SUPPLEMENTARY METHODS

Patient Samples

The OAC cohort comprised of 273 formalin fixed paraffin embedded collected at four centres in the UK (Belfast, Cambridge, Edinburgh and Southampton) as part of the Oesophageal Cancer Clinical and Molecular Stratification consortium. All patients were treated with platinum-based neo-adjuvant chemotherapy prior to surgical resection with either an oesophagectomy or extended gastrectomy. Prior to platinum-based neo-adjuvant chemotherapy all patients were evaluated by computed tomography (CT) of the chest and abdomen, positron emission tomography (PET) and oesophagogastroduodenoscopy with endoscopic ultrasound (EUS) where clinically indicated. Tumour site ranged from 33 (12.1%) oesophageal tumours to 130 (47.6%), 78 (28.6%) and 32 (11.7%) Siewert type 1, 2 and 3 type tumours respectively (Supplementary Table 2). Samples were obtained from the Northern Ireland Biobank and in collaboration with three further clinical sites, University of Cambridge, University of Southampton and University of Edinburgh. Prior to acquiring the retrospective validation cohort a power calculation was performed using a Hazard ratio of 2.0. Using preliminary data we estimated that the frequency of the assay positive group was approximately 21% of the population with an event rate of 0.362. Therefore with 273 patients this would give a study power of between 80% at a significance level of 0.05.

Molecular Profiling of OAC Biopsy samples

Biopsies were reviewed for pathological subtype prior to marking for macrodissection and samples containing at least 50% adenocarcinoma tissue by area were taken forward. Post randomisation, total RNA was extracted from 5x10 µm macrodissected FFPE tissue slides using the Recoverall™ Total Nucleic Acid Isolation Kit for FFPE (Thermo Fisher Scientific, Waltham, MA). RNA was converted into cDNA, amplified and converted into single-stranded form using SPIA® technology of the WT-Ovation™ FFPE RNA Amplification System (NuGEN Technologies Inc., San Carlos, CA, USA). Amplified cDNA was fragmented, biotin-labelled using FL-Ovation™ cDNA Biotin Module (NuGEN Technologies Inc.), and hybridized to the Almac Xcel array. Arrays were scanned using Affymentrix Genechip® Scanner 7G (Affymetrix Inc., Santa Clara, CA, USA).

Process Control

UHR and ES-2 cell lines were included during processing of the OAC clinical validation cohort as process control measures. These controls were assessed initially for the quality of the microarray data followed by Statistical Process Control whereby pre-defined acceptance criteria (using control assay) were used to determine anomalies or drift in the process.

Microarray Quality Control (QC)

Microarray quality control analysis was performed using Robust Multi-array Average RMA[1] pre-processed data, assessing the following metrics:

Array Image Analysis

Analyses Affymetrix® CEL files (containing raw intensity data) and looks for large deviations in background values or unusual patterns in probe intensities that may indicate the presence of artefacts or problems with hybridization (such as leakage from the array, uneven washing)

GeneChip QC

Examines a number of control parameters from the RPT files as supplied by Affymetrix®. Full descriptions of these control parameters are available from the Affymetrix® Statistical Algorithms Description Document.

This assessment facilitates the monitoring of profile quality and allows for the evaluation of assay and hybridization performance. Affymetrix® has specified absolute thresholds (lower limit, upper limit or both limits) for a number of these parameters. In addition, it is expected that for the majority of studies data should be comparable, therefore an assessment of overall profile similarity is performed using thresholds based on median absolute deviation (MAD). Any values outside median +/- 3.5 sigma (sigma defined as 1.4826 times the median absolute deviation [MAD]) for that metric will be flagged as potential QC outliers.

Principal Component Analysis (PCA)

A well-established exploratory analysis method for high-dimensional data and provides low-dimensional summaries of sample and variable properties.

PCA is used in this context to detect outliers and (known or unknown) systematic structures using pre-processed expression data.

Intensity Distribution Analysis

A collection of graphical distribution assessment methods, employed to assess the spread of the data and examine the position of individual profiles, relative to the full profile set. Histogram plots of normalized expression data are constructed to visualize profile spread and differences and expression box-plots are provided to graphically examine median, upper and lower quartile ranges for each profile.

Within a study, it is generally expected that all samples should show a comparable distribution. The program assesses distribution similarity using the Kolmogorov-Smirnov (KS) Test and samples are deemed as outliers based on a predefined KS statistic thresholds.

Analysis of Gene Expression Data

Generation of DDIR Assay Scores

Pre-processing of the OAC independent validation dataset was performed using a refRMA model applied to RMA background corrected data, which applied pre-defined normalisation and summarisation parameters to the probe level data specific for the Xcel™ platform.[1] Prior to DDIR score calculation, datasets were median summarised from probe to probe set levels, followed by a median summary to gene level. DDIR scores were calculated using the DDIR assay parameters (derived from a regression based model) applied as previously described.[2]

An optimised DDIR threshold (identified in an independent technical study of *n*=45 samples prior to profiling of the independent OAC validation cohort) for DDIR on Xcel™ microarray platform of 0.3403 was applied to dichotomise DDIR assay scores of the OAC independent validation samples to classify samples as DDIR positive (>threshold) otherwise DDIR negative (<=threshold).

Functional Enrichment

A moderated t-test (SAMR) was applied to identify differentially expressed genes between the DDIR-positive and negative groups. Using an FDR of <0.05 a list of 707 genes were derived and used as input into Database for Annotation, Visualization and Integrated Discovery (DAVID).[3] The enrichment of Gene Ontology Biological Process terms was analysed. Terms with p-values less than 0.05 were retained and fold enrichment and FDR reported.

Statistical Analysis

All tests of significance were two-sided and performed at a 5% alpha significance level. The *p*-value for the test was calculated for original, log-transformed and rank time scales. Clinical factor association with DDIR status was performed using chi-squared test. Cox proportional hazards regression was used to investigate the prognostic effects of the DDIR signature on relapse-free (RFS) and overall survival (OS) defined as the time from surgical resection to relapse of disease or death from any cause, respectively. The estimated effect of the signature was adjusted for factors available at the time of diagnosis (clinical tumour status, clinical nodal status and tumour grade) by fitting a multivariate model. The proportional hazard

assumption was verified for the Cox model using a formal statistical test based on the Schoenfeld residuals.[4] The p value for the test was calculated for original, log-transformed and rank time scales. Samples with unknown clinical factors were excluded. All tests of significance were two-sided and performed at a 5% alpha level. One way ANOVA was used to test the association between DDIR score and pathological response status defined by TRG.

Whole Genome Sequencing

Matched whole genome sequencing data was available for 44 patients who received neo-adjuvant chemotherapy prior to surgical resection at three OCCAMS centres (Cambridge, Edinburgh and Southampton; Supplemental Table 4). A single library was created for each sample, and 100-bp paired-end sequencing was carried out by Illumina (San Diego, CA) and the Broad Institute (Cambridge, MA) to a typical depth of at least 50X for tumours and 30X for matched normals, with 94% of the known genome being sequenced to at least 8X coverage and achieving a Phred quality of at least 30 for at least 80% of mapping bases. Read sequences were mapped to the human reference genome (GRCh37) using Burrows–Wheeler alignment (BWA) 0.5.9 and duplicates were marked and discarded using Picard 1.105. As part of an extensive quality-assurance process, quality-control metrics and alignment statistics were computed on a per-lane basis. The FastQC package was used to assess the quality-score distribution of the sequencing reads and perform trimming if necessary.

Mutational signatures were identified using the non-negative matrix factorization (NMF) methodology as previously described.[5, 6]

Immunohistochemistry

A 3-µm thick section was deparaffinised and endogenous peroxidase activity was quenched with 0.3% hydrogen peroxide prior to staining using the Ventana Discovery XT® automated immunostainer (Ventana Medical Systems Inc, Tuscon, AZ). Antibodies to CD8 (C8/144B, M7103, Dako) and PD-L1 (SP142, Roche) were used as previously described. Staining for p53 was performed as previously described.[7] Sections were incubated with antibody at 37ºC for 8 minutes prior to use of the Omnimap® anti-rabbit HRP conjugate detection kit (Ventana). Lung adenocarcinoma tissue was used as a positive control and test tissue with no primary antibody was used as a negative control.

1. Irizarry RA, Bolstad BM, Collin F, et al: Summaries of Affymetrix GeneChip probe level data. Nucleic Acids Res 31:e15, 2003

2. Mulligan JM, Hill L a, Deharo S, et al: Identification and validation of an anthracycline/cyclophosphamide-based chemotherapy response assay in breast cancer. J Natl Cancer Inst 106, 2014

3. Huang DW, Sherman BT, Lempicki RA: Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res 37:1–13, 2009

4. GRAMBSCH PM, THERNEAU TM: Proportional hazards tests and diagnostics based on weighted residuals. Biometrika 81:515–526, 1994

5. Alexandrov LB, Nik-Zainal S, Siu HC, et al: A mutational signature in gastric cancer suggests therapeutic strategies. Nat Commun 6:8683, 2015

6. Secrier M, Li X, de Silva N, et al: Mutational signatures in esophageal adenocarcinoma define etiologically distinct subgroups with therapeutic relevance. Nat Genet , 2016

7. Parkes EE, Walker SM, Taggart LE, et al: Activation of STING-Dependent Innate Immune Signaling By S-Phase-Specific DNA Damage in Breast Cancer. J Natl Cancer Inst 109:djw199, 2017