

# AI-powered immune profiling from histopathology slides for chemo-radiotherapy outcome prediction in rectal cancer: a study using clinical trial and real-world cohorts



Zhuoyan Shen,<sup>a,\*</sup> Douglas Brand,<sup>a,b</sup> Mikael Simard,<sup>a</sup> Adam P. Levine,<sup>c,d</sup> Sumeet Hindocha,<sup>a,b</sup> Talisa Mistry,<sup>c</sup> Dahmane Oukrif,<sup>c</sup> Andre Lopes,<sup>e</sup> Rubina Begum,<sup>e</sup> Nicholas P. West,<sup>f,g</sup> Ying Zhang,<sup>a</sup> Gary Royle,<sup>a</sup> Tim S. Maughan,<sup>h</sup> David Sebag-Montefiore,<sup>g</sup> Maria A. Hawkins,<sup>a,b,i</sup> and Charles-Antoine Collins-Fekete<sup>a,i</sup>



<sup>a</sup>Department of Medical Physics and Biomedical Engineering, University College London, London, UK

<sup>b</sup>Department of Radiotherapy, University College London Hospitals NHS Foundation Trust, London, UK

<sup>c</sup>Research Department of Pathology, University College London, London, UK

<sup>d</sup>Department of Cellular Pathology, University College London Hospitals NHS Foundation Trust, London, UK

<sup>e</sup>Cancer Institute, University College London, London, UK

<sup>f</sup>Division of Pathology and Data Analytics, Leeds Institute of Medical Research, School of Medicine, University of Leeds, Leeds, UK

<sup>g</sup>Division of Oncology, Leeds Institute of Medical Research, School of Medicine, University of Leeds, Leeds, UK

<sup>h</sup>Department of Oncology, University of Oxford, Oxford, UK

## Summary

**Background** The impact of the tumour-immune microenvironment on locally advanced rectal cancer (LARC) outcomes remains unclear. This study quantitatively assesses the synergistic influence of tumour-infiltrating lymphocytes (TILs), tumour-associated macrophages (TAMs), mitotic activity, and DNA mutations in predicting outcomes for LARC patients undergoing neoadjuvant chemo-radiotherapy (nCRT).

**Methods** Three cohorts (ARISTOTLE-RC, UCLH-RC, TCGA-CRC) were stratified by densities of AI-quantified TILs, TAMs, and mitotic figures with cut-offs identified on a hold-out subset and integrated with DNA mutations to assess correlations with disease-free survival (DFS) and overall survival (OS). Immune cell dynamics pre- and post-CRT were also evaluated.

**Findings** In ARISTOTLE-RC, TIL<sup>+</sup> patients had significantly improved DFS (HR = 0.59, 95% CI: 0.39–0.90,  $p = 0.013$ ) and OS (HR = 0.42, 95% CI: 0.24–0.73,  $p < 0.005$ ), while TAM<sup>+</sup> was associated with shorter DFS (HR = 1.65, 95% CI: 1.00–2.72,  $p = 0.045$ ). Similar patterns were observed in UCLH-RC and TCGA-CRC. TIL<sup>+</sup>/KRAS<sup>-</sup> patients had significantly improved DFS (HR = 0.41, 95% CI: 0.22–0.75,  $p < 0.005$ ) and OS (HR = 0.28, 95% CI: 0.13–0.62,  $p < 0.005$ ). In TP53-mutated patients, TAM<sup>+</sup> group showed shorter DFS (HR = 1.46, 95% CI: 1.07–2.01,  $p = 0.0151$ ), while among TP53-wild-type patients, no difference in DFS (HR = 1.00, 95% CI: 0.66–1.52,  $p = 0.9930$ ) was observed between the two subgroups. Patients who transitioned after nCRT from TIL<sup>-</sup> to TIL<sup>+</sup> had improved DFS (HR = 0.70, 95% CI: 0.50–0.97,  $p = 0.028$ ) and exhibited a significantly higher pre-treatment mitotic index (mean difference = 9.36, 95% CI: 1.87–16.85,  $p = 0.0385$ ).

**Interpretation** These findings suggest the potential utility of AI-driven immune profiling for clinical decision-making in LARC patients undergoing nCRT.

**Funding** Cancer Research UK (RRNPSF-Jan21/100001, A18745, C7893/A2899), UK Research and Innovation (MR/T040785/1).

**Copyright** © 2025 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

**Keywords:** Colorectal cancer; Tumour immune microenvironment; Chemo-radiotherapy; Artificial intelligence; Digital pathology

\*Corresponding author.

E-mail address: [zhuoyan.shen.18@ucl.ac.uk](mailto:zhuoyan.shen.18@ucl.ac.uk) (Z. Shen).

<sup>i</sup>Senior authors contributed equally.

eBioMedicine

2025;122: 105993

Published Online 17

November 2025

<https://doi.org/10.1016/j.ebiom.2025.105993>

1016/j.ebiom.2025.105993

### Research in context

#### Evidence before this study

We conducted a systematic review of the PubMed database to identify studies examining the prognostic significance of tumour-infiltrating lymphocytes (TILs), tumour-associated macrophages (TAMs), and mitotic activity in colorectal cancer (CRC), covering the period from May 2015 to May 2025. The search included all languages and used the terms “colorectal cancer” AND “tumour infiltrating lymphocytes”, “colorectal cancer” AND “tumour-associated macrophages”, and “colorectal cancer” AND “mitotic activity”. This yielded 1186, 557, and 133 articles, respectively. Evidence shows that higher TIL density is associated with improved outcomes and may predict response to neoadjuvant chemo-radiotherapy (nCRT). In contrast, findings on TAMs were inconsistent across studies. However, most of the studies were conducted with a limited number of patients and the impact of KRAS mutation and nCRT was barely illustrated, nor the pattern dynamics after the nCRT. Studies on mitotic activity were scarce, and rarely used quantitative methods. Additionally, few studies employed artificial intelligence (AI) for automated cell quantification in histopathology, highlighting a gap in scalable, reproducible biomarker analysis.

#### Added value of this study

This study provides evidence that high TIL and low TAM levels are favourable prognosticators for all-stage colorectal cancer patients and locally advanced rectal cancer (LARC)

patients who received nCRT in both a clinical trial and real-world clinical settings. We identified TIL density and KRAS mutation status as independent predictors of survival, and their combined stratification revealed strong prognostic significance. Furthermore, dynamic changes in TIL density following nCRT were associated with better outcomes in immunoreactive patients. We also found that extremely high mitotic activity is linked to poorer survival and may contribute to immune suppression after nCRT. Notably, this study developed an AI-based framework to automatically quantify these markers from routine histopathology slides, offering a rapid and cost-effective approach for biomarker integration in clinical workflows.

#### Implications of all the available evidence

Our findings highlight the potential for AI-assisted quantification of key components in the tumour immune microenvironment to enhance risk stratification in CRC, especially for LARC patients undergoing nCRT. The automated analysis offers a scalable and reliable method for evaluating TIL and TAM density, as well as mitotic index, in clinical workflows, potentially leading to more personalised treatment strategies and improved patient outcomes. Future research should focus on validating these AI-driven markers in larger, diverse cohorts and exploring their integration with other genomic and molecular markers to identify patients whose immune activity can be triggered by nCRT, and further refine patient selection and treatment planning.

### Introduction

Colorectal cancer (CRC) is the third most common cancer and the second leading cause of cancer-related mortality worldwide.<sup>1</sup> The charity Bowel Cancer UK highlighted the importance of integrating genomics, big data science and digital pathology methods to improve the morpho-molecular taxonomy and biological stratification of CRC.<sup>2</sup> Developing biomarkers is crucial to define the optimal curative therapeutic strategy for an individual or group, improving treatment selection and minimising the risk of overtreatment.<sup>3</sup>

Chemo-radiotherapy (CRT) is a common treatment option for rectal cancers, especially locally advanced rectal cancer (LARC).<sup>4,5</sup> LARC patients have a higher risk of recurrence, necessitating more intensive neoadjuvant treatment strategies than early-stage rectal cancer patients.<sup>6</sup> Optimising treatments in LARC involves intensifying treatment (e.g. total neoadjuvant therapy (TNT)) prior to total mesorectal excision (TME).<sup>7</sup> Trials investigating changes in TNT regimes have been published with evidence that impacts clinical practice.<sup>8-10</sup> Although the management of LARC is advancing, the focus has been largely on modifying the sequencing of treatments rather than the selection of treatments themselves. Combining emerging

techniques, including artificial intelligence (AI), digital pathology, and whole genome sequencing (WGS) offers the opportunity to develop advanced signatures for precise treatment selection. Recently, a phase II study reported that LARC patients with mismatch repair deficiency are sensitive to treatment with neoadjuvant PD-1 blockade and can avoid standard CRT and surgery.<sup>11</sup>

The tumour immune microenvironment (TIME) encompasses the whole cellular and acellular environment where tumour cells exist. TIME has a pivotal role in shaping tumour phenotype, evolutionary dynamics, and therapy responses, and its impact on prognosis and treatment response is well-established across human cancers.<sup>12</sup> Tumour-infiltrating lymphocytes (TILs) and tumour-associated macrophages (TAMs) serve as key indicators of immune activity within the TIME. High TIL is associated with favourable outcomes across a variety of cancer types.<sup>13-15</sup> High TIL density is also a prognostic factor for better treatment response and longer disease-free survival (DFS) after neoadjuvant chemo-radiotherapy (nCRT) for rectal cancer.<sup>16,17</sup>

The prognostic significance of TAMs remains less clear. Their impact is influenced by phenotypic diversity, as M1-like TAMs are linked to better prognosis

due to their pro-inflammatory and tumour-suppressive functions, whereas M2-like TAMs promote tumour progression by enhancing angiogenesis, suppressing immune responses, and facilitating metastasis.<sup>18</sup> Studies have reported conflicting findings regarding TAMs in CRC. While some studies suggest that high TAM infiltration is linked to improved prognosis in colon cancer,<sup>19–21</sup> others have reported contradictory results.<sup>22–24</sup> Additionally, it has been linked to poorer outcomes in rectal cancer.<sup>25,26</sup> This discrepancy highlights the differences in molecular phenotypes and treatment regimens. Notably, the common use of CRT in rectal cancer may alter the TIME and influence TAM behaviour.<sup>27,28</sup> However, large-scale studies investigating TAMs in CRT-treated rectal cancer cohorts remain limited.

Tumour proliferation is a critical factor influencing cancer outcomes. The mitotic index (MI), which quantifies mitotic activity within a tumour, is a well-established grading criterion for multiple solid tumours.<sup>29–31</sup> However, limited studies have reported the negative impact of MI or cell proliferation markers on survival in CRC.<sup>32,33</sup> In addition, while some studies suggest an association between mitosis and immune activity, the evidence is inconclusive.<sup>34</sup>

In this study, we aimed to characterise the prognostic utility of AI-derived TIL and TAM density through a phase III trial cohort and a real-world clinical cohort undergoing CRT. Immune cell metrics and MI were quantified using an AI framework applied to digitised whole slide images (WSIs) of pre-treatment biopsies and post-treatment resections. Cut-off values were determined within a hold-out sub-cohort, and DFS and overall survival (OS) were compared among stratified patients to evaluate the impact of the immune cell landscape, CRT, tumour proliferation, and DNA mutations.

## Methods

### Study population

Three cohorts were included in this study:

1. ARISTOTLE-RC: 589 participants with MRI-defined LARC from the ARISTOTLE trial (ISRCTN09351447). They were randomised into two treatment arms: A) Standard arm (SCRT)-capecitabine 900 mg/m<sup>2</sup> orally twice daily Monday to Friday for five weeks with radiotherapy 45Gy in 25 fractions; B) Experimental arm (IrCRT)-irinotecan 60 mg/m<sup>2</sup> intravenous once weekly for four weeks and capecitabine 650 mg/m<sup>2</sup> orally twice daily Monday to Friday for five weeks with radiotherapy 45Gy in 25 fractions. 414 patients with available digitised whole slide images (WSIs) were included and divided into sub-cohorts in this study. A diagram of the inclusion and exclusion criteria is shown in [Figure S1](#). Diagnostic years range from 2011 to 2018. Sex data were collected from clinical records as documented by the treating clinical teams.
2. UCLH-RC: A retrospective cohort of 70 MRI-defined LARC patients who received standard CRT at University College London Hospitals NHS Foundation Trust. Diagnostic years range from 2013 to 2022. Sex data were collected from clinical records as documented by the treating clinical teams.
3. TCGA-CRC: A publicly available cohort of 458 colorectal cancer (CRC) patients across stages I–IV from The Cancer Genome Atlas (TCGA).<sup>35</sup> Diagnostic years range from 1998 to 2013.

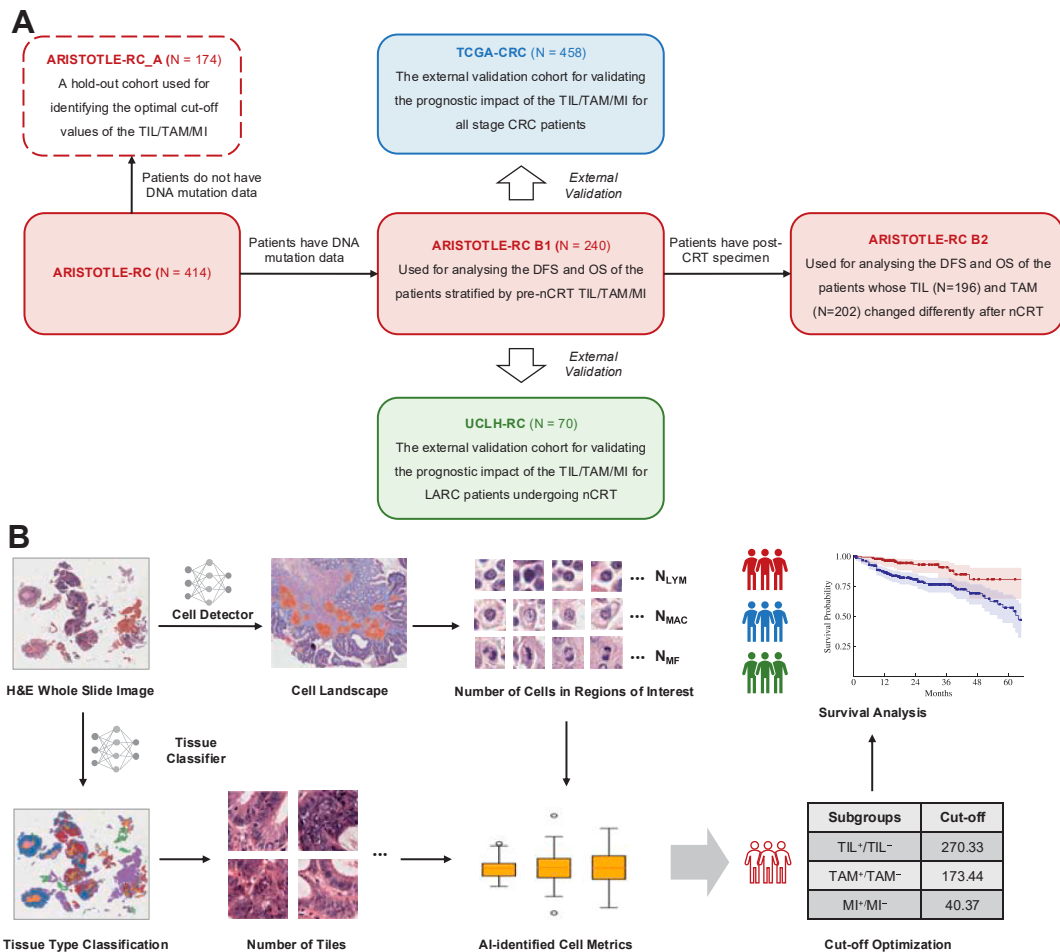
Information on the clinical data and digitised WSIs used in each cohort is listed in [Table S1](#). DFS and OS were used as the outcome measures. The follow-up time is up to five years.

### Ethics

Ethical approval for the ARISTOTLE trial was granted by the NHS Health Research Authority (HRA) (REC reference: 10/H0706/65, IRAS ID: 5406). Ethical approval for the UCLH-RC cohort was provided by the UCL/UCLH Biobank for Studying Health and Disease (study EC30.22) under NHS HRA delegated authority (REC reference: 20/YH/0088, IRAS ID: 272816). Written informed consent was obtained from all participants.

### Study design

The study design workflow is illustrated in [Fig. 1A](#). This study is based on a post hoc analysis of the ARISTOTLE trial and was not conducted according to a pre-specified protocol. Within ARISTOTLE-RC, 174 patients without available DNA mutation data were designated as a hold-out cohort (ARISTOTLE-RC A) to determine optimal cut-off values for TIL and TAM density and MI. The optimal values maximise Youden's index<sup>36</sup> in the receiver operating characteristic (ROC)<sup>37</sup> analysis targeting the three-year disease progression. ARISTOTLE-RC B1 and ARISTOTLE-RC B2 are used as the validation sets to compare the DFS and OS of the patients stratified by the pre-treatment metrics (B1) and the immune dynamics before and after the treatment (B2). The pre-treatment timepoint was defined as the baseline diagnostic biopsy before initiation of nCRT, and the post-treatment timepoint as the surgical resection specimen. Only patients with matched pre- and post-treatment samples and remaining tumour tissue post-treatment were included in ARISTOTLE-RC B2. KRAS and TP53 mutation status were selected for downstream patient stratification for the high mutation rates. Findings from ARISTOTLE-RC were externally validated using the UCLH-RC cohort. In addition, the



**Fig. 1: Illustration of the study design.** A. Description of the cohorts used in this study. B. Workflow of stratifying patients using cell metrics identified by the AI framework.

TCGA-CRC cohort was used to explore the prognostic value of cell-based metrics across all CRC stages.

**AI framework**

An AI framework (Fig. 1B) consisting of a tissue classifier and a cell detector was developed and deployed to automatically detect lymphocytes, macrophages, and mitotic figures in the three cohorts.

The framework consists of two AI models:

- (1) Tissue Classifier. An EfficientNetB0<sup>38</sup> was trained on the NCT-CRC-HE-100 K dataset,<sup>39</sup> consisting of 100,000 non-overlapping tiles of nine types of tissue: adipose (ADI), background (BACK), debris (DEB), lymphocytes (LYM), mucus (MUC), smooth muscle (MUS), normal colon mucosa (NORM), cancer-associated stroma (STR), colorectal adenocarcinoma epithelium (TUM). The size of a single tile is 224 pixels × 224 pixels with a pixel size of 0.5 μm/pixel. The tissues were manually extracted and labelled from pathologist-

annotated regions in the WSIs. The EfficientNet was trained using a single NVIDIA GeForce RTX 3090 GPU for 100 epochs. The training batch size was 256, and the optimiser was AdamW algorithm with a learning rate of 0.0001. Random horizontal flips (p = 0.5) were also used for spatial augmentation.

- (2) Cell Detector. A YOLOv10x model<sup>40</sup> was trained for detecting individual cells of interest. For detecting the immune cells, the model was trained on the Immunocito dataset<sup>41</sup> with 2,282,818 immune cells labelled by multiplex immunofluorescence. For detecting the mitotic figures, the model was trained on the OMG-Octo<sup>42</sup> dataset with 74,620 pan-cancer mitotic figures. The yolov 10 model was trained using 4 NVIDIA GeForce RTX 3090 GPUs for 200 epochs. The training batch size was 48, and the optimiser was AdamW algorithm with a learning rate of 0.01. Spatial augmentation was applied, and the configuration parameters are listed in Table S4.

Colour augmentation was applied to the training data for both models to reduce the impact of the staining variation and increase the robustness of the model. RGB images were first deconvolved into haematoxylin (H) and eosin (E) channels using the stain separation method proposed by Ruifrok and Johnston.<sup>36</sup> To introduce realistic stain variation, we employed the stain concentration perturbation scheme described by Tellez et al.,<sup>43</sup> where perturbations were uniformly sampled ( $\sigma = 0.14$ ) and applied to the deconvolved H&E channels before reconstructing the RGB images. The data size for training, validation and testing, model performance on detecting tumours and different types of cells are listed in [Tables S3 and S5](#). Examples of outputs from the framework are illustrated in [Figure S2](#).

The WSIs are first analysed by the tissue classifier to identify regions of interest (ROIs) for analysing TILs, TAMs, and MI. Lymphocytes and mitotic figures are then detected by the cell detector within tumour regions, while macrophages are detected in both tumour and stromal regions. AI model development and digital pathology data analyses were conducted blinded to clinical outcomes.

### Definition of the measurements and notations

In clinical practice, no robust method has been established for directly quantifying TILs and TAMs in H&E-stained slides. In this study, we reported cell densities per 10 High-Power-Field ( $2 \text{ mm}^2$ ) for evaluating the TIL, TAM, and MI within a WSI, to maintain consistency with our previous work on AI-assisted mitotic index assessment.<sup>42</sup>

The metrics were defined by Equations (1)–(3), where  $N_{lym}$ ,  $N_{mac}$  and  $N_{mf}$  are the number of detected lymphocytes, macrophages and mitotic figures, respectively.  $N_{TUM}$  and  $N_{STR}$  are the number of tumour tiles and the number of stroma tiles,  $A_{10HPF}$  is the area of  $2 \text{ mm}^2$  and  $A_{tile}$  is the area of a single tile ( $1.25 \times 10^{-2} \text{ mm}^2$ ).

$$TIL = \frac{N_{lym}}{N_{TUM} \times A_{tile}} \times A_{10HPF} \quad \text{Equation 1}$$

$$TAM = \frac{N_{mac}}{(N_{TUM} + N_{STR}) \times A_{tile}} \times A_{10HPF} \quad \text{Equation 2}$$

$$MI = \frac{N_{mf}}{N_{TUM} \times A_{tile}} \times A_{10HPF} \quad \text{Equation 3}$$

[Table 1](#) lists the cut-off values for stratifying patients by the density metrics and the notations used to represent different subgroups in this paper. The optimisation of the cut-off values was done by ten-fold cross-validation using the data of ARISTOTLE-RC A.

### Statistics

Survival analysis was conducted using Python (v3.9.13) and the lifelines library (v0.28.0). Kaplan-Meier<sup>44</sup>

Notation	Subgroup	Cut-off (cells per $2 \text{ mm}^2$ )
TIL <sup>+</sup>	Patients with high pre-treatment TIL density	270.33
TIL <sup>-</sup>	Patients with low pre-treatment TIL density	
TAM <sup>+</sup>	Patients with high pre-treatment TAM density	173.44
TAM <sup>-</sup>	Patients with low pre-treatment TAM density	
MI <sup>+</sup>	Patients with high pre-treatment mitotic figure density	40.37
MI <sup>-</sup>	Patients with low pre-treatment mitotic figure density	
TIL/TAM <sup>++</sup>	Patients whose TIL/TAM density remained high after the treatment	-
TIL/TAM <sup>-+</sup>	Patients whose TIL/TAM density remained low after the treatment	-
TIL/TAM <sup>+-</sup>	Patients whose TIL/TAM density changed from high to low after the treatment	-
TIL/TAM <sup>--</sup>	Patients whose TIL/TAM density changed from low to high after the treatment	-

**Table 1:** Notations representing patients stratified by cell density.

estimation was performed to generate survival curves, and differences between subgroups were assessed using the log-rank test.<sup>45</sup> Hazard Ratios (HR) with a 95% confidence interval (CI) were given by the Cox proportional hazard model.<sup>46</sup> Pearson test, Chi-square test,<sup>47</sup> Mann–Whitney test<sup>48</sup> and Kruskal–Wallis test<sup>49</sup> were performed using SciPy library (version = 1.13.1) for correlation and differentiation analysis. All p-values are two-sided, and a p-value of 0.05 or less was deemed significant. This was a post-hoc analysis of the ARISTOTLE trial; sample size was determined by the availability of WSIs. For the UCLH cohort, the number of patients included was based on the availability of pre-treatment biopsies from LARC patients who received neoadjuvant chemoradiotherapy and were archived in the UCL/UCLH Biobank.

### Role of funders

The ARISTOTLE trial sample collection and slide scanning were funded by Cancer Research UK (A18745). The other funders played no direct roles in study design, data collection, data analyses, interpretation, or writing of the report. None of the authors received compensation from the funder or any other entity for writing this article.

## Results

### Patient characteristics

[Table 2](#) shows the clinical characteristics and cell densities of the patients in ARISTOTLE-RC and UCLH-RC. The patient characteristics of TCGA-CRC are listed in [Table S7](#).

### Prognostic value of immune cell infiltration and mitotic index

In ARISTOTLE-RC B1, TIL<sup>+</sup> patients had significantly better DFS (HR = 0.59, 95% CI: 0.39–0.9, p = 0.0139) and OS (HR = 0.42, 95% CI: 0.24–0.73, p = 0.00148)

Characteristic	ARISTOTLE-RC (N = 414)	UCLH-RC (N = 70)
<b>Age</b>		
≤60	207 (50.0%)	21 (30.0%)
>60	207 (50.0%)	49 (70.0%)
<b>Sex</b>		
Female	140 (33.8%)	28 (40.0%)
Male	274 (66.2%)	42 (60.0%)
<b>Treatment</b>		
CRT	213 (51.4%)	70 (100.0%)
IrCRT	201 (48.6%)	0 (0.0%)
<b>MRI T stage<sup>a</sup></b>		
T2	24 (5.8%)	14 (20.0%)
T3	321 (77.5%)	43 (61.4%)
T4	64 (15.5%)	13 (18.6%)
Missing	5 (1.2%)	0 (0.0%)
<b>MRI N stage<sup>a</sup></b>		
N0	104 (25.1%)	17 (24.3%)
N1	184 (44.4%)	31 (44.3%)
N2	121 (29.2%)	22 (31.4%)
Missing	5 (1.2%)	0 (0.0%)
<b>Post-Operation Pathological T stage</b>		
ypT0	12 (2.9%)	0 (0.0%)
ypT1	96 (23.2%)	9 (12.9%)
ypT2	163 (39.4%)	24 (34.3%)
ypT3	15 (3.6%)	5 (7.1%)
ypT4	60 (14.5%)	29 (41.4%)
Missing		
<b>Post-Operation Pathological N stage<sup>b</sup></b>		
ypN0	244 (58.9%)	27 (38.6%)
ypN1	79 (19.1%)	10 (14.3%)
ypN2	30 (7.2%)	3 (4.3%)
ypNx	0 (0.0%)	1 (1.4%)
Missing	61 (14.7%)	29 (41.4%)
Median DFS (months)	51.8	33.6
Median OS (months)	60.2	47.0
5-year disease-free rate	75.5%	62.0%
5-year survival rate	61.2%	64.8%
5-year censoring rate	20.5%	33.8%
TIL density (mean ± std)	214.01 ± 204.26	271.43 ± 293.84
TAM density (mean ± std)	212.26 ± 173.46	229.72 ± 248.50
MI (mean ± std)	17.43 ± 13.75	15.48 ± 11.23

<sup>a</sup>Baseline MRI staging performed before the start of nCRT. <sup>b</sup>"Nx" indicates cases where nodal status could not be assessed (e.g., insufficient pathological data or further evaluation required). "N missing" denotes cases where nodal information was not available due to missing imaging or clinical documentation.

**Table 2: Patient characteristics in the two cohorts.**

compared with TIL<sup>-</sup> patients (Fig. 2A and B). The TAM<sup>+</sup> patients showed significantly shorter DFS (HR = 1.68, 95% CI: 1.02–2.76, p = 0.0397) compared with TAM<sup>-</sup> patients (Fig. 2C). However, the difference in OS is not significant (Fig. 1D). An inverse association between TAM and TIL levels ( $\chi^2 = 8.10$ , p = 0.0044) was found. The MI<sup>+</sup> patients had significantly shorter DFS (HR = 2.43, 95% CI: 1.39–4.24, p = 0.0012) and OS (HR = 0.42, 95% CI: 0.24–0.73, p = 0.0014) (Fig. 2E and F). Those patterns existed in both the SCRT and ICRT arms (Figures S4–S6).

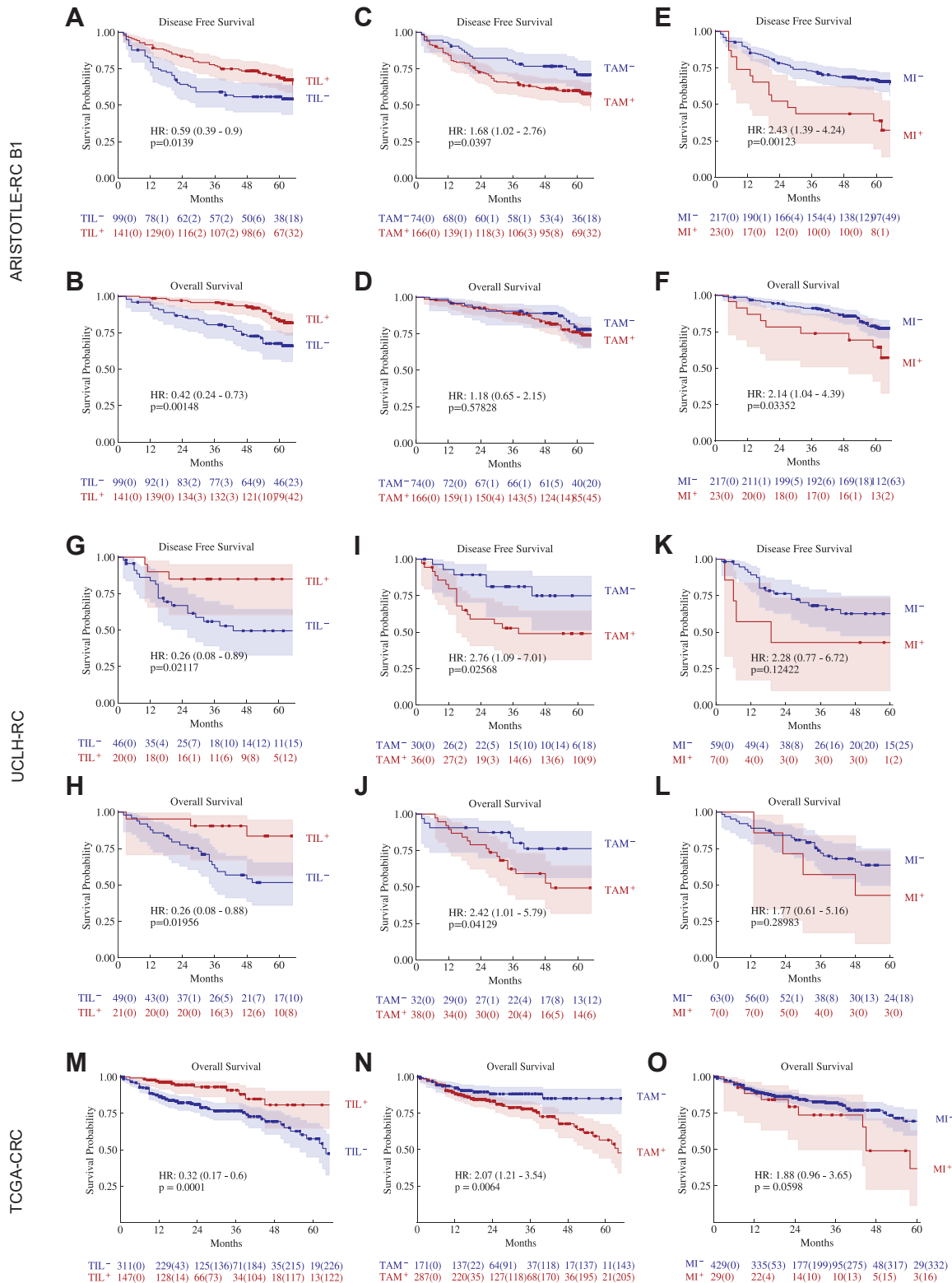
In the UCLH-RC cohort, TIL<sup>+</sup> patients had significantly better DFS (HR = 0.26, 95% CI: 0.08–0.89, p = 0.0212) and OS (HR = 0.26, 95% CI: 0.08–0.88, p = 0.0196) (Fig. 2G and H), while TAM<sup>+</sup> patients had significantly shorter DFS (HR = 2.42, 95% CI: 1.01–5.79, p = 0.0413) and OS (HR = 2.76, 95% CI: 1.09–7.01, p = 0.0257) (Fig. 2I and J). TAM<sup>+</sup> patients are also more likely to exhibit TIL<sup>-</sup> ( $\chi^2 = 5.62$ , p = 0.0177). Since no patients were stratified as MI<sup>+</sup> by the original cut-off value, we designate patients whose mitotic index falls within the top 10% as MI<sup>+</sup>, aligning with the percentage observed in the ARISTOTLE-RC cohort. The MI<sup>+</sup> patients tended to have shorter DFS (HR = 2.28, 95% CI: 0.77–6.72, p = 0.1242) (Fig. 2K and L). The prognostic impact of TIL and TAM in the combined ARISTOTLE-RC and UCLH-RC cohort is presented in Table S9.

In the TCGA-CRC, the TIL<sup>+</sup> patients also showed significantly longer OS (HR = 0.32, 95% CI: 0.17–0.59, p = 0.0001, Fig. 2M). The TAM<sup>+</sup> had significantly shorter OS (HR = 2.16, 95% CI: 1.25–3.73, p = 0.0048) (Fig. 2N). A negative association between TAM and TIL status was also found ( $\chi^2 = 22.04$ , p < 0.0001). The MI<sup>+</sup> patients tended to have shorter OS (HR = 1.88, 95% CI: 0.96–3.65, p = 0.0598, Fig. 2O).

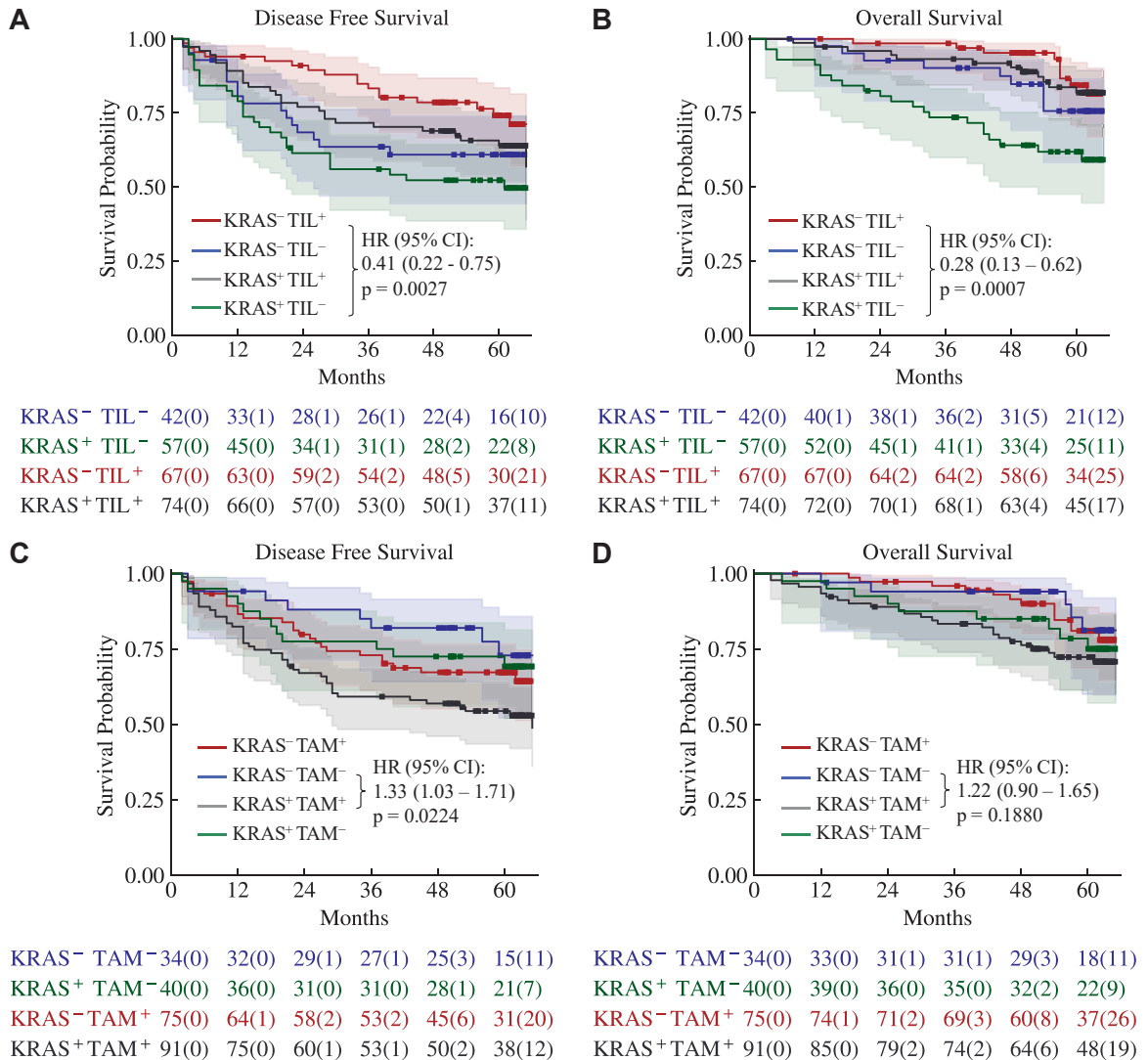
**Composite stratification of immune cell infiltration and mutation status**

In ARISTOTLE-RC, KRAS<sup>+</sup> patients tended to have shorter DFS (HR = 1.45, 95% CI: 0.94–2.23, p = 0.089) and OS (HR = 1.63, 95% CI: 0.92–2.87, p = 0.084) (Figure S7). TIL<sup>+</sup> and KRAS<sup>-</sup> patients had significantly longer DFS (HR = 0.41, 95% CI: 0.22–0.75, p = 0.0027) and OS (HR = 0.28, 95% CI: 0.13–0.62, p = 0.0007), compared to the TIL<sup>-</sup> and KRAS<sup>+</sup> group (Fig. 3A and B). TAM<sup>+</sup> and KRAS<sup>+</sup> patients had significantly shorter DFS (HR = 1.33, 95% CI: 1.03–1.71, p = 0.0224), compared to the TAM<sup>-</sup> and KRAS<sup>-</sup> group (Fig. 3C). No significant association was observed between KRAS mutation status and either TIL density (median difference = 10.15, 95% CI: -95.77 to 93.78, p = 0.8842) or TAM density (median difference = -143.26, 95% CI: -393.96 to 133.75, p = 0.5658).

In the UCLH-RC, the KRAS mutation status is only available for 18/70 patients and does not significantly differ in DFS (HR = 2.31, 95% CI: 0.61–8.68, p = 0.2039) and OS (HR = 1.33, 95% CI: 0.36–4.94, p = 0.6656) (Figure S7). No death or disease progression is observed in TIL<sup>+</sup> and KRAS<sup>-</sup> patients (Figure S8). TAM<sup>+</sup> and KRAS<sup>+</sup> patients had significantly shorter DFS (HR = 1.87, 95% CI: 0.92–3.78, p = 0.0494), compared to the TAM<sup>-</sup> and KRAS<sup>-</sup> group (Figure S8). In the combined cohort comprising ARISTOTLE-RC B1 and 18 patients from UCLH-RC with known KRAS mutation status, TIL<sup>+</sup> and KRAS<sup>-</sup> patients had significantly longer DFS (HR = 0.32, 95% CI: (0.18–0.58), p < 0.0001) and OS (HR = 0.23, 95% CI:



**Fig. 2: The Kaplan-Meier curves of DFS and OS of the patients stratified by TIL, TAM density and MI in the ARISTOTLE-RC B1 (A-K), UCLH-RC (G-L) and TCGA-CRC (M-O). TIL<sup>+/−</sup>, TAM<sup>+/−</sup> and MI<sup>+/−</sup> represent the patients with high (+)/low (−) levels of TIL, TAM densities and MI, respectively. p-values are from log-rank tests. The number of patients at risk in each subgroup is shown below the curves and the number of censored patients is indicated in parentheses.**

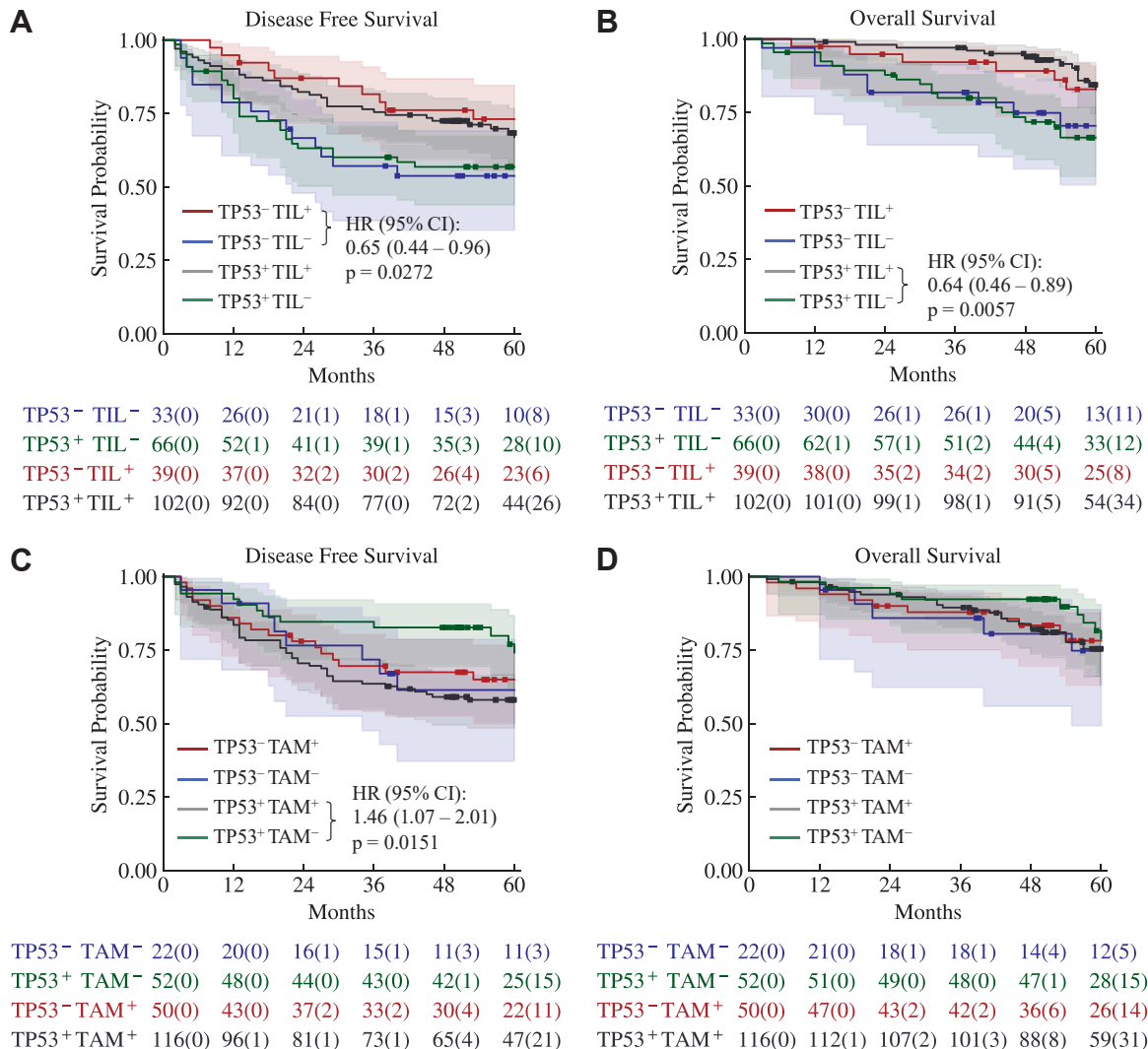


**Fig. 3: The Kaplan-Meier curves of DFS and OS of the four subgroups stratified by KRAS mutation and TIL/TAM density in the ARISTOTLE-RC B1.** A. DFS of the subgroups stratified by KRAS and TIL. B. OS of the subgroups stratified by KRAS and TIL. C. DFS of the subgroups stratified by KRAS and TAM. D. OS of the subgroups stratified by KRAS and TAM. p-values are from log-rank tests. The number of patients at risk in each subgroup is shown below the curves and the number of censored patients is indicated in parentheses.

0.11–0.48,  $p < 0.0001$ ), compared to the TIL<sup>-</sup> and KRAS<sup>+</sup> group (Figure S9a and b). TAM<sup>+</sup> and KRAS<sup>+</sup> patients had significantly shorter DFS (HR = 1.37, 95% CI: 1.08–1.74,  $p = 0.0066$ ) and tended to have shorter OS (HR = 1.26, 95% CI: 0.96–1.66,  $p = 0.0920$ ), compared to the TAM<sup>-</sup> and KRAS<sup>-</sup> group (Figure S9c and D).

We also assessed the prognostic impact of TIL and TAM in the context of TP53 mutation status in ARISTOTLE-RC. No significant difference in DFS (HR = 1.02, 95% CI: 0.64–1.61,  $p = 0.9363$ ) or OS (HR = 0.87, 95% CI: 0.49–1.55,  $p = 0.6368$ ) was observed between TP53<sup>-</sup> and TP53<sup>+</sup> patients (Figure S10). No significant association was found

between TP53 mutation status and either TIL density (median difference = 63.19, 95% CI: -21.67 to 167.83,  $p = 0.2640$ ) or TAM density (median difference = 15.11, 95% CI: -348.63 to 230.74,  $p = 0.9701$ ). TP53<sup>-</sup> and TIL<sup>+</sup> had significantly improved DFS (HR = 0.65, 95% CI: 0.44–0.96,  $p = 0.0272$ ) and tended to have longer OS (HR = 0.66, 95% CI: 0.40–1.10,  $p = 0.0991$ ), compared to TP53<sup>-</sup> and TIL<sup>-</sup> patients. TP53<sup>+</sup> and TIL<sup>+</sup> patients tended to have longer DFS (HR = 0.83, 95% CI: 0.64–1.06,  $p = 0.1355$ ) and had significantly longer OS (HR = 0.64, 95% CI: 0.46–0.89,  $p = 0.0057$ ), compared to TP53<sup>+</sup> and TIL<sup>-</sup> patients (Fig. 4A and B). Notably, TP53<sup>+</sup> and TAM<sup>+</sup> patients showed significantly shorter DFS (HR = 1.46, 95% CI: 1.07–2.01,  $p = 0.0151$ ),



**Fig. 4: The Kaplan-Meier curves of DFS and OS of the four subgroups stratified by TP53 mutation and TIL/TAM density in the ARISTOTLE-RC B1.** A. DFS of the subgroups stratified by TP53 and TIL. B. OS of the subgroups stratified by TP53 and TIL. C. DFS of the subgroups stratified by TP53 and TAM. D. OS of the subgroups stratified by TP53 and TAM. p-values are from log-rank tests. The number of patients at risk in each subgroup is shown below the curves and the number of censored patients is indicated in parentheses.

compared to TP53<sup>+</sup> and TAM<sup>-</sup> patients. In contrast, among TP53 wild-type patients, no difference in DFS (HR = 1.00, 95% CI: 0.66–1.52, p = 0.9930) was observed between TAM<sup>+</sup> and TAM<sup>-</sup> subgroups (Fig. 4C). No significant difference in OS was observed across four subgroups (Fig. 4D). TP53 mutation status is not available in UCLH-RC.

**Multivariable cox regression and added predictive value of TIME markers**

Table 3 present the results of the multivariable Cox regression models incorporating TIL, TAM, and other pre-treatment baseline clinical variables. An additional analysis incorporating postoperative pathological stage,

tumour location, and histological differentiation is presented in Table S8, although including pathological stage violated the proportional hazards assumption.

Table 4 shows that the C-index derived from 10-fold cross-validation improved for both disease-free survival (DFS) and overall survival (OS) upon inclusion of the AI-identified TIME markers (TIL, TAM, and MI) in the Cox regression model. A likelihood ratio test (LRT) was also conducted to compare the full Cox model to a reduced model without the TIME markers. The results demonstrated that inclusion of these variables significantly improved model fit for both DFS (LRT  $\chi^2 = 24.81$ , p < 0.0001) and OS (LRT  $\chi^2 = 22.46$ , p < 0.0001).

Variable	ARISTOTLE-RC		UCLH-RC		Combined Cohort	
	DFS (HR [95% CI])	OS (HR [95% CI])	DFS (HR [95% CI])	OS (HR [95% CI])	DFS (HR [95% CI])	OS (HR [95% CI])
Age (>60)	1.33 (0.85–2.09) p = 0.2144	2.09 (1.13–3.84) p = 0.0182	0.41 (0.12–1.40) p = 0.1558	1.79 (0.57–5.65) p = 0.3208	1.21 (0.82–1.79) p = 0.3397	2.02 (1.23–3.31) p = 0.0052
Sex (Male)	1.06 (0.66–1.71) p = 0.8060	0.70 (0.38–1.28) p = 0.2451	2.31 (0.82–6.50) p = 0.1120	0.94 (0.38–2.31) p = 0.8961	1.16 (0.77–1.75) p = 0.4800	0.79 (0.49–1.27) p = 0.3235
MRI T stage <sup>a</sup>						
T3	1.29 (0.46–3.64) p = 0.6336	0.96 (0.28–3.25) p = 0.9481	0.20 (0.03–1.19) p = 0.0760	1.47 (0.30–7.34) p = 0.6363	0.99 (0.46–2.13) p = 0.9723	0.97 (0.40–2.37) p = 0.9521
T4	1.99 (0.64–6.19) p = 0.2369	1.11 (0.28–4.37) p = 0.8763	0.70 (0.11–4.32) p = 0.6968	1.93 (0.39–9.45) p = 0.4185	1.65 (0.70–3.91) p = 0.2560	1.33 (0.49–3.65) p = 0.5758
MRI N stage <sup>a</sup>						
N1	0.75 (0.42–1.34) p = 0.3336	0.88 (0.40–1.95) p = 0.7552	0.96 (0.26–3.58) p = 0.9507	0.88 (0.29–2.64) p = 0.8134	0.76 (0.46–1.26) p = 0.2904	0.91 (0.49–1.68) p = 0.7632
N2	0.73 (0.39–1.37) p = 0.3302	1.05 (0.45–2.45) p = 0.9172	0.74 (0.16–3.48) p = 0.7079	0.71 (0.21–2.44) p = 0.5901	0.68 (0.39–1.18) p = 0.1737	0.92 (0.47–1.78) p = 0.8029
MRI circumferential Resection Margin (Involved) <sup>a</sup>	0.55 (0.27–1.10) p = 0.0904	0.70 (0.27–1.79) p = 0.4499	1.78 (0.57–5.52) p = 0.3180	2.03 (0.66–6.30) p = 0.2191	0.94 (0.53–1.67) p = 0.8253	1.03 (0.52–2.04) p = 0.9289
MRI extramural vascular invasion (Present) <sup>a</sup>	0.74 (0.24–2.26) p = 0.5935	2.23 (0.59–8.42) p = 0.2370	–	–	–	–
MRI extramural venous invasion (Present) <sup>a</sup>	3.05 (1.00–9.27) p = 0.0498	1.16 (0.31–4.29) p = 0.8250	8.13 (1.83–36.16) p = 0.0059	1.47 (0.48–4.50) p = 0.4980	2.31 (1.53–3.48) p = 0.0001	2.07 (1.26–3.40) p = 0.0039
KRAS (mutated)	1.33 (0.84–2.11) p = 0.2189	1.52 (0.83–2.77) p = 0.1742	–	–	–	–
TP53 (mutated)	1.29 (0.79–2.09) p = 0.3060	1.10 (0.59–2.07) p = 0.7552	–	–	–	–
TIL (high)	0.60 (0.39–0.92) p = 0.0204	0.41 (0.23–0.73) p = 0.0026	0.97 (0.23–4.10) p = 0.9706	0.48 (0.12–1.93) p = 0.3042	0.59 (0.40–0.87) p = 0.0076	0.37 (0.23–0.61) p = 0.0001
TAM (high)	1.47 (0.87–2.48) p = 0.1487	0.86 (0.46–1.62) p = 0.6462	3.60 (1.17–11.09) p = 0.0258	1.77 (0.66–4.77) p = 0.2583	1.55 (0.99–2.43) p = 0.0576	1.00 (0.60–1.66) p = 0.9915
MI (high)	2.83 (1.57–5.11) p = 0.0005	2.65 (1.23–5.73) p = 0.0132	0.71 (0.16–3.11) p = 0.6543	1.76 (0.49–6.36) p = 0.3900	2.69 (1.62–4.49) p = 0.0001	2.45 (1.33–4.52) p = 0.0040

<sup>a</sup>From the baseline MRI performed before the start of nCRT.

**Table 3: Multivariable cox regression results with baseline clinical variables.**

### Characterising dynamics of immune cell density after nCRT

In ARISTOTLE-RC B2, 89 cases showed re-stratification after the treatment, while 31 patients had persistently low immune infiltration (TIL<sup>-</sup>, Fig. 5A) and exhibited a significantly shorter DFS. Patients undergoing treatment-related immune activation (TIL<sup>+</sup>) appeared to have significantly improved DFS (HR = 0.70, 95% CI: 0.50–0.97, p = 0.0280), compared to the TIL<sup>-</sup> patients (Fig. 5B). Patients who experienced treatment-related immune suppression (TIL<sup>+</sup>) also exhibited a significantly higher density of mitotic figures compared to the patient who experienced treatment-related immune activation (TIL<sup>+</sup>) (median difference = 9.36, 95% CI: 1.87–16.85, p = 0.0385, Fig. 5C).

The TAM density also changes dynamically after the treatment, with 44 TAM<sup>+</sup> patients and 23 TAM<sup>-</sup> patients (Figure S11). Significantly shorter DFS was observed in the 121 TAM<sup>+</sup> patients (HR = 6.62, 95% CI: 1.02–48.14, p = 0.0302), compared to the 14 TAM<sup>-</sup> patients. TAM<sup>+</sup> patients also tended to have a higher

density of mitotic figures, while not as significant (median difference = 6.96, 95% CI: 1.99–15.92, p = 0.2272).

### Discussion

We developed an AI framework for quantitatively analysing the tumour immune landscape and mitotic activity in digitised H&E-stained WSIs. The framework consists of a deep-learning-based image classifier to first identify the tumour-associated regions by tissue classification and a deep-learning-based object detector to detect the lymphocytes, macrophages, and mitotic figures. These measurements are used to calculate the density of TIL, TAM and MI. We applied this model to three CRC cohorts to stratify patients based on these features. Using the TCGA-CRC, we confirmed that our AI-derived TIL and TAM densities were associated with patient survival in a manner consistent with previous studies derived from the same cohort. Moreover, we specifically explored the prognostic relevance and post-treatment dynamics of TIME metrics and their interaction

Model	DFS (C-index [95% CI])	OS (C-index [95% CI])	DFS (log-likelihood)	OS (log-likelihood)
Baseline clinical variables <sup>a</sup>	0.60 (0.52-0.68)	0.60 (0.54-0.67)	-591.73	-413.21
Baseline clinical variables + TIL, TAM and MI	0.65 (0.59-0.71)	0.66 (0.61-0.70)	-579.33	-401.98

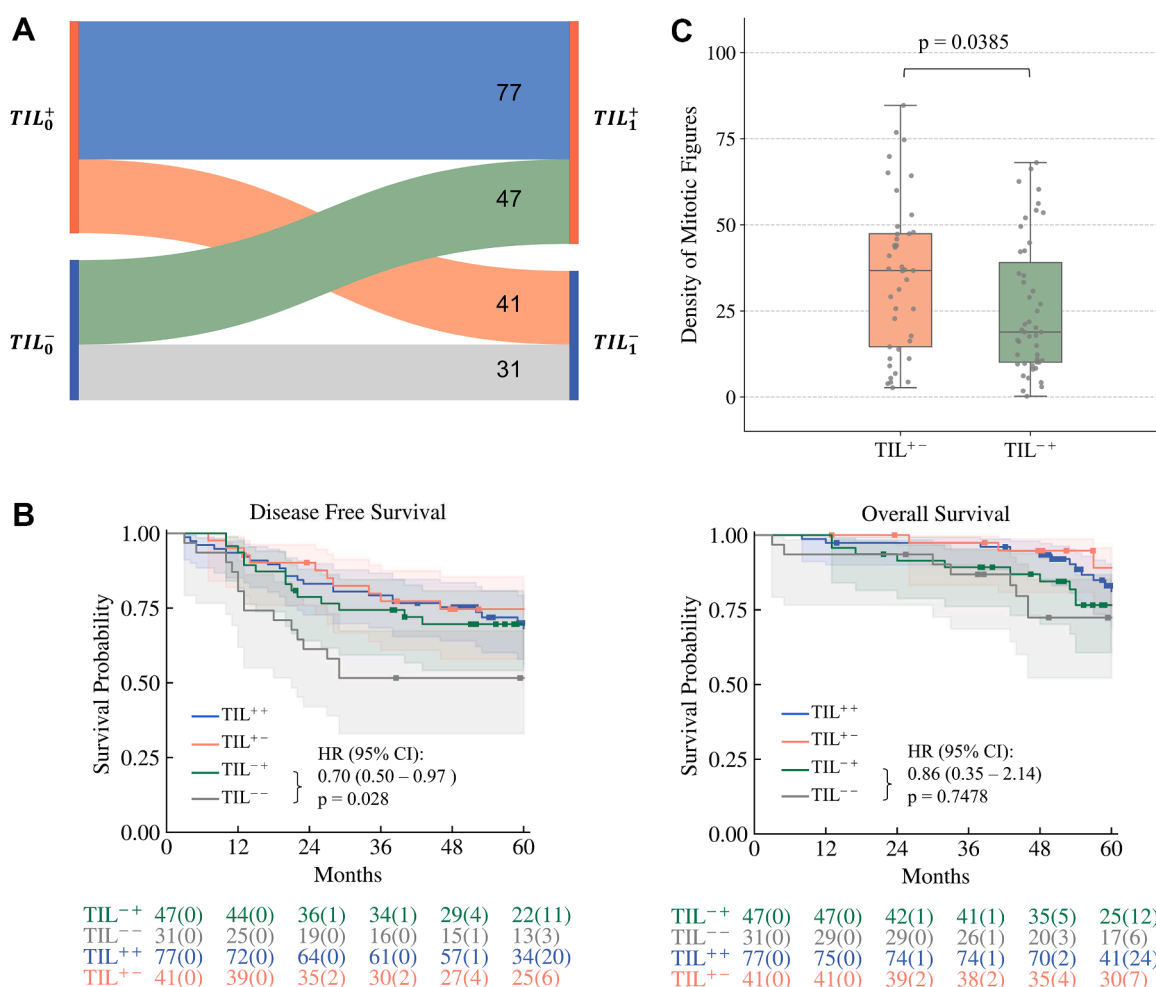
<sup>a</sup>Clinical Variables presented in Table 3.

**Table 4: C-index of cox regression models with versus without TIME markers.**

with DNA mutations in LARC patients following nCRT using a clinical trial cohort (ARISTOTLE-RC) and a real-world hospital cohort (UCLH-RC). To promote accessibility and clinical translation, we have developed a web-based platform (<https://octopath.ai/>) for automatically generating TIME metrics from uploaded slides.

### The prognostic impact of TIL and TAM

The AI-driven immune profiling reveals that both ARISTOTLE and UCLH cohorts displayed a strong association between high pre-treatment TIL density and longer survival in both the univariable and multivariable analyses. The prognostic impact of TIL has been widely reported. High TIL density is associated with



**Fig. 5: The patient re-stratification after the nCRT.** A. The four re-stratification routes and the numbers in each group; B. The Kaplan-Meier curves of DFS (left) and OS (right) of TIL<sup>++</sup>, TIL<sup>+-</sup>, TIL<sup>-+</sup> and TIL<sup>--</sup>. p-values are from log-rank tests. The number of patients at risk in each subgroup is shown below the curves and the number of censored patients is indicated in parentheses. C. The density of mitotic figures in the TIL<sup>+-</sup> and TIL<sup>-+</sup> groups. The densities are presented using box plots with individual data points (N = 41 [left] and 47 [right] independent patients). p-value is from the Mann-Whitney U test.

longer DFS and OS in CRC patients with primary stage,<sup>16</sup> including LARC.<sup>50,51</sup> However, our study did not separate lymphocyte subtypes, for which the impact of their proportions, especially the cytotoxic T-cells, needs to be further investigated.

Macrophage infiltration has emerged as a research hotspot in recent years due to its potential role in tumour prognosis; however, studies in CRC have reported controversial findings.<sup>18–23</sup> Notably, although only a limited number of studies have focused exclusively on rectal cancer cohorts, all of them observed a negative prognostic impact.<sup>25,26,52</sup> One possible explanation is that CRT is commonly used in patients with rectal cancer, yet no study has specifically shown the impact of TAM in CRT-treated cohorts. In this study, similar results were found in the two CRT-treated rectal cancer cohorts, where high TAM density is associated with worse DFS in the ARISTOTLE-RC and OS in the UCLH-RC. The association between TAM and DFS remains significant in the multivariable analysis, indicating the predictive value of TAM on disease progression. We also found that TAM<sup>+</sup> patients had worse OS in the TCGA-CRC, which is consistent with the previous study.<sup>24</sup> Future analyses should include rectal cancer cohorts not treated with CRT to better understand the influence of CRT on TAM behaviour. High infiltration of M1-like macrophages is generally associated with improved survival in CRC, while M2-like macrophages tend to promote tumour progression and are linked to worse outcomes.<sup>53,54</sup> However, the AI framework used in this study detects aggregated macrophages. Future studies are needed to investigate the prognostic significance of specific TAM subtypes.

#### Immune activity and DNA mutation

We found that *KRAS*-mutated patients with low TIL density experience significantly worse outcomes in both DFS and OS, compared to *KRAS*-wildtype patients with high TIL density. While high TIL density has already been established as a favourable prognostic factor, the *KRAS* mutation provides an additive prognostic value. Similarly, *KRAS* mutation status and TAM levels also demonstrated an additive impact on prognosis. *KRAS* mutations occur in approximately 30%–50% of CRC patients and have been associated with poor clinical outcomes.<sup>55</sup> However, few studies have explored the co-impact of the *KRAS* mutation and TIL density in CRC<sup>56,57</sup> and data on DFS and OS stratified by both TIL density and *KRAS* status remain scarce. A study by Liu, H. et al. indicated that the impact of TAMs on OS is more significant in *KRAS*-mutated tumours.<sup>58</sup>

Studies reporting the association between *KRAS* mutation and immune cell densities in CRC are also limited. Fu, X. et al. reported a potential association between *KRAS* mutations and lower TIL density,<sup>59</sup> while Mojarad, E. et al. observed a correlation between *KRAS* mutations and stromal lymphocyte

density, rather than intra-tumoral TILs.<sup>57</sup> Liu, H. et al. found a positive correlation between *KRAS* mutation and TAM expression.<sup>58</sup> However, in our cohorts, neither TAM nor TIL densities were associated with *KRAS* mutation status. Further studies are required to investigate the interactions between *KRAS* mutations and immune activity within TIME.

We also found that high TAM density was associated with significantly worse DFS specifically in *TP53*-mutated patients. This observation can be explained by the findings of Cooks et al.,<sup>60</sup> who reported that mutant p53 proteins, particularly those with gain-of-function (GOF) mutations, can reprogramme macrophages toward a tumour-promoting phenotype. Although our study did not include molecular subtyping of *TP53* mutations (GOF versus non-GOF), we still observed a differential impact of TAM density in *TP53*-mutated versus *TP53*-wildtype subgroups. This supports the hypothesis that *TP53* mutation status may modulate macrophage behaviour, and the combination of *TP53* and the AI-identified TAM could serve as risk stratification factors for LARC patients undergoing nCRT.

The absence of multi-omics data in the UCLH-RC cohort restricts the ability to independently validate molecular and immune-related findings derived from the ARISTOTLE-RC cohort. Future work will involve whole-genome sequencing and transcriptomic profiling of retrospective samples for comprehensive validation.

#### Immune activity, CRT and tumour proliferation

By comparing the TIL density before and after nCRT, we observed that patients with consistently pre-/post-treatment low TIL levels (TIL<sup>-</sup>) had the poorest outcomes. Noticeably, TIL<sup>+</sup> patients, who showed increased immune activity post-treatment, demonstrated significantly improved DFS. No significant difference was found between the TIL<sup>++</sup> and TIL<sup>+-</sup> patients, suggesting that patients with an active immune response before the treatment tend to have better outcomes, even if it is suppressed by the treatment. While only 14 patients exhibited consistently low TAM density (TAM<sup>-</sup>), this group had significantly longer DFS and tended to have better OS, with no deaths recorded.

Several studies have reported that nCRT could activate anti-tumour immunity by increasing the number of CD8<sup>+</sup> TILs, which is associated with better outcomes.<sup>61–63</sup> The CRT-induced immune activity is also related to microsatellite instability (MSI) status, methylation level and HLA class I expression.<sup>64–66</sup> However, all the patients in ARISTOTLE-RC B2 are microsatellite-stable, and only 1% of them show HLA class I expression, leaving us unable to validate those findings in this study. No association was found between the change in TIL and TAM density and methylation value or the DNA mutations of 80 genes (Table S10).

Of note, we found that patients with increased TIL post-treatment had significantly higher mitotic index (MI) compared to those with decreased TIL. A similar pattern was seen in TAM dynamics, with higher mean MI in TAM<sup>+</sup> patients than in TAM<sup>-</sup> patients. High tumour proliferative activity may suppress immune responses after nCRT. Specht, J. reported increased TIL and TAM infiltration in breast cancer with G2/M arrest.<sup>67</sup> However, further research is needed to understand how tumour proliferation interacts with immune activation in CRC.

MI itself appeared prognostic in this study. MI<sup>+</sup> patients had significantly shorter DFS and OS in ARISTOTLE-RC and tended toward shorter OS in TCGA-CRC. A limited number of studies have reported the impact of MI on CRC. The study from Li, B. et al. reported that the cell mitosis was not associated with tumour regression grade in 106 CRC patients after nCRT, while long-term survival was not compared.<sup>68</sup> Notably, the threshold used to define MI<sup>+</sup> status in our study was relatively high, resulting in only 10% and 6% of patients being classified as MI<sup>+</sup> in the ARISTOTLE-RC and TCGA-CRC cohorts, respectively. Furthermore, no MI<sup>+</sup> cases were observed in the UCLH-RC cohort, likely due to the small sample size and the staining and scanning variability across institutions. We have recalibrated the cut-off for the UCLH-RC cohort, designating as MI<sup>+</sup> those patients whose mitotic index falls within the top 10%, mirroring the proportion used in the ARISTOTLE-RC cohort. However, recalibration may not be ideal as this assumes that mitotic activity distributions should be directly comparable despite differences in sample size, tumour biology, or technical factors. Future studies with larger cohorts are planned to further evaluate the prognostic role of tumour mitosis.

### Advantages and disadvantages of AI-powered TIME analysis

Manual quantification of immune cells within TIME is labour-intensive, subject to large intra-observer variability, and time-consuming, posing limitations for large-scale studies.<sup>69,70</sup> In contrast, AI models offer rapid, consistent, and scalable analysis. Additionally, AI-based approaches improve robustness and reproducibility, addressing the variability often observed among pathologists in grading immune cell infiltration or assessing mitotic activity.

Noticeably, the AI-identified macrophage densities in this study were lower than commonly reported values in the literature.<sup>71</sup> This discrepancy can be attributed to two main factors: we computed average densities across the entire tumour and tumour-associated stroma, rather than hotspot regions or the invasive margin, which typically show higher localised macrophage infiltration; many studies used tissue microarrays or selected regions of interest,<sup>72</sup> while our

whole-slide analysis captures both high- and low-density areas, resulting in a more conservative overall estimate.

Distinguishing immune cell subtypes from H&E-stained WSIs is challenging, which restricts the analysis of subtype-specific immune markers. In other studies, specific TIL subsets, such as cytotoxic, helper, and regulatory T cells, have been identified using immunohistochemical staining with markers including CD8, CD4, and FOXP3, respectively.<sup>16</sup> Similarly, efforts have been made to classify M1/M2-like macrophages using molecular markers.<sup>18</sup> Commonly reported markers include CD80, CD86, and iNOS for M1-like macrophages, and CD163, CD206, and Arginase-1 for M2-like macrophages.<sup>73</sup> However, macrophage subtyping remains inherently challenging due to their remarkable plasticity and the existence of a spectrum of activation states, rather than discrete subsets.<sup>74</sup>

Despite this, numerous studies have demonstrated that aggregate immune cell density, particularly total TAM burden, retains prognostic and predictive value even without subclassification.<sup>75</sup> Additionally, total TAM density has also been linked to response to immune checkpoint blockade in various cancers,<sup>76</sup> highlighting the clinical significance of quantifying general TAM density. Moreover, our approach offers a practical and clinically scalable solution for profiling the TIME, especially in resource-limited settings or retrospective cohorts where advanced multiplexed assays are not available. In future studies, we plan to integrate AI with subtype-specific techniques such as multiplex fluorescence imaging or spatial transcriptomics.

The quality control (QC) process in this study relied on manual review, which, while effective for the two LRAC cohorts, is insufficient for large-scale cohorts such as TCGA. Although staining and spatial augmentations were applied during model development to enhance generalisability, the lack of QC in TCGA-CRC may still introduce bias in AI-driven quantification. Scalable automated QC pipelines will be incorporated in future studies. Another limitation is that pre-treatment biopsies may not capture the full extent of tumour heterogeneity, particularly in terms of immune infiltration and spatial architecture. However, pre-treatment pathological assessment typically relies on biopsies, as post-surgical resections are often altered by neoadjuvant therapies, making them less suitable for baseline biomarker evaluation. Future studies will aim to use multiple biopsies from different tumour regions.

In conclusion, this study investigated the prognostic impact of AI-identified immune cell densities and mitotic index in three independent cohorts, comprising one all-stage CRC cohort and two LARC cohorts treated with nCRT. We found that high TIL density and low TAM density were associated with improved DFS and OS, and their impact was modulated by DNA mutation

status. Additionally, patients whose TIL density increased following nCRT had improved DFS. A potential link between tumour proliferation and immune suppression was also discussed. These findings support the integration of immune and proliferative biomarkers into future LARC risk stratification and nCRT treatment planning.

Future work will focus on external validation and clinical translation. To support that, we have developed a web-based platform (<https://octopath.ai/>) that enables users to upload H&E-stained whole-slide images and apply our AI models for TIME assessment. A prospective clinical study to evaluate the real-world impact of this tool in routine pathology workflows is planned. In parallel, we aim to apply our models to immunotherapy-treated cohorts to assess whether AI-identified TIME features from H&E slides can stratify patients by treatment response and identify those most likely to benefit from immunotherapy.

#### Contributors

CACF, MAH and ZS designed the study, with input from the ARISTOTLE trial team (DS, AL, RB). ZS, MS and CACF developed the AI framework. ZS analysed the data and wrote the draft manuscript. DB, SH, MAH contributed to the clinical interpretation. NW provided the digital pathology data of the ARISTOTLE trial cohort. TSM provided the NGS data of the ARISTOTLE trial cohort. APL, TM, DO provided the digital pathology data of the UCLH-RC cohort. All the authors contributed to the manuscript revision and were involved in the scientific discussion. All authors read and approved the final version of the manuscript. ZS, CACF, MAH, DB had full access to and verified the integrity of the underlying data.

#### Data sharing statement

The clinical and pathology data of the trial participants are not publicly available due to restrictions by the ethical approval of this study. To get access to the ARISTOTLE dataset and the UCLH dataset, a signed data access agreement ensuring compliance with ethical guidelines, approval from the institutional review board, and submission of a detailed proposal outlining the intended use of the data will be needed. Requests will be considered on a case-by-case basis and should be directed to and the clinical trial unit ([ctc.aristotle@ucl.ac.uk](mailto:ctc.aristotle@ucl.ac.uk)) and UCL/UCLH Biobank ([ci.bbhad-admin@ucl.ac.uk](mailto:ci.bbhad-admin@ucl.ac.uk)) at University College London. The TCGA data is available via Genomic Data Commons Data Portal (<https://portal.gdc.cancer.gov/>). The codes and parameters of the AI models are available on GitHub: <https://github.com/cacof1/DigitalPathologyAI>. The AI models for cell detection can be inferred via <https://octopath.ai/>. The datasets used for developing the AI models are available on Zenodo: Tissue Classification: <https://zenodo.org/records/1214456>, Immune Cell Detection: <https://zenodo.org/records/11073373> and Mitotic Figure Detection: <https://zenodo.org/records/14246170>. For any other requests and queries relating to data sharing, please contact the corresponding author [zhuoyan.shen.18@ucl.ac.uk](mailto:zhuoyan.shen.18@ucl.ac.uk).

#### Declaration of interests

TSM received a consulting fee from Nordic Pharma and was the Interim Chief Executive of the National Cancer Research Institute (NCRI) in 2022. NPW is supported by grants from Yorkshire Cancer Research, GSK, Pierre Fabre, Roche Diagnostics and NCRI, and received consulting fees from Bristol Myers, Squibb, Astellas, Pfizer, GSK, Amgen, Servier, and Beigene. He also has patents planned, issued, or pending with Roche Diagnostics. SH is supported by the Radiation Research Unit at the Cancer Research UK City of London Centre.

#### Acknowledgements

This research was funded by Cancer Research UK Radiation Research Network-Project Seed Funding (RRNPSF-Jan21/100001), UK Research and Innovation Future Leadership Fellowship (No. MR/T040785/1), the Radiation Research Unit at the Cancer Research UK City of London Centre Award (C7893/A2899) and The Pathological Society of Great Britain and Ireland. We thank the UCL/UCLH Biobank for Studying Health and Disease for the provision of human tissue samples and clinical data. We acknowledge the contribution of all members of the ARISTOTLE Trial Management Group for delivering the trial and providing the data. Maria Hawkins is supported by the National Institute for Health and Care Research University College London Hospitals Biomedical Research Centre. Nicholas West is funded by a Cancer Research UK Radiation Research Network Infrastructure Award (RRNIA-Feb22/100003) and is supported in part by the National Institute for Health and Care Research (NIHR) Leeds Biomedical Research Centre (BRC) (NIHR203331). The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care. David Sebag-Montefiore and Nick West would like to acknowledge Cancer Research UK funding for the Leeds Radiotherapy Research Centre of Excellence (RadNet; RRCOER-Jun24/100004 and C19942/A28832). The ARISTOTLE trial is funded by Cancer Research UK C19942/A10016. The ARISTOTLE trial sample collection and slide scanning were funded by Cancer Research UK (A18745).

#### Appendix A. Supplementary data

Supplementary data related to this article can be found at <https://doi.org/10.1016/j.ebiom.2025.105993>.

#### References

- Bray F, Laversanne M, Sung H, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2024;74(3):229–263.
- Annual reports and strategy documents. Bowel Cancer UK; 2024. Available from: <https://www.bowelcanceruk.org.uk/about-us/documents/>.
- Lawler M, Alsina D, Adams RA, et al. Critical research gaps and recommendations to inform research prioritisation for more effective prevention and improved outcomes in colorectal cancer. *Gut*. 2018;67(1):179–193. Available from: <https://gut.bmj.com/content/gutjnl/67/1/179.full.pdf>.
- Ma B, Gao P, Wang H, et al. What has preoperative radio(chemo)therapy brought to localized rectal cancer patients in terms of perioperative and long-term outcomes over the past decades? A systematic review and meta-analysis based on 41,121 patients. *Int J Cancer*. 2017;141(5):1052–1065.
- Mendis S, To YH, Tie J. Biomarkers in locally advanced rectal cancer: a review. *Clin Colorectal Cancer*. 2022;21(1):36–44.
- de Wilt J, Vermaas M, Ferenschild F, Verhoef C. Management of locally advanced primary and recurrent rectal cancer. *Clin Colon Rectal Surg*. 2007;20(3):255–263.
- Benson AB, Venook AP, Al-Hawary MM, et al. Rectal cancer, version 2.2022. NCCN clinical practice guidelines in oncology. *J Natl Compr Canc Netw*. 2022;20(10):1139–1167. Available from: <https://jncn.org/view/journals/jncn/20/10/article-p1139.xml>.
- Conroy T, Bosset JF, Etienne PL, et al. Neoadjuvant chemotherapy with FOLFIRINOX and preoperative chemoradiotherapy for patients with locally advanced rectal cancer (UNICANCER-PRODIGE 23): a multicentre, randomised, open-label, phase 3 trial. *Lancet Oncol*. 2021;22(5):702–715.
- Bahadoer RR, Dijkstra EA, van Etten B, et al. Short-course radiotherapy followed by chemotherapy before total mesorectal excision (TME) versus preoperative chemoradiotherapy, TME, and optional adjuvant chemotherapy in locally advanced rectal cancer (RAPIDO): a randomised, open-label, phase 3 trial. *Lancet Oncol*. 2021;22(1):29–42. Available from: [https://www.thelancet.com/journals/lanonc/article/PIIS1470-2045\(20\)30555-6/fulltext](https://www.thelancet.com/journals/lanonc/article/PIIS1470-2045(20)30555-6/fulltext).
- Garcia-Aguilar J, Patil S, Gollub MJ, et al. Organ preservation in patients with rectal adenocarcinoma treated with total neoadjuvant therapy. *J Clin Oncol*. 2022;40(23):2546–2556.
- Cercek A, Lumish M, Sinopoli J, et al. PD-1 blockade in mismatch repair-deficient, Locally Advanced Rectal Cancer. *N Engl J Med*. 2022;386(25).

- 12 Baghban R, Roshangar L, Jahanban-Esfahlan R, et al. Tumor microenvironment complexity and therapeutic implications at a glance. *Cell Commun Signal*. 2020;18(1):59. Available from: <https://biosignaling.biomedcentral.com/articles/10.1186/s12964-020-0530-4>.
- 13 Ruffini E, Asiola S, Filosso PL, et al. Clinical significance of tumor-infiltrating lymphocytes in lung neoplasms. *Ann Thorac Surg*. 2009;87(2):365–372 [cited 2023 May 8] Available from: <https://www.sciencedirect.com/science/article/abs/pii/S0003497508022856>.
- 14 Santoiemma PP, Powell DJ. Tumor infiltrating lymphocytes in ovarian cancer. *Cancer Biol Ther*. 2015;16(6):807–820. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4622931/>.
- 15 Stanton SE, Disis ML. Clinical significance of tumor-infiltrating lymphocytes in breast cancer. *J Immunother Cancer*. 2016;4(1):59.
- 16 Kong JC, Guerra GR, Pham T, et al. Prognostic impact of tumor-infiltrating lymphocytes in primary and metastatic colorectal cancer. *Dis Colon Rectum*. 2019;62(4):498–508.
- 17 Orhan A, Khesrawi F, Tvilling Madsen M, et al. Tumor-infiltrating lymphocytes as biomarkers of treatment response and long-term survival in patients with rectal cancer: a systematic review and meta-analysis. *Cancers*. 2022;14(3):636.
- 18 Noy R, Pollard J. Tumor-associated macrophages: from mechanisms to therapy. *Immunity*. 2014;41(1):49–61.
- 19 Pinto ML, Rios E, Durães C, et al. The two faces of tumor-associated macrophages and their clinical significance in colorectal cancer. *Front Immunol*. 2019;10:1875. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6710360/>.
- 20 Chaput N, Srcek M, Aupérin A, et al. Tumour-infiltrating CD68+ and CD57+ cells predict patient outcome in stage II–III colorectal cancer. *Br J Cancer*. 2013;109(4):1013–1022.
- 21 Edin S, Wikberg ML, Dahlin AM, et al. The distribution of macrophages with a M1 or M2 phenotype in relation to prognosis and the molecular characteristics of colorectal cancer. *Najbauer J*, ed. *PLoS One*. 2012;7(10):e47045.
- 22 Herrera M, Herrera A, Domínguez G, et al. Cancer-associated fibroblast and M2 macrophage markers together predict outcome in colorectal cancer patients. *Cancer Sci*. 2013;104(4):437–444.
- 23 Shibutani M, Maeda K, Nagahara H, et al. The peripheral monocyte count is associated with the density of tumor-associated macrophages in the tumor microenvironment of colorectal cancer: a retrospective study. *BMC Cancer*. 2017;17(1):404.
- 24 Liu Q, Liao L. Identification of macrophage-related molecular subgroups and risk signature in colorectal cancer based on a bioinformatics analysis. *Autoimmunity*. 2024;57(1):2321908 [cited 2024 May 23] Available from: <https://pubmed.ncbi.nlm.nih.gov/38466182/>.
- 25 Shabo I, Olsson H, Sun X, Svanvik J. Expression of the macrophage antigen CD163 in rectal cancer cells is associated with early local recurrence and reduced survival time. *Int J Cancer*. 2009;125(8):1826–1831.
- 26 Nishi M, Yamashita S, Takasu C, et al. Role of mast cell in locally advanced rectal cancer treated with neoadjuvant chemoradiotherapy. *BMC Cancer*. 2025;25(1):99.
- 27 Jain SM, Nagainallur Ravichandran S, Murali Kumar M, et al. Understanding the molecular mechanism responsible for developing therapeutic radiation-induced radioresistance of rectal cancer and improving the clinical outcomes of radiotherapy - a review. *Cancer Biol Ther*. 2024;25(1):2317999.
- 28 Nagtegaal ID, Marijnen C, Kranenbarg EK, et al. Short-term pre-operative radiotherapy interferes with the determination of pathological parameters in rectal cancer. *J Pathol*. 2002;197(1):20–27 [cited 2024 Apr 17] Available from: <https://pubmed.ncbi.nlm.nih.gov/12081199/>.
- 29 Cianfrocca M, Goldstein LJ. Prognostic and predictive factors in early-stage breast cancer. *Oncologist*. 2004;9(6):606–616.
- 30 Demicco EG, Park MS, Araujo DM, et al. Solitary fibrous tumor: a clinicopathological study of 110 cases and proposed risk assessment model. *Mod Pathol*. 2012;25(9):1298–1306.
- 31 McCall CM, Shi C, Cornish TC, et al. Grading of well-differentiated pancreatic neuroendocrine tumors is improved by the inclusion of both Ki67 proliferative index and mitotic rate. *Am J Surg Pathol*. 2013;37(11):1671–1677.
- 32 Sinicrope FA, Hart J, Hai-An H, Lemoine M, Michelassi F, Stephens LC. Apoptotic and mitotic indices predict survival rates in lymph node-negative colon carcinomas. *Clin Cancer Res*. 1999;5(7):1793–1804.
- 33 Hilska M, Collan YU, O Laine VJ, et al. The significance of tumor markers for proliferation and apoptosis in predicting survival in colorectal cancer. *Dis Colon Rectum*. 2005;48(12):2197–2208.
- 34 Ullah M, Aoudjehout W, Pimpie C, Pocard M, Mirshahi M. Mitosis in cancer cell increases immune resistance via high expression of HLA-G and PD-L1. *Cancers*. 2020;12(9):2661.
- 35 GDC. Cancer.gov. Available from: <https://portal.gdc.cancer.gov/>; 2019.
- 36 Ruifrok AC, Johnston DA. Quantification of histochemical staining by color deconvolution. *Anal Quant Cytol Histol*. 2001;23(4):291–299.
- 37 Bewick V, Cheek L, Ball J. Statistics review 13: receiver operating characteristic curves. *Crit Care*. 2004;8(6):508–512.
- 38 Tan M, Le QV. EfficientNet: rethinking model scaling for convolutional neural networks. In: *36th International Conference on Machine Learning, ICML 2019*. 2019.
- 39 Kather JN, Halama N, Marx A. *100,000 histological images of human colorectal cancer and healthy tissue*. Zenodo; 2018.
- 40 Wang A, Chen H, Liu L, et al. YOLOv10: real-time end-to-end object detection. *arXiv.org*; 2024. Available from: <https://arxiv.org/abs/2405.14458>. Not peer-reviewed.
- 41 Simard M, Shen Z, Hawkins MA, Collins-Fekete CA. Immunoto: a massive immune cell database auto-generated for histopathology. *arXiv (Cornell University)*. 2024. Not peer-reviewed.
- 42 Shen Z, Simard M, Brand D, et al. A deep learning framework deploying segment anything to detect pan-cancer mitotic figures from haematoxylin and eosin-stained slides. *Commun Biol*. 2024;7(1):1674 [cited 2025 Mar 25] Available from: <https://www.nature.com/articles/s42003-024-07398-6>.
- 43 Tellez D, Balkenhol M, Otte-Holler I, et al. Whole-slide mitosis detection in H&E breast histology using PHH3 as a reference to train distilled stain-invariant convolutional networks. *IEEE Trans Med Imaging*. 2018;37(9):2126–2136.
- 44 Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *J Am Stat Assoc*. 1958;53(282):457–481. Available from: <https://www.jstor.org/stable/2281868?seq=1>.
- 45 Bland JM, Altman DG. The logrank test. *BMJ*. 2004;328(7447):1073.
- 46 Cox DR, Oakes D. *Analysis of survival data*. *Analysis of survival data*. 2018.
- 47 Plackett RL. Karl Pearson and the chi-squared test. *Int Stat Rev*. 1983;51(1):59.
- 48 Mann HB, Whitney DR. On a test of whether one of two random variables is stochastically larger than the other. *Ann Math Stat*. 1947;18(1):50–60. Available from: <https://projecteuclid.org/journals/annals-of-mathematical-statistics/volume-18/issue-1/On-a-Test-of-Whether-one-of-Two-Random-Variables/10.1214/aoms/1177730491.full>.
- 49 McKight PE, Najab J. *Kruskal-Wallis test*. *The corsini encyclopedia of psychology*; 2010. Available from: <https://onlinelibrary.wiley.com/doi/10.1002/9780470479216.corpsy0491>.
- 50 Anitei MG, Zeitoun G, Mlecnik B, et al. Prognostic and predictive values of the immunoscore in patients with rectal cancer. *Clin Cancer Res*. 2014;20(7):1891–1899.
- 51 Koukourakis IM, Platoni K, Tiniakos D, Kouloulis V, Zygogianni A. Immune response and immune checkpoint molecules in patients with rectal cancer undergoing neoadjuvant chemoradiotherapy: a review. *Curr Issues Mol Biol*. 2023;45(5):4495–4517.
- 52 Zhu M, Li X, Cheng X, et al. Association of the tissue infiltrated and peripheral blood immune cell subsets with response to radiotherapy for rectal cancer. *BMC Med Genomics*. 2022;15(Suppl 2):107.
- 53 Zhao Y, Ge X, Xu X, Yu S, Wang J, Sun L. Prognostic value and clinicopathological roles of phenotypes of tumour-associated macrophages in colorectal cancer. *J Cancer Res Clin Oncol*. 2019;145(12):3005–3019.
- 54 Li J, Li L, Li Y, et al. Tumor-associated macrophage infiltration and prognosis in colorectal cancer: systematic review and meta-analysis. *Int J Colorectal Dis*. 2020;35(7):1203–1210.
- 55 Andreyev HJN, Norman AR, Cunningham D, et al. Kirsten ras mutations in patients with colorectal cancer: the “RASCAL II” study. *Br J Cancer*. 2001;85(5):692–696.
- 56 Kocián P, Šedivcová M, Drgáč J, et al. Tumor-infiltrating lymphocytes and dendritic cells in human colorectal cancer: their relationship to KRAS mutational status and disease recurrence. *Hum Immunol*. 2011;72(11):1022–1028.
- 57 Nazemalhosseini-Mojarad E, Mohammadpour S, Torshizi Esafahani A, et al. Intratumoral infiltrating lymphocytes correlate with improved survival in colorectal cancer patients: independent of oncogenetic features. *J Cell Physiol*. 2018;234(4):4768–4777.

- 58 Liu H, Liang Z, Zhou C, et al. Mutant *KRAS* triggers functional reprogramming of tumor-associated macrophages in colorectal cancer. *Signal Transduct Target Ther*. 2021;6(1):144.
- 59 Fu X, Wang X, Duanmu J, Li T, Jiang Q. *KRAS* mutations are negatively correlated with immunity in Colon cancer. *Aging*. 2020;13(1):750–768.
- 60 Cooks T, Pateras IS, Jenkins LM, et al. Mutant p53 cancers reprogram macrophages to tumor supporting macrophages via exosomal miR-1246. *Nat Commun*. 2018;9(1):771.
- 61 Matsutani S, Shibutani M, Maeda K, et al. Significance of tumor-infiltrating lymphocytes before and after neoadjuvant therapy for rectal cancer. *Cancer Sci*. 2018;109(4):966–979.
- 62 Lim SH, Chua W, Cheng C, et al. Effect of neoadjuvant chemoradiation on tumor-Infiltrating/Associated lymphocytes in locally advanced rectal cancers. *Anticancer Res*. 2014;34(11):6505–6513.
- 63 Teng F, Mu D, Meng X, et al. Tumor infiltrating lymphocytes (TILs) before and after neoadjuvant chemoradiotherapy and its clinical utility for rectal cancer. *Am J Cancer Res*. 2015;5(6):2064–2074.
- 64 Noshu K, Baba Y, Tanaka N, et al. Tumour-infiltrating T-cell subsets, molecular changes in colorectal cancer, and prognosis: cohort study and literature review. *J Pathol*. 2010;222(4):350–366.
- 65 Hamada T, Soong TR, Masugi Y, et al. TIME (Tumor Immunity in the MicroEnvironment) classification based on tumor *CD274* (PD-L1) expression status and tumor-infiltrating lymphocytes in colorectal carcinomas. *Oncol Immunol*. 2018;7(7):e1442999.
- 66 Han LY, Fletcher MS, Urbauer DL, et al. HLA class I antigen processing machinery component expression and intratumoral T-Cell infiltrate as independent prognostic markers in ovarian carcinoma. *Clin Cancer Res*. 2008;14(11):3372–3379 [cited 2019 Apr 30] Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3426216/>.
- 67 Specht J, Pusztai L, Forero-Torres A, et al. 286PD - post-treatment biopsies show evidence of cell cycle arrest and immune cell infiltration into tumors of ladiratuzumab vedotin-treated advanced breast cancer patients. *Ann Oncol*. 2018;29(suppl\_8).
- 68 Li B, Chen L, Huang Y, et al. Are the tumor microenvironment characteristics of pretreatment biopsy specimens of colorectal cancer really effectively predict the efficacy of neoadjuvant therapy: a retrospective multicenter study. *Medicine*. 2024;103(35):e39429.
- 69 Ivison S, Malek M, Garcia RV, et al. A standardized immune phenotyping and automated data analysis platform for multicenter biomarker studies. *JCI insight*. 2018;3(23):e121867.
- 70 Shen J, Choi YL, Lee T, et al. Inflamed immune phenotype predicts favorable clinical outcomes of immune checkpoint inhibitor therapy across multiple cancer types. *J Immunother Cancer*. 2024;12(2):e008339 [cited 2024 Feb 22] Available from: <https://jitc.bmj.com/content/12/2/e008339>.
- 71 Karjula T, Elomaa H, Väyrynen SA, et al. Multiplexed analysis of macrophage polarisation in pulmonary metastases of microsatellite stable colorectal cancer. *Cancer Immunol Immunother*. 2024;73(3):59.
- 72 Väyrynen JP, Haruki K, Lau MC, et al. The prognostic role of macrophage polarization in the colorectal cancer microenvironment. *Cancer Immunol Res*. 2021;9(1):8–19.
- 73 Zhang K, Guo J, Yan W, Xu L. Macrophage polarization in inflammatory bowel disease. *Cell Commun Signal*. 2023;21:367.
- 74 Mosser DM, Edwards JP. Exploring the full spectrum of macrophage activation. *Nat Rev Immunol*. 2008;8:958–969.
- 75 Zhang QW, Liu L, Gong CY, et al. Prognostic significance of tumor-associated macrophages in solid tumor: a meta-analysis of the literature. *PLoS One*. 2012;7(12):e50946.
- 76 Zhang H, Yue X, Chen Z, et al. Roles of tumor-associated macrophages in anti-PD-1/PD-L1 immunotherapy for solid cancers. *Mol Cancer*. 2023;22:159.