



OPEN ACCESS

Original research

# Understanding the malignant potential of gastric metaplasia of the oesophagus and its relevance to Barrett's oesophagus surveillance: individual-level data analysis

Emily L Black ,<sup>1</sup> Emma Ococks,<sup>1</sup> Ginny Devonshire,<sup>2</sup> Alvin Wei Tian Ng,<sup>1,2</sup> Maria O'Donovan,<sup>3</sup> Shalini Malhotra,<sup>3</sup> Monika Tripathi,<sup>3</sup> Ahmad Miremadi,<sup>3</sup> Adam Freeman,<sup>1</sup> Hannah Coles,<sup>1</sup> Oesophageal Cancer Clinical and Molecular Stratification (OCCAMS) Consortium, Rebecca C Fitzgerald <sup>1</sup>

► Additional supplemental material is published online only. To view, please visit the journal online (<http://dx.doi.org/10.1136/gutjnl-2023-330721>).

<sup>1</sup>Early Cancer Institute, Department of Oncology, University of Cambridge, Cambridge, UK

<sup>2</sup>Cancer Research UK Cambridge Institute, University of Cambridge, Cambridge, UK  
<sup>3</sup>Department of Histopathology, Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK

## Correspondence to

Professor Rebecca C Fitzgerald, Early Cancer Institute, Department of Oncology, University of Cambridge, Cambridge, CB2 0XZ, UK; [rcf29@cam.ac.uk](mailto:rcf29@cam.ac.uk)

Received 18 July 2023

Accepted 1 November 2023

Published Online First

21 November 2023



► <http://dx.doi.org/10.1136/gutjnl-2023-331458>



© Author(s) (or their employer(s)) 2024. Re-use permitted under CC BY. Published by BMJ.

**To cite:** Black EL, Ococks E, Devonshire G, et al. *Gut* 2024;**73**:729–740.

## ABSTRACT

**Objective** Whether gastric metaplasia (GM) of the oesophagus should be considered as Barrett's oesophagus (BO) is controversial. Given concern intestinal metaplasia (IM) may be missed due to sampling, the UK guidelines include GM as a type of BO. Here, we investigated whether the risk of misdiagnosis and the malignant potential of GM warrant its place in the UK surveillance.

**Design** We performed a thorough pathology and endoscopy review to follow clinical outcomes in a novel UK cohort of 244 patients, covering 1854 person years of follow-up. We complemented this with a comparative genomic analysis of 160 GM and IM specimens, focused on early molecular hallmarks of BO and oesophageal adenocarcinoma (OAC).

**Results** We found that 58 of 77 short-segment (<3 cm) GM (SS-GM) cases (75%) continued to be observed as GM-only across a median of 4.4 years of follow-up. We observed that disease progression in GM-only cases and GM+IM cases (cases with reported GM on some occasions, IM on others) was significantly lower than in the IM-only cases (Kaplan-Meier,  $p=0.03$ ). Genomic analysis revealed that the mutation burden in GM is significantly lower than in IM ( $p<0.01$ ). Moreover, GM does not bear the mutational hallmarks of OAC, with an absence of associated signatures and driver gene mutations. Finally, we established that GM found adjacent to OAC is evolutionarily distant from cancer.

**Conclusion** SS-GM is a distinct entity from SS-IM and the malignant potential of GM is lower than IM. It is questionable whether SS-GM warrants inclusion in BO surveillance.

## INTRODUCTION

Barrett's oesophagus (BO) is a precursor lesion for oesophageal adenocarcinoma (OAC) and provides an opportunity to improve outcomes by detecting dysplasia and cancer early. BO is characterised by the replacement of squamous oesophageal epithelium with columnar-lined epithelium with a crypt architecture. Metaplastic glands in BO can be gastric or intestinal in nature, and a non-dysplastic

## WHAT IS ALREADY KNOWN ON THIS TOPIC

- ⇒ The UK guidelines are distinct from most countries, by not requiring goblet cells (intestinal metaplasia (IM)) for a diagnosis of Barrett's oesophagus.
- ⇒ IM can be misdiagnosed as gastric metaplasia (GM) if sampling is insufficient.

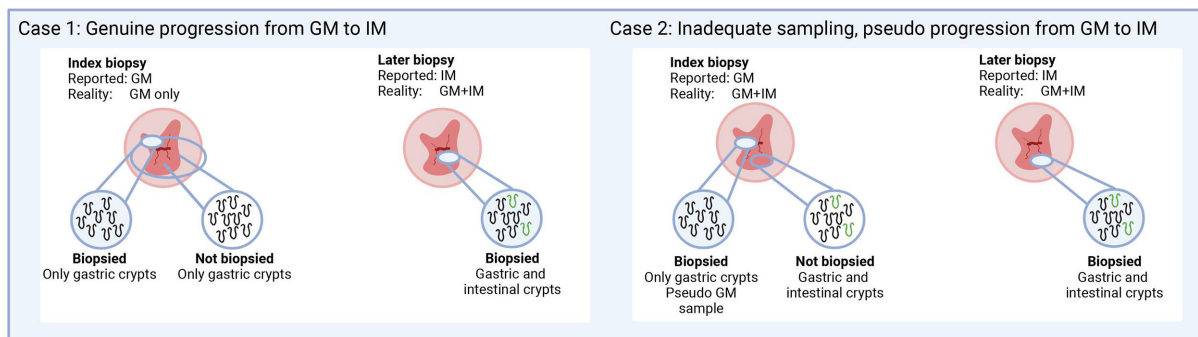
## WHAT THIS STUDY ADDS

- ⇒ Short-segment GM can be molecularly distinguished from short-segment IM.
- ⇒ The risk of progression is substantially lower in GM than in IM.
- ⇒ GM does not bear the genomic hallmarks of oesophageal adenocarcinoma.
- ⇒ GM is not evolutionarily close to oesophageal adenocarcinoma, even when found adjacent to the cancer.

## HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

- ⇒ The low risk of progression in GM suggests that UK Barrett's oesophagus surveillance guidelines could be updated to require the presence of goblet cells in short segments.
- ⇒ This would reduce oversurveillance, resulting in better quality of life for these patients and improved focusing of resources on higher risk patients.
- ⇒ It would also bring the UK in line with the international community. This would reduce confusion, and facilitate comparisons, for research and clinical practice.

BO (NDBO) segment will typically be a mosaic of different gland types. A segment with only gastric-type glands (no goblet cells) is called gastric metaplasia (GM); identification of any intestinal-type glands leads to a designation of intestinal metaplasia (IM) (online supplemental figure 1). NDBO can progress to low-grade and high-grade dysplasia (LGD, HGD) and then potentially to OAC.



**Figure 1** Illustration of the challenge of misdiagnosis due to undersampling of GM+IM segments. BO, Barrett's oesophagus; GM, gastric metaplasia; IM, intestinal metaplasia.

Most patients with BO will never develop OAC though, with annual progression rates to OAC of 0.1%–0.5% for NDBO.<sup>1–3</sup> The UK surveillance guidelines stand apart in their definition of BO. In the UK, GM of at least 1 cm is sufficient for a BO diagnosis<sup>4</sup>; elsewhere, presence of IM is required.<sup>5–8</sup> This distinction may be substantial: a study suggested that removing the US requirement for goblet cells would lead to a 147% increase in BO diagnoses, with little impact on OAC diagnoses.<sup>9</sup>

However, there remains conflicting clinical evidence about the malignant potential of GM. While there is evidence supporting GM having very low risk of progression, and lower risk than IM,<sup>9–11</sup> several studies reached an opposing conclusion. Specifically, that most GM cases progressed to IM<sup>12–13</sup> and then on to OAC at a similar rate as index-IM cases.<sup>14</sup> Much of the discrepancy is likely due to sampling. In patients with gastric and intestinal glands (GM+IM), goblet cells are not uniformly distributed,<sup>15</sup> so areas of the segment can look like GM and be misdiagnosed (figure 1).

Frequent co-occurrence of GM and OAC has been used to argue the malignant potential of GM and to contest the view that OAC typically arises from IM.<sup>16</sup> However, overgrowth of IM and sampling error could also explain the absence of visible IM.<sup>17–18</sup> Recent computational analysis provided new evidence that most OACs do arise via IM, despite its frequent absence.<sup>19</sup> Furthermore, it has been shown that IM develops from undifferentiated gastric cardia cells, and that BO is a unifying pathway to OAC.<sup>20</sup> Transcriptional differences suggest that true GM is more akin to normal gastric cells than IM.<sup>20–23</sup>

The molecular hallmarks of OAC are single base substitution (SBS) signatures 17a and 17b, *TP53* mutations, genomic instability, a plethora of lower frequency mutations and larger scale structural rearrangements.<sup>24–26</sup> Evidence of genomic differences between GM and IM is poorly delineated and tends to be based on a small number of samples and a subset of the genome.<sup>27–28</sup> Reported phylogenies vary from cases where GM shares little evolutionary history with OAC,<sup>29</sup> to a case where OAC had developed directly from gastric BO glands.<sup>30</sup>

This study examines the relative malignant potential of GM compared with IM, to assess whether the distinction between the two is clinically relevant. We performed a thorough review of endoscopy and pathology reports across a novel cohort of 244 patients (figure 2), covering 1854 person years of follow-up. We separately assessed outcomes in GM-only, IM-only and GM+IM cases. We complemented clinical outcome analysis with genomic analysis assessing molecular hallmarks of OAC in GM and IM, in both BO only and BO adjacent to OAC contexts.

## METHODS

### Clinical outcome cohort

A cohort was selected from a database of patients under BO surveillance at Addenbrooke's Hospital (Cambridge University Hospitals NHS Foundation Trust) who consented prospectively to research participation. Strict selection criteria were applied to pathology, endoscopy and medical histories (online supplemental table S1). All oesophagogastrroduodenoscopies (OGDs) comprised of multiple biopsies, each with known pathology. The centre follows a systematic biopsy protocol, in-line with the Seattle protocol. Each OGD had a single classification of IM (if any biopsy had IM) or GM (if no biopsy had IM). Patients needed to have an index diagnosis of GM or IM. The index segment length needed to be at least 1 cm, with no upper limit for index-GM cases, but less than 3 cm for index-IM cases. From a total of 5625 patients in the database, 244 were selected: 77 index-GM, short-segment cases; 23 index-GM, long segment; 144 index-IM, short segment (figure 2).

### Surveillance analysis

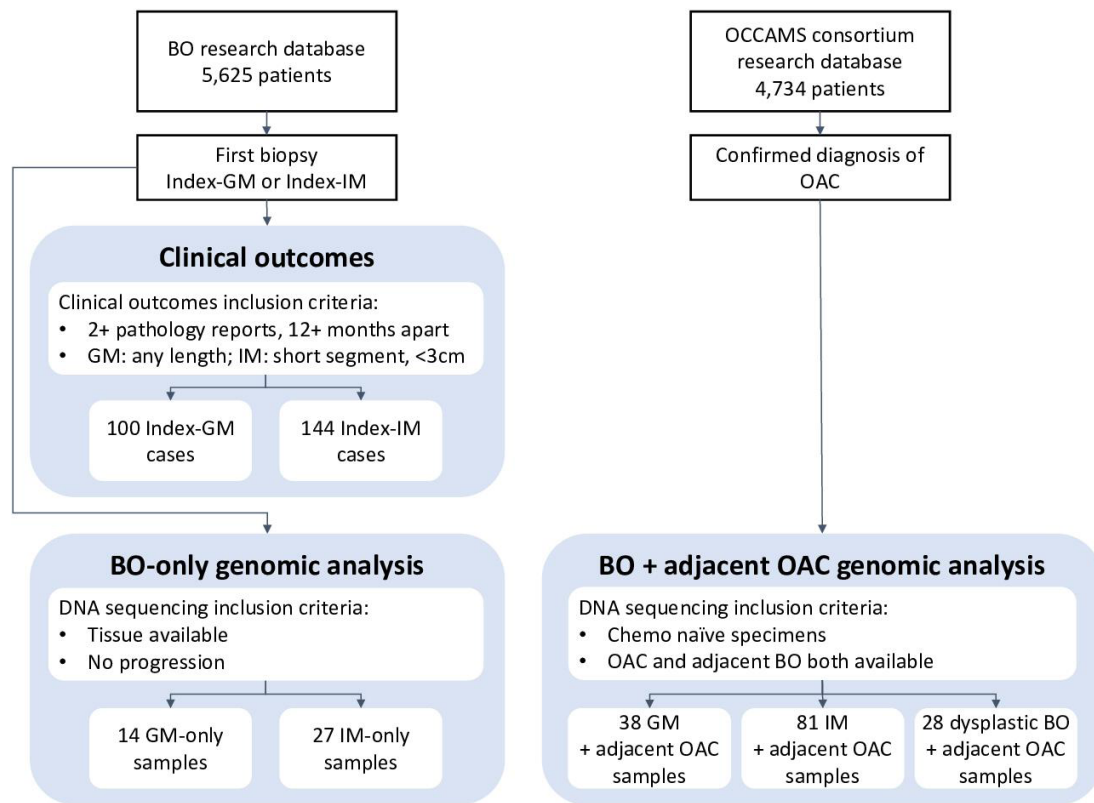
Cases with no follow-up after 1 December 2016 were presumed to have been discharged. This represented 6 years at the time of analysis. IM biopsies were classified into one of three groups: focal IM, widespread IM or unspecified, based on pathologist judgement. The terms 'focal' or 'minimal' were considered for focal IM, and 'widespread' or 'extensive' for widespread IM. Cases were classified as GM-only or IM-only, if OGD diagnoses were consistent; GM+IM, if there were both GM and IM diagnoses in the case history. If an otherwise IM-only case-reported GM followed by IM less than a year later, this was considered IM-only.

### BO-only genomic cohort

Tissue samples were available for patients in the Cell Determinants Biomarker study (REC 01/149), an observational study to identify biomarkers of the development and progression of BO. GM-only patients had to have only ever had GM biopsies and were chosen to include a range of ages and both male and female; 14 patients were selected.<sup>21</sup> Twenty-seven 'non-dysplastic, non-progressor' cases formed an IM-only comparative set<sup>31</sup> (figure 2).

### Adjacent BO and OAC genomic cohort

A cohort was selected from patients with tissue samples available from OAC resections or biopsies. The patients were sourced from the Oesophageal Cancer Classification and Molecular



**Figure 2** Study design. Flow chart showing creation of clinical outcomes cohort and sequencing cohorts. BO, Barrett's oesophagus; GM, gastric metaplasia; IM, intestinal metaplasia; OAC, oesophageal adenocarcinoma.

Stratification (OCCAMS) consortium in the UK. Individual informed consent was provided by all subjects (REC 07/H0305/52 and 10/H0305/1). Whole genome sequencing (WGS) was performed on 53 IM, 4 GM and 28 dysplasia samples and on OAC adjacent to each. Whole exome sequencing (WES) was performed on 34 GM and 28 IM samples, from 54 patients (8 patients had GM and IM), along with adjacent OAC, as well as 42 OAC samples with no adjacent BO (figure 2). Patients had to be chemo-naïve at the time of sampling. Patient characteristics are included in the online supplemental table S2.

### DNA extraction, library preparation and sequencing

For WGS, genomic DNA from samples and germline controls was extracted and processed as previously reported.<sup>21 31</sup> For WES, samples were prepared from formalin-fixed paraffin-embedded

(FFPE) slides. Pathologists identified areas of distinct tissue type with sufficient cellularity for sequencing (online supplemental figure 5), which were macrodissected and processed as described in the online supplemental methods.

Library preparation and enrichment for the WES samples were performed using Illumina DNA Prep with Enrichment, with 6-plex pooling. Sequencing was performed on an Illumina NovaSeq 6000 or HiSeq 4000, at the Cancer Research UK Cambridge Institute. Mean sequencing depth was 150×, with germline samples sequenced to at least 33× and BO or OAC samples to at least 53×.

### Variant calling and copy number alterations

For the WGS samples, variants, copy number alterations (CNAs) and mutational signatures were called as previously described.<sup>21 31</sup> For the WES samples, analysis focused on

**Table 1** Outcomes across all biopsy results for each of the three index biopsy subgroups

All biopsy results	Index GM-SS	Index GM-LS	Index IM-SS	Total cases	Mean OGDs
GM only	58 (75%)	9 (39%)		67	3.28
of which progressed to LGD	0	0		0	
of which progressed to HGD or IMC	0	0		0	
GM+IM	19 (25%)	14 (61%)	44 (31%)	77	4.84
of which progressed to LGD	1	0	1	2	
of which progressed to HGD or IMC	0	0	0	0	
IM only			100 (69%)	100	4.20
of which progressed to LGD			6	6	
of which progressed to HGD or IMC			5	5	
Total	77	23	144	244	4.15
Median follow-up years	6.1	9.7	7.8	7.2	
Total follow-up years	473	220	1161	1854	

GM, gastric metaplasia; HGD, high-grade dysplasia; IM, intestinal metaplasia; LGD, low-grade dysplasia; OGDs, ophagogastrroduodenoscopies.

single-nucleotide variants (SNVs) and small insertion and deletions (indels). FASTQ files were aligned to GRCh37 using BWA-MEM, with duplicates marked by Picard V2.9.5. Variant calling was performed using GATK Mutect2 V4.1.7.0,<sup>32</sup> using multisample and FFPE settings. Mutation filtering and copy number analysis are detailed in the online supplemental methods.

### Driver gene analysis

The set of OAC driver genes used was the 76 genes identified by Frankell *et al.*<sup>25</sup> We classed a subset of these as 'early IM/OAC genes': *TP53*, *CDKN2A*, *ARID1A*, *SMARCA4*, *MUC6*. These genes are frequently mutated in IM,<sup>31 33 34</sup> and under selection in IM.<sup>35</sup> Only non-silent mutations were included in driver gene analysis. Where multiple samples of the same tissue type were extracted for the same patient, a single sample was chosen at random to be included in analyses.

### Phylogenetic analysis

Clustering of SNVs was performed in PyClone V0.13.1,<sup>36</sup> integrating CNAs. Identified clusters with at least 10 mutations and median variant allele frequency (VAF) of at least 0.05 were used in ClonEvol<sup>37</sup> to create phylogenetic trees, with indels assigned to clusters post hoc. Shared mutations between the BO and OAC, with cancer cell fraction of at least 0.3 in OAC, were set as the founding cluster, even if fewer than 10 mutations.

### Statistical analysis

Statistical analysis was performed in R V4.0.3. The 'survival' and 'survminer' packages were used for Kaplan-Meier analysis and survival plots; with p values from log-rank tests. For mutational burden and somatic chromosomal alteration (SCA) load, means were compared using unpaired Wilcoxon's tests. A Pearson's  $\chi^2$  test tested association between focal and widespread IM and outcomes. For all statistical analyses, a p value <0.05 was considered significant. STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) cohort reporting guidelines were used.<sup>38</sup>

### Patient and public involvement

Patient group Heartburn Cancer UK review and co-design all our patient facing materials for the Cell Determinants Biomarker and OCCAMS studies. A lay friendly version of this paper will be shared with patients and the public.

## RESULTS

### Short-segment GM is a distinct state, not only mis-sampling of IM

To understand the prevalence of true GM compared with mis-sampling of GM+IM, we reviewed the full pathological history for 100 index-GM cases. Seventy-seven were short segments (SS, <3 cm) on index endoscopy and 23 were long segments (LS, ≥3 cm). Of the 77 index-GM-SS cases, over a median follow-up of 6.1 years, the majority (58 cases, 75%) were found to have a GM-only outcome (table 1). These GM-only outcomes confirm that persistent GM can be a distinct state in short segments.

In contrast to short segments, in the majority of index-GM-LS cases (14/23 cases, 61%), IM was detected in at least one later biopsy. Later detection of IM could be due to sampling at index or progression, and these two scenarios cannot be distinguished here. Four of the 14 GM+IM cases were first diagnosed before the Seattle protocol was recommended. A GM+IM outcome was also noted in 44 of the 144 (30%) index-IM-SS cases (table 1).

In some pathology reports, IM was noted as focal or widespread. This was examined to understand if focal IM contributed to GM+IM outcomes. A Pearson's  $\chi^2$  test showed that a GM+IM versus IM-only outcome was significantly dependent on focal or widespread IM (online supplemental table S3).

The sequence of diagnoses within the 33 index-GM, GM+IM outcome cases, was reviewed for patterns that could inform clinical practice, but no consistent patterns were observed (figure 3B, online supplemental table S4). Eighteen cases (55%) required three or more OGDs before observing IM. If a GM+IM outcome is due to sampling, a single repeat OGD is not always sufficient to overcome this challenge.

The surveillance history of the index-IM, GM+IM cases was limited, as 17 of the 44 patients (39%) had no further OGDs after the first GM or irregular Z-line observation (figure 3C). The surveillance histories of GM-only and IM-only cases are included for reference (figure 3A and D).

### Progression to dysplasia and cancer is lower in GM and GM+IM than IM

The progression rates to dysplasia and OAC were compared between GM, GM+IM and IM to understand the malignant risk. In total, this covered 1854 person-years of follow-up, with a median of 4.9 years in GM-only cases, 10.1 years in GM+IM cases and 7.5 years in IM-only cases. There was no progression to LGD, HGD or cancer in GM-only cases. Progression to LGD was identified in one index-GM, GM+IM case (1%), but no progression to HGD or cancer. This case that progressed was not treated after the LGD diagnosis, showed no abnormal p53 expression, and LGD was not seen again on later biopsies (figure 3B). Re-review by an expert upper gastrointestinal pathologist upheld the LGD diagnosis. Progression was higher within index-IM cases. Seven index-IM cases (5%, one GM+IM, six IM-only) progressed to LGD but not further. Of these seven cases, three were treated (as now recommended by guidelines) and four were not. Recent biopsies for the untreated cases did not show LGD, consistent with the difficulty in diagnosing this grade of dysplasia. A further five index-IM cases (3%, all IM-only) progressed to HGD or IMC (table 1).

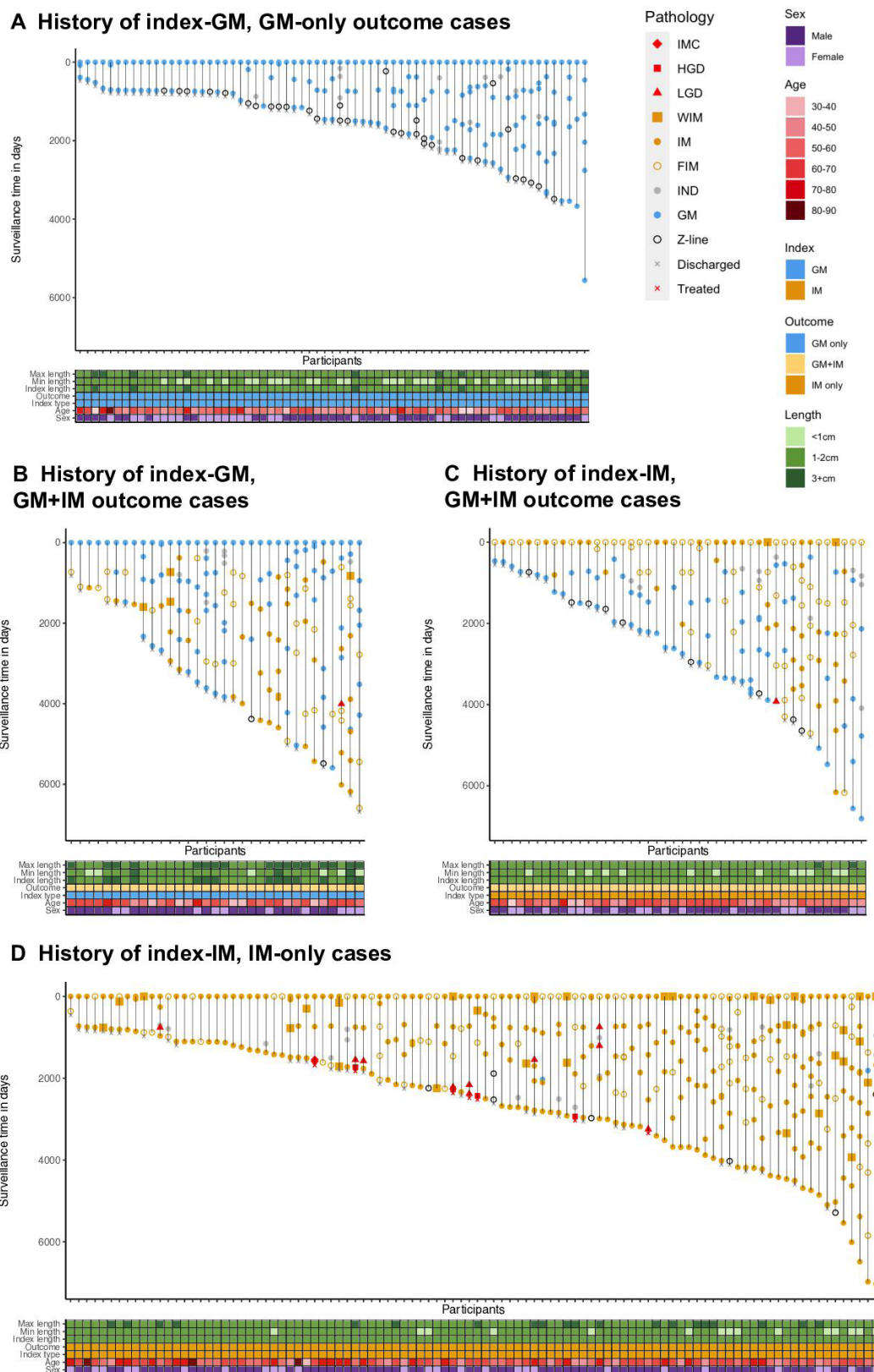
Kaplan-Meier analysis compared progression to LGD and HGD in GM-only, GM+IM and IM-only cases (figure 4). For both LGD and HGD, there was a significance in the difference between progression rate by metaplasia type (p=0.029 LGD, p=0.031 HGD), with progression lower in GM and GM+IM than in IM.

Finally, to understand consistency in length assessment, the endoscopy history was reviewed for each case. Strikingly, nearly all index GM-LS cases had a shorter length on a later endoscopy. The converse pattern was not observed: segments were rarely reported as a longer length in later OGDs (online supplemental figure 2). This suggests a trend over time to shorter assessments, rather than natural variation in length assessment.

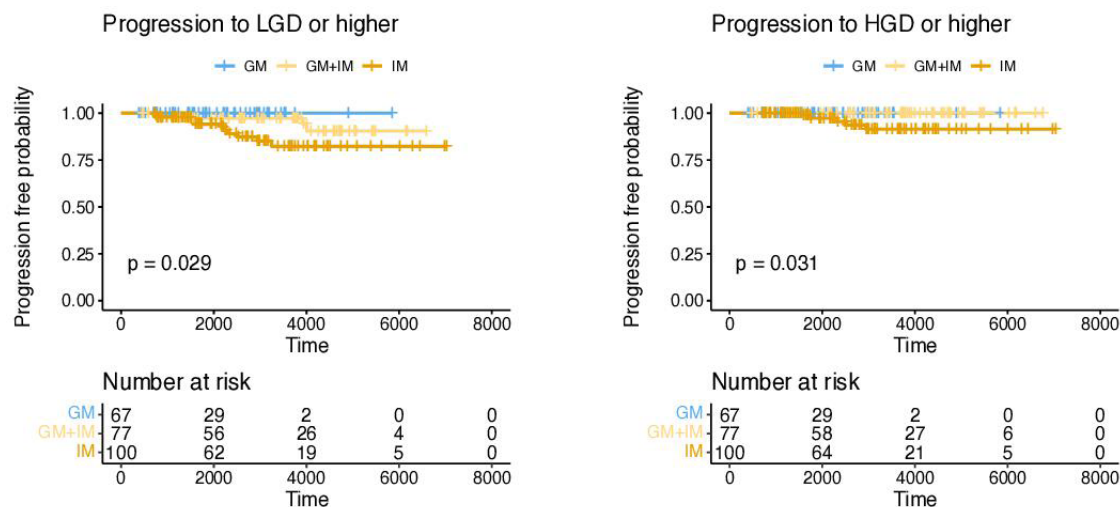
### GM bears few genomic hallmarks of OAC

It is well known that IM bears genomic hallmarks of OAC, such as driver gene mutations and signatures SBS17a/SBS17b. Less has been reported on the genomics of GM, particularly in a GM-only context. In the WGS of 14 GM-only samples and 27 IM-only samples, a significantly lower mutational burden in GM compared with IM was evident, even when the samples were partitioned into short and long segments (figure 5A). The difference in burden was also not explained by the ages in the two groups (figure 5B). SBS17a/b were detected in nearly all IM





**Figure 3** GM is a distinct state, not mis-sampling of IM. Progression to dysplasia and cancer is lower in GM than IM. (A) Surveillance history of 67 index-GM, GM-only outcome cases. (B) Surveillance history of 33 index-GM, GM+IM outcome cases. (C) Surveillance history of 44 index-IM, GM+IM outcome cases. (D) Surveillance history of 100 index-IM, IM-only outcome cases. For (A–D), each vertical trajectory is an individual case history, with the marker shapes denoting the histopathological assessment for each OGD. Each case has its surveillance time indexed to baseline, with the length of the vertical line representing total surveillance time. BO, Barrett's oesophagus; FIM, focal intestinal metaplasia; GM, gastric metaplasia; HGD, high-grade dysplasia; IM, intestinal metaplasia; IMC, intramucosal carcinoma; IND, indefinite for dysplasia; LGD, low-grade dysplasia; WIM, widespread intestinal metaplasia.



**Figure 4** Progression to dysplasia and cancer is lower in GM than IM. Kaplan-Meier curves for progression to LGD or higher and to HGD or higher, separated by classification. GM, gastric metaplasia; HGD, high-grade dysplasia; IM, intestinal metaplasia; LGD, low-grade dysplasia.

samples, but only in one GM sample (figure 5C). SBS1, a signature associated with ageing, was present in all GM samples, and all but one IM sample.

All samples were assessed for non-silent mutations and CNAs in genes associated with IM and OAC. Only one GM sample (7%) had an alteration in an early IM/OAC gene, *SMARCA4*. By contrast, 16 IM samples (59%) had an alteration in at least one early IM/OAC gene (figure 5D). The SCA load—the length of the genome with a copy number gain, loss or loss of heterozygosity—was also assessed. There was no difference in SCA load between GM and IM non-progressors (figure 5E), and both were an order of magnitude lower than in OAC samples (online supplemental figure 3). No driver amplifications or deletions were seen in the GM samples.

### Co-occurring GM and OAC are genomically distant from one another

To better understand the evolutionary relationship between different types of BO and OAC, the mutations found in adjacent BO and tumours were assessed using WES and WGS. The mutational burden in tumours was independent of the presence or type of BO (figure 6A). In adjacent BO, the mutational burden varied by type of BO, with the burden significantly lower in adjacent GM than adjacent IM ( $p=0.022$ ) (figure 6B).

We found that one adjacent GM sample (3%) had a mutation in *TP53* and one (3%) in *MUC6*, but no other GM samples had any mutations in early IM/OAC genes (figure 6C). By contrast, 41 (51%) IM samples and 19 (68%) dysplastic samples had mutations in early IM/OAC genes. Across the full set of OAC driver genes, on average, GM had mutations in 0.5 drivers, compared with 1.6 in IM and 2.0 in dysplasia. The one GM sample with a *TP53* mutation did not have a second hit on *TP53* and had a low VAF of 0.11. Of the 9 IM samples with *TP53* mutations, 4 (44%) had a second hit in the form of LOH, as did 10 of the 14 (71%) *TP53*-mutated dysplastic samples.

There were very few shared driver gene mutations between the 38 GM samples and their adjacent OAC: 4 shared, 16 unique to GM, 132 unique to OAC (figure 6D). In the WGS samples, there were no driver amplifications or deletions called in GM. It is notoriously difficult to robustly call CNAs in WES, but using a joint segmentation approach, one shared amplification was called between GM and OAC (online supplemental figure 4),

although in a GM sample with no *TP53* mutation. Mutational overlap was also low between IM and OAC, but much higher between dysplasia and OAC (27 shared, 32 unique to dysplasia, 74 unique to OAC, figure 6D). This suggests that GM is distantly related to the OAC, and most mutations in driver genes in GM are not drivers of oncogenesis in this context. Phylogenetic trees for the matched samples, with branch length representing number of mutations, also demonstrate the relative distance of GM from OAC compared with dysplasia and OAC (figure 6E). The example tree with GM and OAC had a very short trunk of shared mutations and no shared OAC driver gene mutations. By contrast, the HGD and OAC tree had a long trunk, including four shared mutations in OAC driver genes.

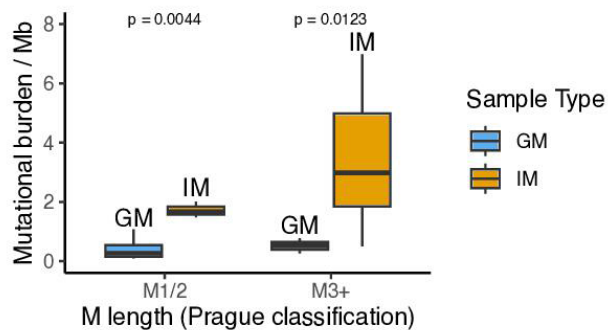
### DISCUSSION

We present clinical and genomic evidence for the lower malignant potential of GM. We followed clinical outcomes in a novel cohort of 244 patients, uniquely representing recent clinical practice in a country where GM is still surveilled. We also analysed sequencing of 41 GM and IM samples from non-progressors and 119 GM and IM samples from patients with OAC.

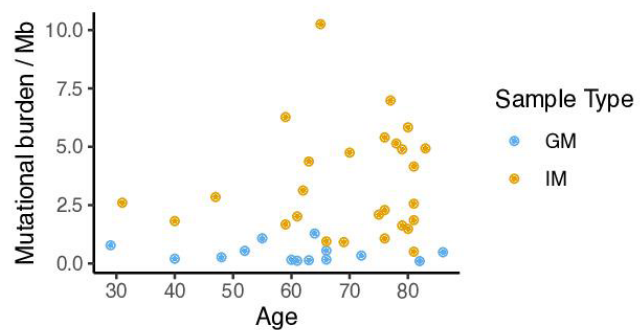
The clinical data demonstrate that the progression in GM and GM+IM is extremely low and lower than in IM (0% incidence per year of HGD or cancer in index-GM, 0.43% in index-IM). This is consistent with a large-scale population study,<sup>10</sup> which observed a lower rate of progression in 3179 GM cases compared with 3917 IM cases (0.07% vs 0.38% incidence of HGD or cancer per year) but lacked length data and a systematic biopsy protocol. The results are also consistent with the US study by Westerhoff *et al*,<sup>9</sup> which saw no progression to OAC across 379 patients with GM but did not have a focus on patients with GM+IM. While it cannot be said that GM will never progress to OAC, we expect that the progression rate is more akin to that of the general population than that of patients with IM.

The results may appear inconsistent with the study by Evans *et al*, which categorised BO segments on a microscopic level, based on diversity of gland phenotypes.<sup>39</sup> They reported that segments with higher gland diversity had higher progression risk, which might be taken to suggest higher progression in GM+IM cases. However, there is no link between our longitudinal outcome categories and microscopic diversity of glands; IM-only segments

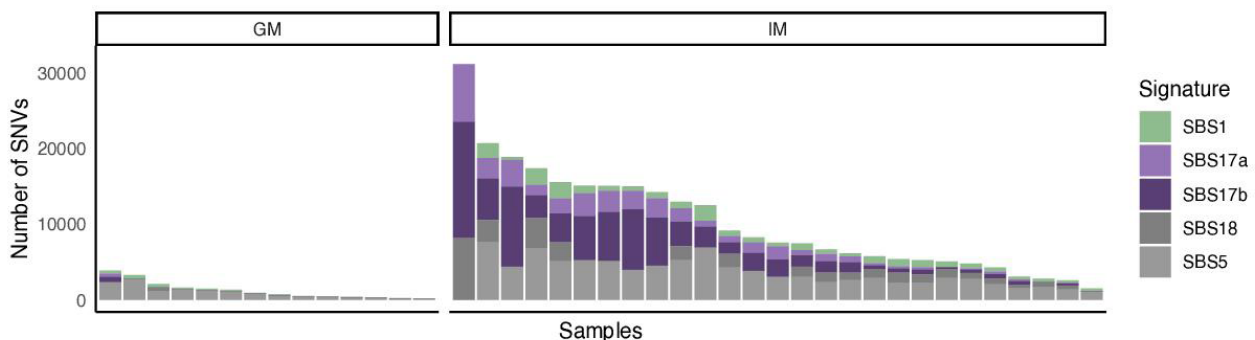
## A Mutational burden, by length



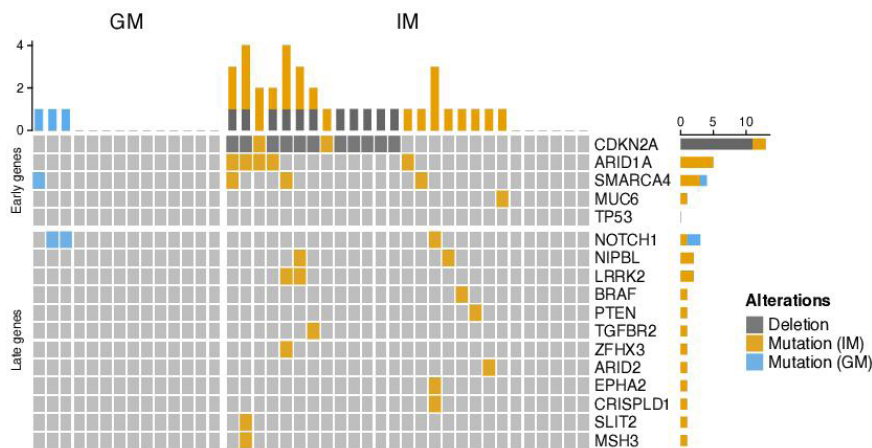
## B Mutational burden, by age



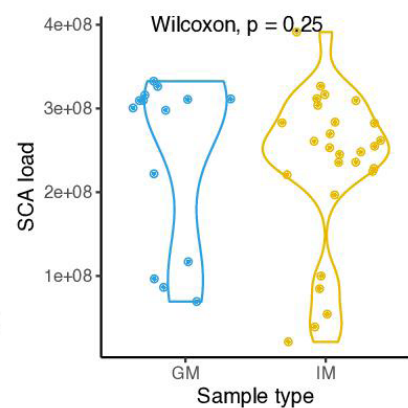
## C Signature contribution by sample



## D Mutations in OAC associated genes



## E SCA load



**Figure 5** GM does not bear the genomic hallmarks of OAC. (A) Mutational burden by sample type and BO segment length. (B) Mutational burden by age. (C) SBS signature contribution by sample. (D) Mutations in OAC associated genes, split by genes typically mutated early and late in the progression of OAC. Only genes with mutations in this cohort, along with TP53, are shown. (E) SCA load by sample, taken as the length of genome altered by a copy gain, loss or copy neutral loss of heterozygosity. BO, Barrett's oesophagus; GM, gastric metaplasia; IM, intestinal metaplasia; OAC, oesophageal adenocarcinoma; SCA, somatic chromosomal alterations; SNV, single nucleotide variant.

can have higher diversity of glands than GM+IM segments, especially given IM is so often focal within GM+IM cases.

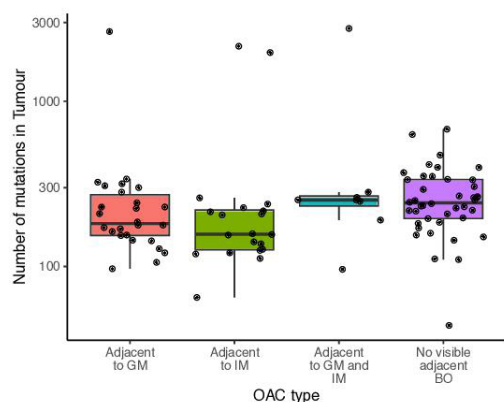
It is well known that IM can be missed in GM+IM samples, particularly if the IM is focal. However, most short GM segments continued to receive a GM diagnosis, including cases with over five OGDs. This gives confidence that these are true GM segments throughout the surveillance period, not under-sampled GM+IM. This is consistent with the findings of Chandrasoma *et al*,<sup>11</sup> who found that under a robust sampling regime,

it was possible to identify true GM cases, and Westerhoff *et al*, who found that 88% of purported patients with GM continued to have no goblet cells.<sup>9</sup> While other studies have shown much higher rates of later IM in index-GM cases,<sup>12–14</sup> much can be explained by inadequate sampling, prior to the introduction of the Seattle protocol.

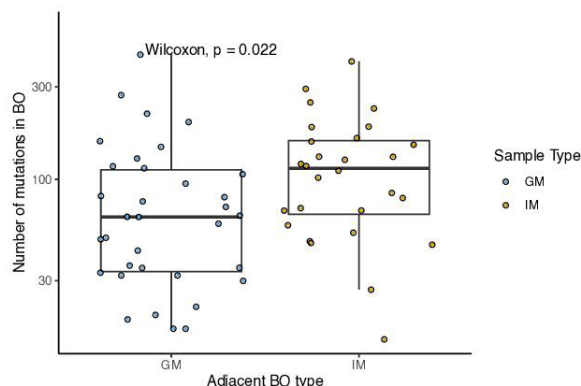
By contrast, the majority of the long GM segments were found to contain IM in a later biopsy, consistent with a study where 7 out of 11 LS-GM were found to have IM later.<sup>12</sup> Goblet cell



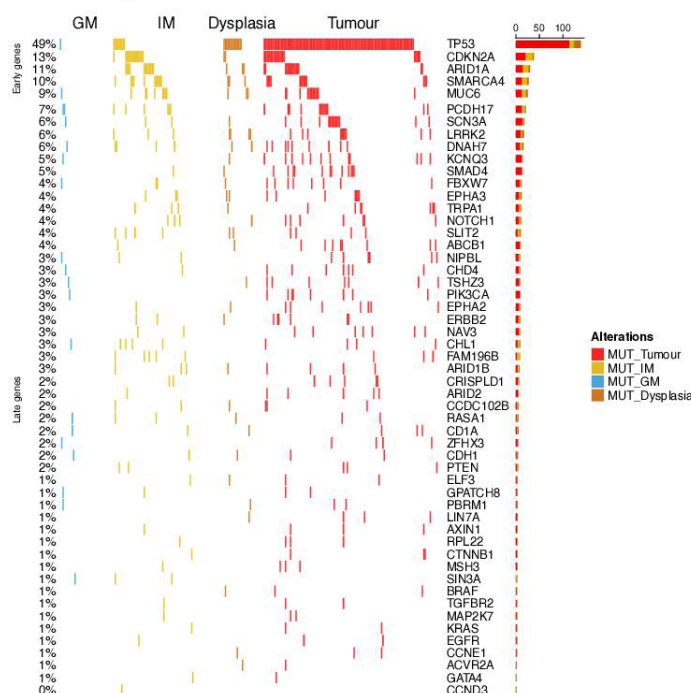
## A Mutation burden in tumours



## B Mutation burden in BO adjacent to OAC

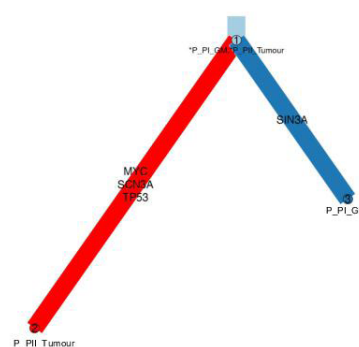


## C Driver gene mutations

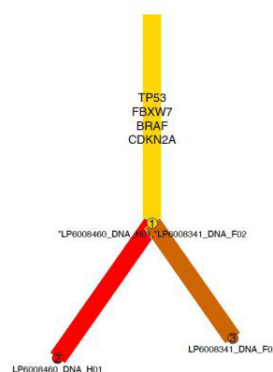


## E Example phylogenetic trees

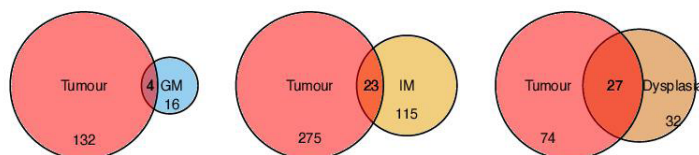
### Phylogenetic tree: GM and adjacent OAC



### Phylogenetic tree: HGD (in IM background) and adjacent OAC



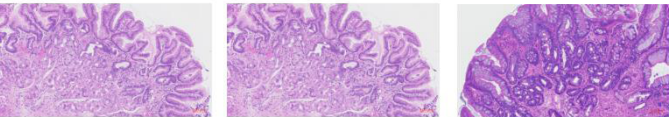
## D Shared driver gene mutations



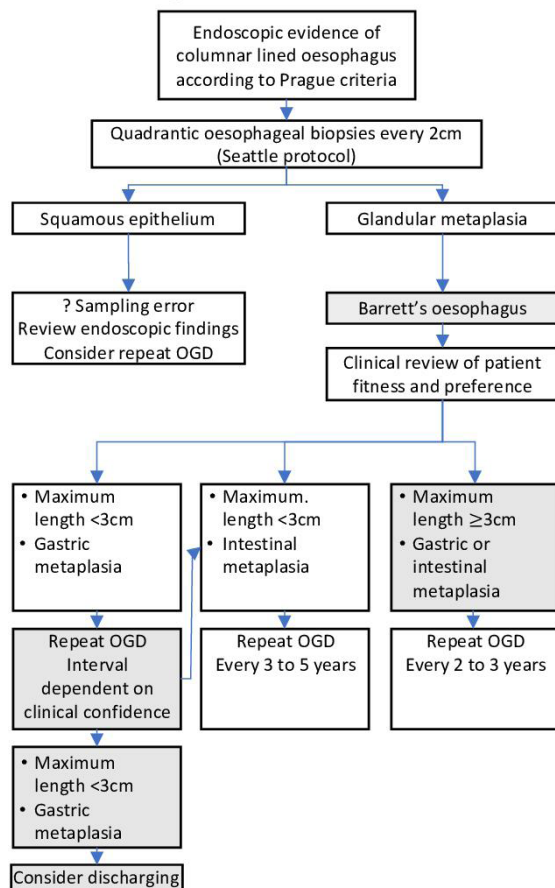
**Figure 6** Co-occurring GM and OAC are evolutionarily distant from one another. (A) Mutation burden in OAC tumours, by type of adjacent BO. (B) Mutation burden in BO adjacent to OAC. (C) Mutations in driver genes by sample, across BO samples and adjacent OAC. Only genes with a mutation in a BO sample are shown. (D) Venn diagrams showing the degree of overlap in mutations between different types of BO and the adjacent OAC. (E) Example phylogenetic trees for one case with GM and adjacent OAC, and one case with HGD in an intestinal background, and adjacent OAC. Branch length represents number of mutations. BO, Barrett's oesophagus; GM, gastric metaplasia; IM, intestinal metaplasia; OAC, oesophageal adenocarcinoma.



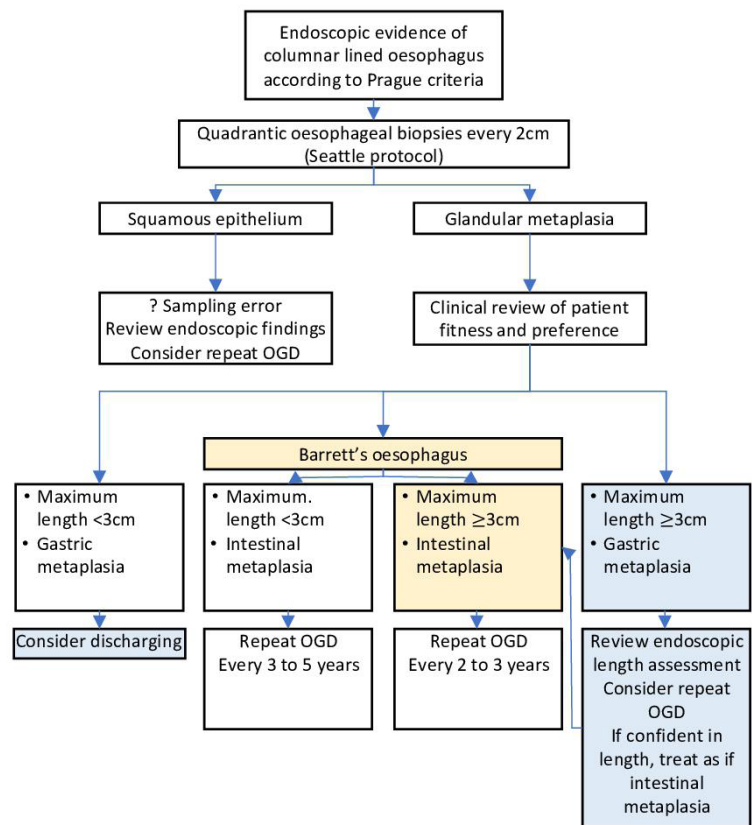
## A Summary of features

	Short-segment GM	Long-segment GM	Short-segment IM (no dysplasia)
<b>Length</b>	<3cm	3cm+	<3cm
<b>Histopathology</b>			
Goblet cells	No	No	Yes
<b>Expected molecular characteristics</b>			
Aberrant P53	No	No	No
Aging signature (SBS1)	Yes	Yes	Yes
Typical OAC signature (SBS17a/b)	No	No	Yes
OAC driver gene mutations	0	0	1-2

## B Current guidelines



## C Recommendations



**Figure 7** Short-segment gastric metaplasia does not warrant surveillance. (A) Summary of features of GM and IM. (B) Current British Society of Gastroenterology guidelines for management of non-dysplastic BO.<sup>4</sup> Grey boxes denote areas of recommended change. (C) Recommended flow chart for non-dysplastic BO. Yellow boxes denote changes to BO definition, blue boxes denote changes to GM management. BO, Barrett's oesophagus; GM, gastric metaplasia; IM, intestinal metaplasia; OGD, oesophagogastrroduodenoscopy.

density being higher proximally<sup>15</sup> also ties with a higher likelihood of IM in a longer segment. These results suggest that more considered management of LS-GM cases is required than that of short-segment GM (SS-GM), as is already the case today.

The WGS of non-progressors showed that mutational burden in GM was significantly lower than that of IM (SS:  $p < 0.01$ , LS:  $p = 0.0123$ ). In GM, there were very few mutations in genes associated with OAC, and signatures SBS17a/b were not prevalent. Altogether, GM does not bear the mutational hallmarks of OAC, but IM does. Similar results were seen in sequencing of GM and IM adjacent to OAC. These results are consistent with the targeted sequencing of Bandla *et al*<sup>27</sup> but strengthen the findings due to examining the whole genome or exome and almost four times as many OAC-associated genes.

Liu *et al* found similar levels of DNA-content abnormalities in GM and IM using image cytometry,<sup>28</sup> and we too found similar SCA loads in GM and IM. Non-progressor IM, with no *TP53* mutation, rarely has substantial CNAs. Therefore, we do not consider this indicative of malignant potential in either metaplasia type.

Of the few driver gene mutations seen in adjacent-GM, there was little overlap with mutations in the tumour. Although Lavery *et al* showed that OAC could arise from gastric glands,<sup>30</sup> this was a single case, with IM. It has not been shown whether this occurs widely or without IM present. Our results suggest that adjacent GM is not routinely evolutionarily closer than IM is to OAC. This is consistent with phylogenies presented by Bao *et al*.<sup>29</sup> Since an argument for the malignant risk of GM is its co-occurrence with OAC, it is important to be clear that even when GM or IM is spatially close to OAC, it does not follow that the GM or IM is evolutionarily close to OAC. Instead, it is likely that the dysplastic clone the OAC directly evolved from has been overgrown or was not sampled.

A few limitations should be considered with the sequencing. First, while the macrodissection was carried out with great care, there remains the possibility that some GM areas could have been sampled from normal cardia. Second, mutations in GM occur at low-variant allele frequencies, so at standard sequencing depths, it is possible that there were undetected mutations in relevant genes. A real strength in the genomic analysis here is that BO-only and BO adjacent to OAC have both been studied, and the results are similar in the two contexts.

Many GM studies, including ours, suffer from small sample sizes, primarily due to few GM patients being enrolled in studies. A strength of our study is the number of GM+IM cases, as low progression in these should reassure clinicians concerned by misdiagnosing GM in a GM+IM case. As with the macrodissections, we cannot rule out some of the GM diagnoses in the clinical outcome cohort coming from mis-sampled cardia. Finally, we recognise the limitation of the outcome cohort coming from a single centre. However, since this is an expert centre that follows a systematic biopsy protocol and sees many tertiary referrals, we are confident in the robustness of our findings.

While not a focus of the study, the data suggest that focal IM has low progression risk. It has previously been shown that focal IM at the gastro-oesophageal junction, after radiofrequency ablation, does not progress to dysplasia.<sup>40</sup> Confirmation of low progression in focal IM could further improve surveillance targeting. Systematic capture by pathologists, which is missing from our study, perhaps aided by digital pathology and AI, could inform this. Alternatively, the cytosponge device could assist, as quantification of TFF3 expression from a cytosponge sample can identify focal IM pathologies.<sup>41</sup>

In summary, SS-GM is molecularly distinct from SS-IM. The malignant potential of SS-GM is low and lower than that of

IM. We have suggested an alternative framework to help determine who should enter surveillance and would recommend that patients with SS-GM do not enter surveillance (figure 7). While some of these cases may have later been found to be GM+IM, short-segment GM+IM cases still represent low-risk cases. By contrast, there is high likelihood that LS-GM is undersampled GM+IM, and risk of progression is higher in longer cases.<sup>42</sup> Therefore, in the absence of molecular confirmation of GM and reconfirmation of length, we would recommend these relatively rare cases be treated as if they were IM (figure 7).

**Acknowledgements** We sincerely thank Judith Honing for the help and guidance on assessing clinical outcomes, and Massimiliano di Pietro for helpful discussions. We thank Elwira Merry for organising the collection of tissue blocks, and Alex Northrop and Conor Flint for assisting with tissue extraction. We thank Lizhe Zhuang for histopathology images. We thank the Human Research Tissue Bank, which is supported by the UK National Institute for Health Research (NIHR) Cambridge Biomedical Research Centre, from Addenbrooke's Hospital. Figure 1 was created with BioRender.com.

**Collaborators** Rebecca C Fitzgerald, Barbara Nutzinger, Aisling M Redmond, Christine Lorenzo, Sujath Abbas, Adam Freeman, Calvin Cheah, Hannah Coles, Curtis Millington (Early Cancer Institute, University of Cambridge, Cambridge, UK), Paul A W Edwards (Early Cancer Institute, University of Cambridge, Cambridge, UK), Cancer Research UK Cambridge Institute, University of Cambridge, Cambridge, UK), Matthew Eldridge, Maria Secrier, Ginny Devonshire, Sriganesh Jammula (Cancer Research UK Cambridge Institute, University of Cambridge, Cambridge, UK), Maria O'Donovan, Ahmad Miremadi, Shalini Malhotra, Monika Tripathi (Early Cancer Institute, University of Cambridge, Cambridge, UK, Department of Histopathology, Addenbrooke's Hospital, Cambridge, UK), Jim Davies, Charles Crichton (Department of Computer Science, University of Oxford, UK, OX1 3QD), Elizabeth C Smyth, Nick Carroll, Richard H Hardwick, Peter Safranek, Andrew Hindmarsh, Vijayendran Sujendran (Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK, CB2 0QQ), Nicola Grehan (Early Cancer Institute, University of Cambridge, Cambridge, UK, Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK, CB2 0QQ), Yeng Ang (Salford Royal NHS Foundation Trust, Salford, UK, M6 8HD), Wigan and Leigh NHS Foundation Trust, Wigan, Manchester, UK, WN1 2NN, GI science centre, University of Manchester, UK, M13 9PL), Shaun R Preston, Izhar Bagwan (Royal Surrey County Hospital NHS Foundation Trust, Guildford, UK, GU2 7XX), Vicki Save, Richard J E Skipworth (Edinburgh Royal Infirmary, Edinburgh, UK, EH16 4SA), Philippe Taniere, Sonia Puig, Gianmarco Contino (University Hospitals Birmingham NHS Foundation Trust, Birmingham, UK, B15 2GW), Ben L Grace (University Hospital Southampton NHS Foundation Trust, Southampton, UK, SO16 6YD), Timothy J Underwood, Robert C Walker (University Hospital Southampton NHS Foundation Trust, Southampton, UK, SO16 6YD), Cancer Sciences Division, University of Southampton, Southampton, UK, SO17 1BJ), Stephen J Hayes (Salford Royal NHS Foundation Trust, Salford, UK, M6 8HD), Faculty of Medical and Human Sciences, University of Manchester, UK, M13 9PL), Ula Mahadeva (Guy's and St Thomas's NHS Foundation Trust, London, UK, SE1 7EH), Grant Sanders, Richard Berrisford, David Chan (Plymouth Hospitals NHS Trust, Plymouth, UK, PL6 8DH), Ed Cheong, Bhaskar Kumar, L Sreedharan (Norfolk and Norwich University Hospital NHS Foundation Trust, Norwich, UK, NR4 7UY), Simon L Parsons, Irshad Soomro, Philip Kaye (Nottingham University Hospitals NHS Trust, Nottingham, UK, NG7 2UH), Laurence Lovat, Rehan Haidry (University College London, London, UK, WC1E 6BT), Michael Scott (Wythenshawe Hospital, Manchester, UK, M23 9LT), Ted R. Hupp (Edinburgh University, Edinburgh, UK, EH8 9YL), Vicky Goh, Francesca D Ciccarelli (King's College London, London, UK, WC2R 2LS), Jesper Lagergren (Guy's and St Thomas's NHS Foundation Trust, London, UK, SE1 7EH), Karolinska Institute, Stockholm, Sweden, SE-171 77), Sharmila Sothi (University Hospitals Coventry and Warwickshire NHS, Trust, Coventry, UK, CV2 2DX), Suzy Lishman (Cancer Research UK Cambridge Institute, University of Cambridge, Cambridge, UK, Peterborough Hospitals NHS Trust, Peterborough City Hospital, Peterborough, UK, PE3 9GZ), Andrew Beggs (University Hospitals Birmingham NHS Foundation Trust, Birmingham, UK, B15 2GW), Institute of Cancer and Genomic sciences, University of Birmingham, B15 2TT), Andrew Sharrocks (GI science centre, University of Manchester, UK, M13 9PL), George B Hanna, Christopher J Peters, Krishna Moorthy (Department of Surgery and Cancer, Imperial College, London, UK, W2 1NY), Anna Grabowska (Queen's Medical Centre, University of Nottingham, Nottingham, UK), Olga Tucker (University Hospitals Birmingham NHS Foundation Trust, Birmingham, UK, B15 2GW), Heart of England NHS Foundation Trust, Birmingham, UK, B9 5SS), Richard Turkington, Damian McManus, Helen Coleman (Centre for Cancer Research and Cell Biology, Queen's University Belfast, Northern Ireland BT7 1NN), Russell D Petty (Tayside Cancer Centre, Ninewells Hospital and Medical School, Dundee, DD1 9SY), Freddie Bartlett (Portsmouth Hospitals NHS Trust, Portsmouth, PO6 3LY), J Robert O'Neill (Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK, CB2 0QQ), Edinburgh Royal Infirmary, Edinburgh, UK, EH16 4SA, Edinburgh University, Edinburgh, UK,

EH8 9YL), James Gossage, Andrew Davies, Fujun Chang (Guy's and St Thomas's NHS Foundation Trust, London, UK, SE1 7EH, King's College London, London, UK, WC2R 2LS), John Saunders (Salford Royal NHS Foundation Trust, Salford, UK, M6 8HD, Nottingham University Hospitals NHS Trust, Nottingham, UK, NG7 2UH).

**Contributors** RCF and ELB conceived of the study. ELB, EO, AWTN and GD carried out the analyses of the genomic data and statistical analyses. RCF and ELB designed and carried out the clinical analysis. RCF supervised the research. MOD, SM, MT and AM performed the pathology review of the samples. AF, HC and EO coordinated and carried out the processing of samples. OCCAMS Consortium provided tissue and blood samples. RCF obtained funding. RCF and ELB wrote the manuscript. RCF is guarantor. All authors approved the manuscript.

**Funding** This research was supported by the NIHR Cambridge Biomedical Research Centre (NIHR203312). The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care. The laboratory of RCF was funded by a Core Programme Grant from the Medical Research Council (G111260, MR/W014122/1). OCCAMS was funded by a Programme Grant from Cancer Research UK (RG66287). OCCAMS2 was funded by a Programme Grant from Cancer Research UK (RG81771/84119). ELB was funded by the Medical Research Council.

**Competing interests** RCF and MOD are named on patents for Cytosponge and associated technology, licensed to Covidien GI solutions (now Medtronic). RCF and MOD are shareholders of Cyted Ltd., a company working on early detection technology. The other authors declare no competing interests.

**Patient and public involvement** Patients and/or the public were involved in the design, or conduct, or reporting or dissemination plans of this research. Refer to the Methods section for further details.

**Patient consent for publication** Not applicable.

**Ethics approval** This study involves human participants and was approved by Cambridge Central Research Ethics Committee REC 01/149, Cambridge South Research Ethics Committee REC 07/H0305/52, 10/H0305/1. Participants gave informed consent to participate in the study before taking part.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** Data are available upon reasonable request. The sequencing data included in this study have been submitted to European Genome-phenome Archive (EGA; <https://ega-archive.org/>) with accession numbers EGAD00001011188, EGAD00001011187 and EGAD00001011255.

**Supplemental material** This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution 4.0 Unported (CC BY 4.0) license, which permits others to copy, redistribute, remix, transform and build upon this work for any purpose, provided the original work is properly cited, a link to the licence is given, and indication of whether changes were made. See: <https://creativecommons.org/licenses/by/4.0/>.

#### ORCID iDs

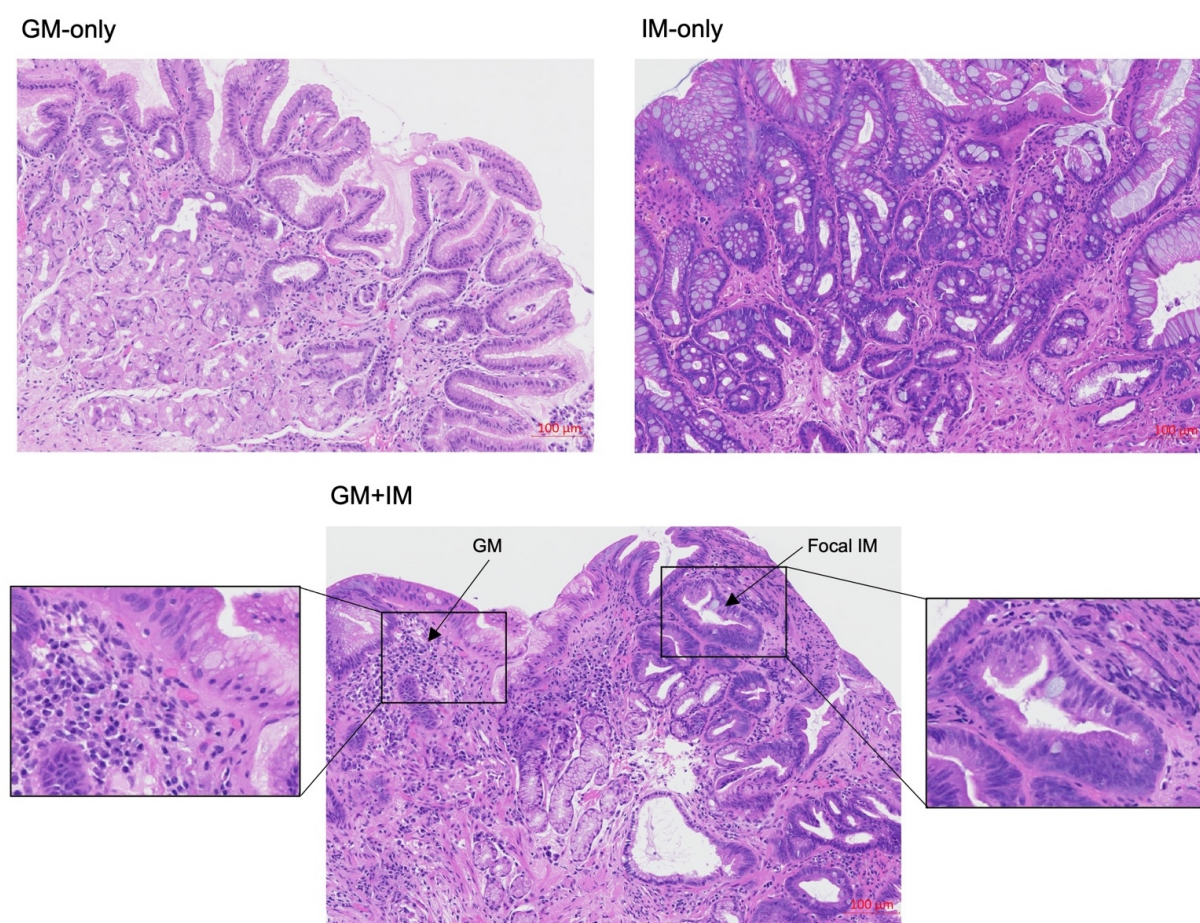
Emily L Black <http://orcid.org/0000-0001-9929-8043>  
Rebecca C Fitzgerald <http://orcid.org/0000-0002-3434-3568>

#### REFERENCES

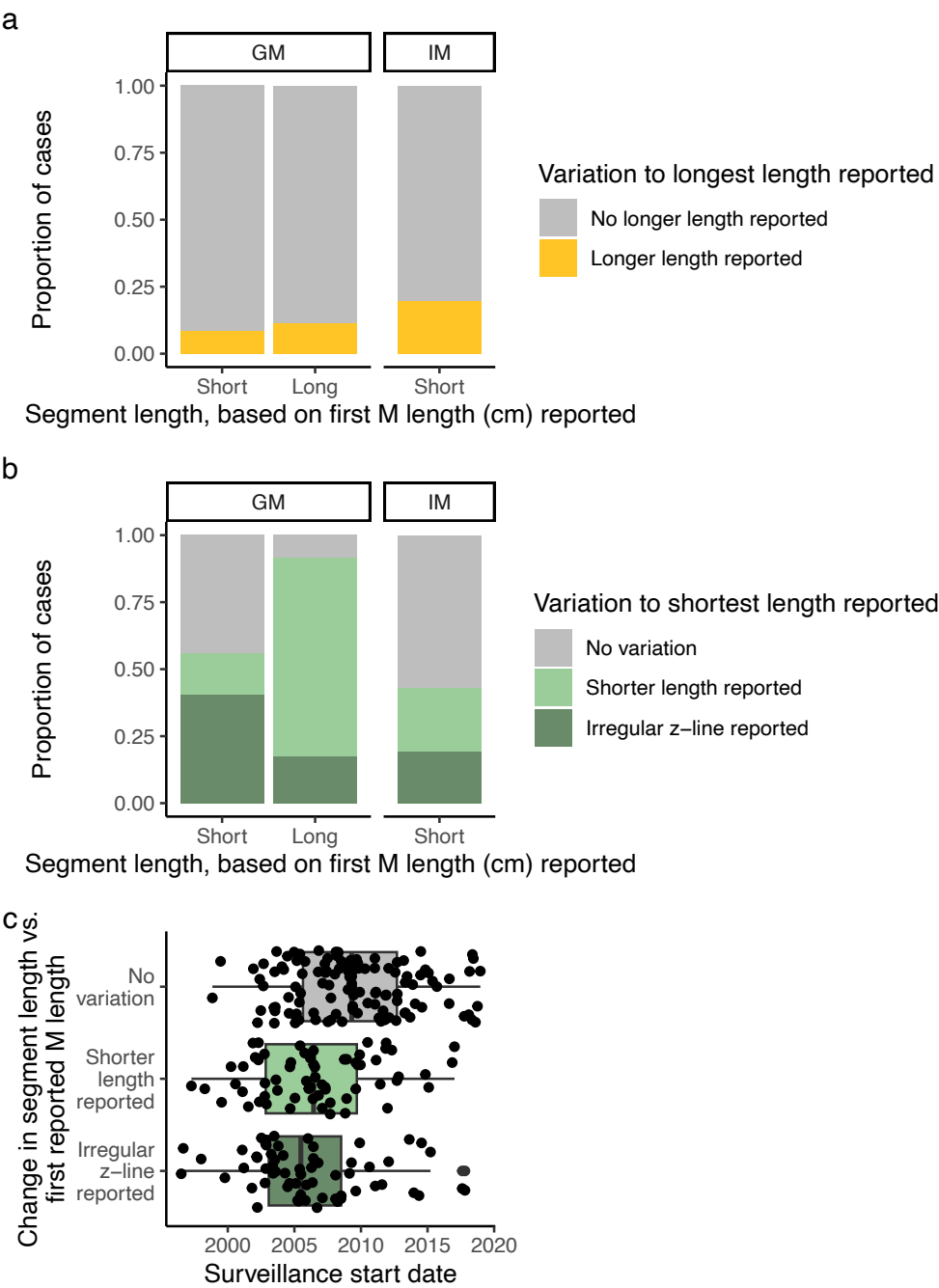
- Rastogi A, Puli S, El-Serag HB, *et al.* Incidence of esophageal adenocarcinoma in patients with Barrett's esophagus and high-grade dysplasia: a meta-analysis. *Gastrointest Endosc* 2008;67:394–8.
- Kastelein F, van Olphen S, Steyerberg EW, *et al.* Surveillance in patients with long-segment Barrett's esophagus: a cost-effectiveness analysis. *Gut* 2015;64:864–71.
- Hvid-Jensen F, Pedersen L, Drewes AM, *et al.* Incidence of adenocarcinoma among patients with Barrett's esophagus. *N Engl J Med* 2011;365:1375–83.
- 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.
- Shaheen NJ, Falk GW, Iyer PG, *et al.* ACG clinical guideline: diagnosis and management of Barrett's esophagus. *Am J Gastroenterol* 2016;111:30–50.
- Whiteman DC, Appleyard M, Bahin FF, *et al.* Australian clinical practice guidelines for the diagnosis and management of Barrett's esophagus and early esophageal adenocarcinoma. *J of Gastro and Hepatol* 2015;30:804–20. 10.1111/jgh.12913 Available: <https://onlinelibrary.wiley.com/toc/14401746/30/5>
- Qumseya B, Sultan S, *et al.* ASGE STANDARDS OF PRACTICE COMMITTEE. ASGE guideline on screening and surveillance of Barrett's esophagus. *Gastrointest Endosc* 2019;90:335–59.
- Săftoiu A, Hassan C, Areia M, *et al.* Role of gastrointestinal endoscopy in the screening of digestive tract cancers in Europe: European society of gastrointestinal endoscopy (ESGE) position statement. *Endoscopy* 2020;52:293–304.
- Westerhoff M, Hovan L, Lee C, *et al.* Effects of dropping the requirement for goblet cells from the diagnosis of Barrett's esophagus. *Clin Gastroenterol Hepatol* 2012;10:1232–6.
- Bhat S, Coleman HG, Yousef F, *et al.* Risk of malignant progression in Barrett's esophagus patients: results from a large population-based study. *J Natl Cancer Inst* 2011;103:1049–57.
- Chandrasoma P, Wijetunge S, DeMeester S, *et al.* Columnar-lined esophagus without intestinal metaplasia has no proven risk of adenocarcinoma. *Am J Surg Pathol* 2012;36:1–7.
- Dias Pereira A, Chaves P. Columnar-lined esophagus without intestinal Metaplasia: results from a cohort with a mean follow-up of 7 years. *Aliment Pharmacol Ther* 2012;36:282–9.
- Gatenby PAC, Ramus JR, Caygill CPJ, *et al.* Relevance of the detection of intestinal Metaplasia in non-Dysplastic columnar-lined Oesophagus. *Scand J Gastroenterol* 2008;43:524–30.
- Kelty CJ, Gough MD, Van Wyk Q, *et al.* Barrett's oesophagus: intestinal Metaplasia is not essential for cancer risk. *Scand J Gastroenterol* 2007;42:1271–4.
- Theodorou D, Ayazi S, DeMeester SR, *et al.* Intraluminal pH and goblet cell density in Barrett's esophagus. *J Gastrointest Surg* 2012;16:469–74.
- Takubo K, Aida J, Naomoto Y, *et al.* Cardiac rather than intestinal-type background in endoscopic resection specimens of minute Barrett adenocarcinoma. *Hum Pathol* 2009;40:65–74.
- Sabel MS, Pastore K, Toon H, *et al.* Adenocarcinoma of the esophagus with and without Barrett mucosa. *Arch Surg* 2000;135:831–5.
- Sawas T, Killcoyne S, Iyer PG, *et al.* Identification of Prognostic phenotypes of Esophageal adenocarcinoma in 2 independent cohorts. *Gastroenterology* 2018;155:1720–8.
- Curtius K, Rubenstein JH, Chak A, *et al.* Computational modelling suggests that Barrett's oesophagus may be the precursor of all oesophageal adenocarcinomas. *Gut* 2021;70:1435–40.
- Nowicki-Osuch K, Zhuang L, Jammula S, *et al.* Molecular phenotyping reveals the identity of Barrett's esophagus and its malignant transition. *Science* 2021;373:760–7.
- Nowicki-Osuch K, Zhuang L, Cheung TS, *et al.* Single-cell RNA sequencing unifies developmental programs of esophageal and gastric intestinal metaplasia. *Cancer Discov* 2023;13:1346–63.
- Huo X, Zhang HY, Zhang XI, *et al.* Acid and bile salt-induced CDX2 expression differs in esophageal squamous cells from patients with and without Barrett's esophagus. *Gastroenterology* 2010;139:194–203.
- Saller J, Al Difalaha S, Neill K, *et al.* CDX-2 expression in esophageal biopsies without goblet cell intestinal metaplasia may be predictive of Barrett's esophagus. *Dig Dis Sci* 2020;65:1992–8.
- Killcoyne S, Fitzgerald RC. Evolution and progression of Barrett's oesophagus to oesophageal cancer. *Nat Rev Cancer* 2021;21:731–41.
- Frankell AM, Jammula S, Li X, *et al.* The landscape of selection in 551 esophageal adenocarcinomas defines genomic biomarkers for the clinic. *Nat Genet* 2019;51:506–16.
- Ng AWT, Contino G, Killcoyne S, *et al.* Rearrangement processes and structural variations show evidence of selection in Oesophageal adenocarcinomas. *Commun Biol* 2022;5:335.
- Bandla S, Peters JH, Ruff D, *et al.* Comparison of cancer-associated genetic abnormalities in columnar-lined esophagus tissues with and without goblet cells. *Ann Surg* 2014;260:72–80.
- Liu W, Hahn H, Odze RD, *et al.* Metaplastic esophageal columnar epithelium without goblet cells shows DNA content abnormalities similar to goblet cell-containing epithelium. *Am J Gastroenterol* 2009;104:816–24.
- Bao C, Tournet RW, Brunette GJ, *et al.* Genomic signatures of past and present Chromosomal instability in Barrett's esophagus and early Esophageal adenocarcinoma. *Nat Commun* 2023;14:6203.
- Lavery DL, Martinez P, Gay LJ, *et al.* Evolution of oesophageal adenocarcinoma from metaplastic columnar epithelium without goblet cells in Barrett's oesophagus. *Gut* 2016;65:907–13.
- Katz-Summercorn AC, Jammula S, Frangou A, *et al.* Multi-omic cross-sectional cohort study of pre-malignant Barrett's esophagus reveals early structural variation and retrotransposon activity. *Nat Commun* 2022;13:1407.
- Van der Auwerda GA, O'Connor BD. *Genomics in the Cloud: Using Docker, GATK, and WDL in Terra (1st Edition)*. O'Reilly Media, 2020.
- Ross-Innes CS, Becq J, Warren A, *et al.* Whole-genome sequencing provides new insights into the clonal architecture of Barrett's esophagus and esophageal adenocarcinoma. *Nat Genet* 2015;47:1038–46.
- Weaver JMJ, Ross-Innes CS, Shannon N, *et al.* Ordering of mutations in preinvasive disease stages of esophageal carcinogenesis. *Nat Genet* 2014;46:837–43.

- 35 Paulson TG, Galipeau PC, Oman KM, *et al.* Somatic whole genome Dynamics of Precancer in Barrett's esophagus reveals features associated with disease progression. *Nat Commun* 2022;13:2300.
- 36 Roth A, Khattra J, Yap D, *et al.* Pyclone: statistical inference of clonal population structure in cancer. *Nat Methods* 2014;11:396–8.
- 37 Dang HX, White BS, Foltz SM, *et al.* ClonEvol: clonal ordering and visualization in cancer sequencing. *Ann Oncol* 2017;28:3076–82.
- 38 von Elm E, Altman DG, Egger M, *et al.* The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *BMJ* 2007;335:806–8.
- 39 Evans JA, Carlotti E, Lin M-L, *et al.* Clonal transitions and Phenotypic evolution in Barrett's esophagus. *Gastroenterology* 2022;162:1197–209.
- 40 Januszewicz W, Subhash VV, Waldock W, *et al.* The utility of a methylation panel in the assessment of clinical response to radiofrequency ablation for Barrett's esophagus. *EBioMedicine* 2020;58:102877.
- 41 Berman AG, Tan WK, O'Donovan M, *et al.* Quantification of Tff3 expression from a non-endoscopic device predicts clinically relevant Barrett's Oesophagus by machine learning. *EBioMedicine* 2022;82:104160.
- 42 Hamade N, Vennelaganti S, Parasa S, *et al.* Lower annual rate of progression of short-segment vs long-segment Barrett's esophagus to Esophageal adenocarcinoma. *Clin Gastroenterol Hepatol* 2019;17:864–8.

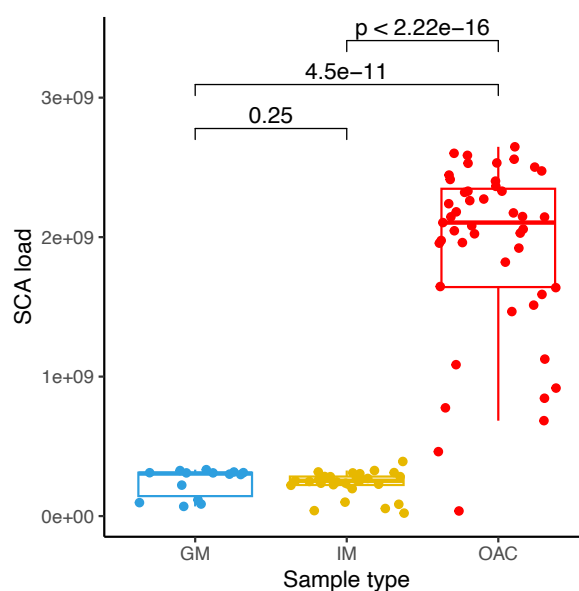




**Supplementary Figure 1:** Representative histopathology of GM-only, GM+IM and IM-only columnar lined oesophageal epithelium. IM is characterised by the presence of goblet cells. Example areas of GM and focal IM are indicated on the GM+IM image. All three main images are shown at the same level of magnification. GM, gastric metaplasia; IM, intestinal metaplasia.

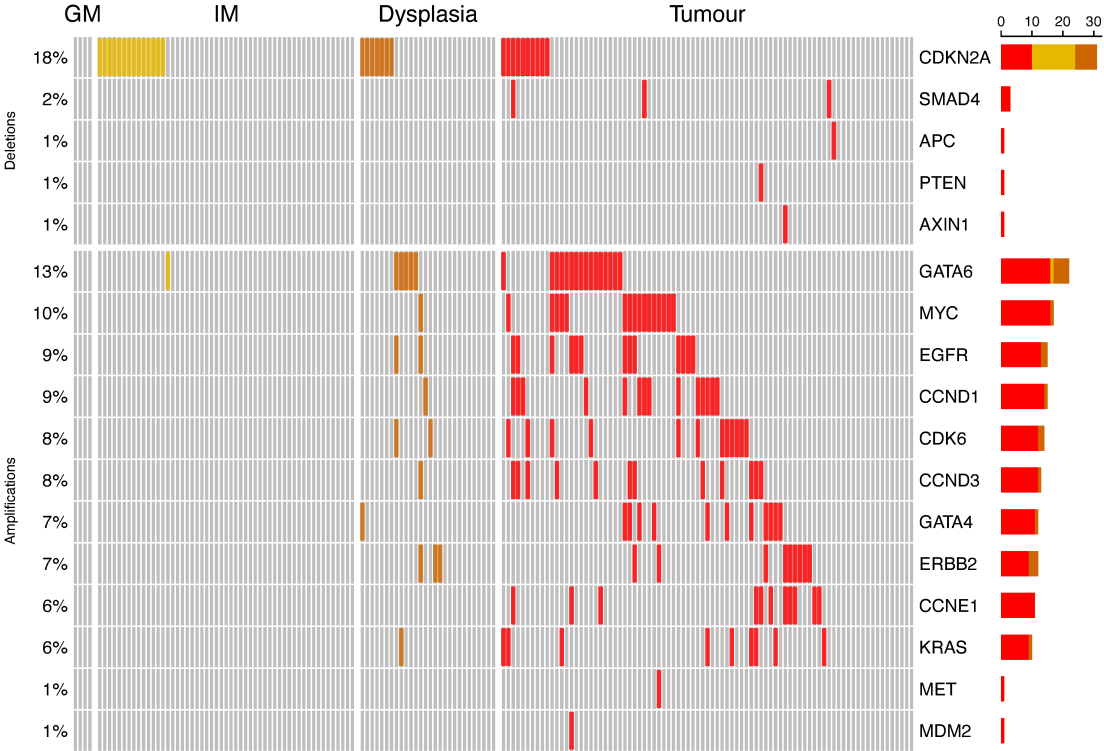


**Supplementary Figure 2:** Variation in length in subsequent OGDs. Proportion of cases with a) shorter or b) longer lengths than index-length reported in subsequent OGDs. It is possible for a case to have a shorter length reported in one follow-up and a longer length reported in a different follow-up, so the two are not mutually exclusive. c) Surveillance start date for segments with shorter lengths or irregular z-line reported, compared with those with no shorter length reported. GM, gastric metaplasia; IM, intestinal metaplasia; OGD, oesophagogastroduodenoscopy.

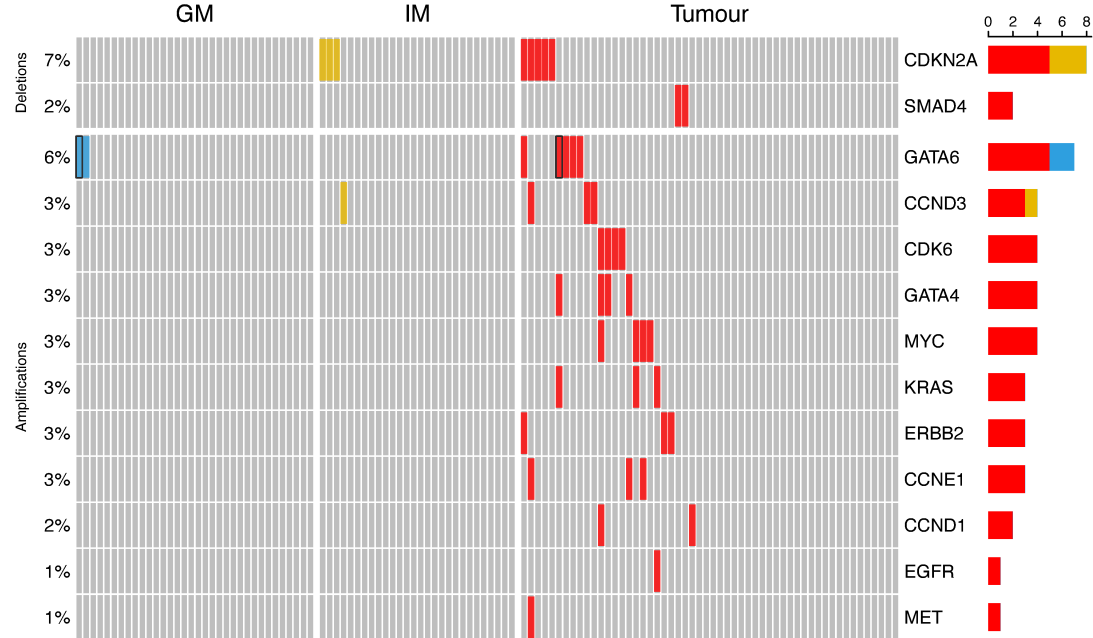


**Supplementary Figure 3:** SCA load comparison between OAC, GM and IM. SCA load taken as the length of genome altered by a copy gain, loss or copy neutral loss of heterozygosity; GM, gastric metaplasia; IM, intestinal metaplasia; OAC, oesophageal adenocarcinoma; SCA, somatic chromosomal alteration.

a) Copy number events in driver genes, WGS samples

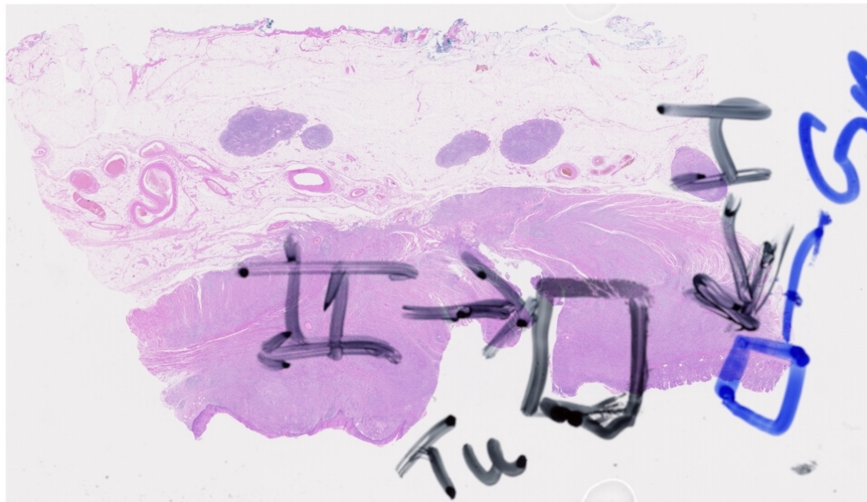
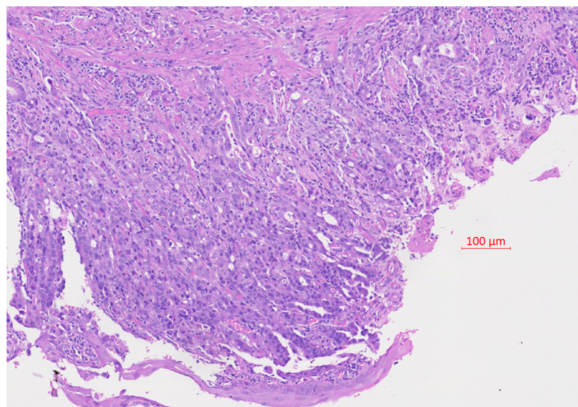
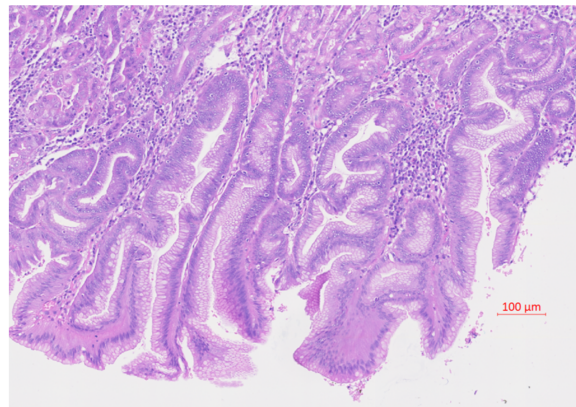


b) Copy number events in driver genes, WES samples



**Supplementary Figure 4:** Amplifications and deletions in copy number driver genes, in a) WGS samples and b) WES samples. Shared event between GM and Tumour is highlighted with a black outline. WGS, whole genome sequencing; WES, whole exome sequencing; GM, gastric metaplasia; IM, intestinal metaplasia.



**a) Whole slide****b) Part of Tumour (Black - Tu) region****c) Part of gastric metaplasia (Blue – GM) region**

**Supplementary Figure 5:** Example of distinct tissue types identified by pathologists, which were then macro-dissected for sequencing. GM area is marked by blue box labelled 'GM' and 'I', tumour area is marked by black box labelled 'Tu' and 'II'. GM, gastric metaplasia.

Criterion type	Criterion	Starting number	# failing criterion	# after criterion applied
Pathology	Must have at least two complete pathology records	5,625	3,744	1,881
Pathology	Pathology records must cover more than a year	1,881	401	1,480
Pathology	First pathology record must be GM or IM, or IND with subsequent GM	1,480	697	783
Pathology	First pathology record must clearly confirm and distinguish GM and IM	783	39	744
Pathology	No change in diagnosis within first year	744	3	741
Pathology	Not missing most recent pathology report	741	1	740
Endoscopy	Length data available from at least one report	740	20	720
Endoscopy	Length data available from a report matching index pathology	720	40	680
Endoscopy	Index length SS GM/IM or LS GM, including not z-line	680	421	259
Medical	No other cancer	259	9	250
Medical	No prior treatment for BO/OAC	250	5	245
Final checks	Length data conceivably consistent between endoscopies	245	1	244

**Supplementary Table S1:** Inclusion criteria for clinical outcomes cohort. GM, gastric metaplasia; IM, intestinal metaplasia; SS, short segment; LS, long segment; IND, indefinite for dysplasia; BO, Barrett's oesophagus; OAC, oesophageal adenocarcinoma.

PatientID	BOType	Sex	Age	TStage	NStage	MStage	IM present (dysplasia)
P1	GM	F	59	T3	N0	M0	
P2	GM	M	78	T1b	N2	Mx	
P3	GM	F	69	T3	N2	M0	
P4	GM	M	66	T1a	N0	M0	
P5	GM	F	82	T1b	N0	M0	
P6	GM	F	81	T3	N2	M0	
P7	GM	M	78	T1a	N0	M0	
P8	GM	M	80	T2	N0	M0	
P9	GM	M	80	T1b	N0	M0	
P10	GM	M	71	T3	N1	M0	
P11	GM	F	73	T3	N0	M0	
P12	GM	M	72	T3	N3	M0	
P13	GM	M	65	T4a	N3	M0	
P14	GM	M	75	T3	N3	M0	
P15	GM	M	77	T3	N2	M0	
P16	GM	F	80	T3	N1	M0	
P17	GM	M	58	T2	N1	Mx	
P18	GM	M	77	T3	N1	Mx	
P19	GM	M	58	T2	N0	M0	
P20	GM	M	52	T1b	N0	M0	
P21	GM	M	78	T3	N1	M0	
P22	GM	M	63	T3	N2	M0	
P23	GM	M	67	T2	N0	M0	
P24	GM	M	74	T2	N0	M0	
P25	GM	M	72	T3	N3	M1	
P26	GM	M	76	T1a	N0	Mx	
P27	GM	M	56	T0	N0	M0	
P28	GM	M	57	T0	N0	M0	
P29	GM	M	68	T1a	N0	Mx	
P30	GM	F	86	T2	N1	M0	
P31	GM, IM	M	78	T3	N2	M0	
P32	GM, IM	M	74	T1a	N1	Mx	
P33	GM, IM	M	83	T2	N1	Mx	
P34	GM, IM	M	82	T2	N0	Mx	
P35	GM, IM	M	73	T2	N0	M0	
P36	GM, IM	M	62	T2	N0	M0	
P37	GM, IM	M	79	T3	N2	M0	
P38	GM, IM	M	70	T2	N1	Mx	
P39	HGD	M	50	T3	N3	Mx	Not recorded
P40	HGD	M	67	T1	N0	Mx	Not recorded
P41	HGD	M	76	T2	N1	M0	Yes
P42	HGD	M	69	T1b	N0	Mx	Yes
P43	HGD	F	66	T3	N1	Mx	Yes
P44	HGD	M	74	T3	N1	M0	Yes

PatientID	BOType	Sex	Age	TStage	NStage	MStage	IM present (dysplasia)
P45	HGD	M	66	T3	N0	Mx	Yes
P46	HGD	M	60	T1a	N0	M0	Yes
P47	HGD	M	74	T1b	N0	Mx	Not recorded
P48	HGD	M	87	T3	N1	M0	Not recorded
P49	HGD	M	67	T3	N1	M0	Not recorded
P50	IM	M	71	T1a	N0	M0	
P51	IM	M	72	T1b	N0	M0	
P52	IM	F	52	T1b	N0	M0	
P52	LGD	F	52	T1b	N0	M0	Yes
P53	IM	M	80	T2	N0	M0	
P54	IM	M	82	T4a	N3	M0	
P55	IM	M	70	T1	N0	M0	
P56	IM	M	65	T3	N1	M0	
P57	IM	M	82	T3	N3	Mx	
P58	IM	M	83	T3	N2	Mx	
P59	IM	M	56	T1b	N0	Mx	
P60	IM	M	Unknown	T2	N0	M0	
P61	IM	M	85	T1b	N0	Mx	
P62	IM	M	54	T1a	N0	M0	
P63	IM	M	83	T2	N0	M0	
P64	IM	M	76	T2	N0	M0	
P65	IM	M	69	T1b	N1	M0	
P66	IM	M	74	T2	N0	Mx	
P67	IM	M	79	T1b	N0	M0	
P68	IM	F	80	T3	N0	M0	
P69	IM	M	77	T3	N2	Mx	
P70	IM	F	80	T1b	N2	M0	
P71	IM	F	79	T2	N1	M0	
P72	IM	M	74	T3	N0	Mx	
P73	IM	M	65	T1	N0	Mx	
P74	IM	M	69	T1b	N0	Mx	
P75	IM	M	68	T3	N3	M0	
P76	IM	M	84	T3	N1	M1	
P77	IM	F	50	T1b	N1	Mx	
P78	IM	F	62	T3	N0	M0	
P79	IM	M	58	T4a	N3	M1	
P80	IM	M	66	T3	N1	Mx	
P81	IM	M	76	T3	N3	M1	
P82	IM	M	73	T3	N3	M1	
P83	IM	M	62	T3	N3	M0	
P84	IM	M	67	T3	N2	M0	
P85	IM	M	61	T2	N1	M1	
P86	IM	M	79	T3	N1	M0	
P87	IM	M	66	T2	N0	Mx	



PatientID	BOType	Sex	Age	TStage	NStage	MStage	IM present (dysplasia)
P88	IM	M	82	T3	N0	Mx	
P89	IM	M	59	T1b	N0	M0	
P90	IM	F	58	T1a	N0	M0	
P91	IM	M	68	T1b	N0	Mx	
P92	IM	F	56	T2	N1	Mx	
P93	IM	M	61	T2	N1	Mx	
P94	IM	M	83	T3	N2	Mx	
P95	IM	M	57	T3	N0	Mx	
P96	IM	M	71	T1a	N0	Mx	
P97	IM	M	61	T3	N2	Mx	
P98	IM	M	66	T3	N1	M0	
P99	IM	M	66	T1	N1	Mx	
P100	IM	M	58	T3	N2	M0	
P101	IM	M	67	T3	N2	M1	
P102	IM	M	63	T1	N0	Mx	
P103	IM	M	75	T3	N3	Mx	
P104	IM	M	75	T1a	N0	Mx	
P105	IM	M	72	T1b	N0	M0	
P106	IM	M	77	T2	N0	M0	
P107	IM	M	74	T3	N3	M0	
P108	IM	M	67	T3	N0	M0	
P109	IM	M	71	T3	N1	M0	
P110	IM	M	76	T2	N0	Mx	
P111	IM	M	58	T2	N1	M0	
P112	IM	M	71	T2	N1	M0	
P113	IM	M	85	T2	N0	M0	
P114	IM	M	69	T1b	N1	Mx	
P115	IM	M	82	T3	N0	M0	
P116	IM	M	59	T3	N1	Mx	
P117	IM	M	67	T2	N0	Mx	
P118	IM	M	61	T1b	N1	Mx	
P119	IM	M	67	Tx	N2	M1	
P120	IM	M	56	T3	N2	Mx	
P121	IM	M	83	T1b	N0	Mx	
P122	IM	M	54	T2	N1	M0	
P123	LGD	M	65	Tx	Nx	Mx	Yes
P124	LGD	M	62	T3	N2	M1	Not recorded
P125	LGD	M	79	T1b	N0	Mx	Yes
P126	LGD	M	83	T3	N2	Mx	Yes
P127	LGD	M	76	T1b	N0	M0	Yes
P128	LGD	F	58	T3	N0	M0	No
P129	LGD	M	73	T3	N1	M0	Yes
P130	LGD	M	64	T1a	N0	M0	Yes
P131	LGD	M	64	T0	N0	M0	No

PatientID	BOType	Sex	Age	TStage	NStage	MStage	IM present (dysplasia)
P132	LGD	F	84	T3	N1	M0	No
P133	LGD	M	73	T1b	N0	M0	Not recorded
P134	LGD	M	56	T1b	N0	M0	Not recorded
P135	LGD	M	80	T3	N3	M1	Not recorded
P136	LGD	M	79	T3	N0	M0	Not recorded
P137	LGD	M	63	T2	N0	Mx	Not recorded
P138	LGD	M	72	T3	N3	M0	Not recorded

**Supplementary Table S2:** Characteristics of patients with BO adjacent to OAC. Age is age at diagnosis. BO, Barrett's oesophagus; OAC, oesophageal adenocarcinoma; GM, gastric metaplasia; IM, intestinal metaplasia; LGD, low grade dysplasia; HGD, high grade dysplasia.

All biopsy results	Total	Focal IM	Widespread IM	Unspecified IM
GM+IM (including progressors)	77	51	5	21
IM only (including progressors)	100	39	25	36
Chi-squared test		$p = 0.0003283$		

**Supplementary Table S3:** Incidence of reporting of focal and widespread IM across the GM+IM and IM only cases. GM, gastric metaplasia; IM, intestinal metaplasia.

Index	Outcome	Total surveillance time		Time to first IM OGD	
		Mean (years)	Median (years)	Mean (years)	Median (years)
GM	GM-only	4.60	4.00	--	--
GM	GM+IM	9.75	10.26	3.86	3.09

**Supplementary Table S4:** Total surveillance time for all index-GM cases, and time to first IM OGD for index-GM, GM+IM outcome cases. GM, gastric metaplasia; IM, intestinal metaplasia; OGD, oesophagogastroduodenoscopy.

## SUPPLEMENTARY METHODS

### DNA extraction for WES samples

FFPE slides were deparaffinised and samples then macro-dissected using a needle, to maximise cellularity. Areas measuring approximately 0.5x0.5cm were extracted. An extraction was made of each type of tissue present in sufficient size (OAC, IM or GM), with the IM and GM having to be adjacent to the OAC. Qiagen's Gene Read DNA FFPE kit was used. Blood germline controls were used for all but one case. For this one exception case, the germline control was snap-frozen, normal tissue taken distant from the tumour.

For most blood samples, DNA was extracted using the QIAamp Blood Maxi Kit (Qiagen). The following adaptations were made to the manufacturer's instructions: spin speeds were changed to 2400 rpm for steps 6, 7, 9-11, and 13; spinning times were changed to 6 min, 2 min, 30 min, 4 min, and 10 min, for steps 6,7, 9-11 and 13, respectively; samples were eluted in buffer AE. For a small number of blood samples, DNA was extracted from buffy coats using the QIAamp DNA Mini Kit (Qiagen), with no modifications to the manufacturer's instructions.

### WES mutation filtering

Mutations called in the WES samples had to: fall within exonic regions or splicing regions of coding genes; not present as SNPs in 0.01% or more of the European and World populations database [Supp. Ref 1]; have at least 5 supporting reads; and variant allele frequency of at least 0.05. Additional filtering was performed to address artefact and quality issues associated with FFPE. Specifically, the average fragment length associated with a mutation had to be at least 85bp. If a mutation was present in at least 5 patients (5% of the cohort), the median distance of the mutation from the end of the read had to have a median average deviation of more than 1, and the active region the mutation occurred in had to have a median number of candidate variants of no more than 1.



### WES copy number calling and copy number analysis

Copy number calling was performed using Sequenza [Supp. Ref 2], first using the ‘full’ breaks method, and then re-run using a minimum consistent segmentation across the samples from the same case. Amplifications and deletions were analysed in subsets of genes recurrently amplified (*KRAS*, *MYC*, *ERBB2*, *MET*, *GATA4*, *CCND1*, *GATA6*, *CDK6*, *EGFR*, *CCNE1*, *CCND3*, *MDM2*, *PPM1D*) or deleted (*CDKN2A*, *PTEN*, *SMAD4*, *AXIN1*, *APC*) in OAC. To call an amplification, the total copy number had to be  $\geq 5$  for samples with no WGD and  $\geq 9$  for samples with WGD. WGD was called as per the PCAWG approach [Supp. Ref 3]. Deletions had total copy number of 0. To minimise false positives, amplifications or deletions had to also be called by VarScan2 [Supp. Ref 4] and amplifications had to have a segment length of at least 100kb.

### SUPPLEMENTARY REFERENCES

- [1] The 1000 Genomes Project Consortium. A global reference for human genetic variation. *Nature* 526:68–74, 2015. doi: 10.1038/nature15393.
- [2] Favero F, Joshi T, Marquard AM et al. Sequenza: allele-specific copy number and mutation profiles from tumor sequencing data. *Annals of Oncology*, 26(1):64-70, 2015. doi: 10.1093/annonc/mdu479.
- [3] D'Antonio SC, Leshchiner I, Haase K et al. Characterizing genetic intra-tumor heterogeneity across 2,658 human cancer genomes. *Cell*, 184(8):2239-2254.e39, 2021. doi: 10.1016/j.cell.2021.03.009.
- [4] Koboldt DC, Zhang Q, Larson DE et al. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome research* 22(3):568-76, 2012. doi:10.1101/gr.129684.111