A Decade of the Oesophageal Cancer Clinical and Molecular Stratification Consortium: OCCAMS

C.J. Peters[1], Y. Ang[2], F.D. Ciccarelli[3], H. Coles[4], H.G. Coleman[5], G. Contino[6], T. Crosby[7], G. Devonshire[8], M Eldridge[8], A. Freeman[4], N. Grehan[4], M. McCord[9], B. Nutzinger[4], S Zamani[4], S.L. Parsons[10], R. Petty[11], A.D. Sharrocks[12], R.J.E. Skipworth[13], E.C. Smyth[14], I. Soomro[15], T.J. Underwood[16], OCCAMS Consortium, R.C Fitzgerald[4].

Names of other OCCAMS consortium members not included in authorship.

Paul A.W. Edwards[4,8] , Aisling M Redmond[4], Christine Loreno[4], Sujath Abbas[4], Maria O'Donovan[4], Ahmad Miremadi[4], Shalini Malhotra[17], Monika Tripathi[17], Calvin Cheah[17], Curtis Millington[4], Maria Secrier[8], Sriganesh Jammula[8], Jim Davies[17], Charles Crichton[17], Nick Carroll[18], Richard H.Hardwick[18], Peter Safranek[18], Andrew Hindmarsh[18], Vijayendran Sujendran[18], Stephen J. Hayes[19], Shaun R. Preston[20], Izhar Bagwan[20], Vicki Save[21], Ted R. Hupp[22], J. Robert O'Neill[18], Olga Tucker[23], Andrew Beggs[24], Philippe Taniere[23], Sonia Puig[23], Robert C. Walker[16], Ben L. Grace[25], Jesper Lagergren[26], James A Gossage[27], Andrew R Davies[27], Fuju Chang[27], Ula Mahadeva[28], Vicky Goh[27], Grant Sanders[29], Richard Berrisford[29], David Chan[29], Ed Cheong[30], Bhaskar Kumar[30], L. Sreedharan[30] , Philip Kaye[10], John H Saunders[10], Laurence Lovat[31], Rehan Haidry[31], Michael Scott[32], Sharmila Sothi[33], Suzy Lishman[8], George B. Hanna[1],Krishna Moorthy[1], Anna Grabowska[34], Richard C Turkington[35], Damian McManus[35], Freddie Bartlett[36]

1. Department of Surgery and Cancer, Imperial College London, London, UK
2. Division of Diabetes, Endocrinology and Gastroenterology, University of Manchester, Manchester, UK
3. Comprehensive Cancer Centre, Kings College London, London, UK
4. Early Cancer Institute, University of Cambridge, Cambridge, UK
5. Centre for Public Health, Queen's University Belfast, Belfast, UK.
6. Institute of Cancer and Genomic Sciences, University of Birmingham, Birmingham, UK.
7. Velindre University NHS Trust, Cardiff, UK.
8. Cancer Research UK Cambridge Institute, Cambridge, UK.
9. Heartburn Cancer UK, UK.
10. Nottingham University Hospitals NHS Trust, Nottingham, UK
11. School of Medicine, University of Dundee, Dundee, UK
12. Division of Molecular and Cellular Function, University of Manchester, Manchester, UK
13. Clinical Surgery, University of Edinburgh, Edinburgh, UK
14. Oxford University Hospitals NHS Foundation Trust, Oxford, UK.
15. Nottingham University Hospital, Nottingham, UK
16. Institute for Life Sciences, University of Southampton, Southampton, UK
17. Department of Computer Science, University of Oxford, UK
18. Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK
19. Faculty of Medical and Human Sciences, University of Manchester, UK
20. Faculty of Medical and Human Sciences, University of Manchester, UK
21. Edinburgh Royal Infirmary, Edinburgh, UK
22. Edinburgh University, Edinburgh, UK

23. University Hospitals Birmingham NHS Foundation Trust, Birmingham, UK
24. Institute of Cancer and Genomic sciences, University of Birmingham
25. University Hospital Southampton NHS Foundation Trust, Southampton, UK
26. Karolinska Institute, Stockholm, Sweden
27. King's College London, London, UK
28. Guy's and St Thomas's NHS Foundation Trust, London, UK
29. Plymouth Hospitals NHS Trust, Plymouth, UK
30. Norfolk and Norwich University Hospital NHS Foundation Trust, Norwich, UK
31. University College London, London, UK
32. Wythenshawe Hospital, Manchester, UK
33. University Hospitals Coventry and Warwickshire NHS, Trust, Coventry, UK
34. Queen's Medical Centre, University of Nottingham, Nottingham, UK
35. Centre for Cancer Research and Cell Biology, Queen's University Belfast, UK
36. Portsmouth Hospitals NHS Trust, Portsmouth, UK

Multi-centre collaboration is essential to achieve the sample sizes required for robust and informative research studies for less common medical conditions. Substantial logistical and governance support is needed to ensure that the clinical and molecular data generated are high quality and can benefit the international research community and patients.

The Oesophageal Cancer Clinical and Molecular Stratification (OCCAMS) Consortium was created in 2009 as a collaboration across the United Kingdom to better understand oesophageal adenocarcinoma (OAC). The aims of OCCAMS are: to develop a bioresource of samples with clinical data from oesophageal and gastro-oesophageal junction (GOJ) adenocarcinoma patients; to identify clinical, demographic, and molecular factors affecting development and progression of OAC; to promote use of these data and the OCCAMS network for clinical trials to improve management of this cancer; and to share data with internal and external academic and commercial parties for the benefit of patients.

At the time of writing, OCCAMS has collected and curated a bioresource derived from 4,440 oesophageal cancer patients, representing over 44,000 individual samples with detailed clinical and epidemiological annotations, from 27 UK centres (Figure 1). OCCAMS was a key contributor to the International Cancer Genome Consortium (ICGC), ICGC 25K and ICGC ARGO projects, as well as the Pan-Cancer Analysis of Whole Genomes (PCAWG) study. OCCAMS has also contributed to the Cancer Research UK Grand Challenge Mutographs project and projects run by Genomics England.

We have continued to extend the OCCAMS genomic resource and are increasingly complementing this with other -omics data including single cell technologies and 3D organoids.

The OCCAMS Steering Committee was created at inception and consists of members of the core OCCAMS infrastructure team and representatives from every site that contributes to the consortium. Proposals for new research projects are formally submitted using a proforma for

discussion at regular meetings of the steering committee. This structure allowed OCCAMS to control the use of finite resources, such as blood and tissue, whilst encouraging researchers to apply to use the resource. It also allowed similar projects to be combined, or aligned, to reduce academic wastage and maximise the quality of the scientific output. Bespoke cohorts can be created for a specific project via the Cambridge-based OCCAMS team, and clinical and -omics data are provided so that specific research can be conducted. Sequencing reads and methylation array data have been made available to the wider research community via the International Cancer Genome Consortium and/or the European Genome-phenome Archive (EGA), to take advantage of existing governance structures.

Molecular and clinical data, as well as patient-derived organoids, from OCCAMS have been extensively exploited to understand various aspects of OAC biology within and outside the Consortium. In total, we have established 33 collaborations within OCCAMS and 14 with external groups, including European and US partners from academia and industry, resulting in 34 research publications. There are further studies still underway, which will provide further research outputs for this cancer, which has poor outcomes. Some scientific outputs of OCCAMS to date are described in table 1.

Our key overall lesson drawn from OCCAMS is that multicentre consortia can collect data and tissue (including fresh frozen) from huge numbers of patients, even when studying a cancer like OAC, which has a relatively low incidence. However, care must be taken to create robust and detailed databases, with metadata, to make sure that the clinical information being recorded is accurate and relevant to the molecular questions being asked. Accurate patient outcomes should be linked to national registries. Ethical approval should encompass all the data and tissue required and include agreement from the patient to track patient outcomes across relevant national registries as well as to share data and resources with academic collaborators. During the course of OCCAMS we amended the consent to allow data sharing with commercial collaborators with a specific tick box so that patients can opt in or out. This was included to ensure maximal opportunity for clinical translation of the findings; the majority of patients have consented for commercial as well as academic collaboration.

We allowed flexibility as to which samples each centre collected, for example whether samples were fresh frozen or formalin-fixed paraffin-embedded, or whether ctDNA was double spun on site or collected in Streck tubes. This allowed more centres to get involved, which often led to their infrastructure being expanded to collect a wider range of sample types. This encouraged inclusivity of OCCAMS membership, maximising buy in from the UK clinical community. Clear, standardised protocols were needed to reduce sample wastage and maximise quality.

As the number of molecular analyses using OCCAMS samples increased there was careful consideration of how all the data, which included demographic, clinical, whole genome sequencing, and RNAseq, could be linked. Linkage increases the value of this data and enhances the ability to generate new insights into disease.

The steering group coordinated the academic output of this resource, driving forward research. A sense of community was created by holding yearly in-person meetings to share new findings and discuss future projects, with competitions to encourage junior clinical

researchers to access the data. The goal was to create a sense of ownership for the clinical researchers who contributed patient data to OCCAMS, as this encouraged use of the data for molecular projects and for understanding the clinical features of OAC.

A multicentre collaboration must focus on the patient. Throughout this project we had strong patient representation to ensure that our research focus was relevant and that all materials were clear and consistent. This included discussion between patients and researchers about research priorities as well as co-production of patient facing materials and patient led sessions at our annual symposium. The exchange between patients, public and researchers, especially those new to the field, led to a heightened appreciation of why research in this disease is so important. New discoveries will lead to improved patient outcomes only when the potential clinical impact of the science is constantly borne in mind.

Competing interests

Grant funding for OCCAMS infrastructure was provided by Cancer Research UK. Project specific funding for commercial assay technologies applied to OCCAMS samples was provided by Roche and Natera. R.F. is a shareholder in Cyted Ltd an early diagnosis company.

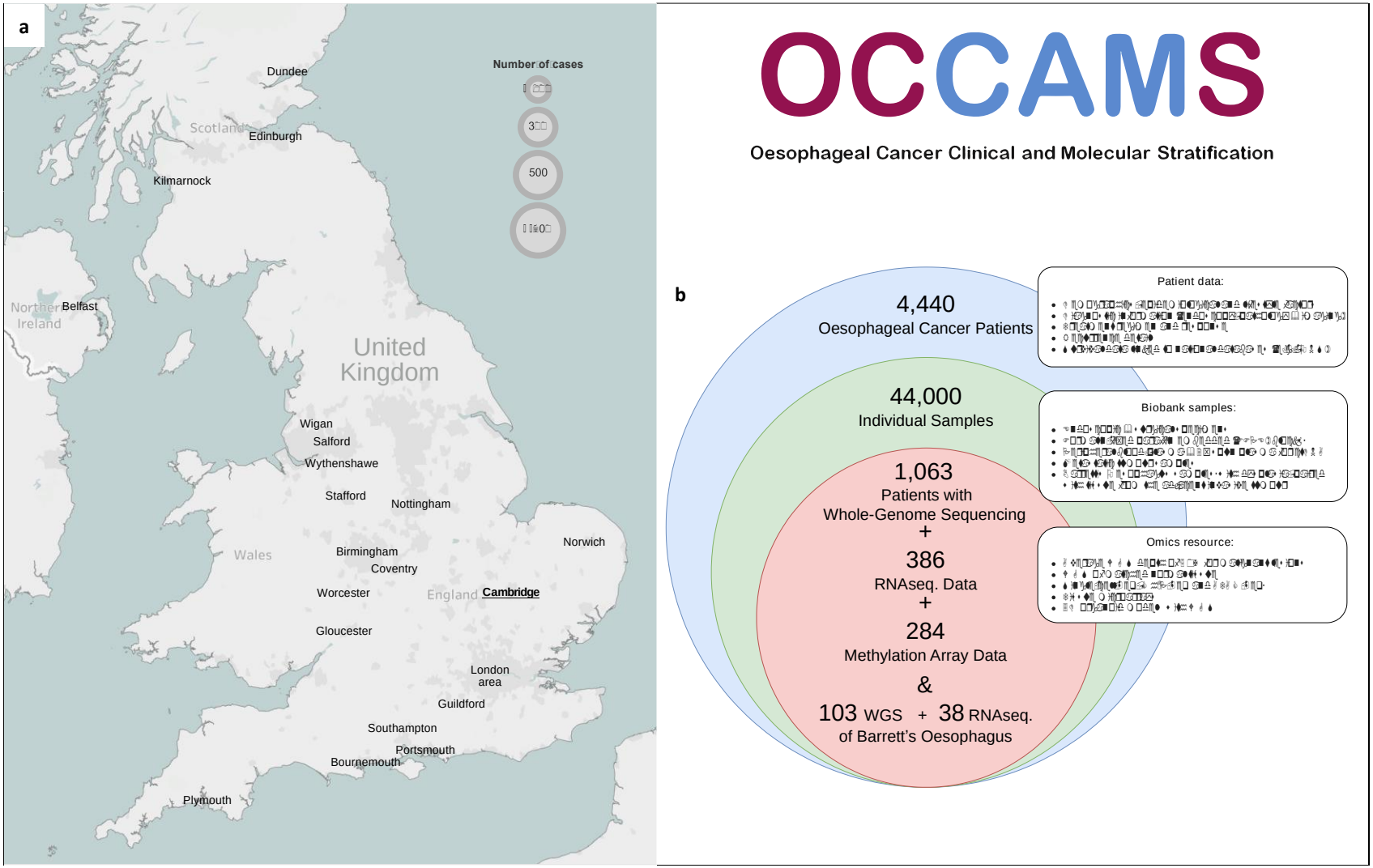

Figure 1: Number of OCCAMS patients recruited, samples collected and molecular data and organoids available.

Study	Data types used	Key findings	Reference
Mutational signatures in esophageal adenocarcinoma define etiologically distinct subgroups with therapeutic relevance	129 WGS with matched germline reference (blood or normal tissue).	Esophageal adenocarcinoma is dominated by copy number changes and frequent large-scale rearrangements. Three distinct molecular groups with therapeutic relevance were identified: Enriched for BRCA signature with defects in homologous recombination pathway, T>G mutation dominant with high mutational load, and C>A/T mutation dominant with evidence of ageing imprint.	1
The landscape of selection in 551 esophageal adenocarcinomas defines genomic	WGS on 551 samples with matched RNA sequencing data in a subset.	4.4 driver events per cancer showed mutual exclusivity or co-occurrence between pathways, suggesting strong functional relationships. New non-coding driver elements identified and SMAD4 and GATA4 shown to be indicators of poor prognosis. 50% of cancers contained sensitizing events for CDK4 and CDK6 inhibitors, highly correlated with data from our organoid panel.	2
Organoid cultures recapitulate esophageal adenocarcinoma heterogeneity providing a model for clonality studies and precision therapeutics	Organoid cultures with matched sequencing from OCCAMS derived primary esophageal adenocarcinoma.	Organoids are able to recapitulate the morphology, copy number changes and mutational signatures of primary esophageal adenocarcinoma. Clonal evolution could be tracked over time and subsequent clonal selection was associated with driver gene status.	3
Patient-specific cancer genes contribute to recurrently perturbed pathways and establish therapeutic vulnerabilities in esophageal adenocarcinoma	WGS data from 261 samples with confirmation of findings in another 107 additional cases.	Identification of helper genes which work alongside well-known drivers to promote cancer. These helpers are rare or patient specific but converge towards perturbing well known cancer processes and help cluster OACs in six groups with differing molecular and clinical features.	4
Machine learning to predict early recurrence after oesophageal cancer surgery	Clinical characteristics and post operative histology from 812 patients undergoing surgery for Oesophageal cancer.	A machine learning approach derived a model that could accurately predict early recurrence after surgery for esophageal adenocarcinoma. Number of positive lymph nodes and lymphovascular invasion were the most important variables.	5

Repurposing of KLF5 activates a cell cycle signature during the progression from a precursor state to esophageal adenocarcinoma	Integration of Gene expression and chromatin accessibility profiles in esophageal adenocarcinoma and Barrett's Oesophagus	The transcription factor KLF5 is implicated in the transition from Barrett's to Adenocarcinoma but its level is unchanged and instead it is redistributed across chromatin. The KLF5 regulome represents a set of new targets with prognostic significance as higher levels equate to worse outcomes.	6
Pan-cancer analysis of whole genomes	WGS across 38 tumour types. OCCAMS contributed the esophageal adenocarcinoma cohort WGS and benchmarking for the pipelines.	Cancer genomes contain 4-5 driver mutations on average. Esophageal adenocarcinoma were ranked 3rd for the number of single nucleotide variants with a high level of structural variants and highest for retrotranspositions.	7
Molecular phenotyping reveals the identity of Barrett's esophagus and its malignant transition	Single cell approach to investigate the cell of origin of Barrett's.	Analysis of healthy oesophageal tissue, mutational lineage tracking and organoid models suggested that Barrett's oesophagus arises from the gastric cardia and esophageal adenocarcinoma arises from undifferentiated Barrett's Oesophagus cells.	8
Longitudinal tracking of 97 esophageal adenocarcinoma using liquid biopsy sampling	Analysis of ctDNA in 245 blood samples from 97 patients through their cancer treatment pathway.	The presence of ctDNA in the plasma following surgery for esophageal adenocarcinoma is prognostic for relapse and can stratify patients into high and low risk groups for intensification or de-escalation of adjuvant chemotherapy.	9
Extrachromosomal DNA in the cancerous transformation of Barrett's oesophagus	WGS including 206 esophageal adenocarcinoma and Barrett's. Focus on extrachromosomal DNA (ecDNA) oncogene amplification.	Frequency of extrachromosomal DNA increased between early (25%), including pre-invasive disease, and late stage esophageal adenocarcinoma (43%). Extrachromosomal DNA progressively evolves under positive selection.	10

Table 1: 10 exemplar studies arising from the OCCAMS collaboration highlight the breadth of data types and insights obtained.

References

1. Secrier, M.*, et al. Nat Genet* **48**, 1131-1141 (2016).
2. Frankell, A.M.*, et al. Nat Genet* **51**, 506-516 (2019).
3. Li, X.*, et al. Nat Commun* **9**, 2983 (2018).
4. Mourikis, T.P.*, et al. Nat Commun* **10**, 3101 (2019).
5. Rahman, S.A.*, et al. Br J Surg* **107**, 1042-1052 (2020).
6. Rogerson, C.*, et al. Elife* **9**(2020).
7. Consortium, I.T.P.-C.A.o.W.G. *Nature* **578**, 82-93 (2020).
8. Nowicki-Osuch, K.*, et al. Science* **373**, 760-767 (2021).
9. Ococks, E.*, et al. Ann Oncol* **32**, 522-532 (2021).
10. Luebeck, J.*, et al. Nature* **616**, 798-805 (2023).