CANCER RESEARCH HORIZONS

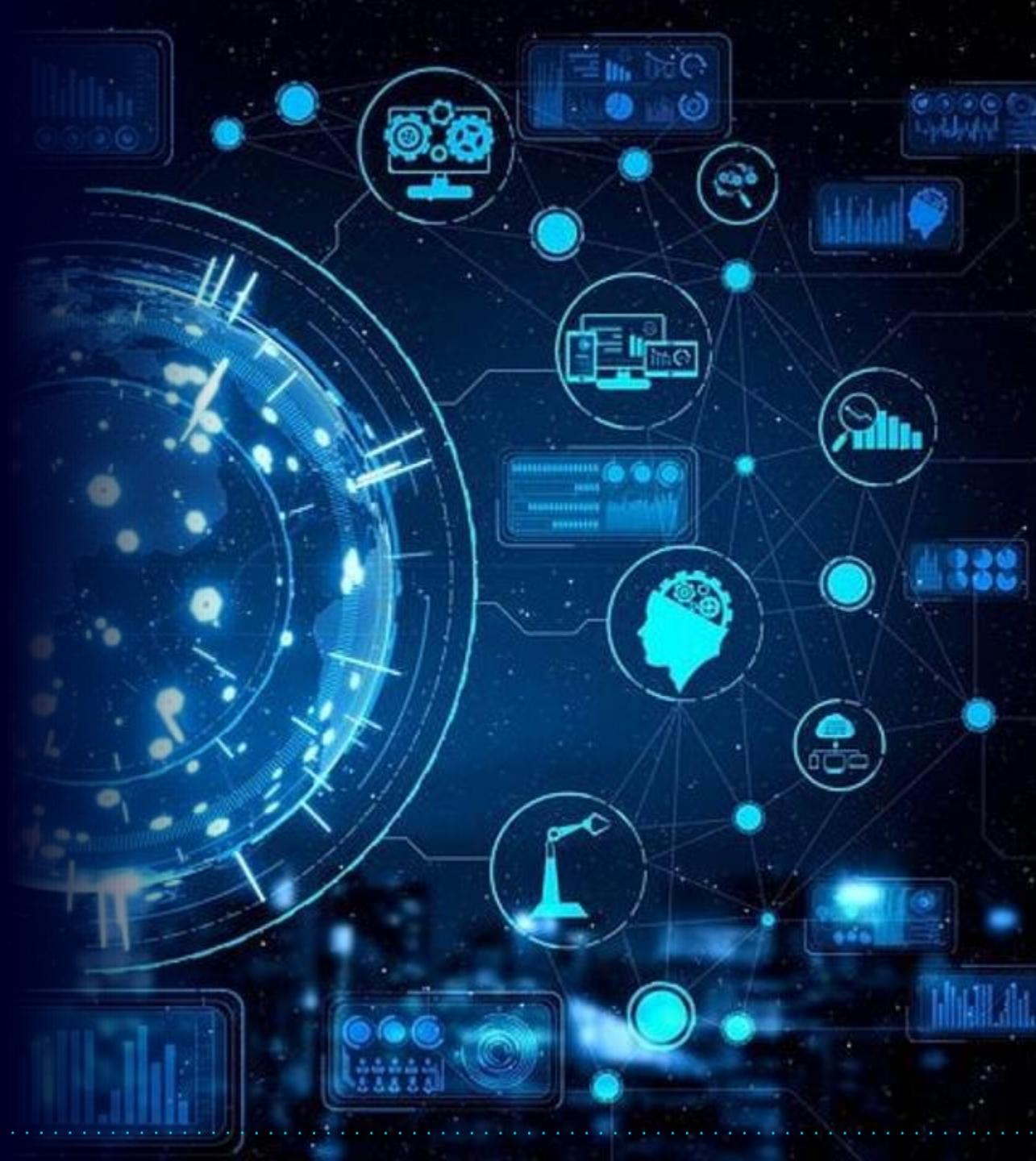
FURTHER FASTER TOGETHER

S:CORT

Stratification in COloRecTal cancer

S:CORT DATABASE

A clinically linked multi-omic dataset in colorectal cancer



KEY VALUE DIFFERENTIATORS

Diverse range of treatments covered

- Covers patients from 11 different cohorts
- Representative of most treatment options available to CRC patients in the UK
- Mainly late-stage patients with a small early stage & pre-cancerous cohort

Linked multi-omics

- 5 different data types captured and linked giving an overview of pathology, genetics, transcriptomics and methylation providing a comprehensive overview of molecular profile of tumours
- Linked to clinical outcome data

Clinical trial standard clinical data

- 6 of the 11 cohorts are from clinical trials with clinical outcome data collected under these studies ensuring detail and accuracy

Standardised & Curated

- Data generation occurred through a single consortium ensuring data is standardised across the different cohorts
- Extensive curation has gone in to ensuring the data is linked and easily accessible

A CLINICALLY LINKED MULTI-OMIC DATASET IN DIVERSE COLORECTAL CANCER COHORTS

	Patients	Samples	Data
Late-Stage CRC	Colon, Colorectal and Rectal cancer patients treated with: • Radiotherapy (n =462) • Chemotherapy* (n=385) • Irinotecan (n=233) • Cetuximab (n=233) **Total = 1413**	• Primary tumour resections • Pre-treatment biopsies • Matched post treatment biopsies for a Radiotherapy & chemotherapy cohort • Matched liver mets from cetuximab treated patients **Total = 1645**	• RNA Array • Mutational profile • Copy number alteration • Methylation analysis • Digitised IHC • Digitised H&E **Total Data points = 9060**
Early CRC	• Adenomas (n=71) • Early invasion (n=271) **Total = 342**	• Adenomas - proximal TVA, distal TVA, SSA, TVA • Early invasion - resections of primary tumours **Total = 364**	• RNA Array • Mutational profile • Copy number alteration • Methylation analysis • Digitised H&E • RNA-seq (Adenomas only) **Total Data points = 1902**

*chemotherapy = a collection of cohorts in which patients are treated with either FOLFOX or 5FUFA

THE AVAILABLE DATA

Numbers are representative of the number of samples profiled

Disease Stage ->		Late Stage									Early stage		
Treatment type ->		Chemotherapy		Radiotherapy				Irinotecan		Cetuximab			
Data type		FOCUS (-Irinotecan)	FOXTROT	ARISTOTLE (-Irinotecan)	GRAMPIAN	TREC	COPERNICUS	FOCUS (+Irinotecan)	ARISTOTLE (+Irinotecan)	New EPOC	Adenomas	PT1	SPINAL
Genomics	Mutational analysis from NGS panel sequencing	354	123	141	189	37	21	79	145	358	82	29	225
	Copy Number alteration	354	124	141	189	37	-	79	149	361	82	29	225
Transcript-omics	RNA Array	375	179	148	223	37	46	75	152	351	-	41	239
	RNA-seq	-	-	-	-	-	-	-	-	-	80	-	-
Imaging	Digitised H&E slides	751	363	301	461	74	101	164	309	670	82	31	480
	Digitised IHC	381	-	-	-	-	-	-	-	-	-	-	-
Epigenetics	Methylation analysis	325	111	62	119	20	-	78	62	241	72	-	205

SUMMARY OF COHORTS – LATE-STAGE PATIENTS

Treatment	Trial	Cohort Description	Patients	Samples
Chemo (Oxaliplatin)	FOCUS (-Irinotecan)	Patients with advanced colorectal cancer, mostly primary resections taken from Groups A, B-ir, B-ox (all treated with **first line 5FUFA**, 305) and C-ox (treated with **first line FOLFOX**, 80)	385	385
	FOXTROT	Patients from Foxtrot clinical trial (arm A). 100 patients with stage 3 colon cancer treated with **3 cycles of FOLFOX in the neoadjuvant** setting. Pretreatment biopsies with some matched post treatment resections	100	188
Radiotherapy	ARISTOTLE (-Irinotecan)	Patients with locally advanced rectal cancer entered into the multicenter randomized Aristotle clinical trial. Pretreatment biopsies before **neoadjuvant radiotherapy and capecitabine**.	138	155
	GRAMPIAN	Patients with rectal cancer from a community cohort from the Aberdeen area. Pretreatment biopsies before neoadjuvant **radiotherapy +/- additional chemotherapy**.	231	231
	TREC	Patients with early rectal cancer. Pretreatment biopsies from TREC clinical trial	37	37
	COPERNICUS	Patients with rectal cancer treated with **8 weeks FOLFOX and short course pre-operative radiotherapy**. Pretreatment biopsies and some matched post treatment resections taken from Copernicus clinical trial	39	51
Irinotecan	FOCUS (+Irinotecan)	Patients with advanced colorectal cancer entered into the national multicenter FOCUS trial and randomized to the group C-Ir and treated with **first line Irinotecan (FOLFIRI)**	82	82
	ARISTOTLE (+Irinotecan)	Patients with locally advanced rectal cancer entered into the multicenter randomized Aristotle clinical trial. Pretreatment biopsies before **neoadjuvant radiotherapy, capecitabine + Irinotecan.**	132	151
Cetuximab	New EPOC	Patients with synchronous, resectable colorectal liver metastases entered into the national multicenter New EPOC trial. Resections of matched primary and liver metastasis from patients randomised to **chemotherapy +/- cetuximab** prior to liver resection from New Epoc clinical trial	233	365

SUMMARY OF COHORTS – EARLY-STAGE PATIENTS

Treatment	Cohort Description	Patients	Samples
Adenoma	Formalin-Fixed Paraffin-Embedded (FFPE) adenomas - proximal TVA, distal TVA, SSA, TVA from clinical care collection	71	82
PT1	Resections of primary pT1 tumours with or without recurrence taken from clinical care collections.	30	41
Early invasion	Sample set from patients with colon and rectal cancer, enriched for earlier cancers with recurrent or metastatic disease. Resections of primary tumours without previous treatment, enriched for nodal/metastatic disease. Samples from clinical care and the COIN clinical trial	241	241

DATA GENERATION TECHNIQUES – ALL COHORTS

Genomics

Raw and processed data from NGS Gene Panel using panel of 80 key CRC driver mutations.
Including per segment and per gene copy number assessments.
Sequencing performed using Illumina Hiseq200.
MSI scores provided based on analysis of NGS panel results

H&E Digitised slides

Typically, two slides are digitised for each sample to depict the slice used for RNA analysis and the slice used for sequencing and methylation.
Annotation show area actually used for profiling.
Digitisation was performed using Aperio scanner at a magnification of 20x

Transcriptomics

Raw and processed RNA Array data (almac Xcel array) showing mean, median and max per gene expression with accompanying probe list.
110,000 probesets covering 24,000 genes.

Epigenetics

Methylation status was profiled using EPIC Array from Illumina and presented as median methylation value per cpg island.
Covers 700,000 CpG islands and 15,000 genes

DATA GENERATION TECHNIQUES – SELECTED COHORTS

Immuno-histochemistry

For the FOCUS (-Irinotecan) cohort only IHC data was generated using HALO or QuPath. IHC looked at expression of proteins associated with immune environment (PD, PDL-1 CD3, CD4, CD8 etc) profileration (Ki67) and DDR (MLH1)

RNA-seq

For the Adenoma cohort only, RNA-seq data was generated using the 5' RNAseq approach. This data is available only raw format and has not been log2 processed.

CLINICAL DATA – AN OVERVIEW

- Follow up PFS/OS data available for the below trials:
  - FOCUS
  - New EPOC
  - COPERNICUS
  - COIN (*part of the early invasion cohort*)
- Pathological assessment of response after treatment – currently no long term follow up
  - ARITSTOTLE
  - FOXTROT
  - TREC
- Grampian and non-COIN trial early invasion cohorts have some survival information, but this will have been collected through routine practice in the clinic and therefore will not be as complete or reliable as that from the trials listed above
- PT1 cohort has information on whether recurrence occurred but not on timelines meaning survival timelines cannot be calculated
- The Adenoma cohort does not have survival information as most of the patients should have remained cancer free. The clinical information here is pathological type of the adenoma

APPLYING FOR ACCESS

- Cancer Research Horizons has exclusive rights to the S:CORT Database
- Licenses are:
    - Non-exclusive
    - Time-limited (typical term is 3 years)
    - For a defined purpose
    - In line with Cancer Research Horizons Guiding Principles for Commercial Data Partnerships (https://www.cancerresearchhorizons.com/our-guiding-principles-data-partnerships)
- To access the S:CORT Database, companies must submit an application to the S:CORT Data Access Committee that outlines:
    - A summary of the research to be performed using the dataset in both technical and lay terms
    - A description of the potential patient benefit generated from the dataset
    - The specific data that you require access to both in terms of data types and cohorts
- Data access will be facilitated through provision of access to a private instance of C-BIO Portal

CANCER RESEARCH HORIZONS

FURTHER FASTER TOGETHER

S:CORT

Stratification in COloRecTal cancer

THANK YOU

Contact Joe Day (joseph.day@cancer.org.uk) for more information