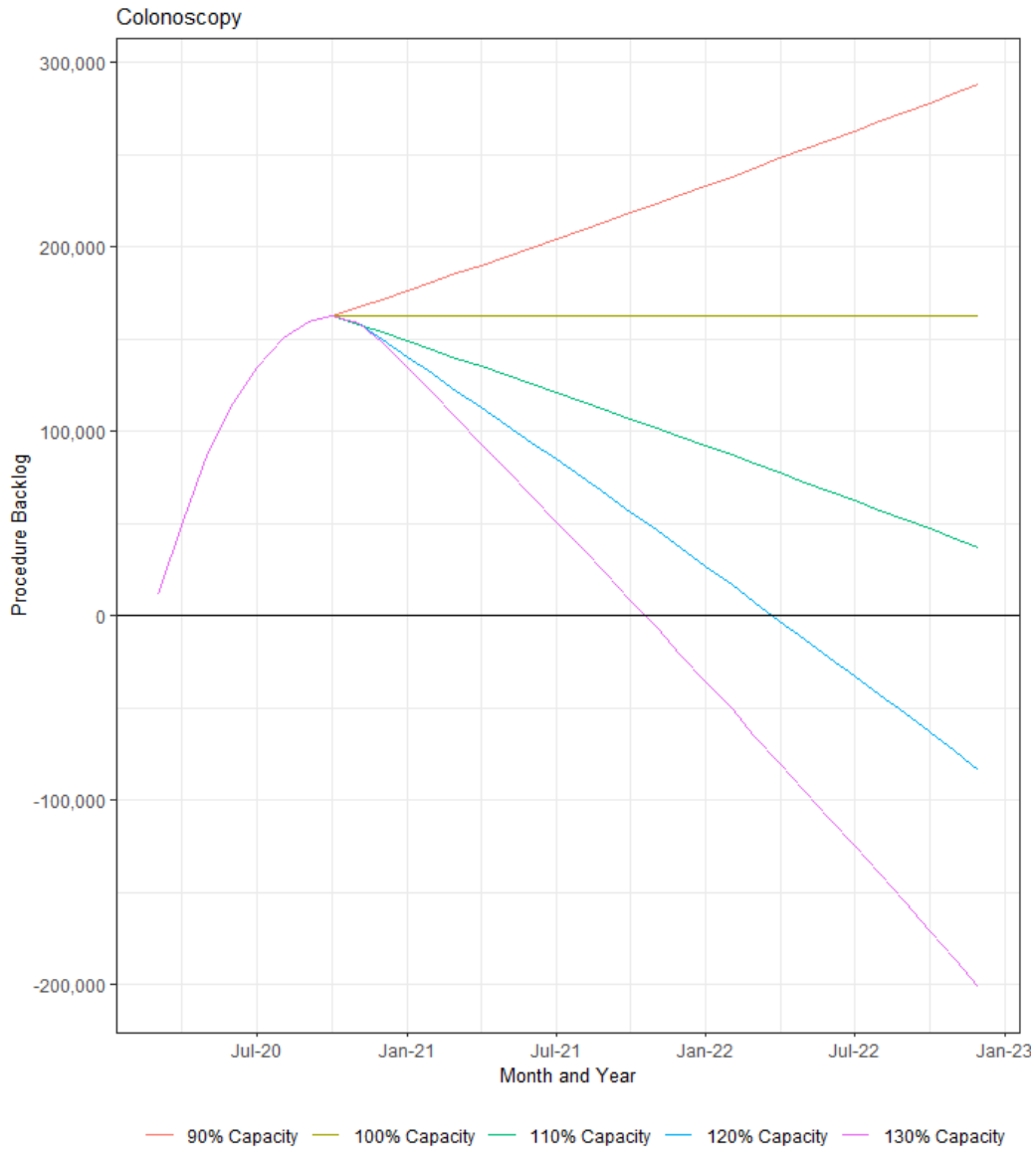
5.4.5 Future Projections for Potential Strategies and Solutions

Scenario 1 – Sustained recovery to a plateau

At the time of analysis in the end of 2020, I created five hypothetical scenarios whereby endoscopic capacity would recover at a steady rate based on the current trajectory, until plateauing at a fixed capacity. These were set at 90%, 100%, 110%, 120% and 130% of what would be expected capacity in the absence of COVID-19. These states would be reached between October 2020 and April 2021 depending on procedure and capacity level (Table 5-10). For colonoscopy, flexible sigmoidoscopy and gastroscopy, I estimated there would be a residual backlog of 162,735 (95% CI: 143,775-181,695), 119,025 (95% CI: 417,398-130,651) and 194,087 (95% CI: 172,564-215,611) procedures respectively attributable to the pandemic, prior to recovery at 100% capacity (Figure 5-11, Figure 5-12 and Figure 5-13). Importantly, this backlog would remain long term unless there was additional intervention. Even if capacity could be increased to 130%, it would take until November 2021, June 2022 and December 2021 to catch up with the pandemic associated backlog for colonoscopy, flexible sigmoidoscopy and gastroscopy respectively. Conversely, recovery to only 90% capacity would have added an additional 4,551 (95% CI: 4,249-4,852) colonoscopies, 2,453 (95% CI: 2,259-2,646) flexible sigmoidoscopies and 5,213 (95% CI: 4,871-5,555) gastroscopies per month to the backlog, with these figures increasing in line with the background upward trend in demand.

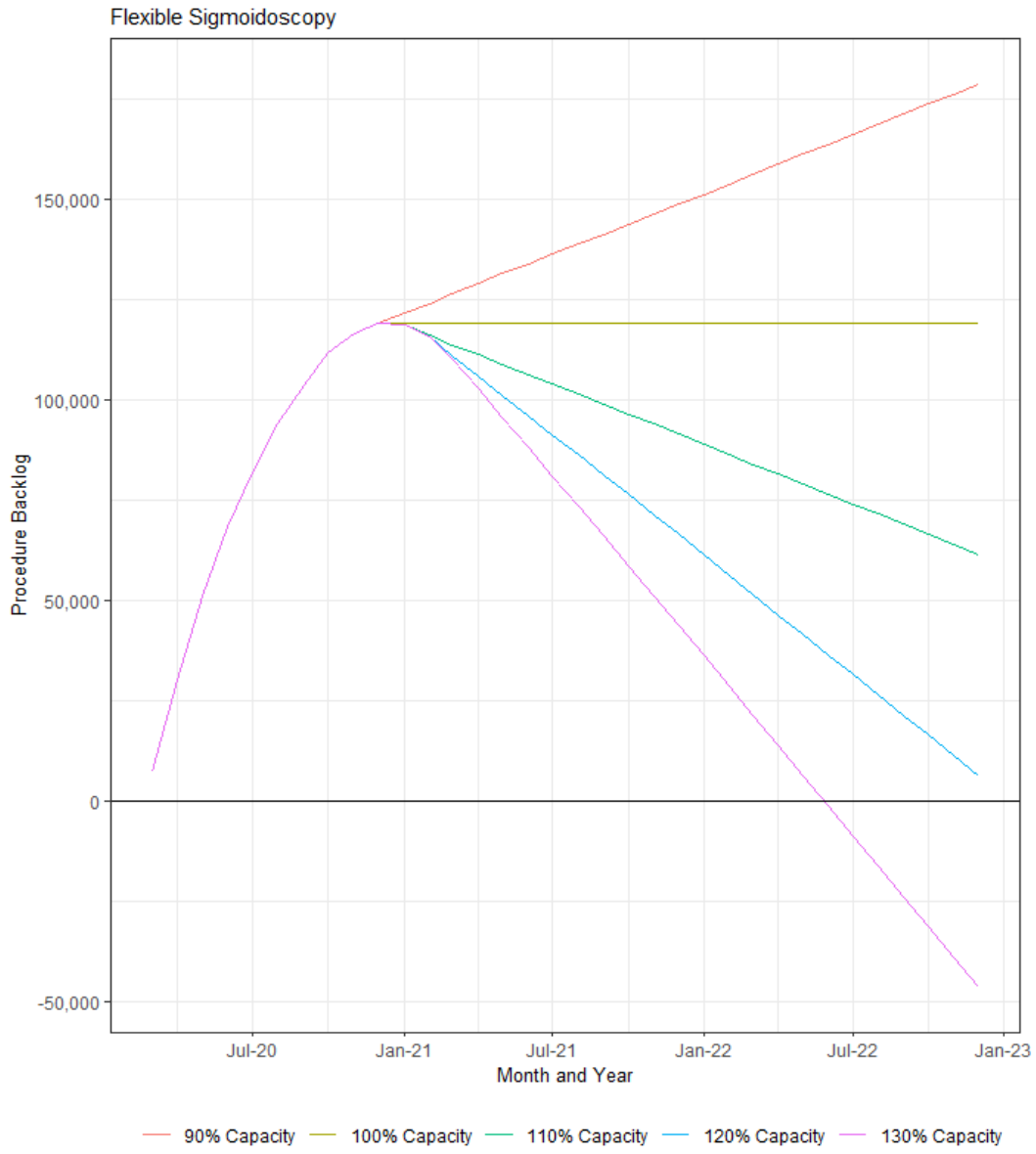
Capacity Level	Colonoscopy	Flexi Sigmoidoscopy	Gastroscopy
90%	October 2020	December 2020	October 2020
100%	November 2020	January 2021	November 2020
110%	December 2020	February 2021	December 2020
120%	December 2020	March 2021	December 2020
130%	January 2021	April 2021	January 2021

Table 5-10: Expected month and year when different capacity levels under scenario 1 are reached.



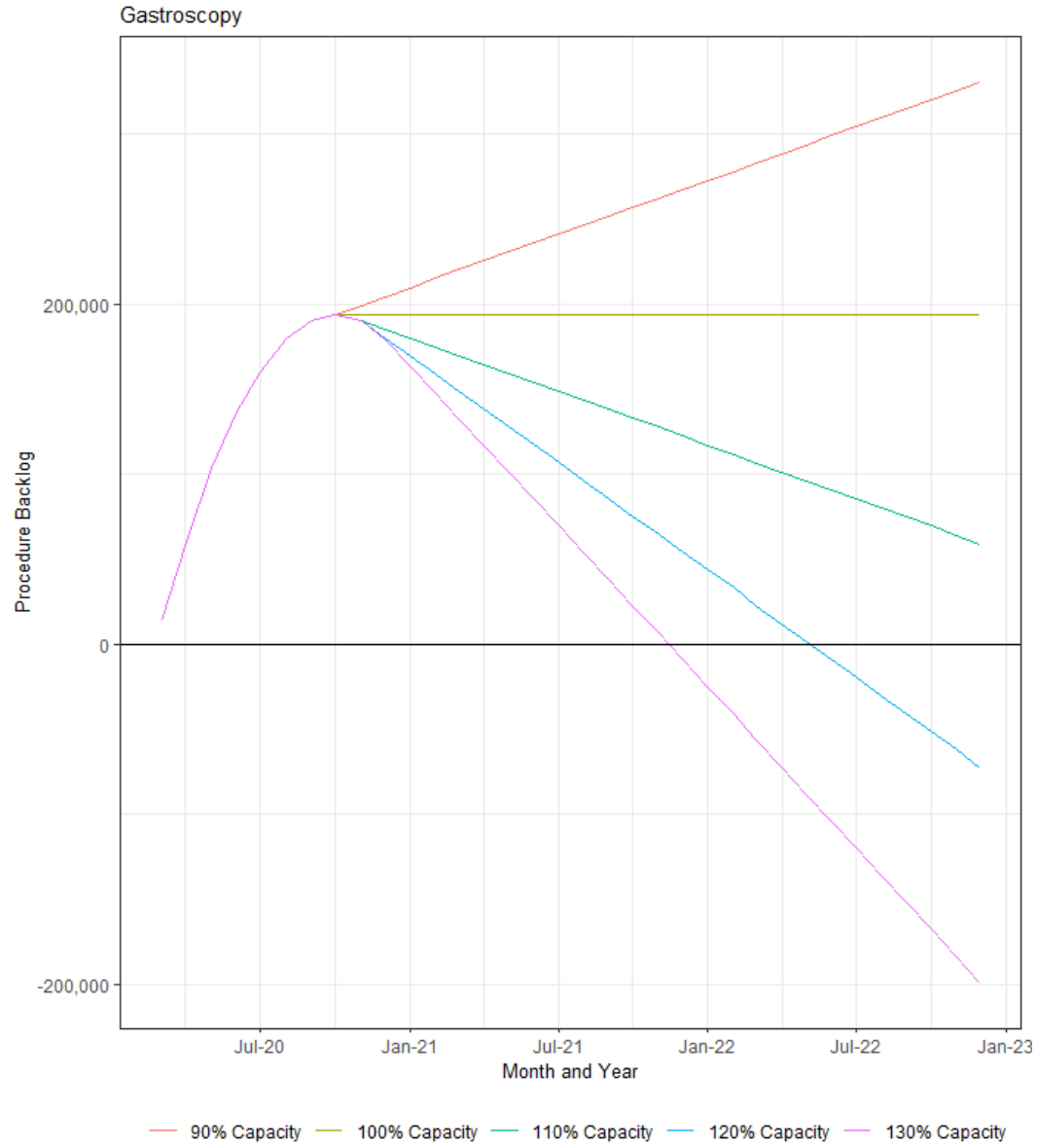
Capacity Level	Backlog in Jan 2021 n (95%CI)	Backlog in Jan 2022 n (95%CI)	Backlog in Jan 2023 n (95%CI)
90%	176,387 (156,522-196,251)	233,239 (208,668-257,809)	293,682 (262,641-324,723)
100%	162,735 (143,775-181,695)	162,735 (143,775-181,695)	162,735 (143,775-181,695)
110%	149,736 (132,568-166,904)	92,884 (80,422-105,347)	32,441 (26,450-38,433)
120%	140,610 (124,057-157,163)	26,906 (19,764-34,048)	Backlog Cleared
130%	134,880 (119,927-149,833)	Backlog Cleared	Backlog Cleared

Figure 5-11: Estimation of colonoscopy procedural backlog over time based on current trajectory before reaching a plateau at a set capacity level, with estimated backlog in January 2021, January 2022 and January 2023 shown in sub-table.



Capacity Level	Backlog in Jan 2021 n (95%CI)	Backlog in Jan 2022 n (95%CI)	Backlog in Jan 2023 n (95%CI)
90%	121,689 (109,219-134,158)	151,362 (135,867-166,858)	181,425 (161,768-201,082)
100%	119,025 (107,398-130,651)	119,025 (107,398-130,651)	119,025 (107,398-130,651)
110%	118,727 (107,888-129,567)	89,054 (81,240-96,868)	58,991 (55,339-62,644)
120%	118,727 (107,888-129,567)	61,459 (57,386-65,532)	1,334 (-2,916-5,583)
130%	118,727 (107,888-129,567)	36,247 (35,823-36,670)	Backlog Cleared

Figure 5-12: Estimation of flexible sigmoidoscopy procedural backlog over time based on current trajectory before reaching a plateau at a set capacity level, with estimated backlog in January 2021, January 2022 and January 2023 shown in sub-table.



Capacity Level	Backlog in Jan 2021 n (95%CI)	Backlog in Jan 2022 n (95%CI)	Backlog in Jan 2023 n (95%CI)
90%	209,726 (187,175-232,276)	272,619 (244,727-300,511)	336,055 (300,817-371,293)
100%	194,087 (172,564-215,611)	194,087 (172,564-215,611)	194,087 (172,564-215,611)
110%	179,969 (159,894-200,045)	117,076 (102,343-131,810)	53,640 (46,252-61,028)
120%	169,540 (150,163-188,917)	43,754 (35,060-52,447)	Backlog Cleared
130%	163,709 (145,442-181,976)	Backlog Cleared	Backlog Cleared

Figure 5-13: Estimation of gastroscopy procedural backlog over time based on current trajectory before reaching a plateau at a set capacity level, with estimated backlog in January 2021, January 2022 and January 2023 shown in sub-table.

Scenario 2 – Further lockdown with reduction of endoscopic activity

I created five hypothetical scenarios where there would be a sudden decrease in performed procedures before a slow recovery, returning to 100% capacity, mimicking the pattern seen early on in the pandemic. This ranged from a drop to 60% capacity for 2-month lockdown to 20% capacity for a 6-month lockdown, which overall would represent a less severe slowdown in services, with units having learnt from their previous experience and improved contingency planning (also see Table 5-11).

Capacity	2-month lockdown	3-month lockdown	4-month lockdown	5-month lockdown	6-month lockdown
Dec-2020	60%	50%	40%	30%	20%
Jan-2021	80%	60%	50%	40%	30%
Feb-2021	100%	80%	60%	50%	40%
Mar-2021	100%	100%	80%	60%	50%
Apr-2021	100%	100%	100%	80%	60%
May-2021	100%	100%	100%	100%	80%

Table 5-11: Hypothetical scenarios of reduction of endoscopy capacity, secondary to a further lockdown.

I estimated that a 2-month lockdown could add an additional 15% (73,359 cases) to the total backlog, while 4-month and 6-month lockdowns could add 44% (208,269 cases) and 83% (392,796 cases) to the total backlog (Table 5-12).

	Colonoscopy		Flexible Sigmoidoscopy		Gastroscopy		Total	
Months of reduced activity	**Estimated additional procedures (95%CI)**	**% of potential backlog**	**Estimated additional procedures (95%CI)**	**% of potential backlog**	**Estimated additional procedures (95%CI)**	**% of potential backlog**	**Estimated additional procedures (95%CI)**	**% of potential backlog**
2	27,354 (25,521-29,187)	16.8	14,721 (13,542-15,899)	12.4	31,284 (29,204-33,365)	16.1	73,359 (68,404-78,314)	15.4
3	50,257 (46,844-53,669)	30.9	26,999 (24,805-29,194)	22.7	57,371 (53,498-61,245)	29.6	134,627 (125,403-143,852)	28.3
4	77,835 (72,482-83,188)	47.8	41,744 (38,302-45,187)	35.1	88,690 (82,613-94,766)	45.7	208,269 (193,797-222,740)	43.8
5	110,113 (102,446-117,781)	67.7	58,958 (54,027-63,889)	49.5	125,243 (116,539-133,948)	64.5	294,315 (273,586-315,043)	61.9
6	147,117 (136,749-157,485)	90.4	78,643 (71,975-85,310)	66.1	167,037 (155,267-178,806)	86.1	392,796 (364,768-420,825)	82.5

Table 5-12: Estimated number of additional procedures added to the backlog, and also percentage of total potential backlog (denominator: colonoscopy: 162,735; flexible sigmoidoscopy: 119,025; gastroscopy: 194,087; total: 475,847) secondary to a further reduction of endoscopic activity.

5.4.6 Discussion

In this analysis, I demonstrated that there is an urgent gap in endoscopy service provision as an indirect effect of the COVID-19 pandemic. The pandemic directly led to a decrease in the number of endoscopic procedures performed, with a partial recovery as COVID-19 infections and hospitalisations started to fall. Furthermore, potential solutions are challenging even if above normal capacity could be achieved.

Endoscopic activity was at its lowest in April 2020, decreasing to 9.5% of procedures compared to April 2019. I also noted regional differences in endoscopy provision as well as changes in the case mix of procedures. A similar overall pattern was seen in the NED database, which recorded the trough in the week beginning 30 March 2020, although no regional differences were seen.[145] This is likely because the NED analysis only included data until the end of May and regional differences may not have been apparent then.

As a comparison, a national study from the Netherlands, which entered a nationwide lockdown on 12 March 2020, demonstrated a drop to 29% for colonoscopy and 37% for gastroscopy when comparing between January and April 2020.[156] Meanwhile, a global web based survey carried out between April and May 2020 covering 252 centres from 55 countries suggested an average of 83% reduction when comparing activity between baseline and during the COVID-19 pandemic.[157] These results would suggest that endoscopy in England was impacted to a greater degree compared to other countries around the world, even though there was broad consensus among worldwide guidelines.[158] A host of reasons, ranging from decision making at local level, lack of referral for colonoscopy, understandable hesitancy at accessing cancer services to preserving personal protective equipment (PPE), staff redeployment and staff sickness with COVID-19 may have been the cause of this difference.[159,160]

My analysis also demonstrated that recovery of endoscopy services from the 1st wave of the COVID-19 pandemic was somewhat slow and that there were some significant inter-regional differences. From the peak of the pandemic in April 2020, waiting lists started to fall but approximately 50% more patients were on the list in October 2020 compared to a year ago. Meanwhile, capacity in the same month remained at 84.5% compared to a year ago, with a slower recovery for planned/surveillance procedures (60% of October 2019 levels) compared to waiting list procedures (87% of October 2019 levels). In addition, there was near normalisation of unscheduled procedures by June 2020 (87% of June 2019), which was faster than planned/surveillance procedures and waiting list procedures. This may have been a manifestation of altered health behaviour during the first peak of the pandemic, with delayed presentation of medical conditions subsequently leading to emergency complications.[143] In addition, COVID-19 led to excess all-cause mortality worldwide.[161] This could have influenced the recovery of endoscopic procedures, as this may have led to reduced demand due to potential patients dying early or becoming too unwell for investigation.

These findings demonstrated the need for a targeted approach regionally to aid recovery with appropriate resource allocation, with attention also placed on planned/surveillance procedures to ensure their recovery did not lag too far behind waitlist procedures.

Endoscopy services also faced further challenges in increasing capacity back to pre-pandemic levels. Staffing was a major concern: redeployment, staff absence due to self-isolation or shielding and additional administrative burden due to COVID-19 mitigation measures such as telephone triaging and pre-procedural SARS-CoV-2 testing all created additional barriers in service recovery.[153,162] In addition, the requirement for PPE, especially with gastroscopy, which is an aerosol-generating procedure and infection control measures such as deep

cleaning and leaving a time gap between procedures all led to decreased room utilisation and efficiency.[153,163]

My work on future projections, based on extrapolation of historical pre-pandemic demand, also demonstrated that potentially there may be a backlog of nearly half a million endoscopic procedures attributable to the pandemic. I also quantified the scale of the challenge faced in the recovery phase: if only 90% capacity was reached, which may have been realistic given the constraints of performing endoscopy, each month will have added an additional 12,217 endoscopic procedures to the potential backlog. Furthermore, with a further surge in COVID-19 cases, such as during a 2nd wave of infection, a short 2-month disruption could have added an additional 73,359 cases to the potential backlog.

5.4.7 Tackling the Backlog

During the pandemic several strategies was used nationwide to tackle the backlog:

Private Sector

One option was to use the private sector to provide this additional capacity for a short period, or to create additional capacity during evenings and weekends. However, a UK study from 2017 showed only 55% of endoscopy units were meeting cancer wait targets, with shortages of endoscopists and nursing staff cited as reasons for missing the target.[164] 82% of English NHS trusts already performed ad-hoc weekend work, hence ability to increase capacity further for a sustained period was likely not to have been practicable.[164]

Reallocation of Resources

Temporary increases in waiting list and unscheduled capacity could have been achieved by reducing cancer screening programmes and reallocating this capacity. As an example, it was expected for BowelScope to have performed 256,000 flexible sigmoidoscopies in 2020.[165] Reallocating this capacity would have created an additional 87% capacity per month.

However, this decision needed to be balanced against the risk of an overall increase in preventable deaths.[166] Furthermore, the latest BSG guidelines for post-polypectomy surveillance suggested that adoption would lead to decrease to 20% of the level of polyp surveillance workload in 2019.[167] 10.4% of colonoscopies in a Dutch series had an indication of adenoma surveillance, so theoretically adherence to these guidelines could have generated an extra 8.3% of capacity.[168] In the end, BowelScope was discontinued in early 2021, although bowel cancer screening colonoscopies are still performed today.

Improving Triaging of Referrals

Enhanced vetting of referrals was employed in 77% of recovery plans and has helped to reduce demand.[167] One such example is the Edinburgh Dysphagia Score to triage upper GI referrals, which was also included in guidance published by the British Society of Gastroenterology.[162,169–171]

Alternatives to Endoscopy

An alternative to colonoscopy was to use CT colonography (CTC) instead. Advantages of CTC include decreased PPE use, improved ability to socially distance compared to colonoscopy and also a shorter patient visit time.[172] However, ensuring there was adequate CT capacity and expertise to perform and report the procedures were barriers to implementation.[173] Indeed, local data at UCLH did not see an appreciable increment in the use of CTC as an alternative to colonoscopy in the recovery phase of the pandemic.

Other touted alternatives included Cytosponge and colon capsule endoscopy in lieu of gastroscopy and colonoscopy, although neither were used in routine clinical practice, and never gained widespread adoption.[174,175]

Wider Implications

The challenges faced during the COVID-19 pandemic was not unique to gastroenterology. Other procedure-heavy specialties such as cardiology and surgery also faced similar challenges.[176,177] Of particular concern was the impact on cancer waiting lists and pathways; it was estimated that diagnostic delays may lead to a 16% increase in colorectal cancer deaths and 5.9% increase in oesophageal cancer deaths over 5 years.[120] Four years on from the peak of the pandemic, latest data suggest there are some 7.6 million people on the national waiting list, with 40.9% waiting for more than 18 weeks for treatment.[178]

5.4.8 Limitations

In this analysis, I assumed that endoscopy will eventually return to pre-pandemic levels of practice and demand, and that over time referral patterns will even out back to pre-pandemic trends. This does not appear to have been the case and nationwide waiting lists have yet to demonstrate sustained decreases since the onset of the pandemic.[178] Beside the direct impact of the pandemic on disruption of healthcare services, the effects of altered health behaviour leading to delayed presentation of medical conditions may still be ongoing, contributing to rising waiting lists despite recovery of services post pandemic.[143]

Secondly, this study excluded endoscopic retrograde cholangio-pancreatographies (ERCP), although local data from UCLH did not suggest significant decreases due to the pandemic. The data also did not include procedures performed as part of the BCSP and BowelScope, which were suspended at the height of the pandemic. Additionally, data on the number of referrals to endoscopy was not available, which would have been a better measure of demand. Finally, although data on the number of patients on the waiting list was available per procedure, this only provided a snapshot of the number of people at the end of a month at a given time point, and also excluded unscheduled or planned/surveillance procedures, hence it was not used as a surrogate for referrals.

5.4.9 Conclusion

This work demonstrated the enormous strain the COVID-19 pandemic placed on NHS endoscopy in England, which is still being felt at present. Unfortunately, there are no simple solutions, but innovative ways of managing demand may be a way forward, as I will discuss in Chapter 6: Creating a Risk Prediction Model.

5.4.10 Media Interest

The publication in the *Lancet Gastroenterology and Hepatology* also generated some media interest. This included being featured as an exclusive article in the *i* newspaper as well as several press releases (Figure 5-14). Press releases were published on websites of UCL, Bowel Cancer UK and DATA-CAN, part of Health Data Research UK.

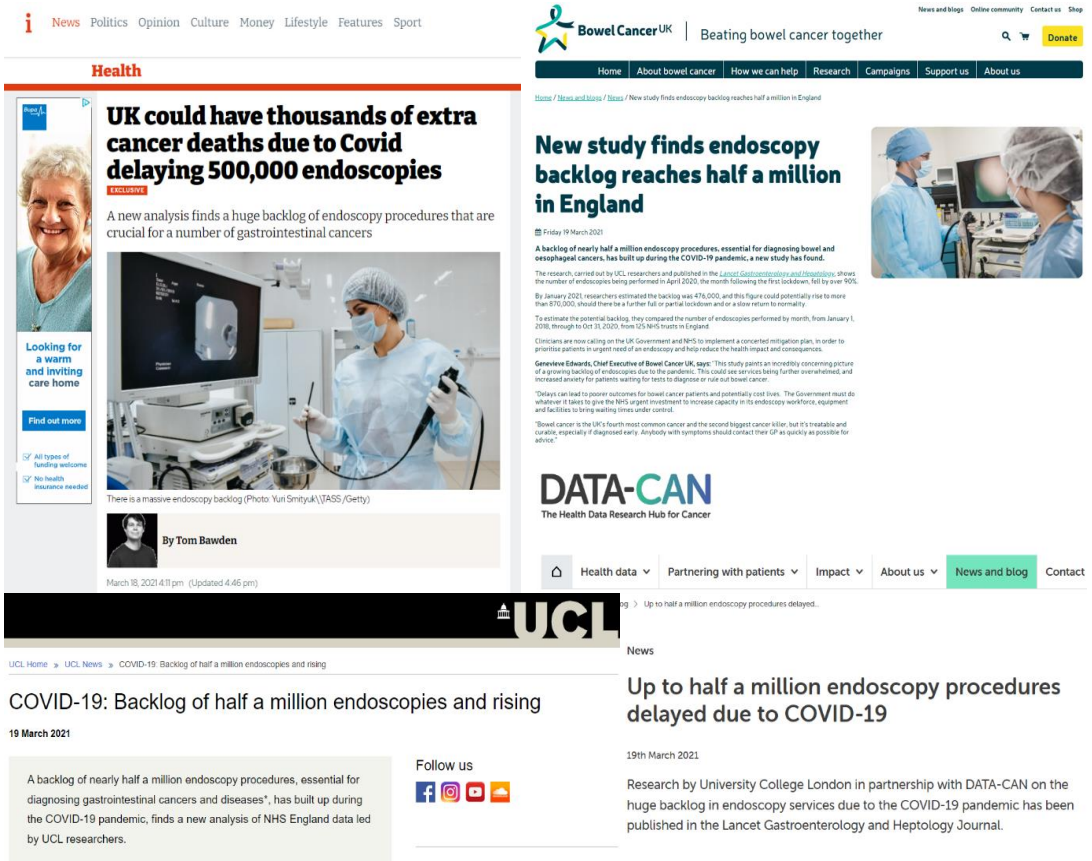


Figure 5-14: Clockwise from top left: article in the *i* newspaper; press release on Bowel Cancer UK website; press release on DATA-CAN website, part of Health Data Research UK; press release on UCL website

5.5 Comparison between local and national data: endoscopic procedures for UCLH

During research for this chapter, I had access to endoscopy procedure data for UCLH, both from the local data analytics department and nationally via the NHS England's Diagnostics Waiting Times and Activity dataset. I therefore compared the two sets of data between January 2020 to November 2020 to see what differences there may be (Table 5-13). This period was specifically chosen due to data availability reasons and also coincided with the UK's first lockdown.

	Colonoscopy			Flexible Sigmoidoscopy			Gastroscopy		
Month	UCLH	NHS	% Diff	UCLH	NHS	% Diff	UCLH	NHS	% Diff
Jan-20	515	454	11.8	236	239	-1.3	632	51	91.9
Feb-20	517	454	12.2	196	196	0.0	598	43	92.8
Mar-20	309	270	12.6	123	126	-2.4	379	33	91.3
Apr-20	23	11	52.2	9	12	-33.3	72	10	86.1
May-20	83	72	13.3	29	31	-6.9	162	13	92.0
Jun-20	189	150	20.6	62	59	4.8	264	14	94.7
Jul-20	349	335	4.0	105	101	3.8	438	28	93.6
Aug-20	419	397	5.3	124	123	0.8	415	412	0.7
Sep-20	509	456	10.4	123	124	-0.8	570	548	3.9
Oct-20	568	507	10.7	109	113	-3.7	475	442	6.9
Nov-20	497	433	12.9	109	105	3.7	488	458	6.1

Table 5-13: Number of colonoscopy, flexible sigmoidoscopy and gastroscopy procedures documented on the UCLH local database and NHS England's Diagnostics Waiting Times and Activity dataset.

% Diff = percentage difference, calculated with the UCLH figure as the denominator

Some discrepancies were noted between the two datasets. The UCLH dataset consistently recorded a higher number of procedures performed compared to the NHS England dataset for colonoscopies and gastroscopies. This was more variable for flexible sigmoidoscopies, where conversely for 6 out of 11 months the NHS England dataset recorded a higher number of procedures performed compared to the UCLH dataset. The greatest discrepancy appeared in the number of gastroscopies performed, where between the months of January to July 2020 the UCLH dataset figure was at least 85% greater than the figure recorded from NHS

England. However, across all three procedure types, the percentage difference after August 2020 was less pronounced, with a maximum of 12.9% for colonoscopies in November 2020.

There could be several explanations for the discrepancies between the two datasets. As mentioned previously, procedures performed under the BCSP would not be recorded on the NHS England dataset. In addition, there may be additional therapeutic procedures, such as planned polypectomy or oesophageal stent insertion that may not be recorded. It would also appear that gastroscopy data from January to July 2020 may have had some quality issues given that large percentage differences was observed. It is likely there was an issue with coding of performed procedures which could have affected the integrity of the NHS England data, hence NHS England data presented is likely to be an underrepresentation of actual endoscopy activity at UCLH, and possibly nationwide too.

5.6 Chapter Conclusions

In this chapter, I have described some of the changes to gastrointestinal services both at a local and national level because of the COVID-19 pandemic. I have demonstrated some of the advantages of local level data, such as having greater access to specific data requests and having the option of delving into patient's record to understand more about individual circumstances. However, some procedure numbers were relatively low making it difficult to draw meaningful conclusions. In addition, I have demonstrated some of the data quality issues that I encountered, and provided possible explanations.

Chapter 6: Creating a Risk Prediction Model

Contents of this chapter have previously been presented and published as detailed below:

Presentations

Ho KMA, Rosenfeld A, Hogan Á, et al. Development and validation of a multivariable risk factor questionnaire to detect oesophageal cancer in 2-week wait patients. *Clin Res Hepatol Gastroenterol* 2023; **47**: 102087.[6]

Publications

Ho KMA, Rosenfeld A, Hogan Á, et al. O5 Using machine learning to develop models for the prediction of upper gastrointestinal cancers. *Gut* 2022; **71(S1)**: A3.[12]

- Presented at British Society of Gastroenterology (BSG) Annual Meeting, Birmingham, UK (oral presentation).
- Awarded "Best Oesophageal Oral Presentation" at the conference.

Ho KMA, Rosenfeld A, Hogan Á, et al. P105 Validation of a machine learning model in a prospective cohort: the RISQ study. *Gut* 2022; **71(S1)**: A92–A92.[13]

- Presented at BSG Annual Meeting, Birmingham, UK (poster presentation).

Ho KMA, Rosenfeld A, Hogan Á, et al. PTH-72 A symptom and risk factor questionnaire accurately predicts upper gastrointestinal cancer. *Gut* 2021; **70(S4**): A136.[14]

- Presented at BSG Annual Meeting 2021, virtual (poster presentation).

Ho KMA, Rosenfeld A, Hogan, Á, McBain H, et al. P0074 A symptom and risk factor questionnaire accurately predicts upper gastrointestinal cancer. *United Eur Gastroenterol J* 2021; **9(S8)**: 302.[15]

- Presented at United European Gastroenterology Week 2021, virtual (poster presentation).

6.1 Chapter Introduction

The previous chapter (Chapter 5: Effect of COVID-19 on Gastrointestinal Services) demonstrated the effect of the COVID-19 pandemic on the provision of gastroenterology services. In this and the next chapter I will explore some of the solutions that could be employed in tackling endoscopic backlogs. This includes creating a novel risk prediction model which could be used to triage upper gastrointestinal (UGI) endoscopy referrals.

6.2 Background

6.2.1 The Clinical Problem

As discussed in preceding chapters, oesophageal cancer represents the seventh most common cause of cancer morbidity and the sixth most common cause of cancer-related death worldwide.[17] There are two major histological subtypes: oesophageal squamous cell carcinoma (OSCC) and oesophageal adenocarcinoma (OAC). OSCC comprises 90% of oesophageal cancer cases worldwide and is predominantly found in Central Asia, East Asia and East Africa. Conversely, OAC make up the remainder but is the dominant histological subtype in the Western world, including the United Kingdom (UK).[20,23,31] Crucially, oesophageal cancer is often diagnosed late; in the United Kingdom 48% of cases with available staging information are diagnosed at stage IV, while 10-year survival stands at 12%, significantly worse than other cancer types.[18,23] Although gastro oesophageal junction (GOJ) cancers represent a heterogenous entity, they have common risk factors with oesophageal cancers. As such, historically they have been included in studies with oesophageal cancer.[179] However, improving early detection of cancer remains challenging due to cancer growth characteristics.

While upper gastrointestinal (UGI) endoscopy remains the gold standard in the diagnosis of oesophageal and GOJ cancers, it is expensive, uncomfortable for the patient and has a low yield for cancer.[180] A UK series of over 580,000 patients demonstrated that only 2.1% of

patients undergoing UGI endoscopy were subsequently found to have cancer, while other serious pathology such as peptic ulceration were found in a further 11.6%.[180] These figures suggest that better selection of patients who are likely to develop oesophageal cancer could help to prioritise higher risk patients, better manage demand for endoscopy and improve overall patient experience.

Several research groups have created scoring systems to try to improve detection of oesophageal cancer. The Edinburgh Dysphagia Scale (EDS) categorises patients into low or high risk for groups using a combination of both patient characteristics and symptoms. In a validation cohort it achieved a sensitivity of 98.4% but specificity was low at 9.3%.[169,171] This would potentially imply that a large proportion of normal or benign cases would go on to have an endoscopy before yielding a cancer case.

Increasingly, machine learning (ML) methods, which apply mathematical approaches to generating computerised algorithms, have been used to develop triaging models, which could optimise use of resources. These models can calculate an individual's risk of having a disease.[181] Our group previously used an ML approach to develop a risk prediction model for Barrett's oesophagus with the aim of improving selection of patients who should be referred for UGI endoscopy for confirmation.[181] Similar approaches have been used in acute UGI bleeding for risk stratification and it even outperformed standard of care clinical scoring systems.[182]

6.2.2 COVID-19 Pandemic

In Chapter 5:Effect of COVID-19 on Gastrointestinal Services, I demonstrated that as a result of the COVID-19 pandemic there was an estimated procedural backlog of nearly half a million endoscopic procedures in England.[3] In addition, during the first 6-months of the pandemic there were decreases in pathological diagnoses of Barrett's oesophagus and OAC.[183] There is potential for worse patient outcomes as a direct consequence of delayed diagnosis due to

the disruption of endoscopy services, on top of the already poor patient outcomes for oesophageal cancer. Modelling studies have demonstrated that disruption to National Health Service (NHS) cancer pathways could lead to excess deaths and life years lost due to delays in diagnosis.[120,121]

During the COVID-19 pandemic, many health systems globally became overwhelmed, with an urgent need to reconfigure health delivery systems and prioritise those who were sickest.[184] There was significant reconfiguration of NHS services and as part of this guidance was issued by the British Society of Gastroenterology (BSG) in March 2020 during the first wave of the COVID-19 pandemic, where all non-emergency endoscopy was recommended to be halted.[144] There was a further recommendation of case by case triaging of cases. However at the initial outset, likely due to emergent nature of the situation, no overt recommendation was made as to which triaging tools should be used. Subsequent to this, the BSG published further guidance in April 2020 during the deceleration and early recovery phase of the first wave of the pandemic where it recommended "re-triage and prioritisation" by senior decision makers. In particular, for UGI endoscopy, it had the following guidance:[153]

- Dysphagia: patients who were referred for dysphagia should have this verified by a clinician at triage and be assessed with the Edinburgh Dysphagia Score (Table 6-1).[169,171] Patients who scored ≥3.5 were recommended for urgent UGI endoscopy if appropriate and fit.
- Dyspepsia: patients who were >55 years old and with new dyspepsia (symptom duration of <6 months) and either weight loss or anaemia should be referred to urgent UGI endoscopy when COVID-19 restrictions were lifted.

Factor	Points
Age Group	
0-39	0
40-49	4
50-59	5
60-69	6
70-79	7
80-89	8
90-99	9
Weight Loss	
≤3kg	0
>3kg	2
Duration of Symptoms	
<6 months	0
≥6 months	-1.5
Sex	
Male	0
Female	-1
Localisation of Dysphagia	
Anywhere except neck	0
Neck	-2
Acid Reflux	
Absent	0
Present	-1
Total score obtained by addition of points in each category Total score ≥3·5 identifies high-risk patients	

Table 6-1: Edinburgh Dysphagia Score[169,171]

Other patients, such as for Barrett's surveillance, or for dysphagia or dyspepsia symptoms that did not meet the criteria either were recommended to undergo alternative investigations, such as computerised tomography (CT) scanning or were to have their investigation deferred.

As mentioned above, BSG guidance during the first wave of COVID-19 recommended the use of the Edinburgh Dysphagia Score (EDS) for triaging. However, the score is focussed on the detection of oesophageal cancer.[169,171] This additional recommendation is not found in National Institute for Health and Care Excellence (NICE) upper gastrointestinal (UGI) cancer referral guidelines, which primarily relies on alarm features to guide referral from primary to secondary care.[185] This is perhaps not surprising as alarm features alone are very non-specific

for the presence of UGI cancers within a population cohort, with a positive predictive value (PPV) of 0.1%.[186]

In contrast, a diagnostic scoring system was not recommended for dyspepsia, although questionnaires such as the gastro oesophageal reflux disease (GORD) questionnaire (GERQ) and the GORD impact scale do exist and were initially designed for diagnosing and assessing GORD.[114,187] Furthermore, apart from the EDS, other scoring systems have been developed to predict incident cases of OAC or GOJ adenocarcinoma. These scoring systems were derived using regression based techniques and were validated by Rubenstein et al. in a head to head analysis.[188–193] Rubenstein et al. demonstrated that in particular the Kunzmann tool, developed on the prospective UK Biobank cohort, outperformed GORD alone and two other tools in predicting the development of cancer.[188,193]

Other scoring systems include one derived by Hippisley-Cox et al. This used Cox proportional hazard models to develop an eight variable algorithm to predict risk of developing gastro-oesophageal cancers within 2 years, based on primary care data.[194] The model contained features including age of patient, one lifestyle factor (smoking), five currently experienced symptoms (presence of dysphagia, abdominal pain, loss of appetite haematemesis and weight loss) as well as one laboratory parameter (anaemia, defined as a haemoglobin of <110g/L within the past year).[194] Within the validation cohort, setting a cut off for the top 10% for risk would equate to a sensitivity of 76.5%, specificity of 91.0%, positive predictive value (PPV) of 1.2% and negative predictive value (NPV) of 100%. Although this model performed better than alarm features alone with a 10-fold increase in PPV, this would still mean at the top 10% risk threshold over 20% of eventual cancers would be initially misclassified as negative.

6.2.3 Model Development

For the risk stratification model, I wanted to focus on finding cases of cancer, especially in times of increased demand and reduced resources in the aftermath of the pandemic. In an ideal scenario, I would envisage the tool to target the following:

- Population – patients referred to secondary care under a 2ww pathway
- Intervention – tool would classify patients into lower or higher risk groups, with higher risk groups prioritised for investigations (i.e. stratification of patients), especially at a time of resource constraints
- Comparator – standard of care, where all patients are referred under the same time frame
- Outcome – assessment of whether overall outcomes such as time to diagnosis and cancer stage at diagnosis (early versus late stage) differ between lower and higher risk groups and standard of care

Notably, the tool may need adaptation to the population in more normal times where there are less resource constraints; this could be through adjustment for cut-offs for risk categories.

6.3 Questionnaire Approach

I was interested to assess if whether using a novel machine learning approach I could improve on the accuracy of patients being referred to endoscopy. A questionnaire-based approach, which could be completed by patients in advance of their appointment with their doctor, either on an electronic device or on paper while awaiting their appointment, could give additional information for the doctor during the consultation and further inform clinical decision making.

6.3.1 Aims

The overall aim was to use a machine learning approach to train, test and then independently validate a risk stratification tool which could be used to predict the risk of detecting oesophageal and GOJ cancers in unselected patients referred through the NHS 2-week wait (2ww) suspected UGI cancer pathway. In addition, I aimed to trial the tool in a limited number of patients with gastric cancer, as both cohorts of patients require the same endoscopic investigation for diagnosis, to assess whether its remit could be expanded.

6.3.2 Methods

Participant Selection and Dataset Description

Patients were selected from two separate prospective cross-sectional studies; the Saliva to Predict rIsk of disease using Transcriptomics and epigenetics (SPIT) study (ISRCTN: 11921553) and predicting RIsk of diSease using detailed Questionnaires (RISQ) study (ISRCTN: 74930639). Both of these studies have been introduced earlier in Chapter 4: Datasets and Data Curation.

For the SPIT study, participants were recruited from those referred for UGI endoscopy through the NHS 2ww suspected UGI cancer pathway from 19 UK hospitals between September 2017 and May 2022. Major inclusion criteria were age over 18 and ability to provide informed consent. Exclusion criterion was pregnancy. For the RISQ study, participants were recruited from 2 UK hospitals between January 2020 and May 2022 using identical inclusion and exclusion criteria. Patients in both studies completed a symptom and risk factor questionnaire either independently or with support from a research nurse immediately before undergoing diagnostic endoscopy.

The SPIT study questionnaire consisted of 209 questions; it was multidimensional with different domains that are known to impact disease risk (Table 6-2). Major symptoms for

oesophageal cancer such as dysphagia, odynophagia and weight loss were included.[195] Questions on duration and severity of acid reflux symptoms were adapted from both the gastro oesophageal reflux disease (GORD) impact scale and GORD questionnaire.[114,196,197] Our research group included wider questions on medication use, food intake, anxiety and depression, loneliness and local engagement, as these have been found to affect health outcomes. In particular, I was interested to see if these factors may add additional richness to any generated models.[198–200]

It is important to note that patients recruited within both the SPIT and RISQ studies comprised of a subset of the population of patients who could potentially have cancer. The studies only included patients who were referred to secondary care under a 2ww pathway. This means that patients who did not present to primary care, either because they are asymptomatic or they choose not to seek medical care and patients not referred to secondary care would have automatically been excluded. Furthermore, patients who did not consent to the study and recruited patients with missing data would have also not been represented in final analyses, although their overall impact on the results would have been less.

Question Group	Number of Questions (SPIT)	Number of Questions (RISQ)
Personal Questions	8	2
Dental Health	9	
Heartburn/Acid Reflux	55	4
Dysphagia	8	3
Unexplained Weight Loss	4	2
Nausea and Vomiting	6	
Smoking History	14	1
Alcohol	9	
Obesity	4	
Diet	19	1
Physical Exercise	6	
Family History	12	1
Mental Wellbeing	14	1
Medical History	26	1
General Questions	15	1

Table 6-2: Breakdown of questions in each group for SPIT and RISQ questionnaires.

Our research group used existing validated questionnaires in several sections of the questionnaire, including the Hospital Anxiety and Depression Score (HADS), the University of California Los Angeles (UCLA) 3 item loneliness scale and the dysphagia score by Mellow and Pinkas.[128,201,202]

The RISQ questionnaire was a shortened version of the SPIT questionnaire with 17 questions (Table 6-2). The derivation of the RISQ questionnaire was from my work presented within this chapter (6.3.3 Results: Selection of Features). Subsequent endoscopic and histological results were linked to the patient's questionnaire responses. As participants underwent endoscopic investigation as part of their routine care, endoscopists and histopathologists were both blinded to questionnaire responses. Questionnaires and outcomes were recorded using a bespoke electronic software programme (TrialSense, London, UK).

Data Handling

All data handling and analysis was completed using R software version 4.1.2 (R Core Team, Vienna, Austria).[93] Prior to data analysis the dataset was manually cleaned for data input issues and errors. I excluded any fields which had greater than 20% of responses missing. I

performed missing data imputation using the 'missForest' package.[203] The process flow for my study is shown in Figure 6-1.

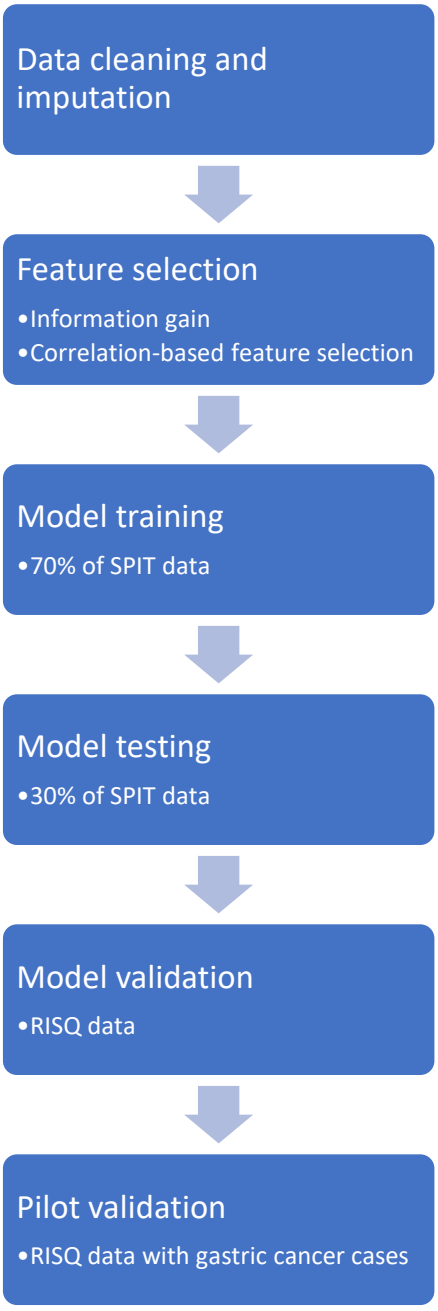


Figure 6-1: Process flow for model training, testing and validation.

Feature Selection

I used feature analysis to determine the most important predictors for oesophageal cancer. I used both information gain and chi-squared correlation-based feature selection from the

'FSelector' package, with a 50:50 weighting given to each feature before producing a final ranking.[181,204] In brief, information gain is an ML method where each feature is compared separately as to its correlation with the variable of interest. Correlation based feature selection assesses multiple features; some of the highly correlated set of features are removed, leaving a single representative feature.[181] Using these methods I was then able to select the most discriminating features to predict the presence of oesophageal cancer.

Model Training and Testing

I split the SPIT dataset in a 70:30 ratio for model training and testing respectively. I performed 10-fold cross-validation during model training. I used seven supervised ML methods from the R software 'caret' package ('caret' function in parenthesis):[205]

- Linear discriminant analysis (lda)
- Classification and regression tree (cart)
- K-nearest neighbour (knn)
- Support vector machines (svm)
- Random forest (rf)
- Logistic regression (glm)
- Regularised logistic regression (glmnet)

In particular, regularised logistic regression (glmnet) applies either ridge or lasso regularisation, which automatically introduces a degree of bias to the ML model and ultimately penalises the creation of more complex, overfitted models with multiple variables.[206] This is done by the addition of a regularisation coefficient, known as lambda (λ), to the overall model.[206] This additional step compared to standard logistic regression leads to reduced variance when the model is applied to new data and thus prevents overfitting.[206] The main difference between ridge and lasso regularisation is in its methodology: ridge

regularisation will square the values of the variables whereas lasso regularisation will take the absolute values of the variables. An elastic net mixing parameter termed alpha (α) can be adjusted during model development to allow both regularisation methods to be used without needing to commit to a method beforehand.[206]

I assessed the performance of the model using receiver operating characteristic curves (ROC) and calculated the area under the ROC curve (AUC) using the 'pROC' package.[207] Finally, to ensure that the model was weighted such that there was an increased penalty for misclassification of cancers, I applied a cost function to my final model using the 'ROCR' package.[208] The penalty for false negatives (i.e., missing a cancer case) was set at 50 times greater than false positives. This was then used to determine the ideal threshold above which the model would predict the presence of cancer. I assessed the performance of the cost function and the associated threshold using sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV).

Model Validation

Patients in the validation dataset were from the RISQ study. Our research group additionally enriched this cohort with further patients with confirmed oesophageal cancer. I also collected data from patients with confirmed gastric cancer to assess the performance of the model in this group.

Statistical Analysis

For risk prediction models, there is no generally accepted approach to estimate sample size requirements for derivation and validation of risk prediction models. Discrete variables are presented as numbers and percentages, while continuous variables are presented as mean and standard deviation (SD) or median and inter-quartile range (IQR). To compare between normal and cancer groups, I used t-tests or chi-squared (X^2) tests depending on the variable. A p-value of ≤0.05 was taken as significant.

6.3.3 Results

Demographics

A total of 807 patients were included in model training (566 patients) and testing (241 patients), while a further 294 patients were included in the independent model validation dataset. Full demographic information and breakdown is presented in Table 6-3. A total of 65, 27 and 42 oesophageal and gastro oesophageal junction (GOJ) cancer cases were included in the training, testing and validation datasets respectively. TNM staging information, where available, is presented in Table 6-4. 80%, 93% and 83% of cancers in the training, testing and validation datasets respectively were OACs. There was no statistically significant difference in the distribution on either the cancer site (X^2 = 2.21, p=0.33) and or the histology (X^2 = 5.06, p=0.28) within the three datasets. Notably, cancer patients were older and more likely to have had some quantified weight loss across the three datasets. Cancer patients across all three datasets were less likely to suffer from psychological disorders compared to non-cancer patients.

	Training Data (n=566)					Testing Data (n=241)					Validation Data (n=294)				
	Cancer Present	%	Cancer Absent	%	p value	Cancer Present	%	Cancer Absent	%	p value	Cancer Present	%	Cancer Absent	%	p value
n	65	11	501	89		27	11	214	89		42	14	252	86	
Age															
Mean	70.4		63.8		<0.001	70.6		63.5		0.002	70.8		52.1		<0.001
SD	10.1		13.3		TT	10.2		14		TT	9.96		17.1		TT
Sex															
Male (1)	45	69	208	42	<0.001	18	67	118	55	0.351	35	83	137	54	<0.001
Female (2)	20	31	293	58		9	33	96	45		7	17	115	46	
Cancer Site															
Oesophageal	52	80				25	93				35	83			
Gastro Oesophageal Junction	13	20				2	7				7	17			
Cancer Histology															
Adenocarcinoma	54	83				21	78				39	93			
Squamous Cell Carcinoma	5	8				3	11				3	7			
Other/Unknown	6	9				3	11				0	0			
Ethnicity															
White British/Irish/European	63	97	457	91	0.529	26	96	198	93	0.845	37	88	190	75	0.049
Mixed Race	1	2	1	0	FET	0	0	1	0	FET	0	0	5	2	FET
Asian	1	2	14	3		0	0	5	2		0	0	10	4	
Asian Other	0	0	8	2		0	0	1	0		4	10	13	5	
Black	0	0	11	2		1	4	5	2		1	2	12	5	
Other	0	0	10	2		0	0	4	2		0	0	22	9	
Current/Previous Smoking															
No (0)	26	40	265	53	0.061	10	37	101	47	0.416	10	24	137	54	<0.001
Yes (1)	39	60	233	47		17	63	112	52		32	76	115	46	
Unknown	0	0	3	1		0	0	1	0		0	0	0	0	

	Training Data (n=566)					Testing Data (n=241)					Validation Data (n=294)				
	Cancer Present	%	Cancer Absent	%	p value	Cancer Present	%	Cancer Absent	%	p value	Cancer Present	%	Cancer Absent	%	p value
Smoking Pack Years															
Mean	13.4		6.8		0.050	12.8		6.5		0.150	14.3		7.1		0.037
SD	22.9		14.6		TT	20.4		11.4		TT	15.6		15.0		TT
Unknown	15		89			3		39			17		14		
Body Mass Index (BMI)															
Mean	25.6		27.2		0.008	26.6		26.4		0.911	28.6		30.9		0.648
SD	4.5		6.0		TT	7.5		5.2		TT	8.5		10.2		TT
Unknown	2		19			2		4			38		223		
Chest Pain Begin															
No Chest Pain (0)	31	48	258	51	0.018	21	78	111	52	0.173	28	67	123	49	0.008
Less than 6 months (1)	20	31	69	14	FET	4	15	30	14	FET	7	17	13	5	FET
6 months to 1 year (2)	4	6	44	9		1	4	18	8		2	5	18	7	
1 to 5 years (3)	3	5	60	12		0	0	22	10		3	7	37	15	
5 to 10 years (4)	1	2	29	6		0	0	15	7		1	2	24	10	
10 to 20 years (5)	1	2	16	3		0	0	10	5		0	0	23	9	
More than 20 years (6)	1	2	16	3		0	0	5	2		1	2	14	6	
Unknown	4	6	9	2		1	4	3	1		0	0	0	0	
Regurgitation Begin															
No Regurgitation (0)	35	54	267	53	0.041	17	63	110	51	0.050	21	50	122	48	0.005
Less than 6 months (1)	15	23	54	11	FET	3	11	28	13	FET	9	21	11	4	FET
6 months to 1 year (2)	2	3	40	8		4	15	19	9		3	7	19	8	
1 to 5 years (3)	2	3	46	9		1	4	24	11		4	10	40	16	
5 to 10 years (4)	2	3	36	7		0	0	10	5		2	5	26	10	
10 to 20 years (5)	2	3	33	7		0	0	13	6		0	0	20	8	
More than 20 years (6)	2	3	20	4		1	4	6	3		3	7	14	6	

	Training Data (n=566)					Testing Data (n=241)					Validation Data (n=294)				
	Cancer Present	%	Cancer Absent	%	p value	Cancer Present	%	Cancer Absent	%	p value	Cancer Present	%	Cancer Absent	%	p value
Unknown	5	8	5	1		1	4	4	2		0	0	0	0	
Sour Taste Present															
No (0)	32	49	172	34	0.014	16	59	74	35	0.005	20	48	80	32	0.067
Yes (1)	25	38	279	56		6	22	117	55		22	52	172	68	
Unknown	8	12	50	10		5	19	23	11		0	0	0	0	
Sour Taste Frequency															
Never (0)	39	60	226	45	0.091	20	74	99	46	0.028	20	48	81	32	0.538
Annually or less (1)	5	8	23	5	FET	2	7	6	3	FET	0	0	4	2	FET
Few times a year (2)	3	5	48	10		2	7	11	5		4	10	40	16	
Few times a month (3)	5	8	47	9		1	4	26	12		6	14	38	15	
Few times a week (4)	9	14	78	16		1	4	37	17		7	17	45	18	
Daily (5)	3	5	67	13		1	4	29	14		5	12	44	17	
Unknown	1	2	12	2		0	0	6	3		0	0	0	0	
Frequency of Symptoms Preventing Eating and Drinking															
Never (0)	24	37	209	42	0.003	13	48	102	48	0.039	9	21	83	33	0.029
Few times a year (1)	3	19	25	5	FET	1	4	3	1	FET	4	10	30	12	FET
Few times a month (2)	5	8	110	22		2	7	50	23		1	2	34	13	
Few times a week (3)	5	8	69	14		1	4	30	14		9	21	33	13	
Daily (4)	21	32	82	16		7	26	26	12		19	45	72	29	
Unknown	7	11	6	1		3	11	3	1		0	0	0	0	
Swallowing Difficulty Present															
No (0)	19	29	256	51	0.002	10	37	109	51	0.398	18	43	154	61	0.040
Yes (1)	44	68	242	48		15	56	104	49		24	57	98	39	
Unknown	2	3	3	1		2	7	1	0		0	0	0	0	
Dysphagia Score															
No dysphagia (0)	19	29	256	51	0.003	10	37	109	51	0.486	18.0	43	155	62	0.538

	Training Data (n=566)					Testing Data (n=241)					Validation Data (n=294)				
	Cancer Present	%	Cancer Absent	%	p value	Cancer Present	%	Cancer Absent	%	p value	Cancer Present	%	Cancer Absent	%	p value
Dysphagia to solids (1)	19	29	120	24		8	30	47	22	FET	17	40	61	24	FET
Dysphagia to solids and semi-solids (2)	14	22	49	10		5	19	28	13		1	2	15	6	
Dysphagia to solids, semi-solids and liquids (3)	11	17	64	13		2	7	23	11		6	14	20	8	
Unknown	2	3	12	2		2	7	7	3		0	0	1	0	
Swallowing Pain Present															
No (0)	33	51	317	63	0.003	17	63	138	64	0.455	27	64	222	88	<0.001
Yes (1)	27	42	110	22		8	30	41	19	TT	15	36	30	12	
Unknown	5	8	74	15		2	7	35	16			0	0	0	
Unexplained Weight Loss Present															
No (0)	28	36	369	74	<0.001	17	63	148	69	0.466	18	43	200	79	<0.001
Yes (1)	36	55	128	26		10	37	64	30		24	57	52	21	
Unknown	1	2	4	1		0	0	2	1		0	0	0	0	
Weight Loss (kg)															
Mean	5.9		1.8		<0.001	2.6		2.5		0.867	6.0		1.5		0.009
SD	7.8		4.6		TT	4.6		5.5		TT	7.0		3.6		TT
Unknown	3		36			1		12			21		33		
Fruit and Vegetable Frequency															
Rarely or never (0)	3	5	8	2	0.003	0	0	8	4	0.516	8	19	11	4	0.538
Few times each month (1)	6	9	14	3	FET	0	0	2	1	FET	0	0	9	4	FET
Few times each week (2)	18	28	114	23		7	26	45	21		6	14	53	21	
Few times each day (3)	27	42	198	40		13	48	85	40		18	43	109	43	
5 a day or more (4)	10	15	162	32		5	19	71	33		10	24	70	28	
Unknown	1	2	5	1		2	7	3	1		0	0	0	0	
Known Psychological Disorders															
No (0)	52	80	324	65	0.002	24	89	137	64	0.008	32	76	147	58	0.043

	Training Data (n=566)					Testing Data (n=241)					Validation Data (n=294)				
	Cancer Present	%	Cancer Absent	%	p value	Cancer Present	%	Cancer Absent	%	p value	Cancer Present	%	Cancer Absent	%	p value
Yes (1)	7	11	153	31		1	4	63	29		10	24	105	42	
Unknown	6	9	24	5		2	7	14	7		0	0	0	0	
Butterflies Feeling															
Not at all (0)	25	38	153	31	0.362	15	56	80	37	0.257	26	62	107	42	0.076
Occasionally (1)	26	40	219	44	FET	8	30	81	38	FET	7	17	82	33	FET
Quite often (2)	4	6	47	9		0	0	13	6		2	5	24	10	
Very often (3)	0	0	14	3		0	0	7	3		3	7	23	9	
Unknown	10	15	68	14		4	15	33	15		4	10	16	6	
Read Local Newspaper															
No, never (0)	16	25	101	20	0.048	6	22	41	19	0.888	16	38	131	52	0.270
Yes, rarely (1)	17	26	86	17		5	19	46	21		7	17	31	12	
Yes, sometimes (2)	10	15	133	27		5	19	54	25		8	19	49	19	
Yes, often (3)	14	22	157	31		9	33	68	32		11	26	41	16	
Unknown	8	12	24	5		2	7	5	2		0	0	0	0	

Table 6-3: Demographic and questionnaire responses for included patients in training, testing and validation data sets of top ranked features.

Numbers in brackets following feature (e.g. no, never (0)) denote coding used for model development. p-values are for chi-squared tests unless otherwise stated. FET= Fisher's Exact Test. TT=Two sample t-test

	Training Data (n=65)		**Testing Data (n=27)**		**Validation Data (n=42)**	
	n	%	n	%	n	%
T Staging						
T1	7	11	2	7.4	2	4.8
T2	4	6.2	0	0	1	2.4
T3	10	15	3	11	6	14
T4	3	4.6	1	3.7	4	9.5
Unknown	41	63	21	78	29	69
N Staging						
N0	8	12	2	7.4	7	17
N1	9	14	3	11	3	7.1
N2	3	4.6	1	3.7	3	7.1
N3	2	3.1	0	0	0	0
Unknown	43	66	21	78	29	69
M Staging						
M0	16	25	4	15	11	26
M1	4	6.2	1	3.7	2	4.8
Unknown	45	69	22	82	29	69

Table 6-4: TNM staging information for oesophageal and gastro-oesophageal junction (GOJ) cancer cases.

Selection of Features

251 features were available for analysis, which includes data from the questionnaire, endoscopy and histology results. Prior to imputation I removed 101 features which had data missing in more than 20% of patients. I subsequently selected 17 features which were top ranked for the prediction of oesophageal and GOJ cancers (Table 6-5). Top ranked features were multidimensional and included demographic information, symptom, psychological and food related variables. These top ranked features were selected for ML model development.

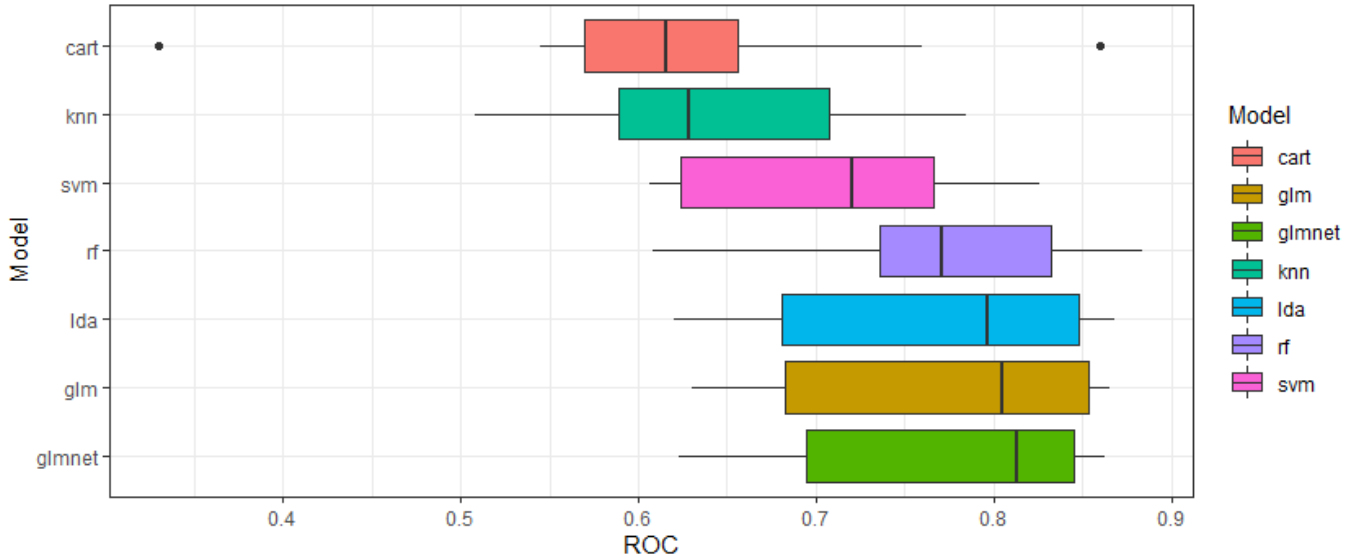
Feature	Information Gain	Correlation Based Feature Selection	Overall Rank
Frequency of Symptoms Preventing Eating and Drinking	0.015	0.175	1=
Unexplained Weight Loss (kg)	0.014	0.184	1=
Age	0.016	0.167	3
Unexplained Weight Loss Present	0.013	0.167	4
Known Psychological Disorders	0.013	0.149	5
Sour Taste Frequency	0.012	0.152	6
Chest Pain Begin	0.012	0.147	7=
Sex	0.011	0.148	7=
Fruit and Vegetable Frequency	0.009	0.136	9
Regurgitation Begin	0.008	0.127	10
Sour Taste Present	0.007	0.122	11=
Swallowing Pain Present	0.007	0.125	11=
Read Local Newspaper	0.006	0.113	13=
Dysphagia Score	0.006	0.116	13=
Butterflies Feeling	0.006	0.106	15
Swallowing Difficulty Present	0.005	0.096	16
Current/Previous Smoking	0.003	0.073	17

Table 6-5: Ranking of information gain and correlation-based feature selection to produce a final model.

Development of ML Model

Figure 6-2 demonstrates the distribution of area under the receiver operating characteristic curve (AUC) after 10-fold cross-validation in the training dataset, as well as median and IQR for AUC, sensitivity and specificity for the seven different ML methods. The best performing model with the highest median AUC was regularised logistic regression (AUC: 0.81, IQR: 0.69-0.85). The regularised logistic regression model was associated with parameters of α=0.1 and

λ=0.0149. I selected this model for further testing and validation and for determining appropriate cut offs in the cost function.



Model	Median ROC (IQR)	Median Sensitivity (IQR)	Median Specificity (IQR)
Classification and Regression Tree (cart)	0.62 (0.57-0.66)	0.99 (0.98-0.98	0.14 (0.04-0.17)
k-Nearest Neighbour (knn)	0.63 (0.59-0.71)	1.00 (0.99-1.00)	0.00 (0.00-0.00)
Support Vector Machines (svm)	0.72 (0.62-0.77)	1.00 (0.98-1.00)	0.17 (0.14-0.29)
Random Forest (rf)	0.77 (0.74-0.83)	1.00 (1.00-1.00)	0.07 (0.00-0.17)
Linear Discriminant Analysis (lda)	0.79 (0.68-0.85)	0.98 (0.97-1.00)	0.23 (0.14-0.33)
Logistic Regression (glm)	0.80 (0.68-0.85)	0.98 (0.98-1.00)	0.17 (0.17-0.29)
Regularised Logistic Regression (glmnet)	0.81 (0.69-0.85)	1.00 (0.98-1.00)	0.14 (0.04-0.17)

Figure 6-2: Boxplot of each model and their spread of area under the receiver operating characteristic curve (ROC) for each model. Table demonstrates median ROC, sensitivity and specificity and their respective inter-quartile range (IQR) during model development using 10-fold cross-validation.

Testing and Validation of Model

Figure 6-3 demonstrates the receiver operating characteristic (ROC) curve for regularised logistic regression, with the final model reapplied to the training dataset for reference purposes and subsequently tested on both the testing and independent validation datasets. For the testing data set the model achieved an AUC of 0.71 (95% CI: 0.61-0.81). For validation, the model achieved an AUC of 0.92 (95% CI: 0.88-0.96). Table 6-6 demonstrates the sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) for each dataset with the cost function applied. My final model achieved a sensitivity and specificity of 96.3% and 20.1% for the testing dataset respectively. For the validation dataset, the sensitivity and specificity were 97.6% and 59.1% respectively. The final model with coefficients is presented in Table 6-7.

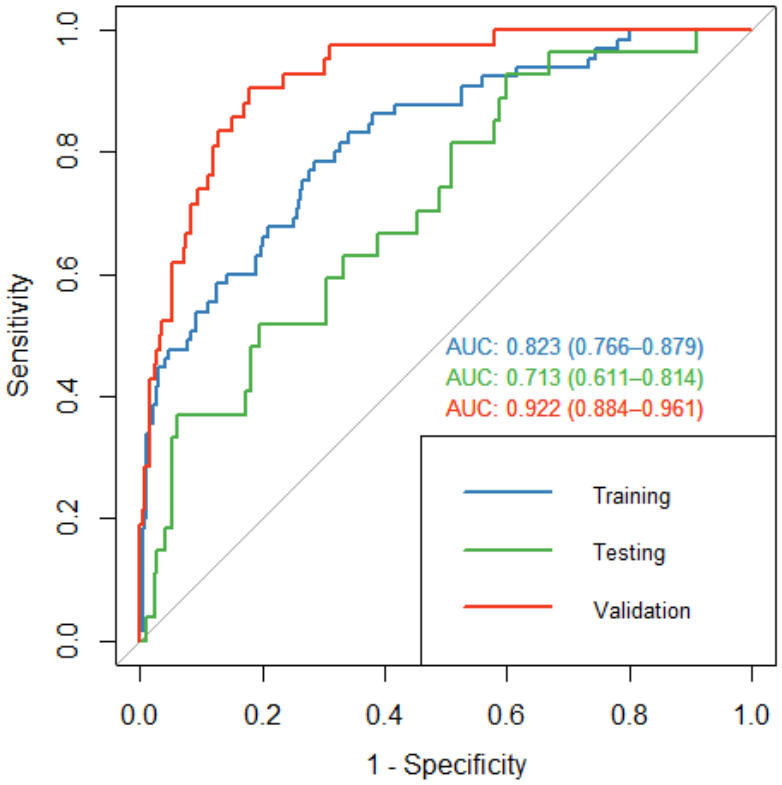


Figure 6-3: ROC curve for training, testing and validation datasets for regularised logistic regression for predicting oesophageal and GOJ cancer

Dataset	Sensitivity (95% CI) (%)	Specificity (95% CI) (%)	Positive Predictive Value (95% CI) (%)	Negative Predictive Value (95% CI) (%)
Training Data	96.9 (89.3-99.6)	24.4 (20.7-28.4)	14.3 (11.1-17.9)	98.4 (94.3-99.8)
Testing Data	96.3 (81.0-99.9)	20.1 (14.9-26.1)	13.2 (8.8-18.7)	97.7 (88.0-99.9)
Validation Data	97.6 (87.4-99.9)	59.1 (52.8-65.3)	28.5 (21.3-36.6)	99.3 (96.3-100.0)
Validation Data (Gastric Cancer)	75.0 (42.8-94.5)	59.1(52.8-65.3)	8.0 (3.7-14.7)	98.0 (94.3-99.6)

Table 6-6: Sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) with 95% confidence intervals (95% CI) for regularised logistic regression. This was set at a cut off of 0.03 (false negatives are weighted 50x more than false positive cases).

Feature	Coefficient
Intercept	*-2.539*
Age	0.440
Female Sex	-0.416
Chest Pain Begin	-0.036
Regurgitation Begin	-0.047
Sour Taste Present	(feature dropped)
Sour Taste Frequency	-0.223
Frequency of Symptoms Preventing Eating and Drinking	0.135
Swallowing Difficulty Present	0.155
Dysphagia Score	0.187
Swallowing Pain Present	0.222
Unexplained Weight Loss Present	0.263
Unexplained Weight Loss (kg)	0.183
Current/Previous Smoking	0.087
Fruit and Vegetable Frequency	-0.228
Known Psychological Disorders	-0.335
Butterflies Feeling	-0.151
Read Local Newspaper	-0.282

Table 6-7: Final model from regularised logistic regression where α=0.1 and λ=0.0149.

Performance of Model by Histological Subtype

I assessed the performance of my model on different histological subtypes of cancer. The model was able to predict correctly all 3 OSCC cases in the testing dataset and all 5 OSCC cases in the validation dataset as cancer. For OAC, all 21 cases were correctly predicted in the testing dataset, while all but one case (38/39; 97.4%) was correctly predicted as cancer in the validation dataset.

Pilot Testing with Gastric Cancer Cases

In a further sub analysis, I assessed if my model was able to predict the presence of gastric cancers. Demographic information for this subgroup of patient is presented in Table 6-8 and includes 12 gastric cancer cases with the same controls as the original validation dataset. Figure 6-4 demonstrates the ROC curve for validation where it achieved an AUC of 0.78 (95% CI: 0.62-0.95). At the same cut off level, this led to a sensitivity of 75.0% (Table 6-6).

	Pilot Data (Gastric) (n=294)				
	Cancer Present	%	Cancer Absent	%	p value
n	12	5	252	95	
Age					
Mean	63.2		52.1		0.032
SD	15.5		17.1		TT
Sex					
Male (1)	35	83	137	54	<0.001
Female (2)	7	17	115	46	
Cancer Diagnosis					
Gastric	12	100			
Cancer Histology					
Adenocarcinoma	12	100			
Squamous Cell Carcinoma	0	0			
Other/Unknown	0	0			
Ethnicity					
White British/Irish/European	9	75	190	75	0.819
Mixed Race	0	0	5	2	FET
Asian	0	0	10	4	
Asian Other	0	0	13	5	
Black	0	0	12	5	
Other	3	25	22	9	

	Pilot Data (Gastric) (n=294)				
	Cancer Present	%	Cancer Absent	%	p value
Current/Previous Smoking					
No (0)	4	33	137	54	0.258
Yes (1)	8	67	115	46	
Unknown	0	0	0	0	
Smoking Pack Years					
Mean	3.4		7.1		0.163
SD	6.4		15.0		TT
Unknown	4		14		
Body Mass Index (BMI)					
Mean	23.5		30.9		0.165
SD	5.8		10.2		
Unknown	9		223		
Chest Pain Begin					
No Chest Pain (0)	12	100	123	49	0.182
Less than 6 months (1)	0	0	13	5	FET
6 months to 1 year (2)	0	0	18	7	
1 to 5 years (3)	0	0	37	15	
5 to 10 years (4)	0	0	24	10	
10 to 20 years (5)	0	0	23	9	
More than 20 years (6)	0	0	14	6	
Unknown	0	0	0	0	
Regurgitation Begin					
No Regurgitation (0)	9	75	122	48	0.726
Less than 6 months (1)	1	8	11	4	FET
6 months to 1 year (2)	0	0	19	8	
1 to 5 years (3)	1	8	40	16	
5 to 10 years (4)	1	8	26	10	
10 to 20 years (5)	0	0	20	8	
More than 20 years (6)	0	0	14	6	
Unknown	0	0	0	0	
Sour Taste Present					
No (0)	4	33	80	32	1.000
Yes (1)	8	67	172	68	FET
Unknown	0	0	0	0	
Sour Taste Frequency					
No (0)	4	33	81	32	0.797
Annually or less (1)	0	0	4	2	FET
Few times a year (2)	1	8	40	16	
Few times a month (3)	3	25	38	15	
Few times a week (4)	3	25	45	18	
Daily (5)	1	8	44	17	
Unknown	0	0	0	0	
Frequency of Symptoms Preventing Eating and Drinking					
Never (0)	6	50	83	33	0.554

	Pilot Data (Gastric) (n=294)				
	Cancer Present	%	Cancer Absent	%	p value
Few times a year (1)	0	0	30	12	FET
Few times a month (2)	2	17	34	13	
Few times a week (3)	2	17	33	13	
Daily (4)	2	17	72	29	
Unknown	0	0	0	0	
Swallowing Difficulty Present					
No (0)	11	92	154	61	0.035
Yes (1)	1	8	98	39	FET
Unknown	0	0	0	0	
Dysphagia Score					
No dysphagia (0)	11.0	92	155	62	0.797
Dysphagia to solids (1)	1	8	61	24	FET
Dysphagia to solids and semi-solids (2)	0	0	15	6	
Dysphagia to solids, semi-solids and liquids (3)	0	0	20	8	
Unknown	0	0	1	0	
Swallowing Pain Present					
No (0)	12	100	222	88	0.371
Yes (1)	0	0	30	12	FET
Unknown	0	0	0	0	
Unexplained Weight Loss Present					
No (0)	3	25	200	79	<0.001
Yes (1)	9	75	52	21	FET
Unknown	0	0	0	0	
Weight Loss (kg)					
Mean	9.0		1.5		0.047
SD	7.0		3.6		TT
Unknown	6		33		
Fruit and Vegetable Frequency					
Rarely or never (0)	2	17	11	4	0.797
Few times each month (1)	0	0	9	4	FET
Few times each week (2)	3	25	53	21	
Few times each day (3)	5	42	109	43	
5 a day or more (4)	2	17	70	28	
Unknown	0	0	0	0	
Known Psychological Disorders					
No (0)	9	75	147	58	0.370
Yes (1)	3	25	105	42	FET
Unknown	0	0	0	0	
Butterflies Feeling					
Not at all (0)	6	50	107	42	0.843
Occasionally (1)	3	25	82	33	FET
Quite often (2)	1	8	24	10	
Very often (3)	0	0	23	9	
Unknown	2	17	16	6	

	Pilot Data (Gastric) (n=294)				
	Cancer Present	%	Cancer Absent	%	p value
Read Local Newspaper					
No, never (0)	5	42	131	52	0.671
Yes, rarely (1)	1	8	31	12	FET
Yes, sometimes (2)	4	33	49	19	
Yes, often (3)	2	17	41	16	
Unknown	0	0	0	0	

Table 6-8: Demographic table for patients diagnosed with gastric cancer included in pilot validation.

Numbers in brackets following feature e.g. no, never (0) denote coding used for model development. p-values are for chi-squared tests unless otherwise stated. FET= Fisher's Exact Test. TT=Two sample t-test

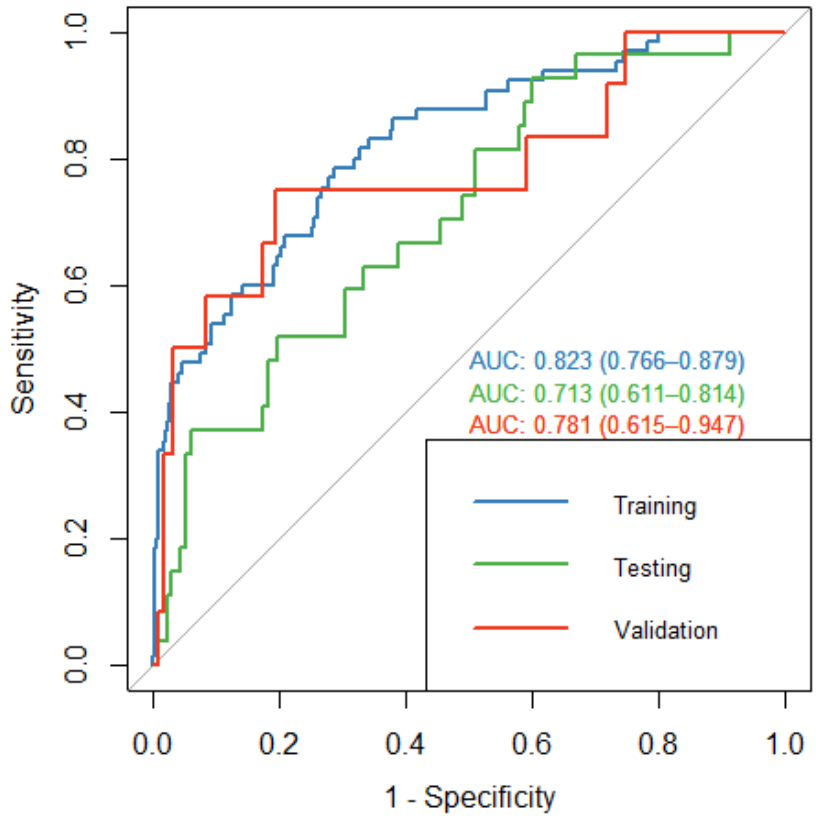


Figure 6-4: ROC curve for training, testing and validation datasets for regularised logistic regression for predicting gastric cancer.

6.3.4 Discussion

Using an ML approach, I created a risk stratification tool using questionnaire only factors which could be used as a diagnostic aid to identify patients who may have oesophageal or

gastro-oesophageal junction (GOJ) cancer. The model is based on demographic factors such as age and sex which are routinely captured, as well as alarm features such as dysphagia and unintentional weight loss which are currently included in National Institute for Health and Care Excellence (NICE) UGI cancer referral guidelines.[185] My model further incorporates additional features, such as known psychological disorders, which gives an additional richness to the model and helps improve its performance. In particular, I have also demonstrated the robustness of the models by both testing and independently validating the model. The model achieved a sensitivity of over 96% for detecting cancer, while specificity ranged from 20% in the testing dataset to 59% in the validation dataset. This large range of results for specificity is likely to be a result of both the relatively small numbers of cancers in both datasets, and also cancer and non-cancer patients were not equally distributed in the datasets. In addition, I also trialled my model on gastric cancer patients. Although the numbers were small it identified 9/12 (75%) gastric cancer patients correctly. It may be possible to enhance accuracy for this condition by capturing symptoms specific to gastric cancer such as early satiety and symptoms of anaemia.[43]

My work adds to previous research; ML algorithms have previously been developed in a Chinese population to predict the risk of UGI lesions.[209] However, the study incorporated a combined endpoint of UGI pathology. This included both gastric cancer, which had a prevalence of 0.3% in the study population, and more benign pathology such as gastric erosions, which had a prevalence of 22%.[209] This model differs by emphasising cancers, which I believe will have a greater utility especially in the triaging of suspected cancer 2ww referrals and prioritising procedures which are likely to have a greater yield for serious pathology.

My model appears to perform well against currently used triaging methods. In the validation dataset the model achieved a PPV of 28.5% and an NPV of 99.3%. This improves upon using alarm features solely, which on their own have PPVs ranging from 0 to 11% for oesophageal

and gastric cancers, with considerable heterogeneity between studies.[83] Even when using a combination of alarm features, PPVs still remain at around 10%.[83] In addition, the model also improves on the EDS.[169,171] During the COVID-19 pandemic the EDS has been used for triaging urgent UGI referrals, and a prospective series demonstrated a sensitivity of 96.7% and a specificity of 32.6%, although the authors also included 10 gastric and one duodenal cancer in the outcome.[170] This equated to a PPV of 9.7% and NPV of 99.3% for the detection of UGI cancer.[170] One major advantage of my approach is that it has a higher PPV compared to the EDS, which could be especially advantageous in reducing the number of normal procedures being classed as urgent or being performed. My model also compares favourably with models recently validated for the detection of incident cases of OAC and GOJ adenocarcinoma. The best model in validation achieved an AUC of 0.73, compared to 0.92 in my study.[188,193]

6.3.5 Limitations of the Study

First, some patients were recruited after the diagnosis of cancer was known. Subsequently, these patients may be subject to recall bias and report higher rates of symptoms. While it would be preferable to perform prospective recruitment, there are challenges associated with this as UGI cancer has a low incidence in the UK.[210] In particular, the validation cohort was specifically enriched with patients with known oesophageal cancer, including some who had late stage diagnoses. This could well explain as to why the validation cohort outperformed training and testing cohorts. As the model was primarily developed in patients who were referred under 2ww pathways into secondary care, this already excludes patients who are earlier in the pathway, such as with dyspepsia in primary care but not referred onwards under 2ww pathway. This would therefore limit the generalisability of the model especially within primary care.

Furthermore, the outcome of the model was binary; 'cancer' compared to 'no cancer'. Although it would be preferable to create models based on stages of cancer, such as early compared to late stage, this was restricted by incomplete data. As shown in Table 6-4, T staging was missing in 63% of training data cases and 78% of testing data cases. There were only 5 patients in the training and testing cohort who had an overall stage of T2N0M0 or less, for which curative surgical resection can usually be offered.[8] Moreover, oesophageal cancer often presents late; 70% present at stage III or stage IV disease.[63] Early disease is often asymptomatic and symptoms such as progressive dysphagia and weight loss are late manifestations. Therefore, there are likely to be distinct differences between early and late stage cancers.[21,63] In addition, there are potential difficulties in identifying individuals with early oesophageal cancer, as they may have yet to present to medical services, which would have affected recruitment into the study. Consequently, later stage cancers were overrepresented in the final models, with the risk of missing features of early stage cancers, reducing its overall clinical utility in early detection of cancer.

Second, I would have preferred to directly test the performance of the EDS on my cohort of patients and therefore create head-to-head comparisons. However, I was unable to extrapolate for one of the elements of the score (dysphagia localises to neck) with my existing data. Third, as this study is based in the UK, where OAC is the predominant subtype, my scoring system may be less applicable to other nations or territories where other histological subtypes are more common. Fourth, my model incorporates a large number of features which increases its complexity during clinical use, although with the rise of electronic health record systems, patients are increasingly prepared to complete health questionnaires online. This can rapidly generate rich datasets for the clinician to review easily and is starting to make its way into routine clinical use. Finally, this model needs further validation in larger cohorts to ensure its performance remains consistent and slight

adjustment to enhance its performance against gastric cancers. This is especially important to ensure greater consistency in the calculation of model performance metrics.

6.3.6 Online Calculator

An online version of the model is available via:

https://endopredict.shinyapps.io/endopredict/

6.4 Epigenetic Factors

6.4.1 Introduction

The development of next generation sequencing has allowed for a large number of patient samples to be investigated for genome wide changes at the same time.[211] One particular area of interest is epigenetic changes which occur in the promoter region of genes, which have a key role in determining gene expression.[211] In particular, these epigenetic changes occur at the 5' end of cytosine in CpG (Cytosine-phosphate-Guanine) islands, whereby after methylation the corresponding gene becomes silenced and is not transcribed.[211] Altered gene expression could be especially important in the case of tumour suppressor genes, which if not transcribed could increase the risk of developing cancer. Indeed, methylation changes have been found in patient with both OAC and OSCC,[212] as well as in the precursor to OAC, Barrett's oesophagus (BO),[213] suggesting a role of epigenetic changes and cancer development. These epigenetic changes thus have the potential to act as biomarkers and allow for the risk stratification of patients who may be at increased risk of developing cancer.

To ensure there is a common nomenclature between different methylation sites, each methylation site has a specific CpG locus cluster ID attached, similar to the identifiers given for single nucleotide polymorphisms (SNPs), where DNA methylation can occur within the human genome.[211] CpG loci can be highly correlated with each other, and thus it may be possible to screen a large number of CpG loci by simply running an assay on a smaller sample

number of CpG loci.[211] In addition, upscaling this technology combined with high throughout methods could mean that a large number of patients could be screened simultaneously, making them potentially suitable for use in either targeted risk group or general population screening.

Previously, analysis of tissue microRNAs within saliva have shown promise as potential biomarkers for oesophageal cancer.[214] Saliva sample collection is very acceptable to patients, and has the advantage that patients could collect their own sample with minimal guidance from their own home. In addition, saliva collection is also non-invasive, unlike UGI endoscopy, which is currently gold standard for the diagnosis and surveillance of BO and diagnosis of OAC and OSCC. Finally, salivary DNA stored within the preservative of self-collection kits can remain stable for at least 8 months at room temperature. This allow ample time for sample collection, and simplifies the process of transportation and handling of patient samples, and also means samples could be stored before being batch processed.[215].Given there is only around 0.5% annual risk of patient with BO progressing to OAC, and that oesophageal cancer often presents late within the disease, there is the potential that patient characteristics or symptom severity, combined with salivary epigenetic analysis could lead to more accurate prediction of a patient's risk. This could lead to more individualised care for patients.[212] There may be huge implications as there could be both better selection of patients to undergo both diagnostic and surveillance UGI endoscopy, as well as improving diagnosis and disease outcomes including cancer survival.

6.4.2 Aims

My aims were:

- Use the same machine learning methods as with the questionnaire data to assess the performance of the different loci in predicting the presence of oesophageal cancer.

- Combine the results of the questionnaire data with epigenetic data to determine if there can be an improvement in model performance in predicting the presence of oesophageal cancer.

6.4.3 Methods

Patient recruitment

Patients were recruited as part of the SPIT study, where on top of completing a symptom risk factor questionnaire patients provided a saliva sample. This was done using the Oragene DNA OG/600 saliva collection kit (DNAGenotek, Ottawa, Canada). Patients were asked to fast, refrain from smoking and chewing gum for an hour prior to sample collection. Patients were given a maximum of 5 minutes to collect their sample. Samples were then stored at -80°C until used for analysis.

Epigenetic Analysis

Epigenetic analysis was carried out by University College London (UCL) Genomics using Illumina EPIC methylation assays. Two methylation array batches were performed in July 2020 and January 2022 respectively (the large time gap was due to COVID-19 related delays). Each array was run in a case control fashion, whereby known cancer cases were appropriately matched with non-cancer cases. To estimate methylation status, Illumina analyses a pair of probes (methylated and unmethylated) and measures the difference between intensities of the methylated and unmethylated alleles at the CpG site.[216] I used the M-value as a metric in measuring methylation levels. This is calculated by taking the log2 ratio of the intensities of methylated probe versus unmethylated probe, where signal intensities range from 0 (complete absence of methylation) and 1 (complete presence of methylation).[216]

The formula is:

$$M\text{-}value = \log_2 \frac{\mathrm{M}}{\mathrm{U} + \mathrm{M}}$$

Where:

- U = signal from unmethylated cytosine
- M = signal from methylated cytosine

The M-value is thought to be more statistically valid compared to calculating the percentage of methylation, also known as the Beta-value.[216]

Quality Control Measures

To ensure consistency and reproducibility, numerous quality control measures were used during epigenetic analysis. These included:

- Quality control measures to ensure the quality of saliva sample and the quantity of DNA in it prior to selection for suitability for analysis.
- Matching of cancer cases to normal cases. Matched features included age, sex, body mass index (BMI), smoking status, alcohol consumption, proton pump inhibitor (PPI) use and presence of heartburn symptoms.
- Running 10 duplicate samples in both array batches.
- Minimisation of batch effects. This is when there is systematic technical variation which could be caused by processing samples on different days, use of different reagent lots or the arrangement of samples across assays.[217] To do this, the research group bioinformatician calculated the average M-value for each row and created a batch row variable. Subsequently he subtracted the batch row variable from the M-value associated with the sample and CpG locus. This then created residual figures

ranging from negative to positive, where the more positive the figure, the greater the methylation associated with that sample and CpG locus.

Statistical Analysis

Identification of the most discriminating methylation sites between healthy individuals and cancer cases was completed by the group bioinformatician. A summary of the analysed data was then provided to me so I would be able to incorporate epigenetic features during model generation. In brief, a representative probe for every gene was found using Weighted Gene Co-Expression Network Analysis (WGCNA), which is designed to find clusters of highly correlated genes.[218] This analysis was repeated multiple times but each time 5 patient samples were deliberately excluded to improve variation and reproducibility. Best scoring probes were then selected and a random forest classifier was used to determine its performance in detecting oesophageal cancers in a testing and validation dataset.

Model Development

I used an ensemble of machine learning model methods including those I previously earlier in this chapter. These were (R software 'caret' function in parenthesis):[205]

- Linear discriminant analysis (lda)
- Classification and regression tree (cart)
- K-nearest neighbour (knn)
- Support vector machines (svm)
- Random forest (rf)
- Logistic regression (glm)
- Regularised logistic regression (glmnet)

As per my previous analysis, I used a 70:30 ratio split between data training and data testing. Unfortunately, my research group did not have enough patient samples available for

validation of this model, although recruitment of the SPIT study is currently ongoing. I used the ROC as the metric for model performance.

As done previously I set a cost function of 50x for false negatives (i.e. it is 50 times more costly to miss a single case of cancer) to determine the optimum cut offs and therefore associated sensitivity and specificity.

6.4.4 Results

141 patients had epigenetic data and questionnaire data which could be used for model generation.

Table 6-9 shows the demographic information and questionnaire responses for these patients. This was split into 100 cases for the training dataset (69 non-cancer patients and 31 cancer patients) and 41 cases for the testing dataset (29 non-cancer patients and 12 cancer patients). Note this is a subset of the patients presented in Table 6-3, although the training and testing data split was done afresh in the new, smaller patient cohort.

	Training Data (n=100)					Testing Data (n=41)				
	Cancer Present	%	Cancer Absent	%	p value	Cancer Present	%	Cancer Absent	%	p value
n	31	31	69	69		12	29	29	71	
Age										
Mean	71.9		66.0		0.007	66.8		64.6		0.513
SD	8.4		12.6		TT	9.54		9.8		TT
Sex										
Male (1)	24	77	46	67	0.396	8	67	21	72	0.721
Female (2)	7	23	23	33		4	33	8	28	FET
Cancer Site										
Oesophageal	20	65				11	92			
Gastro Oesophageal Junction	11	35				1	8			
Cancer Histology										
Adenocarcinoma	30	97				12	100			
Squamous Cell Carcinoma	1	3				0				
Other/Unknown	0					0				
Ethnicity										
White British/Irish/European	30	97	69	100	0.371	12	100	29	100	1.00
Mixed Race	0	0	0	0	FET	0	0	0	0	FET
Asian	1	3	0	0		0	0	0	0	
Asian Other	0	0	0	0		0	0	0	0	
Black	0	0	0	0		0	0	0	0	
Other	0	0	0	0		0	0	0	0	
Current/Previous Smoking										
No (0)	6	19	35	51	0.006	4	33	13	45	0.729
Yes (1)	25	81	34	49		8	67	16	55	FET

	Training Data (n=100)					Testing Data (n=41)				
	Cancer Present	%	Cancer Absent	%	p value	Cancer Present	%	Cancer Absent	%	p value
Unknown	0	0	0	0		0	0	0	0	
Smoking Pack Years										
Mean	20.9		6.57		0.016	7.92		6.77		0.828
SD	26.4		14		TT	14		13.5		TT
Unknown	6		13			2		6		
Body Mass Index (BMI)										
Mean	28.5		26.6		0.044	24.5		27.7		0.094
SD	4.29		4.22		TT	5.66		3.49		TT
Unknown	1		0			0		2		
Chest Pain Begin										
No Chest Pain (0)	15	48	43	62	0.079	11	92	21	72	0.469
Less than 6 months (1)	9	29	4	6	FET	0	0	4	14	FET
6 months to 1 year (2)	1	3	6	9		1	8	0	0	
1 to 5 years (3)	2	6	7	10		0	0	2	7	
5 to 10 years (4)	1	3	2	3		0	0	1	3	
10 to 20 years (5)	1	3	3	4		0	0	0	0	
More than 20 years (6)	2	6	3	4		0	0	1	3	
Unknown	0	0	1	1		0	0	0	0	
Regurgitation Begin										
No Regurgitation (0)	15	48	41	59	0.183	8	67	19	66	0.359
Less than 6 months (1)	7	23	6	9	FET	3	25	1	3	FET
6 months to 1 year (2)	4	13	3	4		0	0	1	3	
1 to 5 years (3)	1	3	6	9		1	8	2	7	
5 to 10 years (4)	0	0	4	6		0	0	4	14	
10 to 20 years (5)	1	3	5	7		0	0	1	3	

	Training Data (n=100)					Testing Data (n=41)				
	Cancer Present	%	Cancer Absent	%	p value	Cancer Present	%	Cancer Absent	%	p value
More than 20 years (6)	2	6	4	6		0	0	1	3	
Unknown	1	3	0	0		0	0	0	0	
Sour Taste Present										
No (0)	16	52	30	43	0.475	8	67	17	59	1.00
Yes (1)	13	42	37	54		4	33	11	38	
Unknown	2	6	2	3		0	0	1	3	
Sour Taste Frequency										
Never (0)	18	58	33	48	1.00	8	67	19	66	1.00
Annually or less (1)	4	13	6	9	FET	1	8	1	3	FET
Few times a year (2)	2	6	9	13		2	17	1	3	
Few times a month (3)	3	10	5	7		0	0	1	3	
Few times a week (4)	3	10	6	9		1	8	4	14	
Daily (5)	1	3	7	10		0	0	3	10	
Unknown	0	0	3	4		0	0	0	0	
Frequency of Symptoms Preventing Eating and Drinking										
Never (0)	14	45	44	64	0.001	7	58	23	79	0.321
Few times a year (1)	3	19	5	7	FET	0	0	0	0	FET
Few times a month (2)	3	10	12	17		1	8	3	10	
Few times a week (3)	1	3	6	9		0	0	2	7	
Daily (4)	10	32	2	3		2	17	1	3	
Unknown	0	0	0	0		2	17	0	0	
Swallowing Difficulty Present										
No (0)	6	19	55	80	<0.001	4	33	20	69	0.045
Yes (1)	25	81	14	20		8	67	9	31	FET
Unknown	0	0	0	0		0	0	0	0	

	Training Data (n=100)					Testing Data (n=41)				
	Cancer Present	%	Cancer Absent	%	p value	Cancer Present	%	Cancer Absent	%	p value
Dysphagia Score										
No dysphagia (0)	6	19	55	80	<0.001	4	33	20	69	0.005
Dysphagia to solids (1)	12	39	5	7	FET	0	0	5	17	FET
Dysphagia to solids and semi-solids (2)	7	23	4	6		4	33	2	7	
Dysphagia to solids, semi-solids and liquids (3)	6	19	5	7		4	33	2	7	
Unknown	0	0	0	0		0	0	0	0	
Swallowing Pain Present										
No (0)	21	68	43	62	0.213	7	58	22	76	0.322
Yes (1)	9	29	8	12		3	25	3	10	FET
Unknown	1	3	18	26		2	17	4	14	
Unexplained Weight Loss Present										
No (0)	18	36	63	91	<0.001	5	42	26	90	0.003
Yes (1)	13	42	6	9		7	58	3	10	
Unknown	0	0	0	0		0	0	0	0	
Weight Loss (kg)										
Mean	4.88		0.69		0.016	6.95		0.98		0.033
SD	8.87		2.83		TT	8.2		4.8		TT
Unknown	1		1			0		1		
Fruit and Vegetable Frequency										
Rarely or never (0)	3	10	0	0	0.039	0	0	1	3	0.026
Few times each month (1)	1	3	1	1	FET	3	25	0	0	FET
Few times each week (2)	4	13	15	22		2	17	5	17	
Few times each day (3)	14	45	23	33		5	42	11	38	

	Training Data (n=100)					Testing Data (n=41)				
	Cancer Present	%	Cancer Absent	%	p value	Cancer Present	%	Cancer Absent	%	p value
5 a day or more (4)	9	29	30	43		1	8	12	41	
Unknown	0	0	0	0		1	8	0	0	
Known Psychological Disorders										
No (0)	31	100	52	75	0.002	9	75	20	69	0.694
Yes (1)	0	0	15	22	FET	2	17	9	31	FET
Unknown	0	0	2	3		1	8	0	0	
Butterflies Feeling										
Not at all (0)	17	55	24	35	0.053	4	33	11	38	0.851
Occasionally (1)	7	23	30	43	FET	4	33	10	34	FET
Quite often (2)	0	0	5	7		1	8	1	3	
Very often (3)	0	0	1	1		0	0	0	0	
Unknown	7	23	9	13		3	25	7	24	
Read Local Newspaper										
No, never (0)	13	42	12	17	0.001	4	33	8	28	0.213
Yes, rarely (1)	7	23	7	10		3	25	2	7	
Yes, sometimes (2)	2	6	25	36		0	0	5	17	
Yes, often (3)	8	26	23	33		4	33	13	45	
Unknown	1	3	2	3		1	8	1	3	

Table 6-9: Demographic and questionnaire responses for included patients in training and testing sets for patients who had both epigenetic and questionnaire date available.

Numbers in brackets following feature (e.g. no, never (0)) denote coding used for model development. p-values are for chi-squared tests unless otherwise stated. FET= Fisher's Exact Test. TT=Two sample t-test

Loci Identification

7 loci with epigenetic changes were found to be the most discriminatory between patients with confirmed OAC and those who did not have cancer. They were denoted as:

1. cg18973389
2. cg23347958
3. cg08267433
4. cg04096519
5. cg27123665
6. cg09771271
7. cg08692104

Notably, the non-cancer cohort included healthy volunteers as well as patients who underwent UGI endoscopy and had normal findings.

Feature Selection

I have retained the same 17 features as in my previous analysis which were most associate with oesophageal cancer. To recap these were:

1. Age
2. Sex (biological)
3. Current/Previous Smoking
4. Chest Pain Begin
5. Regurgitation Begin
6. Acid or Sour Taste Present
7. Acid or Sour Taste Frequency
8. Frequency of Symptoms Preventing Eating or Drinking Frequency
9. Presence of Swallowing Difficulty

10. Composite Dysphagia Score
11. Presence of Swallowing Pain
12. Presence of Unexplained Weight Loss
13. Weight Loss in Kilograms
14. Frequency of Consuming Fruit or Vegetables
15. Previous History of Psychological Disorders
16. Frequency of Having a Butterflies Feeling
17. Frequency of Reading a Local Newspaper

Creation of Machine Learning Models

I aimed to select the best machine learning model to compare under 3 scenarios

1. 7 epigenetic factors only
2. 17 questionnaire features only
3. 7 epigenetic factors and 17 questionnaire features combined

Analysis 1: Epigenetic Only Model

In my first analyses I only incorporated the 7 CpG loci and sex.

Analysis 1: Epigenetic Only Model: Model Generation

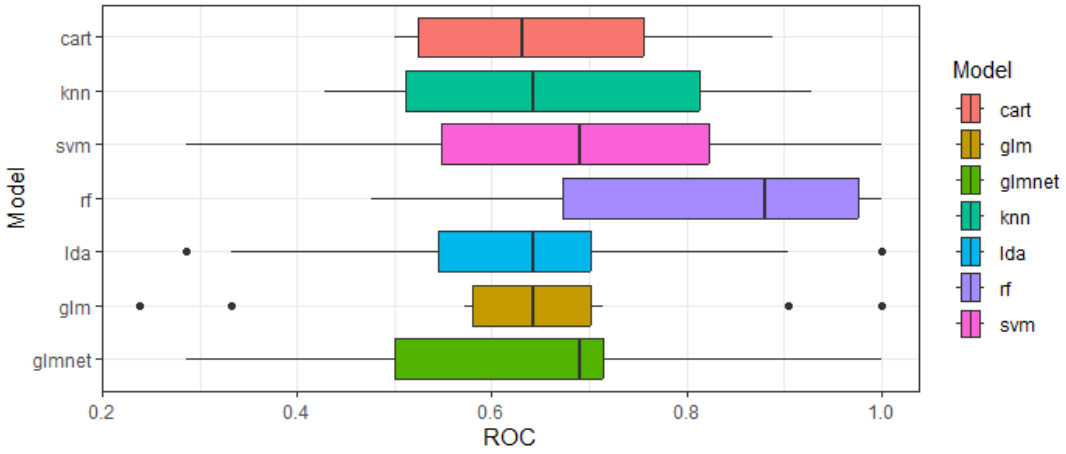


Figure 6-5: Boxplot of the seven machine learning models using epigenetic features in discriminating between non cancer and cancer cases.

Model	Median AUROC (IQR)	Median Sensitivity (IQR)	Median Specificity (IQR)
Classification and Regression Tree (cart)	0.63 (0.52-0.76)	0.86 (0.75-1.00)	0.33 (0.33-0.58)
k-Nearest Neighbour (knn)	0.64 (0.52-0.81)	0.85 (0.85-1.00)	0.33 (0.08-0.58)
Support Vector Machines (svm)	0.69 (0.55-0.82)	1.00 (1.00-1.00)	0.33 (0.00-0.58)
Random Forest (rf)	0.88 (0.67-0.98)	0.86 (0.84-1.00)	0.33 (0.33-0.67)
Linear Discriminant Analysis (lda)	0.64 (0.54-0.70)	0.86 (0.86-1.00)	0.33 (0.00-0.63)
Logistic Regression (glm)	0.64 (0.58-0.70)	0.85 (0.86-1.00)	0.33 (0.08-0.63)
Regularised Logistic Regression (glmnet)	0.69 (0.50-0.71)	0.92 (0.86-1.00)	0.17 (0.00-0.46)

Table 6-10: Table of median ROC, median sensitivity and median specificity of the seven machine learning models using epigenetic features in discriminating between non cancer and cancer cases.

The best model during generation was random forest (Table 6-10 and Figure 6-5) which demonstrated an average median ROC of 0.88 (IQR: 0.67-0.98) across 10-fold cross validation. Applying this model de novo achieved a mean AUC of 1.00 (95% CI: 1.00-1.00) for the training dataset and 0.747 (95%CI: 0.59-0.92) for the testing dataset (Figure 6-6).

Analysis 1: Epigenetic Only Model: Model Testing

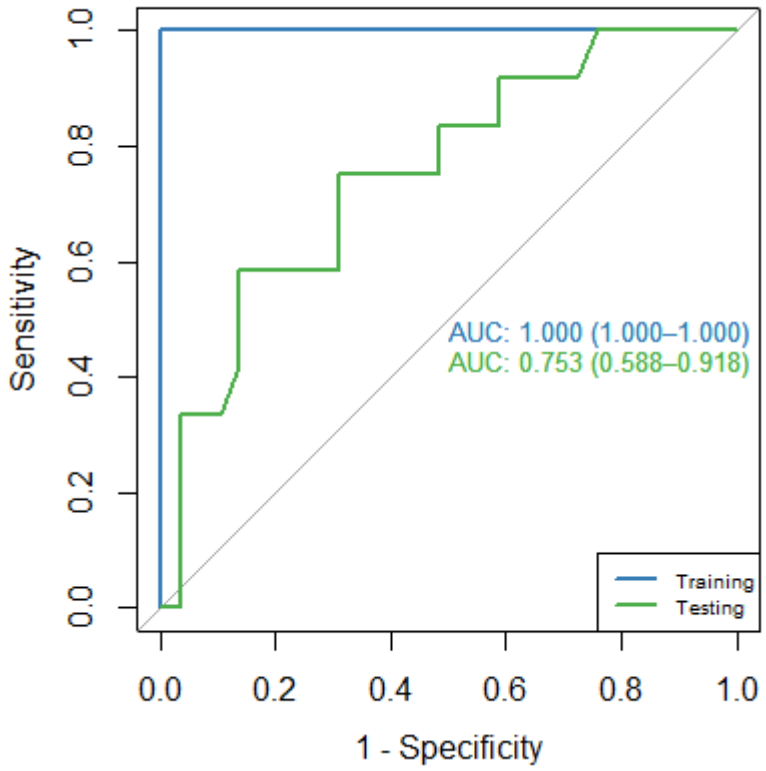


Figure 6-6: Area under the receiver operator curve (AUROC) for the random forest model for the training and testing dataset.

At a cut off of 0.5, the random forest model achieved a sensitivity of 1.00 (95%CI: 0.95-1.00) and a specificity of 1.00 (95%CI: 0.89-1.00) for the training dataset, and a sensitivity of 0.69 (95%CI: 0.49-0.85) and a specificity of 0.67 (95% CI: 0.35-0.90) for the testing dataset. Given the training dataset achieved an AUC of 1.00, it would suggest the data was overfitted; this is backed up by the DeLong test where the training and testing ROC curves demonstrated a statistically significant difference (D statistic = 2.93, df = 40, p-value = 0.006). This would imply the random forest model may be less reproducible and could suggest that the model was overfitted on the training dataset hence its performance was worse on the testing dataset.

To assess if a different model may be more reproducible during training and testing, I also assessed the effect of the second-best performing model: support vector machines. This achieved a median AUC of 0.69 (95% CI: 0.55-0.82) during model generation with 10-fold cross validation (Table 6-10). During model training this achieved an AUC of 0.92 (95% CI: 0.84-0.99) in the training dataset and AUC of 0.83 (95% CI: 0.69-0.96) (Figure 6-7).

At a cut off of 0.22, the support vector machines model achieved a sensitivity of 0.90 (95% CI: 0.74-0.98) and a specificity of 0.29 (95% CI: 0.19-0.41) when reapplied to the training dataset. For the testing dataset, the sensitivity was 1.00 (95% CI: 0.74-1.00) and specificity was 0.31 (95% CI: 0.15-0.51). DeLong's test demonstrated no statistically significant difference (D statistic = 1.10, df = 68.7, p=0.273) between the training and testing ROC curves.

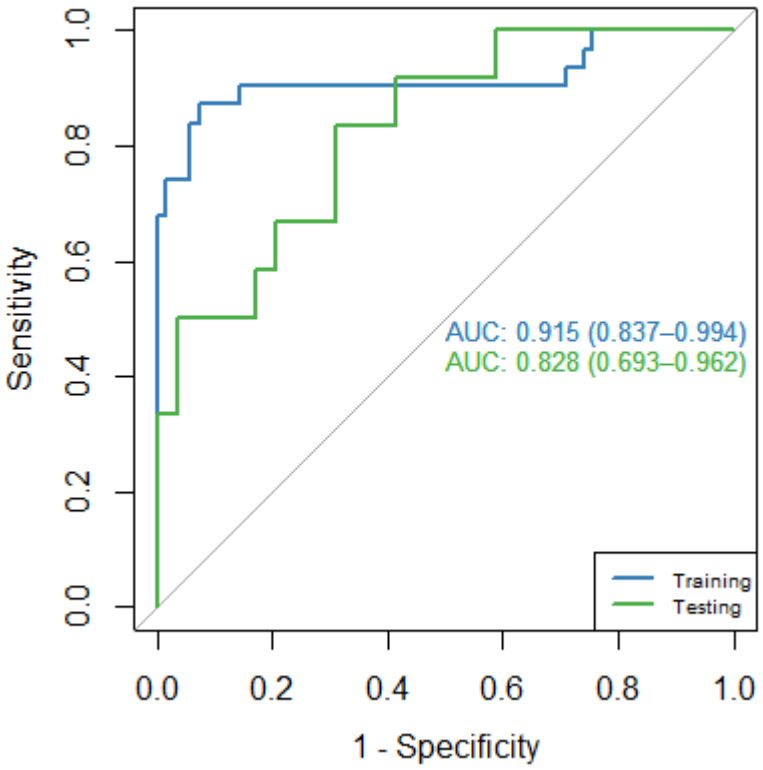


Figure 6-7: Area under the receiver operator curve (AUROC) for the support vector machines model for the training and testing dataset for analysis 1 (epigenetic features only).

Analysis 2: Questionnaire Features Only Model

My next analysis looked at incorporating the 17 questionnaire features only. Although my previous work had identified these features, they were trained on a larger population set hence it was important to reassess model metrics and evaluate whether regularised logistic regression remained the best performing model.

Analysis 2: Questionnaire Features Only Model: Model Generation

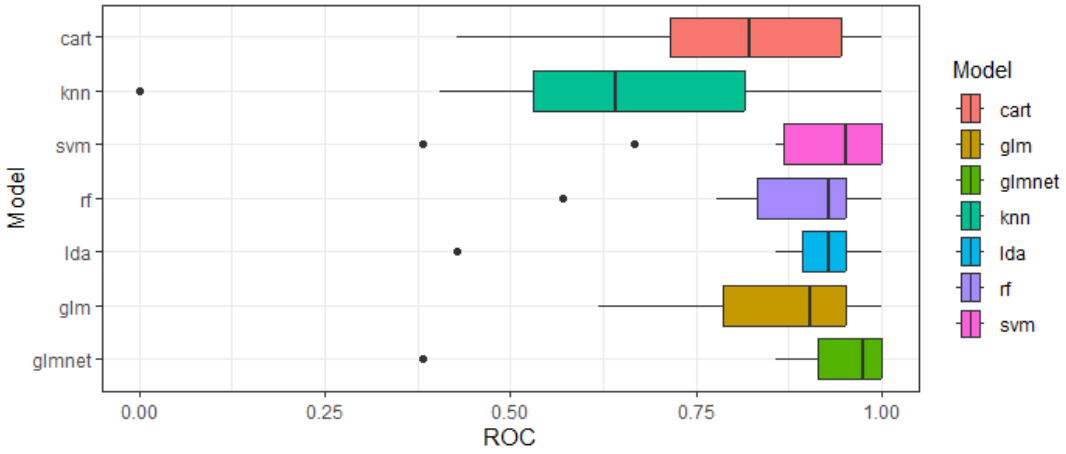


Figure 6-8: Boxplot of the seven machine learning models using questionnaire features in discriminating between non cancer and cancer cases.

Model	Median AUROC (IQR)	Median Sensitivity (IQR)	Median Specificity (IQR)
Classification and Regression Tree (cart)	0.82 (0.71-0.95)	0.93 (0.71-1.00)	0.66 (0.33-0.67)
k-Nearest Neighbour (knn)	0.64 (0.53-0.82)	1.00 (1.00-1.00)	0.13 (0.00-0.33)
Support Vector Machines (svm)	0.95 (0.87-1.00)	0.86 (0.75-1.00)	0.88 (0.67-1.00)
Random Forest (rf)	0.93 (0.83-0.95)	0.86 (0.86-0.86)	1.00 (0.42-1.00)
Linear Discriminant Analysis (lda)	0.92 (0.89-0.95)	0.86 (0.86-1.00)	0.66 (0.66-1.00)
Logistic Regression (glm)	0.90 (0.79-0.95)	0.86 (0.75-0.56)	0.83 (0.67-1.00)
Regularised Logistic Regression (glmnet)	0.98 (0.91-1.00)	0.93 (0.86-1.00)	0.83 (0.67-1.00)

Table 6-11: Table of median AUROC, median sensitivity and median specificity of the seven machine learning models using questionnaire features in discriminating between non cancer and cancer cases.

In contrast to the epigenetic only model, the best performing model with questionnaire only features with the highest median AUROC was regularised logistic regression, with a median AUROC of 0.98 (IQR: 0.91-1.00) (Figure 6-8 and Table 6-11). This was associated with an alpha α=0.1 and λ=0.053. In this model, the following features were dropped:

- Sex
- Acid or Sour Taste Present

- Acid or Sour Taste Frequency
- Presence of Swallowing Pain

Analysis 2: Questionnaire Features Only Model: Model Testing

Applying this model back to the training dataset demonstrated an AUROC of 0.94 (95% CI: 0.89-1.00). For the testing dataset, this model achieved an AUROC of 0.82 (95% CI: 0.67-0.98) (Figure 6-9). DeLong test performed did not show a statistical difference between the two different ROC curves (D statistic = 1.44, df = 49.8, p= 0.158).

At a cut off of 0.07, the model achieved a sensitivity of 0.97 (95%CI: 0.83-1.00) and a specificity of 0.38 (95%CI: 0.26-0.50) for the training dataset and a sensitivity of 0.92 (95%CI: 0.62-1.00) and a specificity of 0.28 (95%CI: 0.13-0.47) for the testing dataset (Table 6-13).

I also assessed the performance of the second best model (support vector machines) on the training and testing datasets. For the training dataset this achieved an AUROC of 0.93 (95% CI: 0.86-1.00); for the testing dataset AUROC was 0.89 (95% CI: 0.77-1.00) (Figure 6-10). DeLong test performed between the training data and testing data ROC curves did not demonstrate a statistically significant difference (D statistic = 0.688, df = 75.8, p = 0.494).

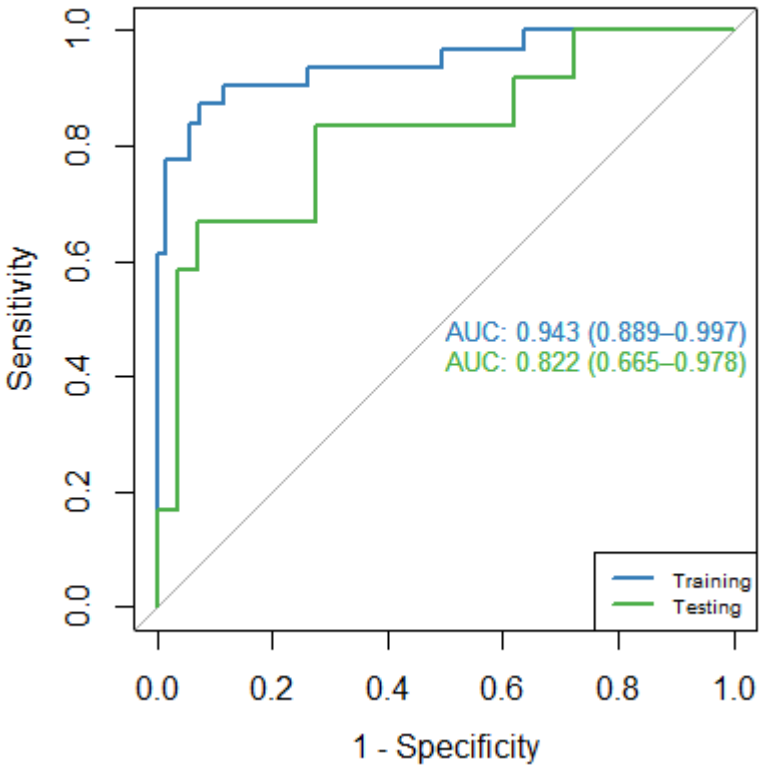


Figure 6-9: Area under the receiver operator curve (AUROC) for the regularised logistic regression model for the training and testing dataset for analysis 2 (questionnaire features only).

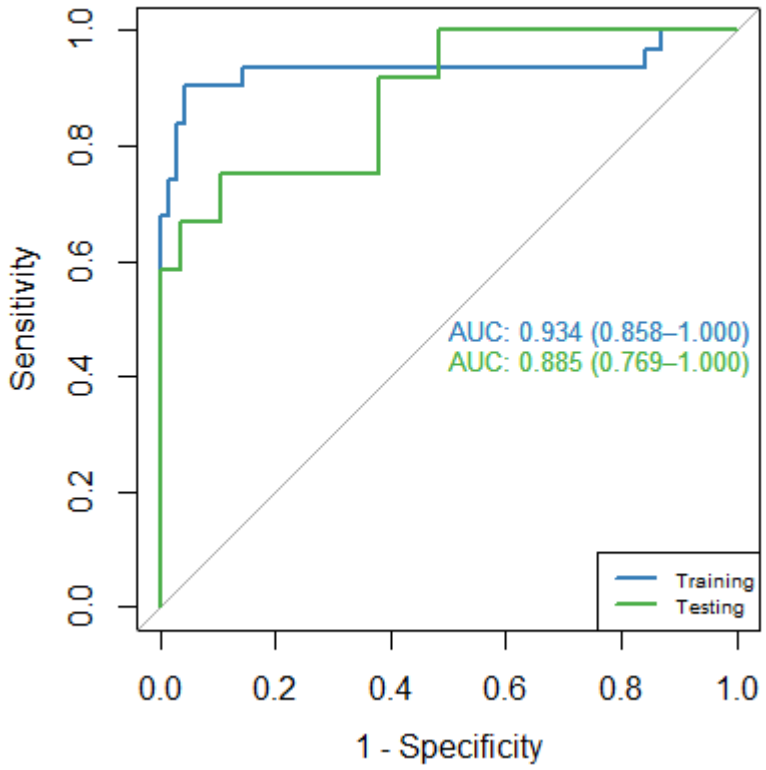


Figure 6-10: Area under the receiver operator curve (AUROC) for the support vector machines model for the training and testing dataset for analysis 2 (questionnaire features only).

I also performed DeLong's test for two correlated ROC curves to assess if there was a difference between the ROC curves of the training and testing datasets between regularised logistic regression and support vector machines. There was no statistically significant difference in both the training dataset (Z = -0.589, p-value = 0.556) or the testing dataset (Z = 1.55, p-value = 0.122).

For the support vector machines model, at a cut off of 0.09, sensitivity was 1.00 (95% CI: 0.89-1.00) and specificity was 0.13 (95% CI: 0.06-0.23) for the training dataset. For the testing dataset sensitivity was 1.00 (95% CI: 0.74-1.00) and specificity 0.21 (95% CI: 0.08-0.40) (Table 6-14).

Analysis 3: Epigenetic and Questionnaire Features Combined Model

Finally, I created a model where epigenetic and questionnaire features were combined to assess if there were additional synergies which could mean creating a more accurate model in prediction.

Analysis 3: Epigenetic and Questionnaire Features Combined Model: Model Generation

Model	Median AUROC (IQR)	Median Sensitivity (IQR)	Median Specificity (IQR)
Classification and Regression Tree (cart)	0.85 (0.69-0.91)	0.71 (0.71-0.96)	0.83 (0.67-1.00)
k-Nearest Neighbour (knn)	0.64 (0.44-0.85)	1.00 (1.00-1.00)	0.13 (0.00-0.33)
Support Vector Machines (svm)	0.95 (0.95-1.00)	0.86 (0.71-0.96)	1.00 (0.67-1.00)
Random Forest (rf)	0.93 (0.87-1.00)	0.86 (0.86-0.96)	0.67 (0.42-0.92)
Linear Discriminant Analysis (lda)	0.92 (0.90-0.96)	0.86 (0.86-0.96)	0.71 (0.67-1.00)
Logistic Regression (glm)	0.77 (0.61-0.94)	0.86 (0.84-0.86)	0.50 (0.33-0.92)
Regularised Logistic Regression (glmnet)	0.95 (0.91-1.00)	0.93 (0.86-1.00)	0.83 (0.42-1.00)

Table 6-12: Table of median AUROC, median sensitivity and median specificity of the seven machine learning models using both epigenetic and questionnaire features in discriminating between non cancer and cancer cases.

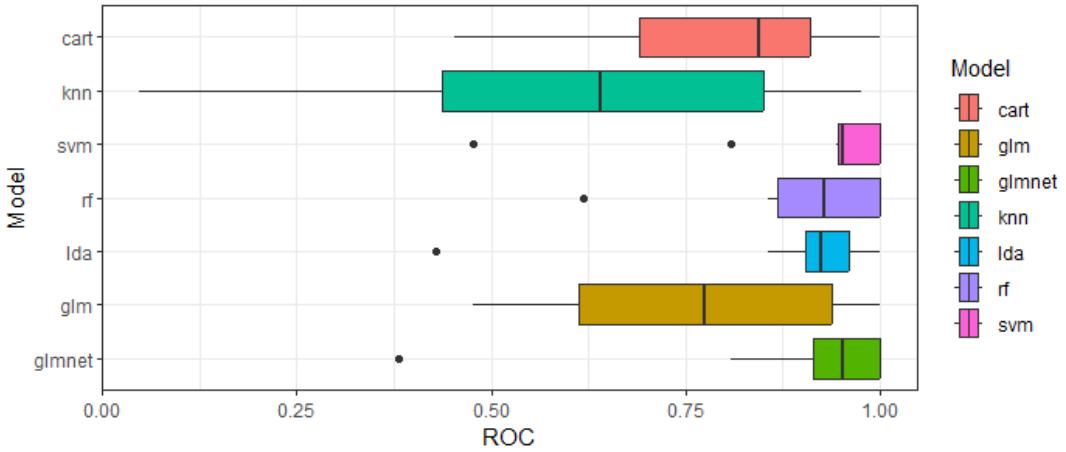


Figure 6-11: Boxplot of the seven machine learning models using both epigenetic and questionnaire features in discriminating between non cancer and cancer cases.

The best performing model during model generation was regularised logistic regression (median AUC: 0.95 (IQR: 0.91-1.00)) Table 6-12 and Figure 6-11). This was associated with

α=0.1 and λ=0.053). In the final regularised logistic regression model, several epigenetic and questionnaire features were dropped. These were:

- cg04096519
- cg27123665
- cg09771271
- Presence of Swallowing Pain
- Frequency of Consuming Fruit or Vegetables

The final model thus incorporated 4 epigenetic factors and 15 questionnaire based factors. Support vector machines also performed similarly (median AUC: 0.95 (IQR: 0.95-1.00)) (Table 6-12 and Figure 6-11), hence was also further evaluated in model testing.

Analysis 3: Epigenetic and Questionnaire Features Combined Model: Model Testing

Applying the regularised logistic regression model de novo back to the training dataset demonstrated an AUC of 0.96 (95% CI: 0.90-1.00) for the training dataset and an AUC of 0.83 (95% CI: 0.67-0.99) for the testing dataset (Figure 6-12). There was no statistically significant difference between the two ROC curves (DeLong test: D statistic = 1.49, df = 49.0, p = 0.143).

Similarly, applying the support vector machines model back to the training dataset demonstrated an AUC of 0.95 (95% CI: 0.89-1.00) for the training dataset and AUC of 0.89 (95% CI: 0.78-1.00) for the testing dataset (Figure 6-13). There was no statistically significant difference between the two ROC curves (DeLong test: D statistic = 1.07, df = 66.8, p = 0.288).

I also performed DeLong test to assess if there was a difference for the ROC curves obtained for the training and testing datasets for the two models. For the training dataset, there was no statistically significant difference between the two different models (DeLong's test for correlated ROC curves: Z = -0.229, p-value = 0.82). Similar results were also obtained for the testing dataset (DeLong's test for correlated ROC curves: Z = 1.13, p-value = 0.260).

At a cut off of 0.04, regularised logistic regression achieved a sensitivity of 0.97 (95% CI: 0.83-1.00) and a specificity of 0.23 (95% CI: 0.14-0.35) for the training dataset. This compares to a sensitivity of 0.97 (95% CI: 0.83-1.00) and a specificity of 0.87 (95% CI: 0.77-0.94) for the training dataset for support vector machines at a cut off of 0.27 (Table 6-13 and Table 6-14).

In contrast, at the same cut offs for the testing data set regularised logistic regression achieved a sensitivity of 1.00 (95% CI: 0.74-1.00) and a specificity of 0.21 (95% CI: 0.08-0.40). This contrasts with support vector machines which achieved a sensitivity of 0.75 (95% CI: 0.43-0.95) and a specificity of 0.72 (95% CI: 0.53-0.87) (Table 6-13 and Table 6-14).

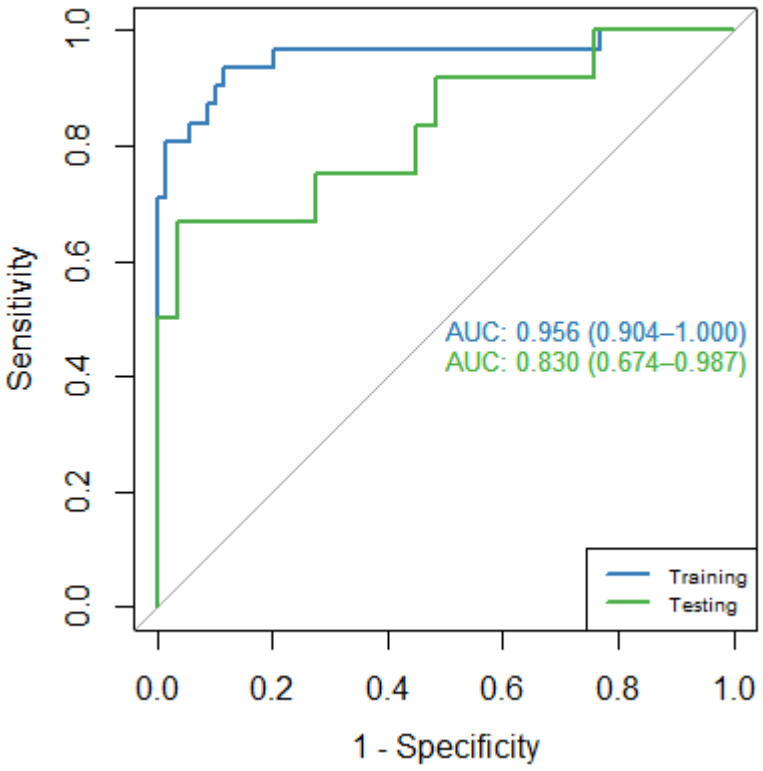


Figure 6-12: Area under the receiver operator curve (AUROC) for the regularised logistic regression model for the training and testing dataset for analysis 3 (epigenetic and questionnaire features combined).

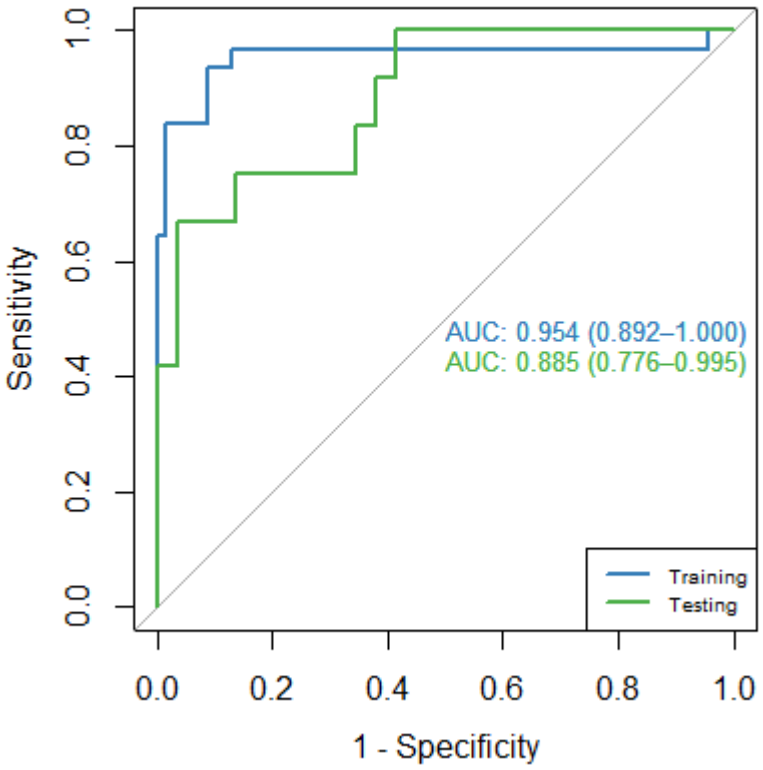


Figure 6-13: Area under the receiver operator curve (AUROC) for support vector machines model for the training and testing dataset for analysis 3 (epigenetic and questionnaire features combined).

	Analysis 1 (Epigenetic Only)		**Analysis 2 (Questionnaire Only)**		**Analysis 3 (Epigenetic and Questionnaire)**	
Best Model	Random Forest		Regularised Logistic Regression		Regularised Logistic Regression	
AUC of 10 fold CV (IQR)	0.88 (0.67-0.98)		0.98 (0.91-1.00)		0.95 (0.91-1.00)	
	Training Data	**Testing Data**	**Training Data**	**Testing Data**	**Training Data**	**Testing Data**
Cut Off	0.05	0.05	0.07	0.07	0.04	0.04
Sensitivity (95% CI)	1.00 (0.95-1.00)	0.67 (0.35-0.90)	0.97 (0.83-1.00)	0.92 (0.62-1.00)	0.97 (0.83-1.00)	1.00 (0.74-1.00)
Specificity (95% CI)	1.00 (0.89-1.00)	0.69 (0.49-0.85)	0.38 (0.26-0.50)	0.28 (0.13-0.47)	0.23 (0.14-0.35)	0.21 (0.08-0.40)
PPV (95% CI)	1.00 (0.95-1.00)	0.47 (0.23-0.72)	0.41 (0.30-0.53)	0.34 (0.19-0.53)	0.36 (0.26-0.47)	0.34 (0.19-0.52)
NPV (95% CI)	1.00 (0.89-1.00)	0.83 (0.63-0.95)	0.96 (0.81-1.00)	0.89 (0.52-1.00)	0.94 (0.71-1.00)	1.00 (0.54-1.00)

Table 6-13: Sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) for the best performing models in all 3 analyses (epigenetic features only, questionnaire features only and combined epigenetic and questionnaire features).

	Analysis 1 (Epigenetic Only)		**Analysis 2 (Questionnaire Only)**		**Analysis 3 (Epigenetic and Questionnaire)**	
Best Model	Support Vector Machines		Support Vector Machines		Support Vector Machines	
AUC of 10 fold CV (IQR)	0.69 (0.55-0.82)		0.95 (0.87-1.00)		0.95 (0.95-1.00)	
	Training Data	**Testing Data**	**Training Data**	**Testing Data**	**Training Data**	**Testing Data**
Cut Off	0.22	0.22	0.09	0.09	0.27	0.27
Sensitivity (95% CI)	0.90 (0.74-0.98)	1.00 (0.74-1.00)	1.00 (0.89-1.00)	1.00 (0.74-1.00)	0.97 (0.83-1.00)	0.75 (0.43-0.95)
Specificity (95% CI)	0.29 (0.19-0.41)	0.31 (0.15-0.51)	0.13 (0.06-0.23)	0.21 (0.08-0.40)	0.87 (0.77-0.94)	0.72 (0.53-0.87)
PPV (95% CI)	0.36 (0.26-0.48)	0.38 (0.21-0.56)	0.34 (0.24-0.45)	0.34 (0.19-0.52)	0.77 (0.61-0.89)	0.53 (0.28-0.77)
NPV (95% CI)	0.87 (0.66-0.97)	1.00 (0.66-1.00)	1.00 (0.66-1.00)	1.00 (0.54-1.00)	0.98 (0.91-1.00)	0.88 (0.68-0.97)

Table 6-14: Sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) for the second best model for epigenetic features only; the data is presented here due to the low reproducibility of the best performing model (random forest) in the testing dataset.

Comparison between best model in each analysis

So far in the three different analyses I have analysed between epigenetic only features questionnaire only features, and also combined questionnaire and epigenetic features in predicting the presence of oesophageal cancer. Although Random Forest was the best performing model during model generation in analysis 1 (epigenetic features only), it was apparent that the model appeared to be overfitted and there was poor reproducibility particularly during model testing. The overall second best model, Support Vector Machines had improved reproducibility between training and testing datasets hence this was selected as the 'best model' in analysis 1 instead for further head to head comparisons.

Using median AUROC of 10-fold cross validation during model generation as the comparator metric, the best performing model was questionnaire features only (analysis 2: median AUROC: 0.98 (IQR: 0.91-1.00), followed by combined epigenetic and questionnaire features (analysis 3: median AUROC 0.95 (IQR: 0.91-1.00)) and epigenetic only (analysis 1: median AUROC 0.69 (IQR: 0.55-0.82)). However, when using DeLong's test to compare the AUROC between the different models in pairwise comparisons, there was no statistically significant difference for both training and testing datasets (Table 6-15).

Comparison	**Analysis 1 v Analysis 2 Epigenetic Only vs Questionnaire Only**		**Analysis 1 v Analysis 3 Epigenetic Only vs Epigenetic & Questionnaire**		**Analysis 2 v Analysis 3 Questionnaire Only vs Epigenetic & Questionnaire**	
Dataset	Training	Testing	Training	Testing	Training	Testing
Z value	-0.53	0.05	-0.79	-0.03	1.03	0.32
p-value	0.59	0.96	0.43	0.98	0.30	0.75

Table 6-15: Head to head comparisons between epigenetic only features only model, questionnaire only features model and combined epigenetic and questionnaire features model, using DeLong's Test (with Z and p-values), for both the training and testing datasets.

6.4.5 Discussion

My analyses demonstrate that both a panel of epigenetic and questionnaire factors could be employed to predict whether an individual may have oesophageal cancer. During model

development, I was able to achieve a median AUROC of 0.98 (for the questionnaire only model) suggesting a high degree of accuracy in my predictions. Interestingly, there was no statistically significant difference between an epigenetic features only model, a questionnaire only model and a combined epigenetic and questionnaire feature model. This would suggest that there is little synergistic effect between epigenetic and questionnaire features in the prediction of oesophageal, or even digestive system cancers.

Epigenetic factors including DNA methylation markers have previously shown promise in the diagnosis of BO, OSCC and OAC as a combined entity, with a quoted AUC of 0.922.[219] Furthermore, epigenetic panels have also been found to be able to detect OAC and high grade dysplasia compared to normal or low grade dysplasia.[220] However, questions remain how reproducible each individual panel may perform between different populations. In addition, there is very limited published data on the combination of epigenetic factors and questionnaire data in the prediction of oesophageal cancer, or indeed gastroenterological cancers. Some limited evidence exists for cancer in general; a large study published in 2021 consisting of 72,284 breast cancer cases and 80,354 controls looked at a polygenic risk score, which contained 313 single nucleotide polymorphisms (SNPs) and questionnaire based breast cancer risk factors.[221] The only statistically significant feature was family history which is perhaps unsurprising. However, further modelling completed by the authors did suggest that there may be a stronger impact on absolute risk reduction of host risk factors for women at higher background genetic risk.[221] Thus, it would appear that epigenetic factors combined with questionnaire data may have limited utility in diagnosis or predicting an individual's risk for cancer. Indeed, I did not detect a statistically significant difference between the 3 analyses I conducted, suggesting that the combination of epigenetic factors combined with questionnaire only features did not confer any additional advantages compared to a questionnaire only approach.

6.4.6 Limitations

The major limitation is that the number of patients in this study is relatively small. In total there were 141 patients in the study, of which 43 patients had the target condition (cancer). As this was a case-control based study, patients were recruited when they were known to have cancer, and hence subject to recall bias in particular when completing the questionnaire. Furthermore, as I did not have enough data available, I was only able to train and subsequently test my machine learning models and not perform any validation. This therefore limits the robustness of my model and also its generalisability and reproducibility to other cohorts and populations.

6.5 Chapter Conclusions

Using a machine learning approach, I was able to create a risk prediction model using a questionnaire approach for the detection of oesophageal cancer on a cohort of 807 patients, which was subsequently validated on a further 294 patients. I was able to achieve an AUC of between 0.71 to 0.92 during testing and validation. I was also able to demonstrate that my model performed better than existing risk triaging models such as the Edinburgh Dysphagia Score. In a pilot study my model also demonstrated an AUC of 0.78 in the detection of gastric cancers, which could suggest there could be scope to extend the remit of the model to detect both oesophageal and gastric cancers. My model now needs further testing in larger, prospective cohorts to ensure its reliability.

In addition, I demonstrated that an epigenetic panel, a targeted questionnaire or both used in combination with each other could be used to predict the presence of oesophageal cancer. However, there did not appear to be a statistically significant difference between these models, although this was tested on a relatively small dataset, and without external validation. While this is potentially an exciting area of development, at present I would advocate using questionnaire based models for risk prediction of oesophageal cancer. These

have the advantage of being relatively cheap to administer and could potentially also be done by patients prior to attendance to either primary or secondary care appointments, with the result being available during the consultation.

Chapter 7: Faecal Immunochemical Testing for Triaging Lower Gastrointestinal Referrals

7.1 Chapter Introduction

Colorectal cancer represents a major health burden globally, where it is the third most common diagnosed cancer and fourth most common cause of cancer related death.[222] Importantly, incidence is rapidly increasing in lower to middle income countries, and by 2030 it is expected there will be 2.2 million cases and 1.1 million deaths as a result of colorectal cancer annually.[222] In the United Kingdom (UK) it is the 4th most common cancer with approximately 42,000 new cases per year.[223] One of the strategies for detecting colorectal cancer early is through the use of population level screening, where asymptomatic individuals are asked to complete a screening test, and higher risk individuals are invited for further investigations. Currently, the UK Bowel Cancer Screening Programme (BCSP) offers individuals aged between 50 and 74 to complete screening using the faecal immunochemical test (FIT).[224]

7.2 Aims

It has been proposed that FIT triaging could be employed to manage the COVID-19 related backlog.[225] Implementation of FIT triaging at 10µg Hb/g of faeces could mitigate 89% of deaths attributable to delays in the 2-week wait (2ww) colorectal pathway, but still reduce immediate demand for colonoscopy by more than 80%.[225] I was therefore interested in assessing the performance of FIT and patient outcomes, such as the development of polyps and cancers, and assessing if safety netting of FIT negative patients could be improved.

7.3 Background of FIT

The original test for faecal haemoglobin was the guaiac faecal occult blood test (FOBT). Faecal haemoglobin is detected by the addition of hydrogen peroxide, where oxidation of haem leads to a colour change.[226] However, the guaiac method typically can only detect faecal haemoglobin concentrations greater than 600µg Hb/g, thus having high false negative rates.[226] This has been superseded with FIT, which represents a major improvement over guaiac FOBT as it is not affected by diet or Vitamin C.[227] It is also easier to collect by the patient compared to FOBT, requiring fewer faecal samples and uses more acceptable stool sampling implements such as brushes and probes.[227] FIT is an immunoassay and forms antigen-antibody complexes with human haemoglobin.[226] The concentration of antigen-antibody complexes can then be quantified using an immunoturbidimetry or light-based methods.[226] Theoretically, more advanced pathology such as cancer would bleed more and lead to higher levels of FIT compared to less advanced pathology such as adenomas.[227] This then also allows the end user to set cut offs and determine sensitivity and specificity of the test and could be set at a threshold which balances the need of reducing the false negative rate but also ensures endoscopy demand is managed.[227]

Several manufacturer kits are available for quantification of FIT. These include the OC Sensor (Eiken Chemical, Japan), HM-JACKarc (Hitachi Chemical Diagnostic Systems Co., Ltd, Japan) and FOB Gold (Sentinel Diagnostics, Italy) which were recommended for adoption into primary care.[228] These systems however are all individually patented, with no standardisation with regards to collection devices, buffer solutions or analytical methods.[229] In addition, there may be globin degradation between sample collection and analysis, unless stabilisers are added to the buffer solution.[226] Notably there is performance variation between the different assays, hence results between different manufacturer kits should not be compared with each other.[229] A recent study compared the OC-Sensor and HM-JACKarc

demonstrated an overall Spearman's Correlation Coefficient rho=0.74, although this reduced to rho=0.26 for measurements of FIT around 10µg Hb/g. There was an agreement rate of 91.7% at a cut off of 10µg Hb/g.[229]

7.3.1 Evidence for FIT in screening for colorectal cancer

Since the introduction of FIT, several studies have been published detecting colorectal cancer or colorectal adenocarcinoma assessing the performance of FIT on a population level. These are summarised in Table 7-1. These studies have been identified from the reference list of two recent meta-analyses published, rather than constitute a de novo systematic review.[230,231]

Lead Author	Year	Journal	Threshold (10µg Hb/g)	Sensitivity % (95% CI)	Specificity % (95% CI)	PPV % (95% CI)	NPV % (95% CI)	Cancers below threshold (FN) (%)
Ayling R[232]	2021	J Clin Path	<10	87.5 (47.4-99.7)	52.6 (45.6-59. 6)	6.6 (5.0-8.7)	99.1 (94.6-99.9)	1/8 (13%)
Bailey S[233]	2021	Br J Cancer	<10	84.3 (71.4-93.0)	85.0 (83.8-86.1)	7.0 (5.1-9.3).	99.8 (99.5-99.9)	8/51 (16%)
Chapman C[234]	2019	BJS Open	<10	87·5 (73·2-95·8)	73·5 (70·2-76·6)	14·6 (10·4-19·8)	99·1 (98·0-99·7)	5/40 (13%)
D'Souza N (NICE-FIT)[235]	2021	Gut	<10	90.9 (87.2-93.8)	83.5 (82.8-84.3	16.1 (14.4-17.8)	99.6 (99.5-99.7)	30/329 (9.1%)
Khan A[236]	2020	BJS Open	<10	85·1 (71·0-93·3)	83·5 (80·8-85·8)	22·6 (16·0-28·3)	99·0 (97·9-99·5)	7/47 (15%)
Laszlo H (qFIT)[231]	2021	Br J Cancer	<10	83.3 (75.6-91.0)	80.1 (78.9-81.4)	9.7 (7.6-11.8)	99.5 (CI not stated)	15/90 (17%)
Mattar R[237]	2020	Arq Gastro	<10	83.3 (36.5-99.1)	86.9 (77.3-92.9)	17.8 (10.8-27.5)	82.2 (72.4-89.2)	1/6 (17%)
Mowat C[238]	2016	Gut	<10	89.3 (CI not stated)	79.1 (CI not stated)	14.2 (CI not stated)	99.5 (CI not stated)	3/28 (11%)
Navarro M[239]	2020	Front Med	<10	94.4 (81.9-98.5)	75.1 (71.8-78.2)	16.5 (12.1-22.2)	99.6 (98.6-99.9)	2/36 (5.6%)
Nicholson B[240]	2019	Frontline Gastro	<10	85.7 (42.1-99.6)	90.5 (85.9-93.9)	21.4 (8.3-41.0)	99.5 (97.4-100)	1/7 (14%)
Nicholson B[241]	2020	Aliment Pharm Ther	<10	90.5 (84.9-96.1)	91.3 (90.8-91.9)	10.1 (8.2-12.0)	99.9 (99.8-100)	12/105 (11%)
Rodríguez-Alonso L[242]	2015	Dig Liv Disease	<10	96.7 (82.2-99.9)	79.8 (77.1-82.2)	15.0 (CI not stated)	99.9 (CI not stated)	1/30 (3.3%)
Tsapournas G[243]	2020	Scan J Gastro	<10	92.3 (77.8-100)	77.3 (71.9-82.7)	18.8 (9.2-28.4)	99.4 (98.2-100)	1/13 (7.7%)
Turvill K[244]	2021	BJGP	<10	87.4 (81.0-92.3)	80.9 (79.7-81.9)	12.4 (10.4-14.5)	99.5 (99.3-99.7)	19/151 (13%)

Table 7-1: Summary of studies which have set a FIT threshold of <10 and their sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) with associated 95% confidence intervals (95% CI). FN=False Negative

Most studies achieved a sensitivity of greater than 80% for the detection of cancer, although there was sometimes a large confidence interval due to the small number of incident cases of cancers detected within a single study. As an example, one study with only six cases of cancer demonstrated a sensitivity of 83.3% (95% CI: 36.5-99.1).[237] Specificity ranged from 52.6% to 91.3%.[232,241] Although the majority of studies were conducted within the UK, one was conducted in Brazil,[237] one in Sweden,[243] and two in Spain.[239,242] A systematic review and meta-analysis of 15 prospective studies consisting of over 28,000 patients found a pooled summary sensitivity of 88.7% (95% CI: 85.2-91.4) and pooled summary specificity of 80.5% (95% CI: 75.3-84.8) for the detection of colorectal cancer at a FIT threshold of <10µg Hb/g.[230] The authors also performed subgroup analyses by assay type. At the same FIT threshold, six studies evaluated HM-JACKarc, while five evaluated OC-Sensor. For HM-JACKarc, sensitivity was 88.4% (95% CI: 84.8-91.3) and specificity was 82.2% (95% CI: 80.3-83.9).[230] Meanwhile, sensitivity was 87.8% (95% CI: 72.9-95.0) and specificity was 82.4% (95% CI: 60.3 to 93.5) for OC-Sensor.[230] This data therefore suggests that a FIT threshold of <10µg Hb/g can be used to exclude colorectal cancer in symptomatic patients. A further use of the FIT could be in risk triaging patients, where patients with a FIT ≥10µg Hb/g could be deemed high risk and have their investigations for colorectal cancer prioritised. However, FIT performs less well for the detection of high risk findings per BSG criteria, which is defined by ≥2 premalignant polyps including 1 advanced polyp or ≥5 premalignant polyps.[245] A multi-centre study by Chandrapalan et al. compared the use of volatile organic compounds (VOC) in urine compared to the FIT in detecting high risk findings.[246] For FIT, the sensitivity was 54% (95% CI: 43-65) while specificity was 79% (95% CI: 73-84).[246] This compares with VOC which had a sensitivity of 94% (95% CI: 88-98) and specificity of 69% (95% CI: 64-75).[246] This would suggest that FIT is inferior to VOC in detecting high risk findings, although the authors recommended that FIT could be used a second step after VOC testing.[246] The study would also suggest that FIT alone should not be relied on for high risk polyp detection.[246] This is confirmed by a UK

based multicentre observational study which also demonstrated low sensitivities of 34.9% and 46.8% for all polyps and high risk polyps respectively at a FIT threshold of 4µg Hb/g.[247]

7.3.2 Use of FIT in the United Kingdom

In 2017, NICE published guidance on the use of FIT for lower risk patients who did not meet more stringent suspected cancer guidelines, known as DG30.[228] This guidance was aimed at patients who did not meet the guidelines for suspected lower gastrointestinal (LGI) cancer, where patients would otherwise be referred under a 2ww pathway. It suggested that patients who were at low risk of colorectal cancer with an absence of red flag symptoms could be offered a FIT for triaging, with the potential that these patients could be managed within primary care. FIT was subsequently introduced as part of the UK BCSP in June 2019, replacing the gFOBT, where patients testing over the higher threshold of 120µg Hb/g of faecal haemoglobin are offered a colonoscopy.[55] The higher threshold was deliberately chosen to avoid BCSP colonoscopy lists being overwhelmed with demand as a lower threshold would have a very low specificity. However, this higher threshold has a sensitivity of only 48.9% for detecting colorectal cancer.[55] This could mean 51% of colorectal cancers present at the time of sampling are missed when compared to a lower threshold of 20µg Hb/g.[55]

As discussed earlier in Chapter 5: Effect of COVID-19 on Gastrointestinal Services, COVID-19 led to significant changes in the delivery of care within gastroenterology, the most significant of which was the decline in number of procedures being performed and the corresponding rise in waiting lists. As the first wave of the pandemic stated to wane in the United Kingdom in June 2020, guidance was issued by the NHS England where FIT testing was extended to cover 2ww suspected lower gastrointestinal (LGI) cancer referrals (NG12).[248] Patients who had a FIT ≥10µg Hb/g of faecal haemoglobin or had high risk features such as iron deficiency anaemia, palpable abdominal mass, rectal bleeding or obstructive symptoms were

recommended to be referred to hospital as usual for further investigations, such as a colonoscopy.[248] However, patients with a FIT <10µg Hb/g were recommended to be placed on a patient tracking list and be safety netted. This in effect meant that FIT could be seen as a diagnostic test rather than a triaging test.[249] There is some real-world evidence of this guidance working in practice; in a case series of 381 patients adoption of FIT triaging reduced the number of colonoscopies being referred through 2-week wait pathways by nearly half, without a detrimental effect on cancer detection rates.[250] Patients with a FIT <10µg Hb/g were safety netted and asked to return to their GP if symptoms returned or persisted.[250] However, The British Society of Gastroenterology (BSG) Endoscopy COVID-19 working group were concerned of 'unintended consequences' if there was premature use of FIT as a diagnostic tool for colorectal cancer.[249] Furthermore, there was concern that given there was a singular focus of diagnosing or excluding cancer, patients with other gastrointestinal diseases may be harmed and need re-referral, adding to the burden of COVID-19 related backlogs.[249] It was also unclear what should constitute safety netting, although duplicate or repeated FIT testing has been touted as a possible solution. Finally, uptake by GPs may be a concern; a survey done in 2018 suggested there was low acceptability of using FIT as a rule out test, with less than half perceiving it to be accurate and preferring over colonoscopy.[251] However, increased awareness, education, and changes to guidelines as a result of COVID-19 are likely to have improved acceptability since the survey was completed.

7.3.3 Piloting A New Safety Netting Pathway

While safety netting of patients who test below the FIT threshold of 10µg Hb/g has been recommended,[235] what is less clear is how this should operate. Some key questions include:

- Should the responsibility remain within primary care or transferred under secondary care, as patients should be placed on patient tracking lists

- What should be safety-netting protocol be? This could range from a follow-up with a clinician, additional tests such as a repeat FIT or blood tests
- What would be the optimum time period for repeat tests? In particular, is there an optimum time period for a repeat FIT?

7.3.4 Problems with a Single FIT Strategy

One major concern of using a single FIT on its own as a means of triaging patients at risk of colorectal cancer is that in some studies up to 10% of patients with colorectal cancer would not be detected if a threshold of FIT <10µg Hb/g was solely used.[235,252] A meta-analysis suggested that there was an 8.7% cancer miss rate pooled across 9 studies from the United Kingdom, although there was significant heterogeneity between the studies.[231] Indeed, if focussing on the number of cancers that occur in patients below a FIT <10µg Hb/g (i.e. false negative rate) (Table 7-1), this was as high as 17% in two studies.[231,237] While incident cases of colorectal cancers are low especially in population-based studies, nonetheless there is a strong argument to ensure sensitivity is as close to 100% as possible, thereby reducing the false negative rate and reducing the impact of delayed diagnose of cancer would have on individuals. One strategy proposed could be a duplicate FIT testing strategy as a means to try to increase the sensitivity of FIT overall.

7.3.5 Duplicate FIT Testing

The evidence for duplicate FIT testing initial came from bowel cancer screening in the Netherlands. It found that patients with two FITs recorded at 8µg Hb/g were 14 times more likely to develop advanced colorectal neoplasia, compared to patients with two FITs recorded at 0µg Hb/g.[253] A further Swedish prospective study with symptomatic patients who completed two FITs a few days apart found that while a single FIT missed one out of 13 cancers, a two FIT strategy identified all cancers.[243] More recently, two UK based studies have assessed the feasibility of a duplicate FIT testing strategy in ruling out colorectal cancer. The

first was performed in England in a population sample of 28,622 patients; two-thirds of patients (18,952) had two FITs below threshold of 10µg Hb/g.[254] While seven patients developed colorectal cancer in the two FIT below threshold group, all had high risk features such as iron deficiency anaemia which would have prompted further investigation regardless of the FIT result.[254] A further study in Scotland consisting of 5,761 patients found that of the 3,487 patients who had two FITs below threshold of 10µg Hb/g, three subsequently tested positive for colorectal cancer, of which two were anaemic, although it was unclear if they had additional high risk features.[255] The authors concluded that duplicate FIT testing, spaced one year apart, could be used as a safety netting strategy.[255] Indeed, there are significant cost benefits for using FIT rather than referring onward for further investigation. A single FIT costs from £1.96 to £6.04 depending on the kit manufacturer, compared to £136 for a CT colonography and £372 for a colonoscopy.[228] However, one major drawback of using duplicate FIT testing is that there can be sampling variation, as blood may not be evenly distributed within a stool sample.[249]

7.4 Retrospective Audit of FIT in Barnet

To understand what may constitute optimum safety-netting a 1-year pilot service evaluation project was started in North Central Thames in April 2021, which was subsequently relaunched in June 2022 due to initial implementation issues. Patients with lower abdominal symptoms and who also had a FIT of <10µg Hb/g were invited to a dedicated follow-up clinic, where they would be reviewed by a hospital clinician 8 to 10 weeks after their referral. They would also complete a second FIT. If a patient had two below threshold FIT results, they would be referred back to primary care with long term safety-netting, with more defined guidance as to what this should entail.[256]

I used data from the EMIS General Practice Database to assess the historical trends of the use of FIT prior to the service evaluation project being rolled out. The dataset has been

formally introduced in Chapter 4: Datasets and Data Curation. Briefly, this contains anonymised data on 23,383 individual patients, with 76,421 FIT results. There was also data available on 2ww lower gastrointestinal (LGI) referrals and subsequent colorectal cancer diagnoses.

7.4.1 Aims

In this audit I aimed to:

1. Audit the historical use of FIT in Barnet from January 2019 onwards
2. Assess the diagnostic performance of FIT
3. Assess if performing duplicate FITs could be used as part of a safety netting strategy

7.4.2 Methods

Data Cleaning

I performed data cleaning prior to analysis. I first removed any duplicate entries for both laboratory results and diagnoses. In addition, some results were recorded as below the lower limit of detection for the assay (e.g. <4µg Hb/g). As these types of readings could not otherwise be analysed in combination with other results, I recoded them without the less than (<) symbol. I also only analysed data from 1 January 2019 onwards; this was because most FIT results prior to this date were qualitative, with results stating either 'positive' or 'negative.' Subsequently, it was unclear what threshold these were used, and it was likely some of these results arose from the Bowel Cancer Screening Programme (BCSP), which has a higher threshold cut off value.

Endpoints

I created two distinct end points:

1. Cancer: this endpoint was defined as lower GI cancer of any histology

2. Cancer and polyp

To ensure that these endpoints occurred after a valid test result, I removed any diagnoses that had occurred before any FIT result. To capture as much information as possible I used a free text search strategy whereby synonyms and subtypes for colorectal cancer (e.g. sigmoid carcinoma, sigmoid adenocarcinoma, rectosigmoid cancer) were grouped together. I completed the same process for IBD diagnoses as well (e.g. Crohn's disease and ulcerative colitis).

Clinical Parameters

Laboratory values including haemoglobin, mean cell volume, platelet count, serum ferritin and serum C-reactive protein were included. I defined anaemia in females as <120 g/dL and in males <130 g/dL, as per World Health Organization (WHO) definitions.[257]

Statistical Analyses

To avoid duplicate counting of cases, I analysed FIT results by patient. In the case of there being multiple FIT results, I took the highest valid reading that the patient had. I performed all statistical analyses and created graphical plots using R software v 4.2.1.[93] I performed chi-squared tests to compare between categorical variables and t-tests between continuous variables. To assess for correlation, I used the rho statistic, also known as Pearson correlation coefficient.

7.5 Results

7.5.1 Demographic Information

The mean age was 60.8 years, with a standard deviation of 17.6. 55.2% of the dataset was female.

7.5.2 FIT Results

After data cleaning and removal of duplicate and qualitative FIT results, 16,082 unique patients with 18,430 FIT results were available for analysis. FIT results were taken from January 2019 to March 2022. Table 7-2 demonstrates the number of FIT readings. Notably, 2,046 patients had at least 2 different FIT readings.

Number of FIT Readings	Number of Patients (% overall)
At least 1	16,082 (87)
At least 2	2,046 (11)
At least 3	250 (1)
At least 4	34 (0)
At least 5	8 (0)
At least 6	*1 (0)*

Table 7-2: Number of FIT readings and the corresponding number of patients.

*One patient had 15 different unique FIT readings.

7.5.3 Distribution of FIT Results by Value

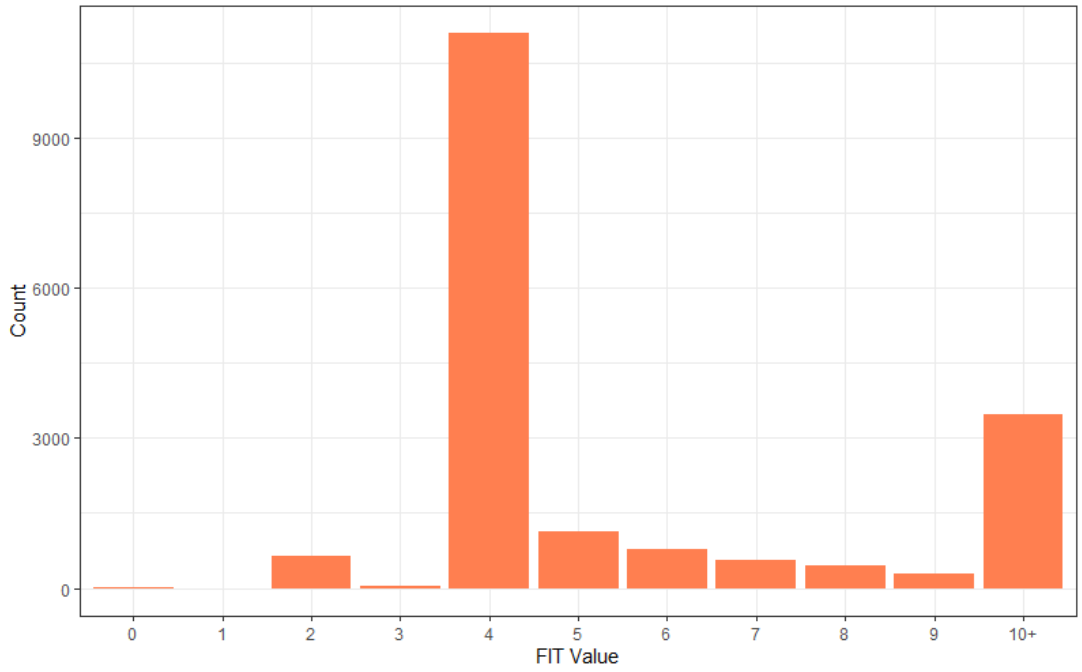


Figure 7-1: FIT values within database

Figure 7-1 shows the distribution of FIT values across the database. 19% (3,464/18,430) results were 10µg Hb/g or above. The most common value was 4µg Hb/g, likely as results marked was below the limit of detection (<4µg Hb/g) were changed to 4 for statistical analysis purposes.

7.5.4 Distribution of FIT Results by Date

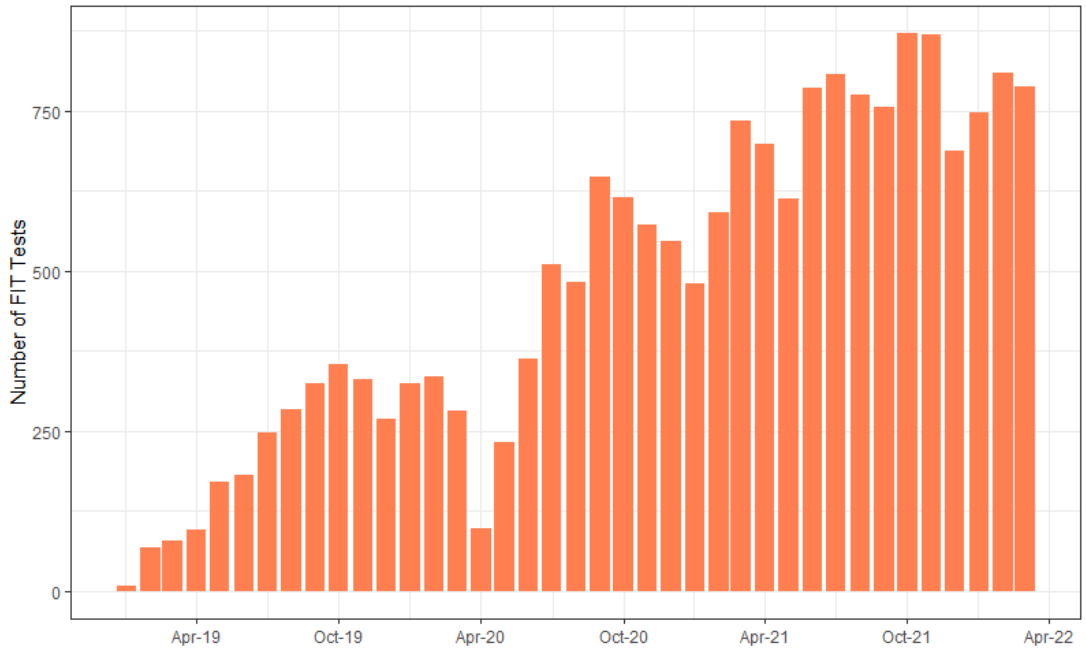


Figure 7-2: FITs performed on a monthly basis from January 2019 to March 2022

Figure 7-2 demonstrates the number of FITs performed on a per month basis across Barnet. There was a noticeable drop in the number of tests carried out at the start of the 1st COVID-19 lockdown in April 2020, with 98 tests performed. This was at 29% of the pre-pandemic peak of 334 tests in February 2020. By June 2020 however, the number of FITs had recovered to pre-pandemic levels (362 tests). Notably, this was at the time when NHS England recommended the use of a FIT to accompany all 2ww LGI referrals.[142] FITs reached a peak in September 2020 (647 tests) before dropping slightly in January 2021 (480 tests), coinciding with the 3rd national lockdown. Since then, there has been a continued upward trend in the number of FITs being done, with smaller drops seen in May and December 2021. The highest number of tests done in a month was 871, achieved in October 2021.

7.5.5 Single FIT and Cancer

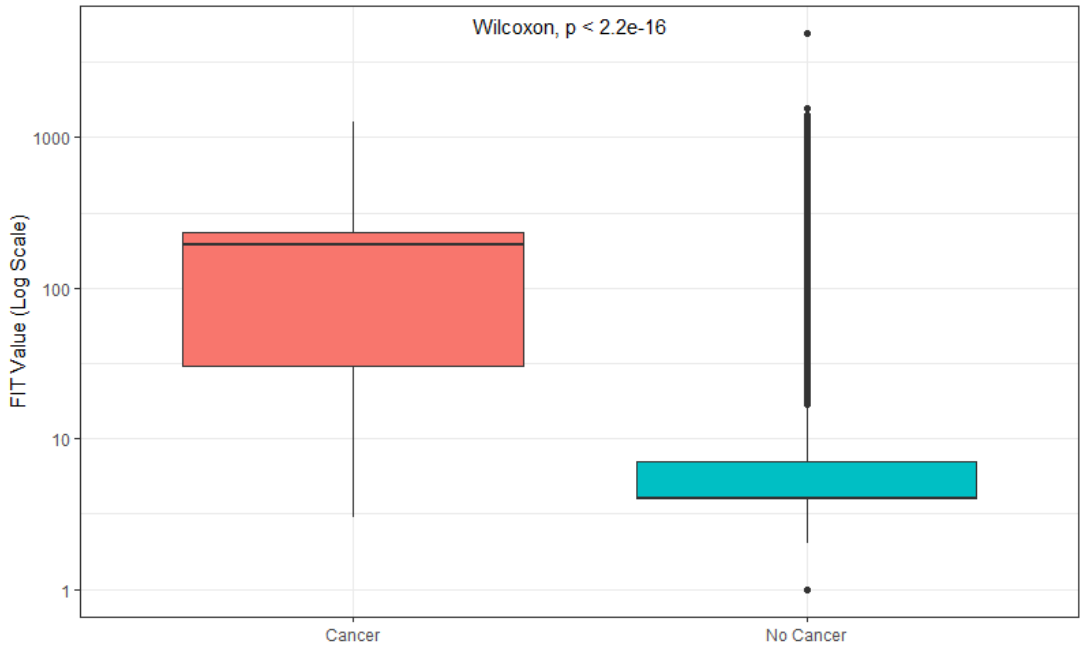


Figure 7-3: Boxplot of maximum FIT value of each patient (on a logarithmic scale) and cancer state

There were 101 new colorectal cancer cases in this dataset. Cancer cases were associated with a higher FIT value than non-cancer cases (cancer median FIT = 194µg Hb/g (interquartile range (IQR) = 204), non-cancer median FIT = 4µg Hb/g (IQR = 3)) (Figure 7-3). This was statistically significant ($p<0.001$).

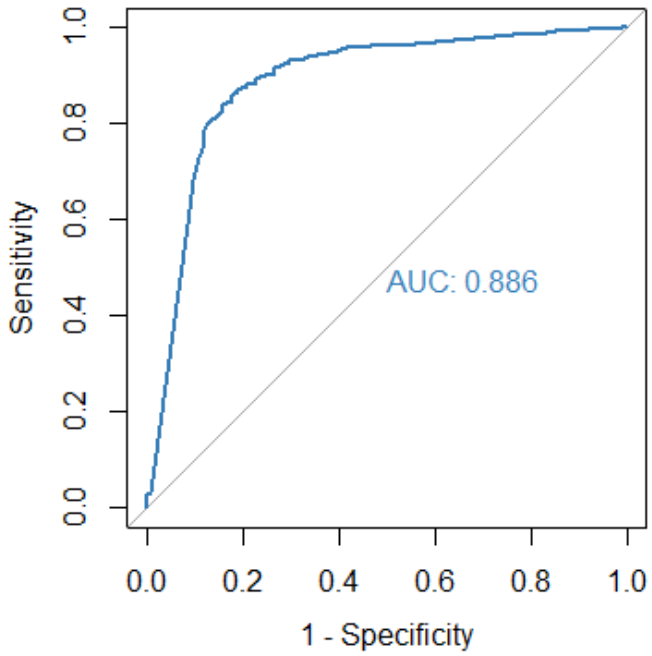


Figure 7-4: Receiver operator curve (ROC) for FIT and the detection of cancer

I also assessed the performance of a single FIT for the detection of cancer. Table 7-4 demonstrates the sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) for different cut offs. The corresponding receiver operator characteristic (ROC) curve is shown in Figure 7-4. This demonstrates that consistent with published literature a FIT threshold of <10µg Hb/ gives an NPV of 100%. However, it is worth noting 13% (13/101) of diagnosed cancers did not have a FIT <10µg Hb/g. These patients are outlined in Table 7-3. Importantly, only 5/13 (38%) patients who had clinical blood results available were actually anaemic, and only one of those patients was microcytic.

Identifier	Age	Sex	Haemoglobin (g/dL)	Mean Cell Volume (fL)	Platelet Count ($\times 10^9$/L)
02a07979-0b36-eff3-3a6e-e08b145a7106	38	Male	119	87.4	230
09739ebb-3afd-5354-4bc6-d3786a24cc90	68	Male	134	99.5	288
30f02d78-c8de-49d3-7c7e-206136663384	72	Female	135	94.8	292
3b12a73f-32ee-fabd-6013-db22fc819a77	89	Male	NA	NA	NA
421a7505-eaa2-c870-2131-32ef7e891218	83	Female	114	84	306
46d70a53-3569-6048-d053-85d4e9cb6041	52	Male	135	85.4	326
80a47201-4e12-60f0-9faa-04c948381833	58	Female	130	88.4	247
bbf94ed8-3365-24d0-5d56-08e99cf0a62d	77	Male	80	93.9	265
c0b514fd-0c79-c037-a6f2-e878d36b733b	70	Male	103	95	284
c9ef6434-5860-fc0b-853e-5e5df64beea9	93	Female	55	58.5	273
df2b3720-1171-53d7-5db1-09ed7ff7b702	61	Male	NA	NA	NA
f69db5c4-905a-328d-6137-c392350bcc71	60	Male	138	82.3	352
fa80bef6-0298-3eeb-0bd5-9f1c6e616a7e	81	Male	145	89.5	302

Table 7-3: Clinical parameters for the 13 patients who had a FIT <10µg Hb/g who were diagnosed with cancer

FIT Threshold (µg Hb/g)	True Positive (Cancers Correct)	True Negative	False Positive	False Negative (Cancer Incorrect)	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	P-value
4	100	502	15,479	1	1.00 (0.99-1.00)	0.01 (0.01-0.01)	0.03 (0.03-0.03)	0.99 (0.95-1.00)	0.380
6	91	10,926	5,055	10	0.90 (0.83-0.95)	0.68 (0.68-0.69)	0.02 (0.01-0.02)	1.00 (1.00-1.00)	<0.001
8	89	12,150	3,831	12	0.88 (0.80-0.94)	0.76 (0.75-0.77)	0.02 (0.02-0.03)	1.00 (1.00 1.00)	<0.001
10	88	12,802	3,179	13	0.87 (0.79-0.93)	0.80 (0.79-0.81)	0.03 (0.02-0.03)	1.00 (1.00 1.00)	<0.001
15	83	13,644	2,337	18	0.82 (0.73-0.89)	0.85 (0.85-0.86)	0.03 (0.03-0.04)	1.00 (1.00 1.00)	<0.001
20	81	13,989	1,992	20	0.80 (0.71-0.87)	0.88 (0.87-0.88)	0.04 (0.03-0.05)	1.00 (1.00 1.00)	<0.001

Table 7-4: Correlation matrices and performance metrics (sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) for different thresholds of FIT and the detection of cancer.

P-values are from Fisher's Exact Test or Chi-squared test comparisons. 95% CI = 95% confidence interval

7.5.6 Single FIT and Cancer and Polyps

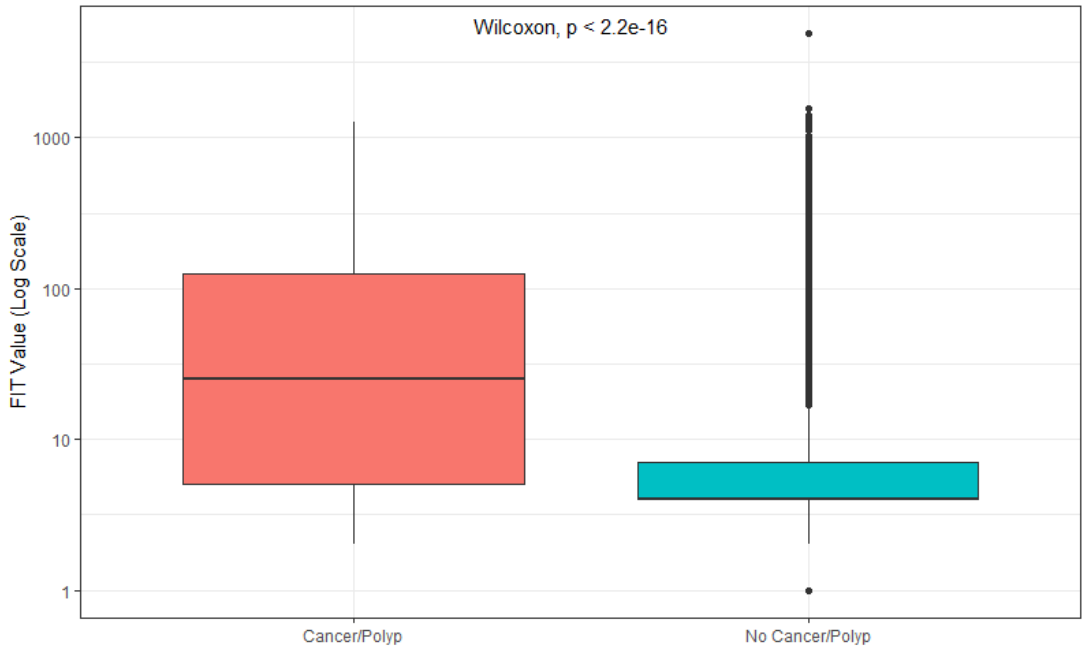


Figure 7-5: Boxplot of maximum FIT value of each patient (on a logarithmic scale) and cancer or polyp state

457/16,802 (2.7%) of patients were diagnosed with the combined endpoint of cancer or polyps. Similar to cancer as a single outcome, a higher FIT was significantly associated with a diagnosis of cancer or polyp. Median FIT was 25µg Hb/g (IQR: 119) for the cancer and polyp group, compared to 4µg Hb/g (IQR: 3) for the no cancer and polyp group (p<0.001) (Figure 7-5). Similar to cancer, a FIT of <10µg Hb/g was also associated with a NPV of 99.9% for excluding the presence of cancer or polyps (Table 7-5). This was associated with an AUC of 0.77 (Figure 7-6).

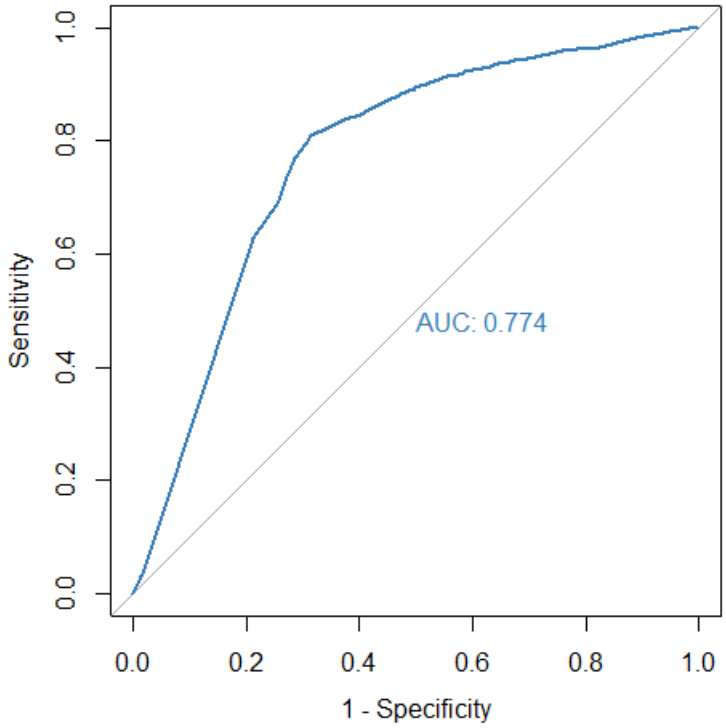


Figure 7-6: Receiver operator curve (ROC) for FIT and the detection of cancer or polyps

FIT Threshold (µg Hb/g)	True Positive (Cancer/Polyp Correct)	True Negative	False Positive	False Negative (Cancer/Polyp Incorrect)	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	P-value
4	450	496	15,129	7	0.98 (0.97-0.99)	0.03 (0.03-0.03)	0.03 (0.03-0.03)	0.99 (0.97-0.99)	0.0640
6	339	10,820	4,805	118	0.74 (0.70-0.78)	0.69 (0.69-0.70)	0.07 (0.06-0.07)	0.99 (0.99-0.99)	<0.001
8	326	12,034	3,591	131	0.71 (0.67-0.75)	0.77 (0.76-0.78)	0.08 (0.07-0.09)	0.99 (0.99-0.99)	<0.001
10	314	12,674	2,951	143	0.69 (0.64-0.73)	0.81 (0.80-0.82)	0.10 (0.09-0.11)	0.99 (0.99-0.99)	<0.001
15	262	13,467	2,158	195	0.57 (0.53-0.62)	0.86 (0.86-0.87)	0.11 (0.10-0.12)	0.99 (0.98-0.99)	<0.001
20	241	13,793	1,832	216	0.53 (0.48-0.57)	0.88 (0.88-0.89)	0.12 (0.10-0.13)	0.98 (0.98-0.99)	<0.001

Table 7-5: Correlation matrices and performance metrics (sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) for different thresholds of FIT and the detection of cancers and polyps as a composite endpoint.

P-values are from Fisher's Exact Test or Chi-squared test comparisons. 95% CI = 95% confidence interval

7.5.7 Duplicate FIT Testing

7.5.8 Characteristics of FITs

Figure 7-7 demonstrates the correlation between 1st and 2nd FIT values, with the axes truncated at 1250µg Hb/g. There was a generally poor positive correlation (rho = 0.2). Figure 7-8 shows the same correlation graph between 1st and 2nd FIT values but stratified by the time interval between the two different tests. There was a general downward trend of rho as the time interval between tests increased: at a time interval of less than 3 months rho equalled 0.29, compared to 0.13 for a time interval of 12 months or more.

7.5.9 Duplicate FIT Testing and Cancer

2,046 patients had at least 2 FIT values available for analysis. There were only four cancers diagnosed in this cohort (Table 7-6). There was only one patient who had a 1st FIT result of <10µg Hb/g and a 2nd FIT ≥10µg Hb/g. In addition, this patient was not anaemic. Therefore, a duplicate FIT testing strategy could theoretically have identified an extra case of colorectal cancer in this dataset.

7.5.10 Duplicate FIT Testing and Cancer or Polyp

98/2,046 (4.8%) patients who had at least 2 FIT values had a composite diagnosis of cancers or polyps. At a FIT threshold of <10µg Hb/g, there was a sensitivity of 59.2%, specificity of 69.9%, PPV of 9.0% and NPV of 97.1%.

When assessing the 98 patients who had a confirmed diagnosis of either cancer or polyps, only 58 (59.2%) had at least one FIT value ≥10µg Hb/g. Furthermore, of the 60 patients who had a below threshold 1st FIT, 19 (31.7%) subsequently had a FIT which was ≥10µg Hb/g. Therefore, a second FIT could potentially identify these patients for investigation.

Identifier	Age	Gender	1st Reading				2nd Reading				Difference in Days
			FIT (µg Hb/g)	Haemo-globin (g/dL)	Mean Cell Volume (fL)	Platelet Count ($\times 10^9$/L)	FIT (µg Hb/g)	Haemo-globin (g/dL)	Mean Cell Volume (fL)	Platelet Count ($\times 10^9$/L)	
71084ed8-dbd8-d5a1-99b9-ab28a85dee59	62	Male	18	108	58.7	299	4	114	59.8	355	169
fa80bef6-0298-3eeb-0bd5-9f1c6e616a7e	81	Male	4	145	89.5	302	4	128	93.4	341	299
fc910182-0366-7ab2-320e-6a199ca49c7c	62	Male	487	NA	NA	NA	875	NA	NA	NA	21
ffd8dc29-514d-99b4-0554-3c39c2a77376	**77**	**Female**	**4**	**131**	**102.2**	**284**	**16**	**139**	**99.8**	**285**	**32**

Table 7-6: Patients with 2 FIT results who were subsequently diagnosed with cancer.

The patient highlighted in bold had a 1st FIT <10µg Hb/g and 2nd FIT ≥10µg Hb/g.

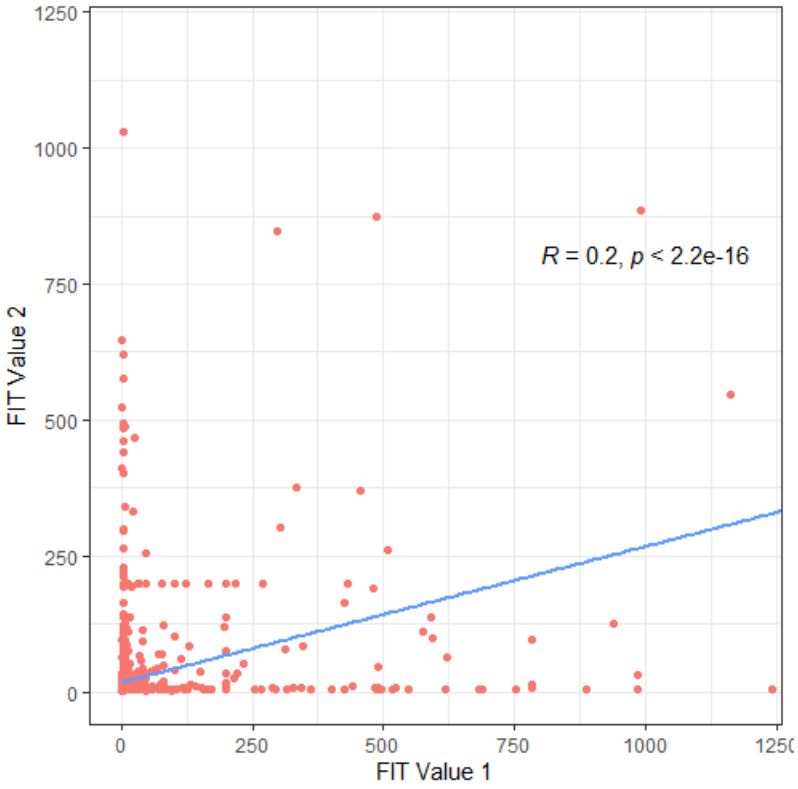


Figure 7-7: Overall correlation between 1st and 2nd FIT value.

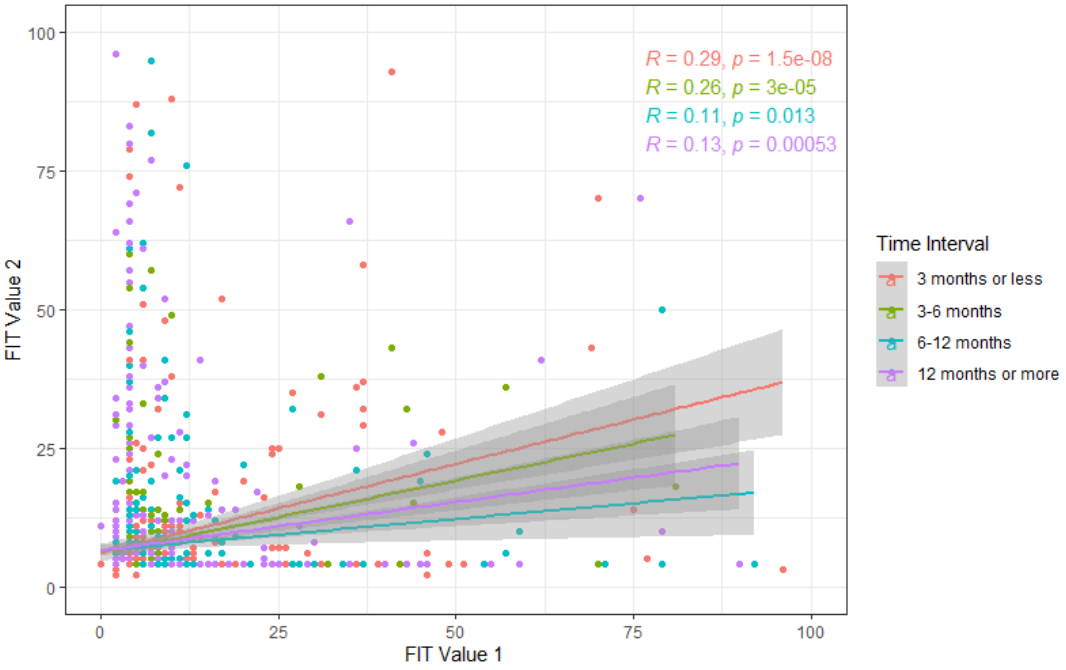


Figure 7-8: Correlation between 1st and 2nd FIT value, stratified by time interval between FITs.

7.6 Discussion

7.6.1 Usage of FITs

Given the large-scale disruption COVID-19 had in the delivery of healthcare, it is unsurprising that there was disruption in the number of FITs being performed in Barnet during the COVID-19 pandemic. However, within three months the number of FITs performed had recovered to pre-pandemic levels and has remained above pre-pandemic levels since then. This is unsurprising given the guidance that was issued by NHS England with regards to FIT as a triaging tool for patient with LGI symptoms.[248,249] It is unfair to make head-to-head comparisons between colonoscopy and FIT testing, as one is designed to be a gold standard diagnostic test while the other is a screening test but FIT has several inherent advantages in the era of COVID-19. The sample collection can be done by patients themselves at home and does not require personal protective equipment, which was in shortage especially in the early phase of the pandemic.[144]

7.6.2 Performance of FITs

My results demonstrate firstly that the endpoints of cancer alone and cancer combined with polyps were associated with a higher median FIT value. Furthermore, a single FIT set at a threshold of 10µg Hb/g has an NPV of 100% for excluding cancer. This is similar to published studies, confirming the effectiveness of using FIT as a rule out test for cancer.[231,235] However, it should be noted that at this threshold using FIT alone as a triaging criterion in this series would have missed 13/101 (13%) of cancers. Moreover, even when incorporating anaemia, 8/101 (8%) of cancer patients still had a below threshold FIT and a normal haemoglobin. This therefore highlights a real world need of safety-netting and ensuring patients do not have their diagnoses delayed.

7.6.3 Duplicate FIT Testing as Safety Netting

Although there is some encouraging data on duplicate FIT testing as a strategy in the literature,[254,255] my data showed mixed results. For the end point of cancer, as there were only 4 patients, it was not possible to draw definite conclusions. In contrast, for the composite cancer/polyp endpoint, 31.7% patients with a first below threshold FIT subsequently had a FIT ≥10µg Hb/g. This would suggest that a safety net strategy with 2 FITs could be cost effective in preventing future cancers, given patients with polyps could be identified earlier and be offered colonoscopy, potentially preventing future colorectal cancer.

One question that remains is what the optimum time period for successive FITs is. My data showed there was a weak positive correlation between successive FIT values in the same patient, even when the time interval is short (i.e. 3 months or less). This observation may be due to sampling variability, although may also point to aspects of the FIT assay which is not fully understood. It is therefore not possible to recommend a time duration when FITs ought to be repeated, although given the underlying concern of cancer this ought to be short. My results also suggest that as part of any safety netting other clinical parameters such as anaemia should be also used to reduce the risk of missed cancers.

7.6.4 Limitations

Firstly, due to the small number patients with duplicate FIT values subsequently diagnosed with cancer, it is not possible to draw definitive conclusions on whether duplicate testing can be part of a safety-netting strategy for cancer, although there is promising data for a polyps or cancer as a combined endpoint. Secondly, it is not clear which FIT assay was used for each individual result. Indeed, it could be possible that FIT results even for the same patient were done on different assays and hence analyses of successive FITs, even for the same patient, may not be valid. Thirdly, it is not clear why some patients had duplicate FITs done; this may

well have been a conscious clinical decision. Equally it could be that a single result could have been logged twice. This perhaps could occur if results were automatically transferred from information systems, but then also added in manually by GP practice staff. Despite careful data curation it is not possible to guarantee the complete absence of any duplicated results in this dataset. Furthermore, it is also unclear as to why a single patient had 15 unique FIT results. Fourthly, some additional clinical results were unavailable for patients, in particular symptomatology or clinical examination results. FIT in addition to these findings may have further increased cancer detection in FIT below threshold patients.

Finally, it is important to consider how the endpoints of this database would have been diagnosed. While LGI cancer is usually diagnosed endoscopically, it would also be possible for patients to be diagnosed non-invasively, such as using an abdominal computerised tomography (CT) scan or a CT colonography. This is especially true for frailer patients who may be unfit for colonoscopy but able to have imaging investigations for diagnostic and prognostic purposes. However, polyps and inflammatory bowel disease are both endoscopic and histological diagnoses. Subsequently, these groups are likely to be underdiagnosed in patients who had a FIT value of <10µg Hb/g, as they would not necessarily have been referred to secondary care for further investigation. This is further compounded by the fact that polyps can be asymptomatic and hence at a time of resource constraints such patients may not have met 2ww referral criteria and hence had a definitive investigation for diagnosis. In addition, I did not have access to histological data for polyps and staging data for cancers. Availability of this data would have enabled me to assess the utility of FIT in detecting benign versus pre-malignant polyps, and early versus late-stage cancers, and allowed for a more comprehensive assessment of the performance of FIT.

7.6.5 Further Work

Ideally, a large-scale prospective observational trial would be set up to assess the impact of duplicate FIT testing. Patients would undergo duplicate FIT testing as part of their standard of care, ideally at differing time intervals, with follow up over a defined interval to assess if they develop cancer or other pathology such as polyps. This data would help to determine the optimal interval for repeating the test, which could have a significant impact on current clinical practice.

In addition, safety netting could incorporate published risk scores. These include the Bristol-Birmingham equation and the CAPER (Cancer Prediction in Exeter) score, both of which have been shown to outperform NICE referral guidelines.[258,259] The Bristol-Birmingham equation includes a mixture of clinical characteristics and laboratory values, while the CAPER score relies on multiple presenting symptoms such as diarrhoea and weight loss.[258,259] As both scores have been developed from primary care databases these could especially be useful for 2ww referrals in association with FIT testing. In addition, several scoring systems have been published in combination with FITs as triaging tools; one example is the Asia-Pacific Colorectal Screening (APCS) scoring system which incorporates age, sex, family history and smoking as risk factors. In combination with FIT, this was estimated to reduce colonoscopy workload by 50%.[260] Several other examples also exist, primarily based in Asia-Pacific demonstrating utility of a risk scoring plus FIT approach rather than FIT alone,[261–263] although a study based in the United Kingdom did not reach the same conclusion.[264]

7.7 FIT Triaging as a Strategy for Overcoming Endoscopic Backlog

7.7.1 Introduction

In Chapter 5: Effect of COVID-19 on Gastrointestinal Services I demonstrated that there could be a nationwide backlog of nearly half a million endoscopic procedures directly as a result of the COVID-19 pandemic. One change that could be implemented quickly was the addition of

FIT triaging in order to reduce referrals and hence the potential backlog. In March 2020, NHS England recommended that patients who had a FIT threshold of ≥10 µg Hb/g would be referred for colonoscopy, with patients under the threshold offered safety netting.[162] I therefore sought to assess how the introduction of FIT triaging would affect the overall backlog of colonoscopy at the time of height of the COVID-19 pandemic in October 2020.

7.7.2 Methods and Results

Loveday et al. suggested that using a FIT cut-off at 10 µg Hb/g could reduce urgent 2ww suspected cancer endoscopies to 18% of usual requirements, assuming the remaining 82% of patients are not offered endoscopy.[225] Unfortunately, data on the proportion of patients referred on an urgent 2ww pathway was not available in the NHS Diagnostics Waiting Times and Activity dataset. I therefore estimated this using data using indications for colonoscopy. Data from the Dutch Gastrointestinal Endoscopy Audit registry showed that 29% of all colonoscopy referrals had an indication of changes in bowel habit, iron deficiency, chronic diarrhoea or abdominal complaints, all of which would trigger a 2ww referral.[168] Furthermore, data provided by the National Endoscopy Database comprising of 92,879 colonoscopies in 2019 showed that 31.8% of procedures had at least one indication of acute or chronic changes in bowel habit, anaemia, abdominal pain or weight loss, although the true figure may be lower, as around 27,000 procedures had more than one indication recorded.[265] Assuming these referrals would be suitable for FIT triaging at 10 µg Hb/g, this could equate to a reduction to 73.9% (1-(0.82 x 0.318) if using NED data, or 76.2% (1-(0.82 x 0.29) if using Dutch data. This equates to 11,866 (95% CI: 11,080-12,652) or 10,821 (95% CI: 10,104-11,538) colonoscopies per month. However, from my logistic regression model, I estimated if implemented in December 2020 it would still take until December 2021 to January 2022 to clear the pandemic related backlog (Figure 7-9).

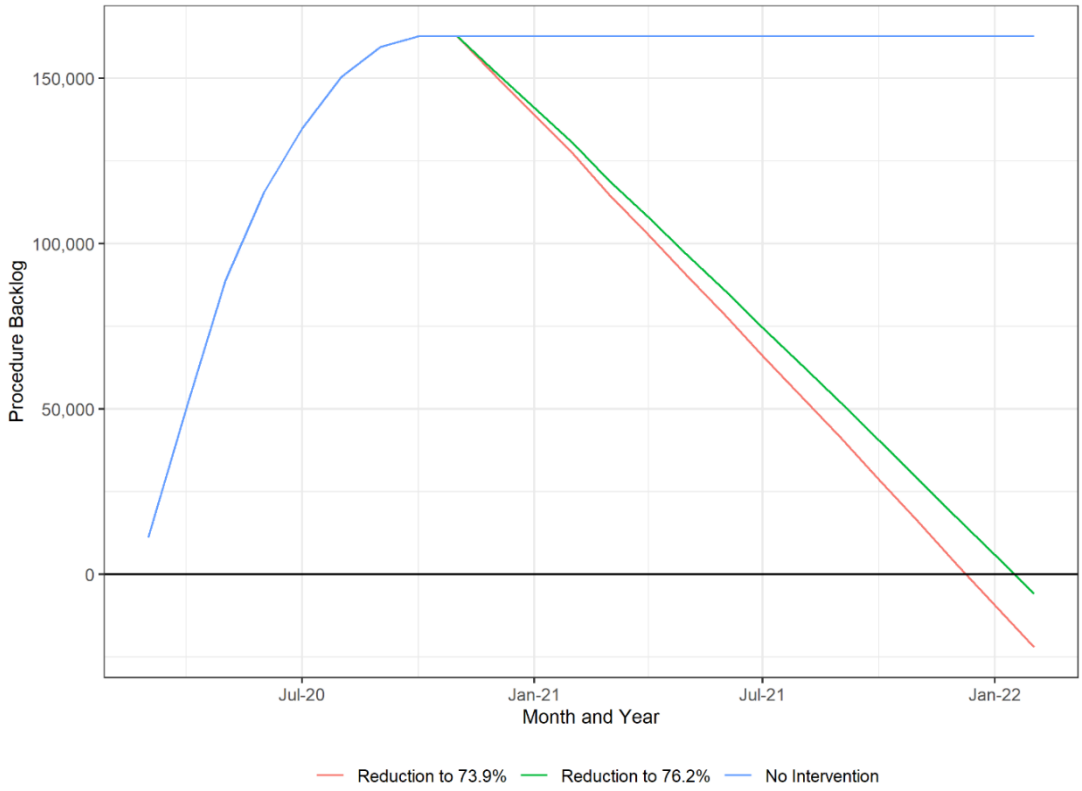


Figure 7-9: Estimated effect of FIT triaging on pandemic related colonoscopy backlog, if introduced from December 2020.

Calculation assumes endoscopy recovers to 100% prior to national introduction of FIT triaging.

7.7.3 Discussion and Conclusion

Through modelling, I demonstrated here that FIT could have reduced the colonoscopy backlog during the COVID-19 pandemic if patients below a set FIT threshold are not offered endoscopy in the early phase of the pandemic. There is some real-world data to support this. Data from a single NHS trust suggested that between September 2020 and February 2021, 74.5% of patients on a suspected lower GI cancer pathway underwent colonoscopy.[266] After FIT triaging was introduced, between March and August 2021 43.6% on the same pathway underwent colonoscopy.[266] Although there was no statistically significant change in colorectal cancer detection, this was also associated with a significant increase in CT scan use.[266] There is a potential risk that introducing FIT triaging may have simply shifted some of the endoscopic demand onto other diagnostic modalities. In addition, it would appear that given backlogs and waiting lists have continued to rise post pandemic it is likely that FIT

triaging would have only had at most a modest impact on overall backlogs.[178] Furthermore, it has also been argued that FIT should be used as a triaging rather than a diagnostic tool, to guide timing of the procedure rather than replace it.[249] Finally, there remains debate as to how patients with a FIT ≤10µg Hb/g ought to be followed up.[267] It therefore seems unlikely that FIT triaging alone will be able to have a significant impact in reducing overall endoscopy demand.

7.8 Chapter Conclusions

This analysis has highlighted some of the advantages and disadvantages of working with routine datasets. I was able to obtain data quickly and at low cost and audited the use of the FIT within Barnet. I described trends that occurred during the COVID-19 pandemic, especially how the number of FITs initially decreased but quickly recovered. However, there are questions about the reliability of the data source, especially with regards to FIT data results. In addition, the lack of suitable data made it difficult to make conclusions and recommendations about duplicate FIT testing. Further research is needed to understand how we can maximise the utility and yield of the FIT within lower GI cancer pathways, especially as it is a cheap and acceptable test, while ensuring there is still appropriate access to endoscopic investigation.

Chapter 8: Societal Impacts of COVID-19 Pandemic

Contents of this chapter have previously been presented and published as detailed below:

Publications

Ho KMA, Baggaley RF, Stone TC, et al. Face mask acceptability for communal religious worship during the COVID-19 pandemic in the United Kingdom: Results from the CONFESS Study. *J Relig Health* 2023; **62**: 608–26.[5]

Ho KMA, Davies H, Epstein R, et al. Spatiotemporal droplet dispersion measurements demonstrate face masks reduce risks from singing. *Sci Rep* 2021; **11**: 1–11.[4]

8.1 Introduction

The COVID-19 pandemic led to a direct impact on the initial plans for this thesis. Recruitment into the SPIT and RISQ studies was halted during the first wave of the pandemic, due to redeployment of research staff back into full time clinical duty and the suspension of routine gastroenterology care. In addition, hospitals and research ethics committees prioritised COVID-19 related studies.[268] While my initial research plans were on hold, an opportunity arose to investigate some of the wider issues that COVID-19 had had on society, with a focus on religious worship.

In the early phase of the pandemic, religious worship and associated activities such as choral singing was associated with mass outbreaks, coined superspreader events. This was because large numbers of people meeting together, sometimes for many hours and often in confined spaces, facilitated transmission of airborne viruses.[269] Religious events which led to large numbers of participants being infected included the Shincheonji Church of Jesus in South Korea and the Sri Petaling mass gathering in Malaysia, which at one point accounted for more than 60% and 35% of cases in their respective countries.[270,271] Furthermore, activities common in worship including group singing was identified as high risk for transmission of SARS-CoV-2. Perhaps the most famous example was the Skagit County choir practice, which led to an attack rate of 53.3%-86.7% after a rehearsal in early March 2020.[272] Further evidence to support this risk came from experimental studies: airborne droplets produced during singing did not appear to settle rapidly, and without adequate ventilation, could lead to an outbreak.[273] Moreover, normal and loud singing produced more aerosol particles compared to the same volume while speaking, which increases the risk of airborne transmission of viruses.[274] Other aspects of religious practice, such as holy communion and touching of the Torah, could also increase the likelihood of viral transmission.[275,276] These factors may explain the observation from the UK that religious faith could have been

associated with increased COVID-19 related mortality, even when adjusted for confounding variables.[277]

These risks associated with communal worship and singing led to onsite religious services being halted in England in March 2020 because of the rising incidence of COVID-19, with many services moving online.[278,279] Services and gatherings onsite resumed in July 2020 but with a ban on congregational singing and chanting, limits on numbers of participants allowed in congregations, and a requirement for social distancing and the wearing of face masks.[280] Relaxation of the rules and singing outdoors only occurred in late March 2021.[281] Meanwhile, with regards to singing, UK government guidance in April 2021 advocated for singing to be "limited to one person where possible" and stated "communal singing should not take place", even in the presence of social distancing or use of face masks.[280] However, guidance from Public Health England published in November 2020 around principles of safer singing conceded that there was a lack of evidence to suggest the degree to which wearing face masks may reduce the transmission of SARS-CoV-2.[282] A study involving 12 singers demonstrated that singing with a surgical face mask reduced the number of generated aerosols to a level similar to normal talking, although this did not reach statistical significance,[274] while wearing face masks could reduce the large variability of droplets being produced when singing, with larger droplets likely to be carrying higher viral loads.[4]

Closure of places of worship during the first lockdown in England and subsequent restrictions on worship significantly changed many people's daily or weekly routines, affecting their ability to pray, enjoy group discussion or take part in singing or chanting. In particular, relaxation of singing restrictions was seen as important as it could vastly improve congregants' worshipping experiences, and restore "a sense of celebration."[281]

8.2 Aims of Study

In this study, practising worshippers of any faith were recruited to complete an online questionnaire to improve understanding of how religious worship has changed during the COVID-19 pandemic in the UK. The primary aim was to understand how acceptable people would find face mask wearing in places of worship and whether they would be prepared to sing or chant whilst wearing them. The secondary aims were to understand the changes in worshipping practice due to COVID-19 and how well places of worship have complied with COVID-19-related safety guidelines.

8.3 Methods

8.3.1 Dataset

I used data generated from the CONFESS (COvid aNd FacE maSkS) study, introduced in Chapter 4.3.3. Briefly, this was a cross-sectional study comprising of an online questionnaire to assess the effect of the COVID-19 pandemic on religious practice. Participation was voluntary and used a convenience sampling technique with targeted advertising through religious institutions, social media and mass media, including being featured in the national news bulletin (Figure 8-1) and a religious affairs programme.[129,283]

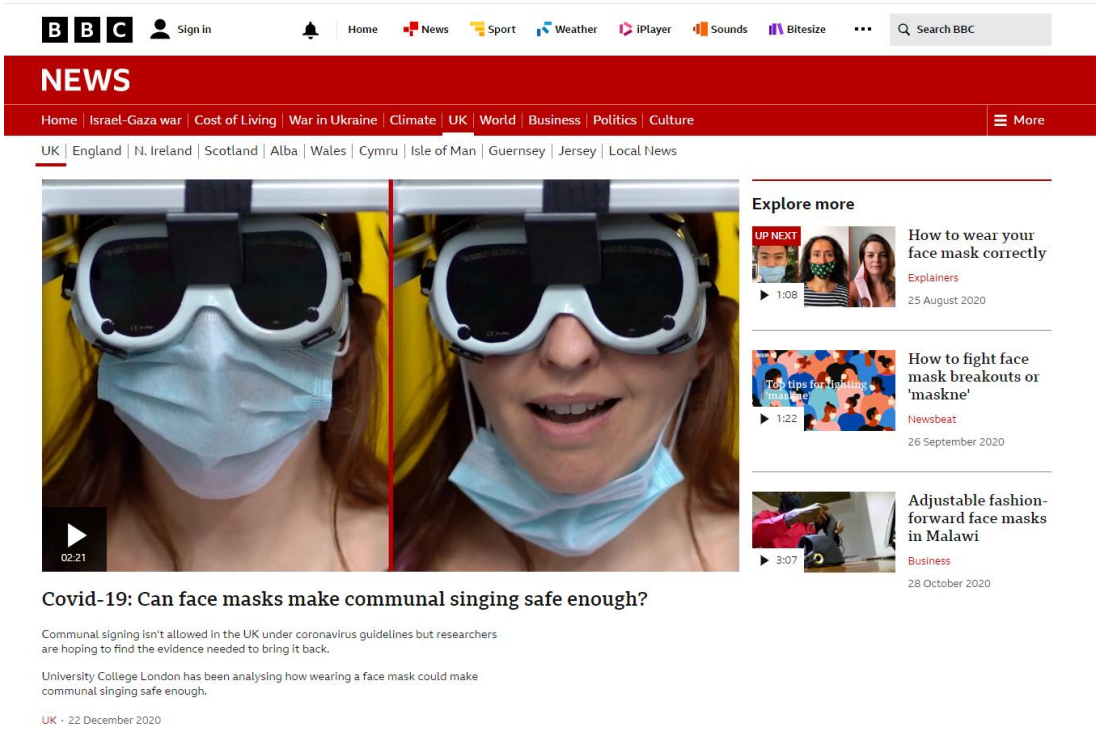


Figure 8-1: screenshot of BBC News website with a clip of the report in the BBC national news bulletin.[283]

Particular efforts were made to recruit participants from a range of religious backgrounds including Christian, Muslim, Jewish, Hindu and Buddhist by contacting specific groups, such as through Facebook (Table 8-1). In addition, religious leaders were approached directly from churches, mosques, synagogues, and Hindu, Buddhist and Hare Krishna temples.

Acton Masjid	Babul Murad Centre	Bristol Hindu Temple	Buddhapadipa Temple
East London Mosque	Friendly Frummers	Hindu Temple Newcastle	Ilford Hindu Temple
Iskon Temple Watford	Islamic Integration Community	Jalaram Mandir Temple	Jamia Masjid
Jamyang Buddhist Centre	Jewish Connections	Junior Anglican Evangelical Conference	Kingsbury Buddhist Temple
London Asian Seventh-Day Adventist Church	London International Christian Church	London International Church @RhemaFaithMinistries	Mayfair Islamic Centre
Methodist Central Hall Westminster	MOO Modern Open Orthodox	Moslem Ali Khan Welfare Foundation	Slough Hindu Temple
Southampton Hindu Temple	Swindon Hindu Temple	Synagogues of London and the UK	UK Christian Events
UK Methodists	UKEvents.net	Vedic Society Hindu Temple	
In addition, our research team advertised to various large mosques and synagogues around London and the UK			

Table 8-1: Facebook groups on which the CONFESS Study was advertised.

8.3.2 Survey Measures

The CONFESS questionnaire included items concerning demographic characteristics, including age, sex, religion and ethnic background; changes in worshipping practice due to COVID-19; importance of religious life and singing; acceptability and comfort of face mask wearing during communal worship and while singing during worship; awareness and understanding of government guidelines regarding COVID-19 and compliance of participants' place of worship with these guidelines. Notably when the questionnaire was first rolled out it was not mandatory to answer all the questions within the questionnaire, due to a technical error. This therefore has led to varying number of responses to each question within the questionnaire.

8.3.3 Data Analysis

Quantitative data were analysed using R software version 4.0.4.[93] Chi-squared tests were used for categorical variables, with a p-value of ≤0.05 taken as statistically significant. I performed univariable and multivariable (adjusted) logistic regression to assess for characteristics which may predict face mask acceptance. Adjustments were made for sex, age (as a continuous variable), highest level of education, religion, ethnicity, relationship status, place of residence and employment status. I calculated with my collaborators both unadjusted and adjusted prevalence odds ratios (OR), with 95% confidence intervals (95% CI), as opposed to prevalence rate ratios, because the variables included in the analysis were long-term characteristics of respondents. In addition, I performed thematic analysis of open-ended questions to complement quantitative analysis.

8.4 Results

8.4.1 Demographics

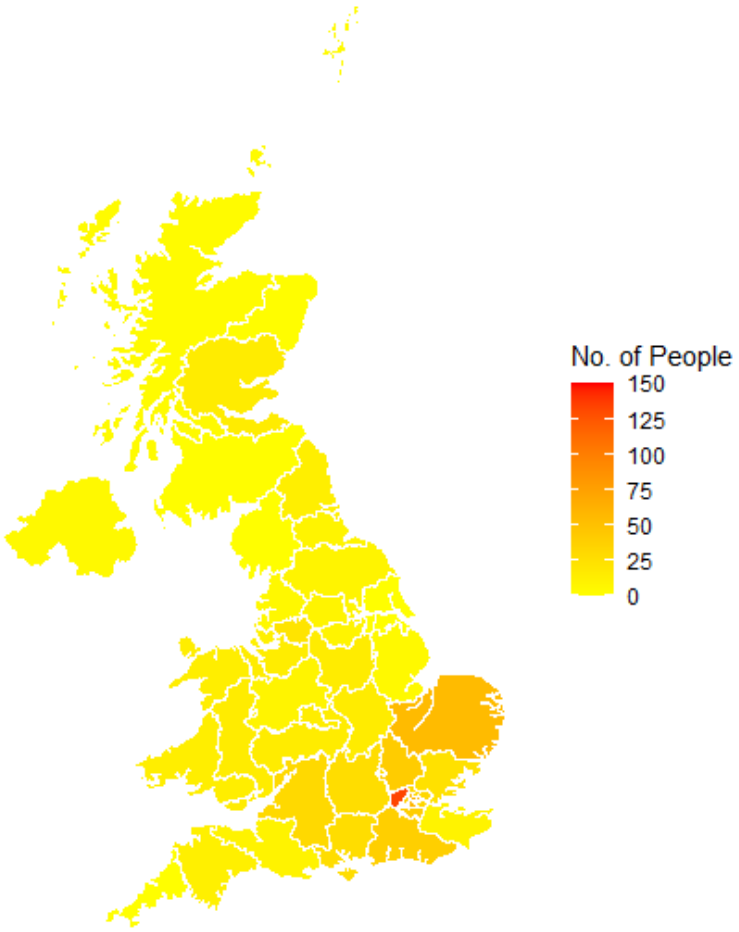


Figure 8-2: Region of UK residence for CONFESS study participants.

In total, 1,063 people volunteered for the study. 939 (88.3%) participants completed the questionnaire and were included in the analysis. Demographic information for included participants is shown in Table 8-3. Median age was 52.7 years and approximately two-thirds were female, while 845 (90.0%) completed at least undergraduate-level education. The majority of participants were Christian (n=758, 80.7%), followed by Jewish (n=145, 15.4%) and of White ethnicity (n=869, 92.5%). Most (789, 84.0%) participants lived in an urban area. Of the 831/939 (88.5%) participants who gave valid postcode data, 305 participants (36.7%)

lived in London, while 122 (14.7%) participants lived in East of England and 105 participants (12.6%) lived in South East England (Figure 8-2). Most participants (779/939, 83.0%) reported neither had suspected nor confirmed COVID-19 infection previously at the time of recruitment (Table 8-2).

		Participants (%)
COVID-19 related characteristics of respondents		
COVID-19 infection status	Yes diagnosed and recovered	27 (2.9%)
	Yes diagnosed and still ill	3 (0.3%)
	Not formally diagnosed but suspected	130 (13.8%)
	Not that I know of / No	779 (83.0%)
Ever had a SARS-CoV-2 virus swab test	Yes – positive result	9 (1.0%)
	Yes – negative result	274 (29.3%)
	No swab test	653 (69.8%)
Ever had a SARS-CoV-2 antibody test	Yes – positive result	33 (3.5%)
	Yes – negative result	72 (7.7%)
	No swab test	834 (88.8%)
Current isolation status	1. Living life as normal	128 (13.6%)
	2. Not "staying at home" but cut down on usual activities as a precaution	471 (50.2%)
	3. Not "staying at home" specifically, but working from home	132 (14.1%)
	4. "Staying at home" but not high risk - worried about spreading to others or getting ill	31 (3.3%)
	5. "Staying at home" to protect a family member/friend/housemate with an existing medical condition/ high risk	21 (2.2%)
	6. "Staying at home" – existing medical condition or categorised as high risk	23 (2.4%)
	7. "Self-isolating" – avoiding contact with all people as much as possible due to a COVID-19 diagnosis or possible COVID-19 infection	0 (0.0%)
	8. "Staying at home" – ordered government/local authority as part of a lockdown	28 (3.0%)
	9. “Staying at home" for a non-COVID-19-related reason e.g., a pre-existing health condition or disability	4 (0.4%)
	Missing	*101 (10.8%)*

Table 8-2: COVID-19 related characteristics of CONFESS study participants.

Characteristic		Participants	Face mask acceptability		Unadjusted odds ratio (95%CI)	p value	Adjusted odds ratio (95%CI)	p value
			Very/somewhat acceptable	Very/somewhat unacceptable				
Demographic characteristics								
Sex	Female	620 (66.1%)	463/564 (82.1%)	101/564 (17.9%)	-	-	-	-
	Male	318 (33.9%)	209/271 (77.1%)	62/271 (22.9%)	0.74 (0.52-1.05)	0.0905	0.70 (0.47-1.05)	0.0795
Age	Median age (range), years	52.7 (18-85)	-	-	0.98 (0.97-0.99)	0.0050	0.98 (0.96-1.00)	0.0218
	18-34 years	161 (17.7%)	128/140 (91.4%)	12/140 (8.6%)	-	-		
	35-64 years	585 (64.1%)	406/523 (77.6%)	117/523 (22.4%)	0.33 (0.17-0.59)	<0.001	0.30 (0.13-0.61)	0.0016
	≥65 years	166 (18.2%)	119/149 (79.9%)	30/149 (20.1%)	0.37 (0.18-0.74)	0.0070	0.38 (0.14-1.02)	0.0619
Highest educational level	Undergraduate degree or professional qualification	470 (50.1%)	345/422 (81.8%)	77/422 (18.2%)	-	-	-	-
	Postgraduate degree	375 (39.9%)	272/334 (81.4%)	62/334 (18.6%)	0.98 (0.68-1.42)	ns	0.97 (0.65-1.44)	ns
	A-levels or equivalent/post-16 vocational course	72 (7.7%)	42/60 (70.0%)	18/60 (30.0%)	0.52 (0.29-0.97)	0.0345	0.52 (0.28-0.99)	0.0403
	GCSE/CSE/O-levels or equivalent/no qualifications	22 (2.3%)	13/19 (68.4%)	6/19 (31.6%)	0.48 (0.18-1.41)	ns	0.58 (0.21-1.73)	ns
Religion	Christian	758 (80.7%)	559/688 (81.3%)	129/688 (18.8%)	-	-		-
	Jewish	145 (15.4%)	92/119 (77.3%)	27/119 (22.7%)	0.79 (0.50-1.28)	ns	0.88 (0.53-1.49)	ns
	Other	36 (2.8%)	21/28 (75.0%)	7/28 (25.0%)	0.69 (0.30-1.79)	ns	0.61 (0.23-1.78)	ns
Ethnicity	White British	792 (84.4%)	564/707 (79.8%)	143/707 (20.2%)	-	-	-	-
	Other White background	77 (8.2%)	57/69 (82.6%)	12/69 (17.4%)	1.20 (0.65-2.41)	ns	1.27 (0.66-2.60)	ns
	Asian/Asian British	26 (2.8%)	20/23 (87.0%)	3/23 (13.0%)	1.69 (0.57-7.24)	ns	1.53 (0.43-7.54)	ns
	Black/African/Caribbean	12 (1.3%)	9/9 (100%)	0/9 (0.0%)	-	ns	-	ns
	Mixed/Multiple ethnic groups	16 (1.7%)	12/13 (92.3%)	1/13 (7.7%)	3.04 (0.59-55.66)	ns	1.96 (0.35-36.99)	ns
	Other	15 (1.6%)	9/13 (69.2%)	4/13 (30.8%)	0.57 (0.18-2.13)	ns	0.62 (0.19-2.46)	ns
Relationship status	In a relationship/married and cohabiting	686 (73.1%)	495/616 (80.4%)	121/616 (19.6%)	-	-	-	-
	In a relationship/married but living apart	36 (3.8%)	25/30 (83.3%)	5/30 (16.7%)	1.22 (0.50-3.68)	ns	0.89 (0.35-2.78)	ns
	Single, divorced or widowed	60 (6.4%)	38/50 (76.0%)	12/50 (24.0%)	0.77 (0.40-1.59)	ns	0.71 (0.35-1.51)	ns
	Single, never married	157 (16.7%)	114/139 (82.0%)	25/139 (18.0%)	1.11 (0.70-1.83)	ns	0.75 (0.44-1.30)	ns

Characteristic		Participants	Face mask acceptability		Unadjusted odds ratio (95%CI)	p value	Adjusted odds ratio (95%CI)	p value
			Very/somewhat acceptable	Very/somewhat unacceptable				
Demographic characteristics								
Place of residence	City/town	789 (84.0%)	563/694 (81.1%)	131/694 (18.9%)	-	-	-	-
	Village/rural dwelling	150 (16.0%)	109/141 (77.3%)	32/141 (22.7%)	0.79 (0.52-1.24)	ns	0.83 (0.52-1.35)	ns
Employment status	Employed (full-time)	374 (39.8%)	273/326 (83.7%)	53/326 (16.3%)	-	-	-	-
	Employed (part-time)	175 (18.6%)	120/158 (75.9%)	38/158 (24.1%)	0.61 (0.38-0.98)	0.0407	0.62 (0.37-1.04)	0.0661
	Self-employed	113 (12.0%)	79/100 (79.0%)	21/100 (21.0%)	0.73 (0.42-1.30)	ns	0.93 (0.51-1.73)	ns
	Retired	193 (20.6%)	139/176 (79.0%)	37/176 (21.0%)	0.73 (0.46-1.17)	ns	0.80 (0.43-1.51)	ns
	Student (university/school)	35 (3.7%)	27/31 (87.1%)	4/31 (12.9%)	1.31 (0.49-4.57)	ns	1.04 (0.33-4.04)	ns
	Other	49 (5.2%)	34/44 (77.3%)	10/44 (22.7%)	0.66 (0.32-1.48)	ns	0.62 (0.28-1.43)	ns

Table 8-3: Characteristics of participants in the CONFESS study.

ns: non-significant, GCSE: General Certificate of Secondary Education, CSE: Certificate of Secondary Education

8.4.2 Face Mask Acceptability

Most (872/939, 92.9%) respondents answered questions about the face mask they most commonly wore. Of those who answered, reusable masks (661; 75.8%) and disposable surgical masks (157; 18.0%) were the commonest mask type worn. The majority (861/939, 91.7%) of respondents answered questions on face mask acceptability and comfort; 346 (40.2%) and 326 (37.9%) respondents to the question found it very acceptable and somewhat acceptable, respectively, to be required to wear a face mask and reduce their singing volume to sing or chant safely. While 421 participants (48.9%) reported never having sung in a face mask, most respondents (610, 70.8%) reported having already worn a face mask for at least an hour.

When asked to provide more details regarding their attitudes to face mask wearing for singing during communal worship as an open-ended question, 564 (60.1%) provided details. The overarching response was that face masks were unpleasant, but they were better than not singing at all (quotes 1-3, Table 8-5). However, there were participants who felt very strongly that they couldn't sing with a face mask (quotes 4-5, Table 8-5). A large number of respondents also expressed concerns about singing volume. This included whether they could sing more quietly if required (quote 10, Table 8-5), and what impact that would have on their enjoyment and spirituality, expressed in terms of restriction and lack of freedom (quotes 14-15, Table 8-5).

Participant characteristics which may affect acceptability of face masks was also assessed (Table 8-3). Univariable and multivariable regression indicated very few predictors of congregants finding face mask use acceptable. I found that increasing age was associated with a lower likelihood of face mask acceptability, both when evaluating age as a continuous variable (unadjusted OR (uOR) for each additional year of age: 0.98 (95% CI: 0.97-0.99) p=0.005, adjusted OR (aOR): 0.98 (95%CI: 0.96-1.00), p=0.022) and as categories (aOR for 35-

64 years compared to 18-34 years: 0.30 (0.13-0.61) p=0.002; aOR for ≥65 years compared to 18-34 years: 0.38 (0.14-1.02) p=0.062 (reaching borderline significance)). Age categories used in Table 8-3 were defined prior to analysis, but stratifying by narrower, 5-year categories suggested that participants aged less than 40 were more likely to be accepting of face masks compared to participants aged 40 and over (X^2 = 12.47, p=<0.001).

In addition to younger age, people educated to A-level and vocational standard were more accepting of face masks compared to those who had received undergraduate education or higher (uOR: 0.52 (95% CI: 0.29-0.97) p=0.035, aOR: 0.52 (95% CI: 0.28-0.99), p=0.040). Furthermore, there was a trend that men found face mask wearing with quieter singing less acceptable than women (uOR: 0.74, 95%CI: 0.52-1.05, p=0.091, aOR: 0.70, 95%CI: 0.47-1.05, p=0.080), although this did not reach statistical significance. I also assessed if having had suspected or confirmed COVID-19 would predict face mask acceptability, but this was not significant (uOR: 0.76, 95% CI: 0.50-1.19, p=0.220).

8.4.3 Face Mask Comfort

Just under half of respondents (428/858, 49.9%) found wearing face masks in general somewhat or very uncomfortable, but increasing numbers found it uncomfortable wearing face masks for speaking and for singing (Figure 8-3). When asked to provide more details on how comfortable the respondent feels when wearing a face mask while singing/chanting, 564/861 (65.5%) provided a response. Respondents frequently raised similar issues with face mask wearing, especially practical problems such as glasses steaming up (quotes 18-19, Table 8-5) and breathing being less comfortable (quotes 20-21, Table 8-5). Face masks were particularly difficult for asthma sufferers (quotes 22-23, Table 8-5). In addition, many reported that face masks inhibit their ability to communicate properly, particularly for those reliant or partly reliant on lip reading (quotes 32-34, Table 8-5). However, it appeared that face mask comfort improved with increasing use (quotes 40-41, Table 8-5) and was better

tolerated in people who worked in occupations which required their use (quotes 42-43, Table 8-5).

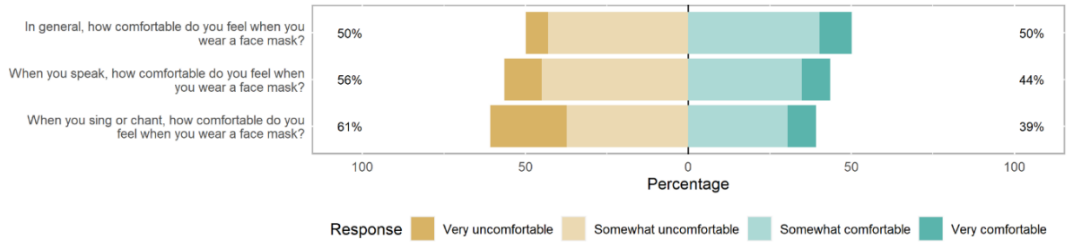


Figure 8-3: Face mask comfort in general, during speaking and singing/chanting.

8.4.4 Changes in Places of Worship During the Pandemic

Respondents reported substantial changes in communal worship post-lockdown. Fewer numbers of people were in attendance, services were shorter, with reduced frequency or no singing/chanting, both on a communal and personal level (Figure 8-4). The notable exception to this was change in frequency of attendance, with over half of respondents reporting no change in their frequency of prayer at their place of worship, although it was unclear whether this represented private prayer or communal worship. The vast majority of participants reported that religious faith was important to them (858/907; 94.6%), and so while fewer congregants were attending services post-lockdown, study respondents may be more willing to attend worship in person than congregants in general.

Most places of worship were reported to be complying with COVID-19-related restrictions: 869/893 (97.3%) reported being aware of COVID-19-related rules in place at their place of worship, with 624/659 (94.7%) and 641/659 (97.3%) reporting that congregants were moderately or very careful to adhere to social distancing and continuously wearing face masks. A large majority (803/887; 90.5%) reported that their place of worship enforced face mask wearing rules and that they were either very or moderately happy with the precautions currently in place (793/887; 89.4%).

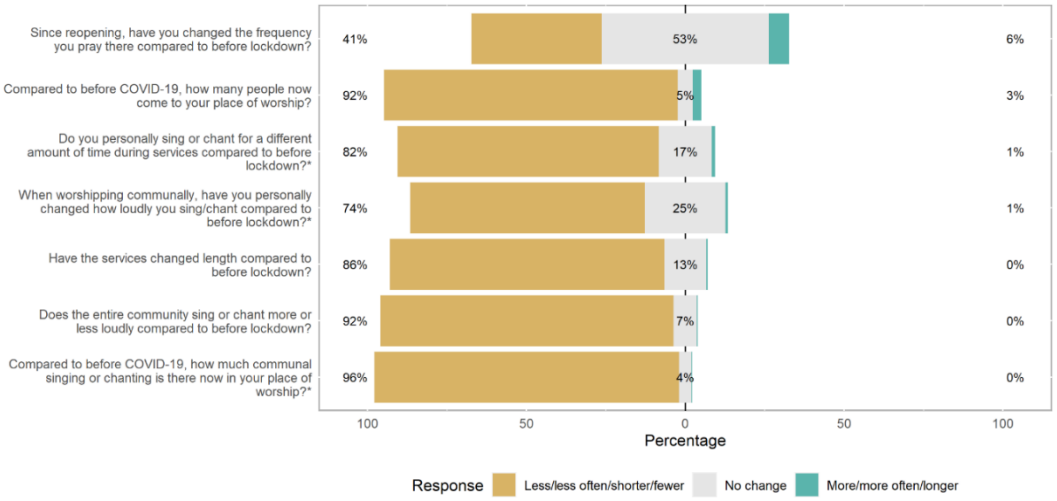


Figure 8-4: COVID-19-related changes to place of worship, comparing post first UK lockdown with pre first UK lockdown. Respondents reporting no singing, either personally or communally, have been excluded from the analysis.

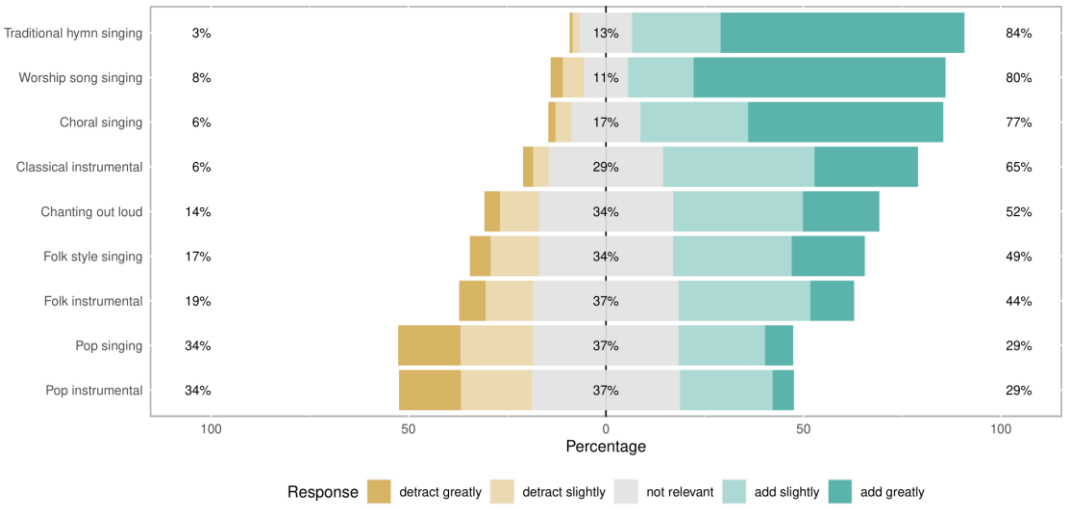


Figure 8-5: Study respondents' views on how different musical styles affect their religious experience.

		Participants (%)			Participants (%)
Compliance of place of worship with government guidelines			**Compliance/satisfaction of congregants with government guidelines**		
			Aware of COVID-19-related rules for places of worship	Yes	869 (97.3%)
				No	24 (2.7%)
Hand sanitiser	Available	825 (87.9%)	**Use of hand sanitiser by congregants upon arrival**	All	402 (61.0%)
	Not available	24 (6.1%)		Most	228 (34.6%)
	Don't know	57 (6.1%)		About half	17 (2.6%)
				A few	11 (1.7%)
				None	1 (0.2%)
Social distancing	Required	850 (95.8%)	**Congregants' care to social distance**	Very careful	400 (60.7%)
	Not required	2 (0.2%)		Moderately careful	224 (34.0%)
	Don't know	35 (3.9%)		Slightly careful	32 (4.9%)
				Not at all careful	3 (0.5%)
Face mask wearing	Enforced	803 (90.5%)	**Congregants' care continuously to wear face masks**	Very careful	546 (82.9%)
	Not enforced	30 (3.4%)		Moderately careful	95 (14.4%)
				Slightly careful	9 (1.4%)
	Don't know	54 (6.1%)		Not at all careful	9 (1.4%)
Air flow from outside (e.g., open doors, windows)	Yes	630 (71.0%)	**Happy with COVID-19 prevention precautions at place of worship**	Very happy	559 (63.0%)
	No	87 (9.8%)		Moderately happy	234 (26.4%)
	Don't know	170 (19.2%)		Slightly happy	51 (5.7%)
				Not at all happy	43 (4.8%)

Table 8-4: Compliance of places of worship and congregants with COVID-19-related restrictions. N.B. many respondents did not answer these questions if they had not visited their place of worship since the start of the COVID-19 pandemic.

8.4.5 Role of Religion and Music

I was also interested in the role of music and how it may affect an individual's religious experience (Figure 8-5). The vast majority of respondents found traditional hymn singing, worship song singing and choral singing most enhanced their religious experience. This contrasted to pop singing and pop instrumental, which was generally felt to be more distracting.

A) Face mask acceptability during singing	
Acceptable if face mask use allowed communal worshippers to sing during services, despite discomfort	
1	*"If the ability to sing is predicated on mask wearing, then what I want is to sing with others, so let's crack on."*
2	*"Any chance to sing would be amazing – even with these limitations!"*
3	*"I would much rather sing with these restrictions than not sing at all."*
Face masks too great a barrier for singing	
4	*"I would not sing while wearing a face mask."*
5	*"Singing with a mask on would not work for me."*
Preference for unrestricted singing online rather than masked singing	
6	*"Still opting for remote singing over masked."*
7	*"We prefer not to attend as the mask wearing and inability to sing together is off putting. We prefer to stay at home and sing as loud as we like with no masks!"*
Face masks distracting / decreasing expression	
8	*"The mask would be distracting and my focus would be changed from worshipping God to feeling distracted by a face covering."*
9	*"Singing and praying together as a community is an integral part of why we worship together. As well as being physically uncomfortable and difficult to focus it creates barriers… we become individuals with no personality, not a community."*
B) Acceptability of reducing singing volume	
Acceptable	
10	*"I don't know if I *can* sing quietly but will have to try!"*
11	*"Quiet singing is far better than no singing, the spiritual words are the most important part of our worship."*
Somewhat unacceptable	
12	*"It was lovely to be part of a socially distanced service but the need to minimise volume of prayer and singing restricted the joy and sense of community."*
13	*"It's better than not singing at all, but I would prefer to be able to sing freely."*
Not acceptable	
14	*"There's no point in singing if you can't do it wholeheartedly."*
15	*"Singing LOUDLY is the best part about singing at church."*
C) Singing as an expression of freedom	
16	*"I hate the idea of anyone singing into a mask. Singing is about freedom."*
17	*"There is usually a freedom in singing which would be affected by these restrictions."*
D) Face mask comfort	
Steaming up glasses	
18	*"Main issue is not being able to wear my glasses as they mist up and I can't see the words without them"*
19	*"Despite following the recommendations for glasses users, my glasses still steam up, if I speak or sing, they steam up even more."*
Breathing difficulties in general	
20	*"It restricts my ability to breathe and be heard slightly"*
21	*"I get too hot and sometimes feel I am not getting enough oxygen."*
Asthma	
22	*"I have asthma, so wearing a face mask is very difficult, and speaking with it even more"*
23	*"I find that occasionally it becomes more of an effort to breathe, and if this happens I need to remove the mask for a couple of minutes. I am asthmatic."*

Breath intake while singing	
24	*"When singing the mask sucks in against my mouth"*
Hot/sweaty	
25	*"Don't like the confinement, stuffiness, impaired contact / expression."*
26	*"I get sweaty and hot if I talk, singing is worse!"*
Sore ears	
27	*"The face masks are relatively comfortable but I get quite hot and they hurt my ears (I wear glasses) if wearing for a long time."*
28	*"I feel smothered, my ears are sore and all this distracts me during worship."*
Face mask movement	
29	*"A face mask… feels like a barrier in my worship to God. I need to continually think about it moving."*
30	*"Masks can tend to slip when mouth movements are made."*
Face masks distracting from worship	
31	*"Church services have been reduced in content, length and numbers, and I have been attending less because of the distress caused by being forced to wear a face covering."*
Communication impairment	
32	*"Mask makes communication more difficult."*
33	*"I can't hear other people when they speak wearing one (as I obviously use mouth signals as well as auditory signals in order to hear)."*
34	*"I find wearing a face mask frustrating as I miss face expressions and they ride up when you speak/sing."*
Impairment of religious expression	
35	*"From the point of view of worship - there is a verse in 2 Corinthians (2:18) that speaks of our 'unveiled faces' reflecting the Lord's glory. While I FULLY accept (and comply with) the need to wear face masks to protect against Covid-19 transmission, at a deep level I feel there is something about wearing a mask that makes my relationship with God in worship less open. I also think in human interaction it somehow de-personalises us and makes communication with one another less 'open'."*
36	*"A mask is restrictive for worship and communication."*
Discomfort increases with duration of use	
37	*"Discomfort increases with time, feeling hot and 'steamed-up'"*
38	*"The length of time I need to wear a mask for affects how comfortable I feel. It becomes more uncomfortable the longer I wear it."*
Discomfort decreases with breaks in use	
39	*"The mask is less than ideal but I get used to it pretty quickly, as long as I can take it off from time to time."*
Discomfort decreases with frequency of use	
40	*"I find the more I wear the mask in different situations, the more comfortable I get with it."*
41	*"Wearing a mask is not particularly comfortable but I have got used to it so it no longer bothers me. I was surprised that it felt OK to sing in a mask but found that it tends to move around and you suck the material in when you breath. I have recently purchased a singer's mask which has more space and fits really securely and is much more comfortable for singing."*
Fewer issues for respondents reporting occupational use of face masks	
42	*"I wear a surgical face mask for healthcare work purposes every weekday - very familiar and comfortable with it.*
43	*"I work in the NHS therefore am used to wearing a mask on a regular basis."*

Table 8-5: Quotations of communal worshippers, entered in free text questions, illustrating the emerging themes.

8.5 Discussion

The study was one of the largest to look at the issue of face masks in the context of religion and worship, and the first in the era of COVID-19. More than half of respondents found wearing of face masks uncomfortable when speaking or singing. They often had practical difficulties such as having sore ears, face mask slippage or steamed up glasses and these issues became more troublesome with age. I also found that younger age and A-level and vocational level education compared to undergraduate education was predictive of better

acceptability of face masks. The latter result is perhaps surprising and could be purely due to chance, but certain vocational occupations such as working in construction may also require routine wearing of face masks and may explain this observation. A large majority of respondents were willing to trade comfort for the ability to sing; 78.1% were prepared to wear face masks and reduce the volume of their singing in order to resume singing or chanting during communal worship. Most of the communal religious worshippers responding to the questionnaire were already used to wearing a face mask for at least an hour. It is likely that since the study started 6 months after the first national lockdown in the UK, attitudes to face masks had started to shift and become part of routine life in the COVID-19 era. In addition, participants may have had fatigue from restrictions on daily life and may be more likely to compromise to return to some semblance of normality. This may be especially true for singing where there is overwhelming evidence to suggest that it enhances worship. Interestingly, respondents already accustomed to wearing face masks such as healthcare workers reported fewer issues with wearing face masks in places of worship.

There was also a stark polarisation of opinion regarding face masks: some believed wearing a face mask was too great a barrier and would not be prepared to sing using one, feeling that it interfered with freedom of religious expression. In contrast, the majority felt that despite the discomfort, it was worth using them to enable worship and singing to continue, although some respondents found them distracting and could be sucked into the mouth.

Previous research in this area is limited, although a cluster-randomised trial did not demonstrate a difference in face mask wearing in the transmission of respiratory viruses during Hajj.[284] However, it should be noted that daily face mask use was low (25% in the intervention group).[284] Reasons given for non-usage of face masks included difficulty in breathing (26%) and discomfort (22%).[284] Several other studies in Hajj pilgrims have cited similar concerns over the non-usage of face masks, although compliance with face masks

increased with increasing perception of effectiveness.[285,286] However, during the COVID-19 pandemic, use of face masks was widespread and in certain scenarios mandatory, hence these previous findings may be less comparable to these results where overall there was good compliance.

It is worth mentioning that face mask wearing was a contentious issue during the pandemic. While there was generally good population-level compliance, there were anti-mask rallies around the world, sometimes associated with violence.[287] The reasoning has often been multifactorial but included discomfort, belief that they were ineffective and violation of civil liberties.[287] Negative attitudes to face mask were also associated with conservative political views.[287] A study from the US suggested that greater religiosity led to reduced adherence to policies to stay at home and reduce social contacts, as it was felt to impinge on personal and religious freedom.[288] However, although several participants mentioned face mask use led to reduced freedom of worship and religious expression, negative attitudes to face mask use and compliance did not appear to be widely held within the cohort.

Finally, the study found that there was good compliance with government guidelines by places of worship in the UK, as well as good overall compliance by congregants. One participant remarked: “I would like to add that of all the places I have been to since the start of COVID-19 my church has by far treated social distancing and disinfecting most seriously. In fact, I feel that going to church has made me more careful because our priest encourages us to abide by the government's recommendations.” There was a high degree of satisfaction with prevention measures, with 89.4% very happy or moderately happy with precautions. It would be worthwhile in future studies to assess how much influence both religious and public health leadership had in encouraging compliance with government guidelines within this cohort.

Religion and collective singing also have wider societal impacts. As an example, collective singing was used to boost morale in Italian cities during lockdown.[289] Religion also helps to create a sense of belonging and can help foster a sense of connection and attachment, reduce feelings of social isolation and improve mental health.[290,291] Religion has been used as a coping mechanism for survival, allowing for a sense of security and hope.[292] Furthermore, religious organisations can play a role in health promotion or provision of welfare, and religious leaders are often seen as pillars of a community, acting as gatekeepers to marginalised or difficult to reach communities.[279,293]

8.5.1 Limitations

The largest limitation to the study was the representativeness of the sample. Despite attempts to maximise inclusivity, the study was biased towards worshippers from London and the South East, of White ethnicity, with university level education and predominantly of Christian, and to a lesser extent Jewish, faith. For speed and ease, as recruitment spanned for just over three months, our research team used a convenience sampling technique, relying on word of mouth and the researchers' local networks to recruit as many participants as possible in the short time. Certain religious groups were underrepresented in the study. For example, while 4.8% of the population in England and Wales identify as Muslim, they only accounted for 0.6% of the study.[294] In contrast, 0.5% of the England and Wales population identify as Jewish, although they accounted for 15.4% of the respondents.[294]. These caveats mean that these results are less generalisable to the UK population. The study was retrospective in design; some questions may be subject to recall bias. The survey was also only available online and shared via digital means, with only English language offered, so participants who do not have access to technology or who are not proficient in English would have been subject to selection bias.

8.6 Conclusions

Although at the time of thesis submission worship in religious buildings had returned back to pre-pandemic practices, the study, one of the largest to date, demonstrated the profound impact the COVID-19 pandemic had on religious worship. These results demonstrated good adherence to COVID-19 guidelines in places of worship and there was a real hunger for a return to normal worship with singing and chanting, even if it meant additional mitigation measures such as wearing a face mask. In addition, face mask discomfort was associated with certain personal factors and health conditions which increase with age, such as wearing of glasses.

Chapter 9: Discussion

In this chapter I will briefly summarise the findings from my thesis, discuss about implications for clinical practice and research and also future directions. I will also discuss limitations to the work before offering final conclusions.

9.1 Overview of Chapters

In Chapter 1: Introduction, I discussed the large global burden of gastrointestinal cancers. Worldwide, gastrointestinal cancers represent over a quarter of all global cancers and over a third of all cancer related deaths.[17] I also discussed some of the risk factors predisposing to oesophageal cancer, gastric cancer and colorectal cancer. Importantly, endoscopy remains a key investigation in the diagnosis of gastrointestinal cancers, which is invasive, unpleasant and costly. Furthermore, the vast majority of endoscopy cases performed do not detect cancer.[62] There is therefore a clinical need to better select patients who undergo endoscopic investigations.

I subsequently outlined my aims and objectives in Chapter 2: Thesis Aims and Objectives.

In Chapter 3: Literature Review of Gastro-Oesophageal Reflux as a Risk Factor for Oesophageal Cancer, I sought to further examine the link between gastro-oesophageal reflux disease (GORD) and the development of oesophageal cancer. In particular, I investigated the link between GORD and oesophageal squamous cell carcinoma (OSCC), the predominant subtype worldwide. While the link for GORD and oesophageal adenocarcinoma (OAC) is well established, I was unable to establish this for OSCC.[81] However, there were only three suitable studies for inclusion for my meta-analysis, which was a major limitation to my research. However, there is some evidence to suggest non-acid gastro-oesophageal reflux may still have a role in the development in oesophageal squamous cell carcinoma.[88]

In Chapter 4: Datasets and Data Curation, I introduced the increasing use of routine health data in clinical research. This has especially been magnified during the COVID-19 pandemic, where rapid analysis and subsequent dissemination of modelling data was instrumental in informing health policy and enacting lockdown.[117] However, there can be issues with data quality and applicability of research questions. I also introduced the datasets being used in the thesis, including both bespoke and routine datasets, and outlined some data quality issues within each dataset.

In Chapter 5: Effect of COVID-19 on Gastrointestinal Services I demonstrated the devastating impact the COVID-19 pandemic had on gastrointestinal services, both locally and nationally. I was able to demonstrate that the COVID-19 led to decreases in the entire patient pathway. This included reductions in the number of referrals received, the number of endoscopic procedures performed, the number gastrointestinal cancer diagnoses made and the number of surgical operations performed. This was most apparent during the month of April 2020. I further demonstrated similar trends nationally for endoscopy in England, where the number of endoscopies performed in April 2020 was at 9.5% of the level in April 2019. Using modelling, I estimated that the total endoscopic backlog reached 476,000 procedures as a direct result of the first wave of the pandemic in January 2021. Furthermore, even with a 30% increase in capacity, it would still take 17 months to overcome the accumulated backlog, while a 2-month interruption of services could add an extra 15.4% to the accumulated backlog.

In Chapter 6: Creating a Risk Prediction Model, I sought to develop solutions to see what could be done to better use already stretched resources to overcome the identified national endoscopic backlog. I used questionnaire data to create a risk prediction model to predict the presence of oesophageal and gastric cancer. Using machine learning methods I created and optimised a model based on logistic regression containing 17 features, with a median

area under the curve (AUC) of 0.81 and interquartile range (IQR) of 0.69–0.85. I was also able to demonstrate that my model outperformed the Edinburgh Dysphagia Score (EDS), which was recommended by the British Society of Gastroenterology (BSG) during the pandemic as a means of triaging upper gastrointestinal referrals.[169–171] In addition, I was interested to see if there would be any synergism between questionnaire factors and epigenetic changes in saliva, although disappointingly my results did not support this.

In Chapter 7: Faecal Immunochemical Testing for Triaging Lower Gastrointestinal Referrals, I assessed the use of the FIT in managing lower GI referrals. Using primary care data, I was able to demonstrate that a raised FIT was associated with either cancer or cancer/polyp as a composite endpoint. This finding was consistent with published studies which I also reviewed at the start of the chapter. I also attempted to assess the value of a duplicate FIT testing strategy to see if it may improve the sensitivity and specificity of the FIT. I was able to demonstrate that a duplicate FIT strategy could potentially detect a further 32% of cases with a composite endpoint of cancer/polyp. However, for cancer alone as an endpoint, there was insufficient data to draw firm conclusions regarding the use of a duplicate FIT strategy.

In Chapter 8:Societal Impacts of COVID-19 Pandemic, I assessed the impact of the pandemic on society, with a focus on religious worship. Overall, I found there was good compliance by places of worship with then enforced government guidelines such as social distancing and wearing of face coverings. However, a significant proportion of respondents found wearing face masks uncomfortable and raised practical concerns such as steaming up of glasses. This chapter allowed me to appreciate some of the wider effects on the pandemic and also afforded me the opportunity to use more qualitative research methods to analyse data.

9.2 Implications for Clinical Practice and Research

The COVID-19 pandemic has been a huge disruptor in the delivery of both emergency and routine healthcare. Many of us would have been shocked by the images of overburdened hospitals and overflowing mortuaries filling news bulletins in the early phase of the pandemic. Society will continue to live with the aftermath of this disruption for some years. In addition, there are additional challenges in the UK such as widening health inequalities, increasing life expectancy, underinvestment in healthcare and increasing burden of long term illnesses which affect recovery of services.[295] These significant challenges mean that new methods of working are needed to best utilise the limited resources that are available. One method could be improved assessment and triage of patients in order to better select patients who should undergo investigation.

In my thesis, I highlighted some of the effects of the COVID-19 pandemic in healthcare, and estimated the endoscopic backlog to be 476,000 procedures directly attributable to the pandemic.[3] To find a potential solution, I used bespoke trial data to create a triaging tool using machine learning, which was demonstrated to outperform currently used scoring systems.[6] The advantage of the tool is that patients can be sent the questionnaire before their consultation or referral to specialist services and have a risk score calculated, which would help prioritise urgency of the referral and even whether the patient should undergo endoscopic examination. In addition, the methodology I have used could be replicated for other datasets and specific questionnaires could be created for different referral pathways based on patient symptomatology.

While in my thesis I have created a risk prediction model using data from hundreds of patients, the potential is to increase the number of patients included and to increase the types of data used in model creation, such as free text notes and laboratory results.[296] This is increasingly possible due to digitalisation of electronic health records and also linkage of

healthcare records between different organisations, such as between primary and secondary care. Many commonly used scoring systems today were historically based on relatively small datasets. As an example, the updated Child-Pugh classification for liver cirrhosis related mortality was based on outcomes for 38 patients.[297] Despite this, the classification was only recently abandoned in treatment algorithms for hepatocellular carcinoma, with authors responsible for the Barcelona Clinic Liver Cancer staging system suggesting that the classification was due for an update.[298] Using big data to create risk prediction models can therefore present an opportunity to create more accurate models, based on real world patient data, thereby improving patient outcomes.

A further hypothetical future direction is whether machine learning models could even replace standard laboratory tests, such as those used in population-based screening programmes. Patients are screened in the UK bowel cancer screening programme using the faecal immunochemical test (FIT) and offered a colonoscopy if their FIT exceeds the threshold. However, processing a large number of FITs and performing screening colonoscopies is resource intensive, requiring specialist equipment and trained healthcare professionals in different fields, e.g. pathology and gastroenterology, to deliver.[224,299] It has been postulated that machine learning approaches could be used in under resourced settings such as Sub-Saharan African for surveillance and early detection of colorectal cancers.[299] Potential methods could include analysis of subtle trends in laboratory tests which a patient may undergo as part of routine clinical care, or automated artificial intelligence (AI) reading of histology from colonoscopy, which would reduce pathology burden.[299] The challenge remains in ensuring that there is no significant difference in cancer detection between machine learning tools and current standard of care. Moreover, there must be timely and widespread adoption of such tools such that it leads to a strong impact across large populations.[299]

For now, the most obvious next steps would be trialling the machine learning tool I have created within Chapter 6: Creating a Risk Prediction Model on a large, multi-centre population so that it could be validated and its generalisability to other populations assessed. In addition, further refinements could be added to it by incorporating features such as laboratory test results extracted from electronic health records. To aid collection of data, patients attending clinic with predefined symptoms or on specific healthcare pathways such as the 2-week wait suspected cancer pathway could be automatically sent the questionnaire to complete prior to the hospital visit, with results potentially available during the consultation to better inform clinical decision-making.

9.3 Limitations

Firstly, the huge increase in interest in machine learning has meant that a large number of models can be generated easily and quickly. However, there is often a lack of head-to-head comparison between models created by different research groups and how easily it can be applied especially within a frontline medical setting. This often limits their utility in routine clinical practice.

Secondly, the quality of any potential model is wholly dependent on the quality of the data used to create the model. Data needs to be reliable and be reflective of everyday clinical practice, and there also needs to be appropriate data cleaning prior to its use in model creation.[296] There may be issues with how transferable models are between different populations, especially if there is evidence of overfitting within a model, which would limit its generalisability to different populations.[300]

Finally, the most significant issue is likely to be based around ethics and governance of the use of AI or machine learning generated tools, especially if they are used autonomously for risk stratification within a population.[301] While AI-clinician cooperation could lead to

improvements in diagnosis and improved organisational efficiency, there needs to be appropriate oversight especially within healthcare where there is strict regulation of medical practitioners and premises.[301] There is a danger that clinicians may become over reliant on them. If harm is caused by a decision support software working autonomously, it is unclear where responsibility would lie at present and whether robust enough mechanisms are in place to prevent similar instances from reoccurring.[301]

9.4 Final Conclusions

The COVID-19 pandemic was described as a "once in a century" event.[302] It affected all facets of society, and we are still living with some of the consequences, with rising waiting lists for medical procedures and treatments. However, there is also an opportunity to reform how healthcare could be delivered. I have demonstrated that AI methods such as machine learning could be used to help improve triaging of referrals and thus improve the selection of patients undergoing specialist tests. The increasing accessibility of routine health data incorporating many more patients and clinical parameters will undoubtedly improve the performance of models and its applicability to different populations. Although, there remains ethical considerations which need to be resolved prior to widespread adoption, increasing use of AI within healthcare is almost inevitable, and will transform the delivery of patient care.

Chapter 10: References

1 Watson C. Rise of the preprint: how rapid data sharing during COVID-19 has changed science forever. *Nat Med* 2022; **28**: 2–5.

2 The Lancet Gastroenterology & Hepatology. About The Lancet Gastroenterology & Hepatology. https://www.thelancet.com/langas/about (accessed May 5, 2024).

3 Ho KMA, Banerjee A, Lawler M, Rutter MD, Lovat LB. Predicting endoscopic activity recovery in England after COVID-19: a national analysis. *Lancet Gastroenterol Hepatol* 2021; **6**: 381–90.

4 Ho KMA, Davies H, Epstein R, *et al.* Spatiotemporal droplet dispersion measurements demonstrate face masks reduce risks from singing. *Sci Rep* 2021; **11**: 24183.

5 Ho KMA, Baggaley RF, Stone TC, *et al.* Face Mask Acceptability for Communal Religious Worship During the COVID-19 Pandemic in the United Kingdom: Results from the CONFESS Study. *J Relig Health* 2023; **62**: 608–26.

6 Ho KMA, Rosenfeld A, Hogan Á, *et al.* Development and validation of a multivariable risk factor questionnaire to detect oesophageal cancer in 2-week wait patients. *Clin Res Hepatol Gastroenterol* 2023; **47**: 102087.

7 Ho KMA, Anandhakrishnan A, Mahay A, Soo Y, Lovat LB, Rochford AP. How COVID-19 has changed the unselected medical take: an observational study. *Clin Med* 2020; **20**: e229–33.

8 Wolfson P, Ho KMA, Bassett P, *et al.* Accuracy of clinical staging for T2N0 oesophageal cancer: Systematic review and meta-analysis. *Dis Esophagus* 2021; **34**: 1–12.

9 Wolfson P, Ho KMA, Wilson A, *et al.* Endoscopic eradication therapy for Barrett's esophagus–related neoplasia: a final 10-year report from the UK National HALO Radiofrequency Ablation Registry. *Gastrointest Endosc* 2022; **96**: 223–33.

10 Maity AK, Stone TC, Ward V, *et al.* Novel epigenetic network biomarkers for early detection of esophageal cancer. *Clin Epigenetics* 2022; **14**: 1–14.

11 Stone TC, Ward V, Hogan A, *et al.* Using saliva epigenetic data to develop and validate a multivariable predictor of esophageal cancer status. *Epigenomics* 2024; **16**: 109–25.

12 Ho KMA, Rosenfeld A, Hogan Á, *et al.* O5 Using machine learning to develop models for the prediction of upper gastrointestinal cancers. *Gut* 2022; **71**: A3.

13 Ho KMA, Rosenfeld A, Hogan Á, *et al.* P105 Validation of a machine learning model in a prospective cohort: the RISQ study. *Gut* 2022; **71**: A92–A92.

14 Ho KMA, Rosenfeld A, Hogan Á, *et al.* PTH-72 A symptom and risk factor questionnaire accurately predicts upper gastrointestinal cancer. *Gut* 2021; **70**: A136.

15 Ho KMA, Rosenfeld A, Hogan, Á, McBain H, *et al.* P0074 A symptom and risk factor questionnaire accurately predicts upper gastrointestinal cancer. *United Eur Gastroenterol J* 2021; **9**: 302.

16 Ho A, Wolfson P, Wilson A, *et al.* P28 Comparison of anxiety and depression scores between 2-week wait and Barrett's surveillance endoscopy referrals. *Gut* 2021; **70**: A56–A56.

17 Arnold M, Abnet CC, Neale RE, *et al.* Global Burden of 5 Major Types of Gastrointestinal Cancer. *Gastroenterology* 2020; **159**: 335-349.e15.

18 Cancer Research UK. Cancer statistics for the UK. Cancer Res. UK. 2025. http://www.cancerresearchuk.org/health-professional/cancer-statistics (accessed May 1, 2025).

19 Kamangar F, Nasrollahzadeh D, Safiri S, *et al.* The global, regional, and national burden of oesophageal cancer and its attributable risk factors in 195 countries and territories, 1990-2017: A systematic analysis for the Global Burden of Disease Study 2017. *Lancet Gastroenterol Hepatol* 2020; **5**: 582–97.

20 Arnold M, Soerjomataram I, Ferlay J, Forman D. Global incidence of oesophageal cancer by histological subtype in 2012. *Gut* 2015; **64**: 381–7.

21 Lagergren J, Smyth E, Cunningham D, Lagergren P. Oesophageal cancer. Lancet. 2017; **390**: 2383–96.

22 Spechler SJ, Souza RF. Barrett's Esophagus. *N Engl J Med* 2014; **371**: 836–45.

23 Coleman HG, Xie SH, Lagergren J. The Epidemiology of Esophageal Adenocarcinoma. *Gastroenterology* 2018; **154**: 390–405.

24 Schneider JL, Corley DA. A review of the epidemiology of Barrett's oesophagus and oesophageal adenocarcinoma. *Best Pract Res Clin Gastroenterol* 2015; **29**: 29–39.

25 Bornschein J, Bird-Lieberman EL, Malfertheiner P. The Rationale and Efficacy of Primary and Secondary Prevention in Adenocarcinomas of the Upper Gastrointestinal Tract. Dig. Dis. 2019; **37**: 381–93.

26 Orloff M, Peterson C, He X, *et al.* Germline mutations in MSR1, ASCC1, and CTHRC1 in patients with Barrett esophagus and esophageal adenocarcinoma. *JAMA - J Am Med Assoc* 2011; **306**: 410–9.

27 Fischbach LA, Nordenstedt H, Kramer JR, *et al.* The association between Barrett's esophagus and Helicobacter pylori infection: A meta-analysis. *Helicobacter* 2012; **17**: 163–75.

28 Corley DA, Kubo A, Levin TR, *et al.* Helicobacter pylori and gastroesophageal reflux disease: A case-control study. *Helicobacter* 2008; **13**: 352–60.

29 Nie S, Chen T, Yang X, Huai P, Lu M. Association of Helicobacter pylori infection with esophageal adenocarcinoma and squamous cell carcinoma: A meta-analysis. *Dis Esophagus* 2014; **27**: 645–53.

30 Jansson C, Johansson ALV, Nyrén O, Lagergren J. Socioeconomic factors and risk of esophageal adenocarcinoma: A nationwide Swedish case-control study. *Cancer Epidemiol Biomarkers Prev* 2005; **14**: 1754–61.

31 Abnet CC, Arnold M, Wei WQ. Epidemiology of Esophageal Squamous Cell Carcinoma. *Gastroenterology* 2018; **154**: 360–73.

32 Domper Arnal MJ, Ferrández Arenas Á, Lanas Arbeloa Á. Esophageal cancer: Risk factors, screening and endoscopic treatment in Western and Eastern countries. *World J Gastroenterol* 2015; **21**: 7933–43.

33 Tran GD, Sun X Di, Abnet CC, *et al.* Prospective study of risk factors for esophageal and gastric cancers in the Linxian General Population Trial cohort in China. *Int J Cancer*

2005; **113**: 456–63.

34 Islami F, Boffetta P, Ren JS, Pedoeim L, Khatib D, Kamangar F. High-temperature beverages and foods and esophageal cancer risk - A systematic review. *Int J Cancer* 2009; **125**: 491–524.

35 Lin Y, Totsuka Y, He Y, *et al.* Epidemiology of esophageal cancer in Japan and China. J. Epidemiol. 2013; **23**: 233–42.

36 Brown LM, Hoover R, Silverman D, *et al.* Excess incidence of squamous cell esophageal cancer among US black men: Role of social class and other risk factors. *Am J Epidemiol* 2001; **153**: 114–22.

37 Kgomo M, Mokoena TR, Ker JA. Non-Acid gastro-oesophageal reflux is associated with squamous cell carcinoma of the oesophagus. *BMJ Open Gastroenterol* 2017; **4**: e000180.

38 Uno K, Iijima K, Hatta W, *et al.* Direct measurement of gastroesophageal reflux episodes in patients with squamous cell carcinoma by 24-h pH-impedance monitoring. *Am J Gastroenterol* 2011; **106**: 1923–9.

39 Islami F, Sheikhattari P, Ren JS, Kamangar F. Gastric atrophy and risk of oesophageal cancer and gastric cardia adenocarcinoma-a systematic review and meta-analysis. Ann. Oncol. 2011; **22**: 754–60.

40 Cook MB. Editorial: Non-Acid reflux: The missing link between gastric atrophy and esophageal squamous cell carcinoma. *Am J Gastroenterol* 2011; **106**: 1930–2.

41 Dar NA, Shah IA, Bhat GA, *et al.* Socioeconomic status and esophageal squamous cell carcinoma risk in Kashmir, India. *Cancer Sci* 2013; **104**: 1231–6.

42 Islami F, Kamangar F, Nasrollahzadeh D, *et al.* Socio-economic status and oesophageal cancer: Results from a population-based case-control study in a high-risk area. *Int J Epidemiol* 2009; **38**: 978–88.

43 Smyth EC, Nilsson M, Grabsch HI, van Grieken NC, Lordick F. Gastric cancer. Lancet. 2020; **396**: 635–48.

44 Kumar S, Metz DC, Ellenberg S, Kaplan DE, Goldberg DS. Risk Factors and Incidence of Gastric Cancer After Detection of Helicobacter pylori Infection: A Large Cohort Study. *Gastroenterology* 2020; **158**: 527-536.e7.

45 Dekker E, Tanis PJ, Vleugels JLA, Kasi PM, Wallace MB. Colorectal cancer. Lancet. 2019; **394**: 1467–80.

46 Krämer HU, Schöttker B, Raum E, Brenner H. Type 2 diabetes mellitus and colorectal cancer: Meta-analysis on sex-specific differences. *Eur J Cancer* 2012; **48**: 1269–82.

47 Song M, Garrett WS, Chan AT. Nutrients, foods, and colorectal cancer prevention. *Gastroenterology* 2015; **148**: 1244-1260.e16.

48 Algra AM, Rothwell PM. Effects of regular aspirin on long-term cancer incidence and metastasis: A systematic comparison of evidence from observational studies versus randomised trials. *Lancet Oncol* 2012; **13**: 518–27.

49 Pashayan N, Pharoah PDP. The challenge of early detection in cancer. *Science (80-)* 2020; **368**: 589–90.

50 Crosby D, Lyons N, Greenwood E, *et al.* A roadmap for the early detection and diagnosis of cancer. Lancet Oncol. 2020; **21**: 1397–9.

51 Granger SP, Preece RAD, Thomas MG, Dixon SW, Chambers AC, Messenger DE. Colorectal cancer incidence trends by tumour location among adults of screening-age in England: a population-based study. *Color Dis* 2023; **25**: 1771–82.

52 NHS England. NHS rolls out lifesaving home testing for bowel cancer to over 50s. NHS Engl. 2025. https://www.england.nhs.uk/2025/01/nhs-rolls-out-lifesaving-home-testing-for-bowel-cancer-to-over-50s/ (accessed Jan 14, 2025).

53 Koo S, Neilson LJ, Von Wagner C, Rees CJ. The NHS bowel cancer screening program: Current perspectives on strategies for improvement. *Risk Manag Healthc Policy* 2017; **10**: 177–87.

54 Al-Zubaidi S, McKigney N, Coyne PE. Bowel cancer screening. Surg. (United Kingdom). 2023; **41**: 22–9.

55 Li SJ, Sharples LD, Benton SC, *et al.* Faecal immunochemical testing in bowel cancer screening: Estimating outcomes for different diagnostic policies. *J Med Screen* 2021; **28**: 277–85.

56 Cancer Research UK. Bowel scope screening to stop in England. 2021. https://news.cancerresearchuk.org/2021/01/14/bowel-scope-screening-to-stop-in-england/ (accessed May 1, 2025).

57 Hirst Y, Stoffel S, Baio G, Mcgregor L, Von Wagner C. Uptake of the English Bowel (Colorectal) Cancer Screening Programme: an update 5 years after the full roll-out. *Eur J Cancer* 2018; **103**: 267–73.

58 Fitzgerald RC, Di Pietro M, Ragunath K, *et al.* British Society of Gastroenterology guidelines on the diagnosis and management of Barrett's oesophagus. *Gut* 2014; **63**: 7–42.

59 Banks M, Graham D, Jansen M, *et al.* British Society of Gastroenterology guidelines on the diagnosis and management of patients at risk of gastric adenocarcinoma. *Gut* 2019; **68**: 1545–75.

60 Monahan KJ, Bradshaw N, Dolwani S, *et al.* Guidelines for the management of hereditary colorectal cancer from the British Society of Gastroenterology (BSG)/Association of Coloproctology of Great Britain and Ireland (ACPGBI)/United Kingdom Cancer Genetics Group (UKCGG). *Gut* 2020; **69**: 411–44.

61 Lamb CA, Kennedy NA, Raine T, *et al.* British Society of Gastroenterology consensus guidelines on the management of inflammatory bowel disease in adults. *Gut* 2019; **68**: s1–106.

62 Vakil N, Talley N, van Zanten SV, *et al.* Cost of Detecting Malignant Lesions by Endoscopy in 2741 Primary Care Dyspeptic Patients Without Alarm Symptoms. *Clin Gastroenterol Hepatol* 2009; **7**: 756–61.

63 van Erp NF, Helsper CW, Slottje P, *et al.* Time to diagnosis of symptomatic gastric and oesophageal cancer in the Netherlands: Where is the room for improvement? *United Eur Gastroenterol J* 2020; **8**: 607–20.

64 Yang H, Wang F, Hallemeier CL, Lerut T, Fu J. Oesophageal cancer. *Lancet* 2024; **404**: 1991–2005.

65 Gupta S, Li D, El Serag HB, *et al.* AGA Clinical Practice Guidelines on Management of Gastric Intestinal Metaplasia. *Gastroenterology* 2020; **158**: 693–702.

66 Collaborative Covids, Nepogodiev D, Omar OM, *et al.* Elective surgery cancellations due to the COVID-19 pandemic: global predictive modelling to inform surgical recovery plans. *Br J Surg* 2020; **107**: 1440–9.

67 Carr A, Smith JA, Camaradou J, Prieto-Alhambra D. Growing backlog of planned surgery due to covid-19. *BMJ* 2021; **372**: n339.

68 Goddard KAB, Feuer EJ, Umar A, Castle PE. Accelerating progress to reduce the cancer burden through prevention and control in the United States. *JNCI J Natl Cancer Inst* 2025; **117**: 20–8.

69 Wyper GMA, Mackay DF, Fraser C, *et al.* Evaluating the impact of alcohol minimum unit pricing on deaths and hospitalisations in Scotland: a controlled interrupted time series study. *Lancet* 2023; **401**: 1361–70.

70 Leja M. Where are we with gastric cancer screening in Europe in 2024? *Gut* 2024; **73**: 2074–82.

71 Cancer Research UK. Be Clear on Cancer. 2016. https://www.cancerresearchuk.org/health-professional/awareness-and-prevention/be-clear-on-cancer (accessed March 21, 2025).

72 Sun D, Mülder DT, Li Y, *et al.* The Effect of Nationwide Organized Cancer Screening Programs on Gastric Cancer Mortality: A Synthetic Control Study. *Gastroenterology* 2024; **166**: 503–14.

73 Von Wagner C, Stoffel ST, Freeman M, *et al.* General practitioners' awareness of the recommendations for faecal immunochemical tests (FITs) for suspected lower gastrointestinal cancers: A national survey. *BMJ Open* 2019; **9**: 25737.

74 NICE. Gastrointestinal tract (lower) cancers - recognition and referral. Natl. Inst. Heal. Care Excell. 2021. https://cks.nice.org.uk/topics/gastrointestinal-tract-upper-cancers-recognition-referral/ (accessed March 21, 2025).

75 Guan WL, He Y, Xu RH. Gastric cancer treatment: recent progress and future perspectives. J. Hematol. Oncol. 2023; **16**: 1–28.

76 Fadlallah H, El Masri J, Fakhereddine H, *et al.* Colorectal cancer: Recent advances in management and treatment. *World J Clin Oncol* 2024; **15**: 1136–56.

77 Davila RE, Rajan E, Adler DG, *et al.* ASGE Guideline: The role of endoscopy in the patient with lower-GI bleeding. *Gastrointest Endosc* 2005; **62**: 656–60.

78 Vakil N, Van Zanten S V, Kahrilas P, Dent J, Jones R. The Montreal Definition and Classification of Gastroesophageal Reflux Disease: A Global Evidence-Based Consensus CME. *Am J Gastroenterol* 2006; **101**: 1900–20.

79 Lagergren J, Bergström R, Lindgren A, Nyrén O. Symptomatic Gastroesophageal Reflux as a Risk Factor for Esophageal Adenocarcinoma. *N Engl J Med* 1999; **340**: 825–31.

80 Chow W-H, Fraumeni JF, Finkle WD, Mclaughlin JK, Frankl H, Ziel HK. The Relation of Gastroesophageal Reflux Disease and Its Treatment to Adenocarcinomas of the Esophagus and Gastric Cardia. *JAMA J Am Med Assoc* 1995; **274**: 474–7.

81 Rubenstein JH, Taylor JB. Meta-analysis: The association of oesophageal adenocarcinoma with symptoms of gastro-oesophageal reflux. *Aliment Pharmacol Ther* 2010; **32**: 1222–7.

82 Islami F, Kamangar F. Helicobacter pylori and esophageal cancer risk: A meta-analysis. *Cancer Prev Res* 2008; **1**: 329–38.

83 Stapley S, Peters TJ, Neal RD, Rose PW, Walter FM, Hamilton W. The risk of oesophago-gastric cancer in symptomatic patients in primary care: A large case-control study using electronic records. *Br J Cancer* 2013; **108**: 25–31.

84 Spechler SJ. Barrett Esophagus and Risk of Esophageal Cancer: A Clinical Review. *JAMA* 2013; **310**: 627–36.

85 Dulai GS, Gornbein J, Kahn KL, Guha S, Weinstein WM. Preoperative prevalence of Barrett's esophagus in esophageal adenocarcinoma: A systematic review. *Gastroenterology* 2002; **122**: 26–33.

86 Eusebi LH, Cirota GG, Zagari RM, Ford AC. Global prevalence of Barrett's oesophagus and oesophageal cancer in individuals with gastro-oesophageal reflux: A systematic review and meta-analysis. *Gut* 2021; **70**: 456–63.

87 Qumseya BJ, Bukannan A, Gendy S, *et al.* Systematic review and meta-analysis of prevalence and risk factors for Barrett's esophagus. *Gastrointest Endosc* 2019; **90**: 707-717.e1.

88 Soroush A, Etemadi A, Abrams JA. Non-Acid Fluid Exposure and Esophageal Squamous Cell Carcinoma. Dig. Dis. Sci. 2022; **67**: 2754–62.

89 Wang S-M, Freedman ND, Katki HA, *et al.* Gastroesophageal reflux disease: A risk factor for laryngeal squamous cell carcinoma and esophageal squamous cell carcinoma in the NIH-AARP Diet and Health Study cohort. *Cancer* 2021; **127**: 1871–9.

90 El-Serag HB, Sweet S, Winchester CC. Update on the epidemiology of gastro-oesophageal reflux disease: a systematic review. *Gut* 2014; **63**: 871–80.

91 National Institute for Health and Care Research. PROSPERO: International prospective register of systematic reviews - Inclusion criteria. 2023; **44**: 1–4.

92 Wells G, Shea B, O'Connell D, Peterson J. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. Ottawa, Ottawa Hosp. Res. Inst. 2000. https://www.ohri.ca//programs/clinical_epidemiology/oxford.asp (accessed May 1, 2025).

93 R Core Team. R: A language and environment for statistical computing. 2021. https://www.r-project.org/.

94 Balduzzi S, Rücker G, Schwarzer G. Statistics in practice How to perform a meta-analysis with R: a practical tutorial. *Evid Based Ment Heal* 2019; **22**: 153–60.

95 Page MJ, McKenzie JE, Bossuyt PM, *et al.* The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021; **372**: n71.

96 Farrow DC, Vaughan TL, Sweeney C, *et al.* Gastroesophageal reflux disease, use of H2 receptor antagonists, and risk of esophageal and gastric cancer. Cancer Causes Control. 2000; **11**: 231–8.

97 Wu AH, Tseng CC, Bernstein L. Hiatal hernia, reflux symptoms, body size, and risk of esophageal and gastric adenocarcinoma. *Cancer* 2003; **98**: 940–8.

98 Anderson LA, Watson RGP, Murphy SJ, *et al.* Risk factors for Barrett's oesophagus and oesophageal adenocarcinoma: results from the FINBAR study. *World J Gastroenterol* 2007; **13**: 1585–94.

99 Crane SJ, Locke GR 3rd, Harmsen WS, *et al.* Subsite-specific risk factors for esophageal and gastric adenocarcinoma. *Am J Gastroenterol* 2007; **102**: 1596–602.

100 Whiteman DC, Sadeghi S, Pandeya N, *et al.* Combined effects of obesity, acid reflux and smoking on the risk of adenocarcinomas of the oesophagus. *Gut* 2008; **57**: 173–80.

101 Pandeya N, Webb PM, Sadeghi S, Green AC, Whiteman DC. Gastro-oesophageal reflux symptoms and the risks of oesophageal cancer: Are the effects modified by smoking, NSAIDs or acid suppressants? *Gut* 2010; **59**: 31–8.

102 Thrift AP, Kendall BJ, Pandeya N, Whiteman DC. A model to determine absolute risk for esophageal adenocarcinoma. *Clin Gastroenterol Hepatol* 2013; **11**: 138-44.e2.

103 Cook MB, Corley DA, Murray LJ, *et al.* Gastroesophageal Reflux in Relation to Adenocarcinomas of the Esophagus: A Pooled Analysis from the Barrett's and Esophageal Adenocarcinoma Consortium (BEACON). *PLoS One* 2014; **9**: e103508.

104 Drahos J, Xiao Q, Risch HA, *et al.* Age-specific risk factor profiles of adenocarcinomas of the esophagus: A pooled analysis from the international BEACON consortium. *Int J Cancer* 2016; **138**: 55–64.

105 Wu AH, Wan P, Bernstein L. A multiethnic population-based study of smoking, alcohol and body size and risk of adenocarcinomas of the stomach and esophagus (United States). *Cancer Causes Control* 2001; **12**: 721–32.

106 Gammon MD, Schoenberg JB, Ahsan H, *et al.* Tobacco, alcohol, and socioeconomic status and adenocarcinomas of the esophagus and gastric cardia. *J Natl Cancer Inst* 1997; **89**: 1277–84.

107 Lu P, Gu J, Zhang N, Sun Y, Wang J. Risk factors for precancerous lesions of esophageal squamous cell carcinoma in high-risk areas of rural China: A population-based screening study. *Medicine (Baltimore)* 2020; **99**: e21426.

108 Zikos TA, Clarke JO. Non-acid Reflux: When It Matters and Approach to Management. Curr. Gastroenterol. Rep. 2020; **22**: 43.

109 Smout AJPM. Review article: The measurement of non-acid gastro-oesophageal reflux. *Aliment Pharmacol Ther* 2007; **26**: 7–12.

110 NCD Risk Factor Collaboration (NCD-RisC). Worldwide trends in underweight and obesity from 1990 to 2022: a pooled analysis of 3663 population-representative studies with 222 million children, adolescents, and adults. *Lancet (London, England)* 2024; **403**: 1027–50.

111 The Nuffield Trust. Obesity. Nuff. Trust. 2024. https://www.nuffieldtrust.org.uk/resource/obesity (accessed April 18, 2025).

112 Dai X, Gakidou E, Lopez AD. Evolution of the global smoking epidemic over the past half century: Strengthening the evidence base for policy action. *Tob Control* 2022; **31**: 129–37.

113 Jones R, Junghard O, Dent J, *et al.* Development of the GerdQ, a tool for the diagnosis and management of gastro-oesophageal reflux disease in primary care. *Aliment Pharmacol Ther* 2009; **30**: 1030–8.

114 Jones R, Coyne K, Wiklund I. The Gastro-oesophageal Reflux Disease Impact Scale: A patient management tool for primary care. *Aliment Pharmacol Ther* 2007; **25**: 1451–9.

115 Banerjee A, Mathew D, Rouane K. Using patient data for patients' benefit. *BMJ* 2017; **358**: j4413.

116 Department of Health & Social Care. Data saves lives: reshaping health and social care with data (Updated 15 June 2022). 2022 https://www.gov.uk/government/publications/data-saves-lives-reshaping-health-and-social-care-with-data (accessed May 1, 2025).

117 Ferguson N, Laydon D, Nedjati Gilani G, *et al.* Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand. 2020 DOI:10.25561/77482.

118 Banerjee A, Pasea L, Harris S, *et al.* Estimating excess 1-year mortality associated with the COVID-19 pandemic according to underlying conditions and age: a population-based cohort study. *Lancet* 2020; **395**: 1715–25.

119 Lai AG, Pasea L, Banerjee A, *et al.* Estimated impact of the COVID-19 pandemic on cancer services and excess 1-year mortality in people with cancer and multimorbidity: Near real-time data on cancer care, cancer deaths and a population-based cohort study. *BMJ Open* 2020; **10**. DOI:10.1136/bmjopen-2020-043828.

120 Maringe C, Spicer J, Morris M, *et al.* The impact of the COVID-19 pandemic on cancer deaths due to delays in diagnosis in England, UK: a national, population-based, modelling study. *Lancet Oncol* 2020; **21**: 1023–34.

121 Sud A, Torr B, Jones ME, *et al.* Effect of delays in the 2-week-wait cancer referral pathway during the COVID-19 pandemic on cancer survival in the UK: a modelling study. *Lancet Oncol* 2020; **21**: 1035–44.

122 Whittaker H, Quint JK. Using routine health data for research: The devil is in the detail. Thorax. 2020; **75**: 714–5.

123 Nicholls SG, Langan SM, Benchimol EI. Routinely collected data: The importance of high-quality diagnostic coding to research. *Cmaj* 2017; **189**: E1054–5.

124 Hemkens LG, Contopoulos-Ioannidis DG, Ioannidis JPA. Routinely collected data and comparative effectiveness evidence: Promises and limitations. *CMAJ* 2016; **188**: E158–64.

125 Tang KL, Lucyk K, Quan H. Coder perspectives on physician-related barriers to producing high-quality administrative data: a qualitative study. *Can Med Assoc Open Access J* 2017; **5**: E617–22.

126 McCormick N, Hamilton CB, Koehn CL, English K, Stordy A, Li LC. Canadians' views on the use of routinely collected data in health research: a patient-oriented cross-sectional survey. *C Open* 2019; **7**: E203–9.

127 Johns Hopkins University. COVID-19 Map. Johns Hopkins Coronavirus Resour. Cent. 2020. https://coronavirus.jhu.edu/map.html (accessed Aug 31, 2021).

128 Zigmond AS, Snaith RP. The Hospital Anxiety and Depression Scale. *Acta Psychiatr Scand* 1983; **67**: 361–70.

129 Schraer R. Covid: Faith groups' singing studied for coronavirus risk. BBC News. 2020. https://www.bbc.co.uk/news/health-54390549 (accessed May 1, 2025).

130 World Health Organization. ICD-10 Version:2019. 2019. https://icd.who.int/browse10/2019/en#/C15-C26 (accessed May 1, 2025).

131 NHS England. Monthly Diagnostic Waiting Times and Activity. https://www.england.nhs.uk/statistics/statistical-work-areas/diagnostics-waiting-times-and-activity/monthly-diagnostics-waiting-times-and-activity/ (accessed July 5, 2021).

132 Ministry of Housing Communities and Local Government. English indices of deprivation. 2019 https://www.gov.uk/government/statistics/english-indices-of-deprivation-2019 (accessed May 1, 2025).

133 London Borough of Barnet. Joint Strategic Needs Assessment. 2024. https://open.barnet.gov.uk/joint-strategic-needs-assessment-2024/ (accessed May 1, 2025).

134 World Health Organization. WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020. 2020. https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020 (accessed Nov 16, 2020).

135 World Health Organization. WHO Director-General's opening remarks at the media briefing on COVID-19 - 13 March 2020. 2020. https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-mission-briefing-on-covid-19---13-march-2020 (accessed Nov 16, 2020).

136 BBC News. Coronavirus: Strict new curbs on life in UK announced by PM. Br. Broadcast. Co. 2020. https://www.bbc.co.uk/news/uk-52012432 (accessed May 1, 2025).

137 Institute for Government. Timeline of UK government coronavirus lockdowns and measures, March 2020 to December 2021. Inst. Gov. 2022.

138 Gandhi RT, Lynch JB, del Rio C. Mild or Moderate Covid-19. *N Engl J Med* 2020; **383**: 1757–66.

139 Berlin DA, Gulick RM, Martinez FJ. Severe Covid-19. *N Engl J Med* 2020; **383**: 2451–60.

140 Wu Z, McGoogan JM. Characteristics of and Important Lessons from the Coronavirus Disease 2019 (COVID-19) Outbreak in China: Summary of a Report of 72314 Cases from the Chinese Center for Disease Control and Prevention. JAMA - J. Am. Med. Assoc. 2020; **323**: 1239–42.

141 Argenziano MG, Bruce SL, Slater CL, *et al.* Characterization and clinical course of 1000 patients with coronavirus disease 2019 in New York: retrospective case series. *BMJ* 2020; **369**: m1996.

142 Wise J. Covid-19: WHO declares end of global health emergency. *BMJ* 2023; **381**: 1041.

143 Hamilton W. Cancer diagnostic delay in the COVID-19 era: what happens next? Lancet

Oncol. 2020; **21**: 1000–2.

144 British Society of Gastroenterology. BSG rationale around current advice to all Endoscopy Units. 2020. https://www.bsg.org.uk/covid-19-advice/bsg-rationale-around-current-advice-to-all-endoscopy-units/ (accessed Nov 19, 2020).

145 Rutter MD, Brookes M, Lee TJ, Rogers P, Sharp L. Impact of the COVID-19 pandemic on UK endoscopic activity and cancer detection: a National Endoscopy Database Analysis. *Gut* 2021; **70**: 537–43.

146 Li SJ, Seedher T, Sharples LD, *et al.* Impact of changes to the interscreening interval and faecal immunochemical test threshold in the national bowel cancer screening programme in England: results from the FIT pilot study. *Br J Cancer* 2022; **127**: 1525–33.

147 Al Saoudi T, Bahri S, Khasawneh F, Bhardwaj N, Garcea G. Two-Week Wait Gastrointestinal (GI) Cancer Pathway: A Single Tertiary Centre Experience During the COVID-19 Pandemic. *Cureus* 2023; **15**. DOI:10.7759/cureus.36857.

148 Walter FM, Emery JD, Mendonca S, *et al.* Symptoms and patient factors associated with longer time to diagnosis for colorectal cancer: results from a prospective cohort study. *Br J Cancer* 2016; **115**: 533.

149 Moffat J, Bentley A, Ironmonger L, Boughey A, Radford G, Duffy S. The impact of national cancer awareness campaigns for bowel and lung cancer symptoms on sociodemographic inequalities in immediate key symptom awareness and GP attendances. *Br J Cancer 2015 1121* 2015; **112**: S14–21.

150 O'Donohoe N, Jamal S, Cope J, Strom L, Ryan S, Nunoo-Mensah JW. COVID-19 recovery: tackling the 2-week wait colorectal pathway backlog by optimising CT colonography utilisation. *Clin Radiol* 2021; **76**: 117–21.

151 Oxford E. The NHS paid private hospitals £2bn in the pandemic: did taxpayers get value for money? *BMJ* 2023. DOI:10.1136/bmj.p329.

152 Araghi M, Arnold M, Rutherford MJ, *et al.* Colon and rectal cancer survival in seven high-income countries 2010-2014: Variation by age and stage at diagnosis (the ICBP SURVMARK-2 project). *Gut* 2021; **70**: 114–26.

153 Rees CJ, East JE, Oppong K, *et al.* Restarting gastrointestinal endoscopy in the deceleration and early recovery phases of COVID-19 pandemic: Guidance from the British Society of Gastroenterology. *Clin Med J R Coll Physicians London* 2020; **20**: 1–7.

154 Morris EJA, Goldacre R, Spata E, *et al.* Impact of the COVID-19 pandemic on the detection and management of colorectal cancer in England: a population-based study. *Lancet Gastroenterol Hepatol* 2021; **6**: 199–208.

155 Longcroft-Wheaton G, Tolfree N, Gangi A, Beable R, Bhandari P. Data from a large Western centre exploring the impact of COVID-19 pandemic on endoscopy services and cancer diagnosis. *Frontline Gastroenterol* 2021; **12**: 193–9.

156 Lantinga MA, Theunissen F, Ter Borg PCJ, Bruno MJ, Ouwendijk RJT, Siersema PD. Impact of the COVID-19 pandemic on gastrointestinal endoscopy in the Netherlands: Analysis of a prospective endoscopy database. *Endoscopy* 2021; **53**: 166–70.

157 Parasa S, Reddy N, Faigel DO, Repici A, Emura F, Sharma P. Global Impact of the

COVID-19 Pandemic on Endoscopy: An International Survey of 252 Centers From 55 Countries. Gastroenterology. 2020; **159**: 1579-1581.e5.

158 Lui RN, Wong SH, Sánchez-Luna SA, *et al.* Overview of guidance for endoscopy during the coronavirus disease 2019 pandemic. J. Gastroenterol. Hepatol. 2020; **35**: 749–59.

159 Verma AM, Nagpal JK. Gastrointestinal cancer pathways may breakdown due to COVID-19. Frontline Gastroenterol. 2020; **11**: 413.

160 Vrdoljak E, Sullivan R, Lawler M. Cancer and coronavirus disease 2019; how do we manage cancer optimally through a public health crisis? Eur. J. Cancer. 2020; **132**: 98–9.

161 Msemburi W, Karlinsky A, Knutson V, Aleshin-Guendel S, Chatterji S, Wakefield J. The WHO estimates of excess mortality associated with the COVID-19 pandemic. / *Nat* / 2023; **613**. DOI:10.1038/s41586-022-05522-2.

162 Penman I, Rees C. British Society of Gastroenterology guidance on recommencing gastrointestinal endoscopy in the deceleration and early recovery phases of COVID-19 pandemic. London, 2020.

163 Bhandari P, Subramaniam S, Bourke MJ, *et al.* Recovery of endoscopy services in the era of COVID-19: Recommendations from an international Delphi consensus. Gut. 2020; **69**: 1915–24.

164 Shenbagaraj L, Thomas-Gibson S, Stebbing J, *et al.* Endoscopy in 2017: A national survey of practice in the UK. *Frontline Gastroenterol* 2019; **10**: 7–15.

165 Beintaris I, Esmaily S, Saunders BP, *et al.* The WASh Trial: water-assisted sigmoidoscopy in the English Bowel Scope Screening Programme: study protocol for a randomized multicenter trial. *Endosc Int Open* 2019; **07**: E1574–82.

166 The Lancet Gastroenterology & Hepatology. Resuming bowel cancer screening post-COVID-19. Lancet Gastroenterol. Hepatol. 2020; **5**: 711.

167 Catlow J, Beaton D, Beintaris I, *et al.* JAG/BSG national survey of UK endoscopy services: Impact of the COVID-19 pandemic and early restoration of endoscopy services. *Frontline Gastroenterol* 2021; **12**: 272–8.

168 de Neree tot Babberich MPM, Ledeboer M, van Leerdam ME, *et al.* Dutch Gastrointestinal Endoscopy Audit: automated extraction of colonoscopy data for quality assessment and improvement. *Gastrointest Endosc* 2020; **92**: 154-162.e1.

169 Rhatigan E, Tyrmpas I, Murray G, Plevris JN. Scoring system to identify patients at high risk of oesophageal cancer. *Br J Surg* 2010; **97**: 1831–7.

170 Kamran U, King D, Banks M, *et al.* Assessment of the role of the Edinburgh dysphagia score in referral triage in a national service evaluation of the urgent suspected upper gastrointestinal cancer pathway. *Aliment Pharmacol Ther* 2022; **55**: 1160–8.

171 Murray IA, Palmer J, Waters C, Dalton HR. Predictive value of symptoms and demographics in diagnosing malignancy or peptic stricture. *World J Gastroenterol* 2012; **18**: 4357–62.

172 Moreno CC, Yee J, Ahmed FS, *et al.* CT colonography's role in the COVID-19 pandemic: a safe(r), socially distanced total colon examination. *Abdom Radiol* 2020; **1**: 3.

173 British Society of Gastrointestinal and Abdominal Radiology. CT Colonography activity

and Covid-19: British Society of Gastrointestinal and Abdominal Radiology guidance. London, 2020 https://www.acpgbi.org.uk/news/joint-acpgbi-bsg-and-bsgar-considerations-for-adapting-the-rapid-access- (accessed Nov 16, 2020).

174 di Pietro M, Modolell I, O'Donovan M, *et al.* Use of Cytosponge as a triaging tool to upper gastrointestinal endoscopy during the COVID-19 pandemic. *Lancet Gastroenterol Hepatol* 2020; **5**: 805–6.

175 MacLeod C, Wilson P, Watson AJM. Colon capsule endoscopy: an innovative method for detecting colorectal pathology during the COVID-19 pandemic? *Color Dis* 2020; **22**: 621–4.

176 Mohamed MO, Banerjee A, Clarke S, *et al.* Impact of COVID-19 on cardiac procedure activity in England and associated 30-day mortality. *Eur Hear J - Qual Care Clin Outcomes* 2021; **7**: 247–56.

177 Sud A, Jones ME, Broggio J, *et al.* Collateral damage: the impact on outcomes from cancer surgery of the COVID-19 pandemic. *Ann Oncol* 2020; **31**: 1065–74.

178 British Medical Association. NHS backlog data analysis. Br. Med. Assoc. 2024. https://www.bma.org.uk/advice-and-support/nhs-delivery-and-workforce/pressures/nhs-backlog-data-analysis (accessed July 25, 2024).

179 Bornschein J, Quante M, Jansen M. The complexity of cancer origins at the gastro-oesophageal junction. *Best Pract Res Clin Gastroenterol* 2021; **50**–**51**: 101729.

180 Shawihdi M, Thompson E, Kapoor N, *et al.* Variation in gastroscopy rate in English general practice and outcome for oesophagogastric cancer: Retrospective analysis of Hospital Episode Statistics. *Gut* 2014; **63**: 250–61.

181 Rosenfeld A, Graham DG, Jevons S, *et al.* Development and validation of a risk prediction model to diagnose Barrett's oesophagus (MARK-BE): a case-control machine learning approach. *Lancet Digit Heal* 2020; **2**: e37–48.

182 Shung DL, Au B, Taylor RA, *et al.* Validation of a Machine Learning Model That Outperforms Clinical Risk Scoring Systems for Upper Gastrointestinal Bleeding. *Gastroenterology* 2020; **158**: 160–7.

183 Turkington RC, Lavery A, Donnelly D, Cairnduff V, McManus DT, Coleman HG. The Impact of the COVID-19 Pandemic on Barrett's Esophagus and Esophagogastric Cancer. Gastroenterology. 2021; **160**: 2169-2171.e1.

184 Tangcharoensathien V, Bassett MT, Meng Q, Mills A. Are overwhelmed health systems an inevitable consequence of covid-19? Experiences from China, Thailand, and New York State. *BMJ* 2021; **372**. DOI:10.1136/bmj.n83.

185 National Institute for Health and Care Excellence. Suspected cancer: recognition and referral. NICE Guideline (NG 12). 2025 www.nice.org.uk/guidance/ng12 (accessed May 1, 2025).

186 Rasmussen S, Haastrup PF, Balasubramaniam K, Christensen RDP, Søndergaard J, Jarbøl DE. Predictive values of upper gastrointestinal cancer alarm symptoms in the general population: A nationwide cohort study. *BMC Cancer* 2018; **18**: 1–7.

187 Locke GR, Talley NJ, Weaver AL, Zinsmeister AR. A new questionnaire for gastroesophageal reflux disease. *Mayo Clin Proc* 1994; **69**: 539–47.

188 Rubenstein JH, Raghunathan T, Doan C, *et al.* Validation of Tools for Predicting

Incident Adenocarcinoma of the Esophagus or Esophagogastric Junction. *Am J Gastroenterol* 2021; **116**: 949–57.

189 Rubenstein JH, Morgenstern H, Appelman H, *et al.* Prediction of barrett's esophagus among men. *Am J Gastroenterol* 2013; **108**: 353–62.

190 Thrift AP, Kendall BJ, Pandeya N, Vaughan TL, Whiteman DC. A clinical risk prediction model for Barrett esophagus. *Cancer Prev Res* 2012; **5**: 1115–23.

191 Xie SH, Ness-Jensen E, Medefelt N, Lagergren J. Assessing the feasibility of targeted screening for esophageal adenocarcinoma based on individual risk assessment in a population-based cohort study in Norway (The HUNT Study). *Am J Gastroenterol* 2018; **113**: 829–35.

192 Thrift AP, Vaughan TL, Anderson LA, Whiteman DC, El-Serag HB. External Validation of the Michigan Barrett's Esophagus Prediction Tool. Clin. Gastroenterol. Hepatol. 2017; **15**: 1124–6.

193 Kunzmann AT, Thrift AP, Cardwell CR, *et al.* Model for Identifying Individuals at Risk for Esophageal Adenocarcinoma. *Clin Gastroenterol Hepatol* 2018; **16**: 1229-1236.e4.

194 Hippisley-Cox J, Coupland C. Identifying patients with suspected gastro-oesophageal cancer in primary care: derivation and validation of an algorithm. *Br J Gen Pract* 2011; **61**: e707–14.

195 Vakil N, Moayyedi P, Fennerty MB, Talley NJ. Limited Value of Alarm Features in the Diagnosis of Upper Gastrointestinal Malignancy: Systematic Review and Meta-analysis. *Gastroenterology* 2006; **131**: 390–401.

196 Liu X, Wong A, Kadri SR, *et al.* Gastro-esophageal reflux disease symptoms and demographic factors as a pre-screening tool for Barrett's esophagus. *PLoS One* 2014; **9**: e94163.

197 Locke GR, Zinsmeister AR, Talley NJ. Can symptoms predict endoscopic findings in GERD? *Gastrointest Endosc* 2003; **58**: 661–70.

198 Steptoe A, Shankar A, Demakakos P, Wardle J. Social isolation, loneliness, and all-cause mortality in older men and women. *Proc Natl Acad Sci U S A* 2013; **110**: 5797–801.

199 Mykletun A, Bjerkeset O, Dewey M, Prince M, Overland S, Stewart R. Anxiety, depression, and cause-specific mortality: The HUNT study. *Psychosom Med* 2007; **69**: 323–31.

200 Friedman GD, Udaltsova N, Chan J, Quesenberry CP, Habel LA. Screening pharmaceuticals for possible carcinogenic effects: Initial positive results for drugs not previously screened. *Cancer Causes Control* 2009; **20**: 1821–35.

201 Hughes ME, Waite LJ, Hawkley LC, Cacioppo JT. A short scale for measuring loneliness in large surveys: Results from two population-based studies. Res. Aging. 2004; **26**: 655–72.

202 Mellow MH, Pinkas H. Endoscopic Laser Therapy for Malignancies Affecting the Esophagus and Gastroesophageal Junction: Analysis of Technical and Functional Efficacy. *Arch Intern Med* 1985; **145**: 1443–6.

203 Stekhoven DJ, Bühlmann P. Missforest-Non-parametric missing value imputation for mixed-type data. *Bioinformatics* 2012; **28**: 112–8.

204 Romanski P, Kotthoff L, Schratz P. FSelector: Selecting Attributes. 2021. https://cran.r-project.org/package=FSelector.

205 Kuhn M. caret: Classification and Regression Training. 2021. https://cran.r-project.org/package=caret.

206 Friedman J, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. *J Stat Softw* 2010; **33**. DOI:10.18637/jss.v033.i01.

207 Robin X, Turck N, Hainard A, *et al.* pROC: An open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics* 2011; **12**: 1–8.

208 Sing T, Sander O, Beerenwinkel N, Lengauer T. ROCR: visualizing classifier performance in R. *Bioinformatics* 2005; **21**: 3940–1.

209 Liu Y, Lin D, Li L, *et al.* Using machine-learning algorithms to identify patients at high risk of upper gastrointestinal lesions for endoscopy. *J Gastroenterol Hepatol* 2021; **36**: 2735–44.

210 Cancer Research UK. Oesophageal cancer statistics. 2022. https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/oesophageal-cancer (accessed July 6, 2022).

211 Lim WJ, Kim KH, Kim JY, Jeong S, Kim N. Identification of DNA-methylated CpG islands associated with gene silencing in the adult body tissues of the ogye chicken using RNA-Seq and reduced representation bisulfite sequencing. *Front Genet* 2019; **10**: 1–15.

212 Kaz AM, Grady WM. Epigenetic biomarkers in esophageal cancer. *Cancer Lett* 2014; **342**: 193–9.

213 Clément G, Braunschweig R, Pasquier N, Bosman FT, Benhattar J. Methylation of APC, TIMP3, and TERT: A new predictive marker to distinguish Barrett's oesophagus patients at risk for malignant transformation. *J Pathol* 2006; **208**: 100–7.

214 Xie Z, Chen G, Zhang X, *et al.* Salivary MicroRNAs as Promising Biomarkers for Detection of Esophageal Cancer. *PLoS One* 2013; **8**. DOI:10.1371/journal.pone.0057502.

215 Nunes AP, Oliveira IO, Santos BR, *et al.* Quality of DNA extracted from saliva samples collected with the Oragene DNA self-collection kit. *BMC Med Res Methodol* 2012; **12**: 65.

216 Du P, Zhang X, Huang CC, *et al.* Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. *BMC Bioinformatics* 2010; **11**: 1–9.

217 Price EM, Robinson WP. Adjusting for Batch Effects in DNA Methylation Microarray Data, a Lesson Learned. *Front Genet* 2018; **9**. DOI:10.3389/fgene.2018.00083.

218 Langfelder P, Horvath S. WGCNA: An R package for weighted correlation network analysis. *BMC Bioinformatics* 2008; **9**: 1–13.

219 Li D, Zhang L, Liu Y, *et al.* Specific DNA methylation markers in the diagnosis and prognosis of esophageal cancer. *Aging (Albany NY)* 2019; **11**: 11640–58.

220 Yu M, Moinova HR, Willbanks A, *et al.* Novel DNA Methylation Biomarker Panel for Detection of Esophageal Adenocarcinoma and High-Grade Dysplasia. *Clin Cancer Res*

2022; **28**: 3761–9.

221 Middha Kapoor P, Mavaddat N, Pal Choudhury P, *et al.* Combined Associations of a Polygenic Risk Score and Classical Risk Factors With Breast Cancer Risk Brief Communication. *JNCI J Natl Cancer Inst* 2021; **113**: 56.

222 Arnold M, Sierra MS, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global patterns and trends in colorectal cancer incidence and mortality. *Gut* 2017; **66**: 683–91.

223 Cancer Research UK. Cancer incidence for common cancers. 2019. https://www.cancerresearchuk.org/health-professional/cancer-statistics/incidence/common-cancers-compared (accessed April 23, 2025).

224 NHS England. Bowel cancer screening: programme overview. Gov.uk. 2024. https://www.gov.uk/guidance/bowel-cancer-screening-programme-overview (accessed April 23, 2025).

225 Loveday C, Sud A, Jones ME, *et al.* Prioritisation by FIT to mitigate the impact of delays in the 2-week wait colorectal cancer referral pathway during the COVID-19 pandemic: A UK modelling study. *Gut* 2021; **70**: 1053–60.

226 Tinmouth J, Lansdorp-Vogelaar I, Allison JE. Faecal immunochemical tests versus guaiac faecal occult blood tests: What clinicians and colorectal cancer screening programme organisers need to know. *Gut* 2015; **64**: 1327–37.

227 Young GP, Cole S. New stool screening tests for colorectal cancer. Digestion. 2007; **76**: 26–33.

228 National Institute for Health and Care Excellence. Quantitative faecal immunochemical tests to guide referral for colorectal cancer in primary care Diagnostics guidance. 2017 www.nice.org.uk/guidance/dg30 (accessed Aug 8, 2022).

229 Chapman CJ, Banerjea A, Humes DJ, *et al.* Choice of faecal immunochemical test matters: Comparison of OC-Sensor and HM-JACKarc, in the assessment of patients at high risk of colorectal cancer. *Clin Chem Lab Med* 2021; **59**: 721–8.

230 Saw KS, Liu C, Xu W, Varghese C, Parry S, Bissett I. Faecal immunochemical test to triage patients with possible colorectal cancer symptoms: Meta-analysis. *Br J Surg* 2022; **109**: 182–90.

231 Laszlo HE, Seward E, Ayling RM, *et al.* Faecal immunochemical test for patients with 'high-risk' bowel symptoms: a large prospective cohort study and updated literature review. *Br J Cancer* 2022; **126**: 736–43.

232 Ayling RM, Machesney M. Service evaluation of faecal immunochemical testing introduced for use in North East London for patients at low risk of colorectal cancer. *J Clin Pathol* 2021; **74**: 163–6.

233 Bailey SER, Abel GA, Atkins A, *et al.* Diagnostic performance of a faecal immunochemical test for patients with low-risk symptoms of colorectal cancer in primary care: an evaluation in the South West of England. *Br J Cancer* 2021; **124**: 1231–6.

234 Chapman C, Bunce J, Oliver S, *et al.* Service evaluation of faecal immunochemical testing and anaemia for risk stratification in the 2-week-wait pathway for colorectal cancer. *BJS open* 2019; **3**: 395–402.

235 D'Souza N, Georgiou Delisle T, Chen M, Benton S, Abulafi M, NICE FIT Steering Group.

Faecal immunochemical test is superior to symptoms in predicting pathology in patients with suspected colorectal cancer symptoms referred on a 2WW pathway: a diagnostic accuracy study. *Gut* 2021; **70**: 1130–8.

236 Khan AA, Klimovskij M, Harshen R. Accuracy of faecal immunochemical testing in patients with symptomatic colorectal cancer. *BJS Open* 2020; **4**: 1180–8.

237 Mattar R, Marques SB, Minata MK, da Silva-Etto JMK, Sakai P, De Moura EGH. Diagnostic accuracy of one sample or two samples quantitative fecal immunochemical tests for intestinal neoplasia detection. *Arq Gastroenterol* 2020; **57**: 316–22.

238 Mowat C, Digby J, Strachan JA, *et al.* Faecal haemoglobin and faecal calprotectin as indicators of bowel disease in patients presenting to primary care with bowel symptoms. *Gut* 2016; **65**: 1463–9.

239 Navarro M, Hijos G, Sostres C, *et al.* Reducing the Cut-Off Value of the Fecal Immunochemical Test for Symptomatic Patients Does Not Improve Diagnostic Performance. *Front Med* 2020; **7**: 1–9.

240 Nicholson BD, James T, East JE, *et al.* Experience of adopting faecal immunochemical testing to meet the NICE colorectal cancer referral criteria for low-risk symptomatic primary care patients in Oxfordshire, UK. *Frontline Gastroenterol* 2019; **10**: 347–55.

241 Nicholson BD, James T, Paddon M, *et al.* Faecal immunochemical testing for adults with symptoms of colorectal cancer attending English primary care: a retrospective cohort study of 14 487 consecutive test requests. *Aliment Pharmacol Ther* 2020; **52**: 1031–41.

242 Rodríguez-Alonso L, Rodríguez-Moranta F, Ruiz-Cerulla A, *et al.* An urgent referral strategy for symptomatic patients with suspected colorectal cancer based on a quantitative immunochemical faecal occult blood test. *Dig Liver Dis* 2015; **47**: 797–804.

243 Tsapournas G, Hellström PM, Cao Y, Olsson LI. Diagnostic accuracy of a quantitative faecal immunochemical test vs. symptoms suspected for colorectal cancer in patients referred for colonoscopy. *Scand J Gastroenterol* 2020; **55**: 184–92.

244 Turvill JL, Turnock D, Cottingham D, *et al.* The Fast Track FIT study: Diagnostic accuracy of faecal immunochemical test for haemoglobin in patients with suspected colorectal cancer. *Br J Gen Pract* 2021; **71**: E643–51.

245 Rutter MD, East J, Rees CJ, *et al.* British Society of Gastroenterology/Association of Coloproctology of Great Britain and Ireland/Public Health England post-polypectomy and post-colorectal cancer resection surveillance guidelines. *Gut* 2020; **69**: 201–23.

246 Chandrapalan S, Khasawneh F, Singh B, *et al.* A Multi-Centre Study to Risk Stratify Colorectal Polyp Surveillance Patients Utilising Volatile Organic Compounds and Faecal Immunochemical Test. *Cancers (Basel)* 2022; **14**: 4951.

247 Bath MF, Malhi A, Ayling RM, *et al.* Faecal immunochemical testing for haemoglobin in detecting bowel polyps in symptomatic patients: multicentre prospective cohort study. *BJS Open* 2023; **7**: 161.

248 NHS Guidelines. Clinical guide for triaging patients with lower gastrointestinal symptoms. NHS. 2020. https://www.nice.org.uk/media/default/about/covid-19/specialty-guides/triaging-patients-with-lower-gi-symptoms.pdf (accessed May 1,

2025).

249 Arasaradnam RP, Bhala N, Evans C, *et al.* Faecal immunochemical testing in the COVID-19 era: balancing risk and costs. *Lancet Gastroenterol Hepatol* 2020; **5**: 717–9.

250 Maclean W, Limb C, Mackenzie P, *et al.* Adoption of faecal immunochemical testing for 2-week-wait colorectal patients during the COVID-19 pandemic: an observational cohort study reporting a new service at a regional centre. *Colorectal Dis* 2021; **23**: 1622–9.

251 von Wagner C, Verstraete W, Hirst Y, Nicholson BD, Stoffel ST, Laszlo H. Public preferences for using quantitative faecal immunochemical test versus colonoscopy as diagnostic test for colorectal cancer: evidence from an online survey. *BJGP open* 2020; **4**: bjgpopen20X101007.

252 Pin-Vieito N, Tejido-Sandoval C, De Vicente-Bielza N, Sánchez-Gómez C, Cubiella J. Faecal immunochemical tests safely enhance rational use of resources during the assessment of suspected symptomatic colorectal cancer in primary care: Systematic review and meta-analysis. *Gut* 2022; **71**: 950–60.

253 Grobbee EJ, Schreuders EH, Hansen BE, *et al.* Association Between Concentrations of Hemoglobin Determined by Fecal Immunochemical Tests and Long-term Development of Advanced Colorectal Neoplasia. *Gastroenterology* 2017; **153**: 1251-1259.e2.

254 Hunt N, Rao C, Logan R, *et al.* A cohort study of duplicate faecal immunochemical testing in patients at risk of colorectal cancer from North-West England. *BMJ Open* 2022; **12**: 59940.

255 Johnstone MS, MacLeod C, Digby J, *et al.* Prevalence of repeat faecal immunochemical testing in symptomatic patients attending primary care. *Color Dis* 2022; **24**: 1498–504.

256 NHS North Central London Cancer Alliance. FIT phases. https://www.nclcanceralliance.nhs.uk/our-work/diagnosis-and-treatment/fit-phases/ (accessed May 1, 2025).

257 Cappellini MD, Motta I. Anemia in Clinical Practice-Definition and Classification: Does Hemoglobin Change With Aging? *Semin Hematol* 2015; **52**: 261–9.

258 Marshall T, Lancashire R, Sharp D, Peters TJ, Cheng KK, Hamilton W. The diagnostic performance of scoring systems to identify symptomatic colorectal cancer compared to current referral guidance. *Gut* 2011; **60**: 1242–8.

259 Hamilton W. The CAPER studies: Five case-control studies aimed at identifying and quantifying the risk of cancer in symptomatic primary care patients. *Br J Cancer* 2009; **101**: S80–6.

260 Chiu H-M, Ching JYL, Wu KC, *et al.* A Risk-Scoring System Combined With a Fecal Immunochemical Test Is Effective in Screening High-Risk Subjects for Early Colonoscopy to Detect Advanced Colorectal Neoplasms. *Gastroenterology* 2016; **150**: 617-625.e3.

261 Liao Y, Li S, Chen H, *et al.* A validated tool of combining risk factors with fecal immunochemical test for selecting colonoscopy in asymptomatic Chinese population. *Eur J Cancer Prev* 2023; **32**: 275–80.

262 Jung YS, Park CH, Kim NH, Park JH, Park D Il, Sohn C Il. Clinical risk stratification model for advanced colorectal neoplasia in persons with negative fecal immunochemical test results. *PLoS One* 2018; **13**. DOI:10.1371/journal.pone.0191125.

263 Jung YS, Park CH, Kim NH, Park JH, Park D Il, Sohn C Il. A combination of clinical risk stratification and fecal immunochemical test is useful for identifying persons with high priority of early colonoscopy. *Dig Liver Dis* 2018; **50**: 254–9.

264 Digby J, Strachan JA, Mowat C, Steele RJC, Fraser CG. Appraisal of the faecal haemoglobin, age and sex test (FAST) score in assessment of patients with lower bowel symptoms: An observational study. *BMC Gastroenterol* 2019; **19**: 213.

265 Lee TJ, Siau K, Esmaily S, *et al.* Development of a national automated endoscopy database: The United Kingdom National Endoscopy Database (NED). *United Eur Gastroenterol J* 2019; **7**: 798–806.

266 Kamel F, Zulfiqar S, Penfold Wi, *et al.* The use of the faecal immunochemical test during the COVID-19 pandemic to triage urgent colorectal cancer referrals. *Color Dis* 2022; **24**: 727–36.

267 Nicholson BD, East JE, Oke J, Roberts NW, James T, Shine B. Letter: extending FIT from DG30 to NG12 patients. Letter: faecal immunochemical testing for adults with symptoms of colorectal cancer – ready for prime time? Authors' reply: a unified approach to safety netting negative FITs is required. Aliment. Pharmacol. Ther. 2020; **52**: 1420–1.

268 Harper L, Kalfa N, Beckers GMA, *et al.* The impact of COVID-19 on research. *J Pediatr Urol* 2020; **16**: 715–6.

269 Quadri SA. COVID-19 and religious congregations: Implications for spread of novel pathogens. Int. J. Infect. Dis. 2020; **96**: 219–21.

270 Kim S, Jeong YD, Byun JH, *et al.* Evaluation of COVID-19 epidemic outbreak caused by temporal contact-increase in South Korea. *Int J Infect Dis* 2020; **96**: 454–7.

271 Mat NFC, Edinur HA, Razab MKAA, Safuan S. A single mass gathering resulted in massive transmission of COVID-19 infections in Malaysia with further international spread. *J Travel Med* 2020; **27**: 1–4.

272 Hamner L, Dubbel P, Capron I, *et al.* High SARS-CoV-2 Attack Rate Following Exposure at a Choir Practice — Skagit County, Washington, March 2020. *MMWR Morb Mortal Wkly Rep* 2020; **69**: 606–10.

273 Bahl P, de Silva C, Bhattacharjee S, *et al.* Droplets and Aerosols Generated by Singing and the Risk of Coronavirus Disease 2019 for Choirs. *Clin Infect Dis* 2020. DOI:10.1093/cid/ciaa1241.

274 Alsved M, Matamis A, Bohlin R, *et al.* Exhaled respiratory particles during singing and talking. Aerosol Sci. Technol. 2020; **54**: 1245–8.

275 Hartley DM, Reisinger HS, Perencevich EN. When infection prevention enters the temple: Intergenerational social distancing and COVID-19. Infect. Control Hosp. Epidemiol. 2020; **41**: 868–9.

276 Anyfantakis D. Holy Communion and Infection Transmission: A Literature Review. *Cureus* 2020; **12**: e8741.

277 Gaughan CH, Ayoubkhani D, Nafilyan V, *et al.* Religious affiliation and COVID-19-

related mortality: A retrospective cohort study of prelockdown and postlockdown risks in England and Wales. *J Epidemiol Community Health* 2021; **0**: 1–6.

278 Bryson JR, Andres L, Davies A. COVID-19, Virtual Church Services and a New Temporary Geography of Home. *Tijdschr voor Econ en Soc Geogr* 2020; **111**: 360–72.

279 Modell SM, Kardia SLR. Religion as a Health Promoter During the 2019/2020 COVID Outbreak: View from Detroit. *J Relig Health* 2020; **59**: 2243–55.

280 Ministry of Housing, Communities & Local Government (2018 to 2021). COVID-19: guidance for the safe use of places of worship during the pandemic from 4 July. 2021. https://www.gov.uk/government/publications/covid-19-guidance-for-the-safe-use-of-places-of-worship-during-the-pandemic-from-4-july (accessed May 1, 2025).

281 Burgess K. Hark! Tomorrow heralds early return of singing. The Times. 2021; published online March 27.

282 Public Health England. Covid-19: suggested principles of safer singing. 2020. https://www.gov.uk/government/publications/covid-19-suggested-principles-of-safer-singing/covid-19-suggested-principles-of-safer-singing (accessed May 1, 2025).

283 BBC News. Covid-19: Can face masks make communal singing safe enough? 2020. https://www.bbc.co.uk/news/av/uk-55414707 (accessed May 1, 2025).

284 Alfelali M, Haworth EA, Barasheed O, *et al.* Facemask against viral respiratory infections among Hajj pilgrims: A challenging clusterrandomized trial. *PLoS One* 2020; **15**. DOI:10.1371/journal.pone.0240287.

285 Alqahtani AS, Sheikh M, Wiley K, Heywood AE. Australian Hajj pilgrims' infection control beliefs and practices: Insight with implications for public health approaches. *Travel Med Infect Dis* 2015; **13**: 329–34.

286 Alqahtani AS, Wiley KE, Tashani M, *et al.* Exploring barriers to and facilitators of preventive measures against infectious diseases among Australian Hajj pilgrims: Cross-sectional studies before and after Hajj. *Int J Infect Dis* 2016; **47**: 53–9.

287 Taylor S, Asmundson GJG. Negative attitudes about facemasks during the COVID-19 pandemic: The dual importance of perceived ineffectiveness and psychological reactance. *PLoS One* 2021; **16**: e0246317.

288 DeFranza D, Lindow M, Harrison K, Mishra A, Mishra H. Religion and Reactance to COVID-19 Mitigation Guidelines. *Am Psychol* 2020. DOI:10.1037/amp0000717.

289 Corvo E, De Caro W. COVID-19 and Spontaneous Singing to Decrease Loneliness, Improve Cohesion, and Mental Well-Being: An Italian Experience. *Psychol Trauma Theory, Res Pract Policy* 2020; **12**: 247–8.

290 Hathaway E. Assisting Faith-Based Organizations Increase Sense of Belonging during the Covid-19 Pandemic. *J Pastoral Care Counsel* 2020; **74**: 226–8.

291 Dutra CCD, Rocha HS. Religious Support as a Contribution to Face the Effects of Social Isolation in Mental Health During the Pandemic of COVID-19. *J Relig Health* 2021; **60**: 99–111.

292 Kowalczyk O, Roszkowski K, Montane X, Pawliszak W, Tylkowski B, Bajek A. Religion and Faith Perception in a Pandemic of COVID-19. *J Relig Health* 2020; **59**: 2671–7.

293 Barmania S, Reiss MJ. Health promotion perspectives on the COVID-19 pandemic: The

importance of religion. *Glob Health Promot* 2020; : 175797592097299.

294 Office for National Statistics. Exploring Religion in England and Wales: February 2020. 2020. https://www.ons.gov.uk/peoplepopulationandcommunity/culturalidentity/religion/articles/exploringreligioninenglandandwales/february2020 (accessed May 1, 2025).

295 Dunn P, Ewbank L, Alderwick H. Nine major challenges facing health and care in England. Heal. Found. 2023. https://www.health.org.uk/publications/long-reads/nine-major-challenges-facing-health-and-care-in-england (accessed May 1, 2025).

296 Ngiam KY, Khor IW. Big data and machine learning algorithms for health-care delivery. Lancet Oncol. 2019; **20**: e262–73.

297 Pugh RNH, Murray-Lyon IM, Dawson JL, Pietroni MC, Williams R. Transection of the oesophagus for bleeding oesophageal varices. *Br J Surg* 1973; **60**: 646–9.

298 Reig M, Forner A, Rimola J, *et al.* BCLC strategy for prognosis prediction and treatment recommendation: The 2022 update. *J Hepatol* 2022; **76**: 681–93.

299 Waljee AK, Weinheimer-Haus EM, Abubakar A, *et al.* Artificial intelligence and machine learning for early detection and diagnosis of colorectal cancer in sub-Saharan Africa. *Gut* 2022; **71**: 1259–65.

300 Zhang X, Chan FTS, Yan C, Bose I. Towards risk-aware artificial intelligence and machine learning systems: An overview. *Decis Support Syst* 2022; **159**: 113800.

301 Morley J, Machado CCV, Burr C, *et al.* The ethics of AI in health care: A mapping review. Soc. Sci. Med. 2020; **260**: 113172.

302 Gates B. Responding to Covid-19 — A Once-in-a-Century Pandemic? *N Engl J Med* 2020; **382**: 1677–9.