





Radiomics signature for dynamic monitoring of tumor inflamed microenvironment and immunotherapy response prediction

Kinga Bernatowicz ¹, Ramon Amat ¹, Olivia Prior,¹ Joan Frigola,¹ Marta Liger ¹, Francesco Grussu,¹ Christina Zatse,¹ Garazi Serna ¹, Paolo Nuciforo,¹ Rodrigo Toledo,¹ Manel Escobar,² Elena Garralda,^{1,2} Enriqueta Felip,^{1,2} Raquel Perez-Lopez^{1,2}

To cite: Bernatowicz K, Amat R, Prior O, *et al.* Radiomics signature for dynamic monitoring of tumor inflamed microenvironment and immunotherapy response prediction. *Journal for ImmunoTherapy of Cancer* 2025;**13**:e009140. doi:10.1136/jitc-2024-009140

► Additional supplemental material is published online only. To view, please visit the journal online (<https://doi.org/10.1136/jitc-2024-009140>).

Accepted 14 November 2024



© Author(s) (or their employer(s)) 2025. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

¹Vall d'Hebron Institute of Oncology, Barcelona, Spain

²Vall d'Hebron University Hospital, Barcelona, Spain

Correspondence to

Kinga Bernatowicz;
kbernaticwz@vhio.net

ABSTRACT

Background The efficacy of immune checkpoint inhibitors (ICIs) depends on the tumor immune microenvironment (TIME), with a preference for a T cell-inflamed TIME. However, challenges in tissue-based assessments via biopsies have triggered the exploration of non-invasive alternatives, such as radiomics, to comprehensively evaluate TIME across diverse cancers. To address these challenges, we develop an ICI response signature by integrating radiomics with T cell-inflamed gene-expression profiles.

Methods We conducted a pan-cancer investigation into the utility of radiomics for TIME assessment, including 1360 tumors from 428 patients. Leveraging contrast-enhanced CT images, we characterized TIME through RNA gene expression analysis, using the T cell-inflamed gene expression signature. Subsequently, a pan-cancer CT-radiomic signature predicting inflamed TIME (CT-TIME) was developed and externally validated. Machine learning was employed to select robust radiomic features and predict inflamed TIME. The study also integrated independent cohorts with longitudinal CT images, baseline biopsies, and comprehensive immunohistochemistry panel evaluation to assess the pan-cancer biological associations, spatiotemporal landscape and clinical utility of the CT-TIME.

Results The CT-TIME signature, comprising four radiomic features linked to a T-cell inflamed microenvironment, demonstrated robust performance with AUCs (95% CI) of 0.85 (0.73 to 0.96) (training) and 0.78 (0.65 to 0.92) (external validation). CT-TIME scores exhibited positive correlations with CD3, CD8, and CD163 expression. Inpatient analysis revealed considerable heterogeneity in TIME between tumors, which could not be assessed using biopsies. Evaluation of aggregated per-patient CT-TIME scores highlighted its promising clinical utility for dynamically assessing the immune microenvironment and predicting immunotherapy response across diverse scenarios in advanced cancer. Despite demonstrating progression disease at the first follow-up, patients within the inflamed status group, identified by CT-TIME, exhibited significantly prolonged progression-free survival (PFS), with some surpassing 5 months, suggesting a potential phenomenon of pseudoprogression. Cox models using

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Traditional biopsies face challenges in assessing the immune microenvironment, prompting the search for non-invasive alternatives like radiomics to evaluate it across different cancers.

WHAT THIS STUDY ADDS

⇒ This study introduces a new method, CT-tumor immune microenvironment (TIME), which combines radiomics and gene expression to assess the immune microenvironment across diverse cancers. It also reveals significant variations in the microenvironment between tumors and highlights potential pseudoprogression in certain patients.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ The CT-TIME method has the potential to impact treatment decisions, patient categorization, and treatment outcomes in immune checkpoint therapies. It may also redefine our understanding of treatment response, particularly in cases of pseudoprogression.

aggregated CT-TIME scores from baseline images revealed a statistically significant reduction in the risk of PFS in the pan-cancer cohort (HR 0.62, 95% CI 0.44 to 0.88, $p=0.007$), and Kaplan-Meier analysis further confirmed substantial differences in PFS between patients with inflamed and uninflamed status (log-rank test $p=0.009$).

Conclusions The signature holds promise for impacting clinical decision-making, pan-cancer patient stratification, and treatment outcomes in immune checkpoint therapies.

BACKGROUND

The emergence of immune checkpoint inhibitors (ICIs) has transformed the treatment landscape for advanced cancers by leveraging the immune system to selectively target tumor cells. Despite these advancements, the overall response rates to ICI remain modest, with less

than 15% of patients experiencing positive responses,¹ alongside instances of severe treatment-related toxicities.² This underscores the urgent need in oncology for the development of precise predictive response biomarkers tailored to ICI therapies.

The efficacy of ICI is intricately tied to the tumor immune microenvironment (TIME).^{3,4} However, characterizing the TIME is complex, given the variability in T cell reactivity within tumors and the nuanced response of tumour-reactive T cells to immunotherapy. Recent investigations revealed key attributes of a T cell-inflamed tumor microenvironment, including active IFN- γ signaling, cytotoxic effector molecules, antigen presentation, and T cell-active cytokines.⁵ The T-cell-inflamed gene-expression

profile (GEP),⁵ serving as an indicator of the activity and condition of the TIME rather than the specific abundance of immune cell types, has demonstrated associations with ICIs response across diverse cancer types and clinical trial settings.^{6,7}

However, the assessment of TIME through tissue sampling presents logistical challenges in clinical practice, primarily due to the invasive character of the procedure. Furthermore, tissue biopsies, limited to small tumor fragments, are susceptible to sampling biases and inherently fail to capture the broader spatiotemporal dynamics of the microenvironment. In contrast, radiomics, a non-invasive approach extracting semi-quantitative features from medical images, offers a

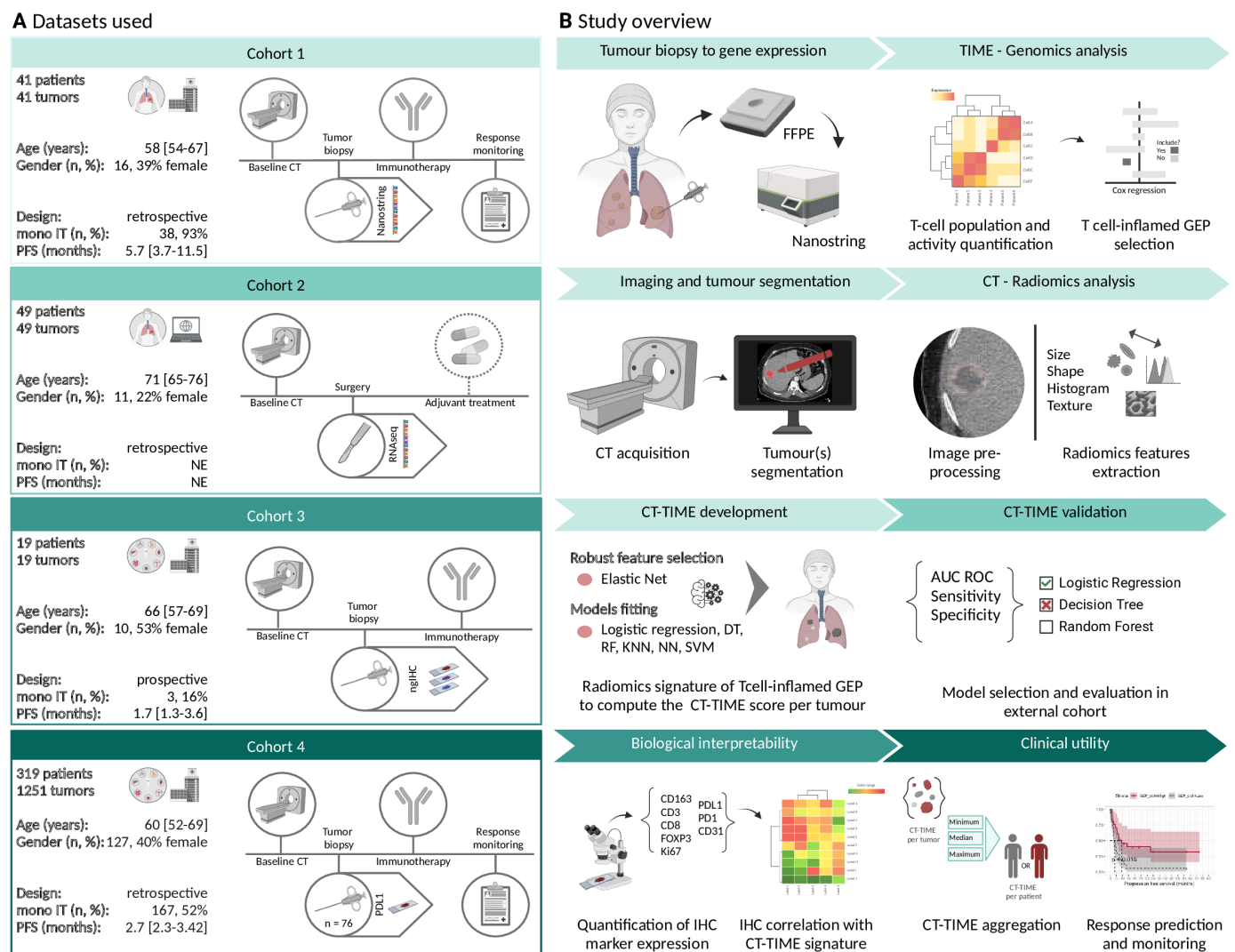


Figure 1 Data and study overview. (A) Dataset overview with demographics and treatment information. Note that patients were treated with a single agent (mono IT) or a combination of immunotherapy agents. Cohort 1 was used for signature development and model training, Cohort 2 for external signature validation, cohort 3 for exploring biological associations, and cohort 4 for analyzing the clinical utility of the signature. (B) Step-by-step study overview with colors corresponding to the analyzed datasets (A). Biopsies and CT images from the training cohort underwent analysis using genomics and radiomics. Robust feature selection was conducted through machine learning, with various models fitted to predict the T cell-inflamed GEP score. The best-performing model was chosen during validation. Biological associations of CT-TIME were explored, along with an investigation into the clinical utility of the aggregated CT-TIME score. AUC, area under the curve; FFE, formalin-fixed paraffin-embedded; GEP, gene-expression profile; IHC, immunohistochemistry; PFS, progression-free survival; RF, random forest; ROC, receiver operating characteristic; SVM, support vector machine; TIME, tumor immune microenvironment.

comprehensive view of tumor heterogeneity, enabling the characterization of an individual patient's tumor phenotype's spatiotemporal landscape to guide personalized treatments.

Radiomic signatures are patterns or profiles derived from radiomic features that can be correlated with specific disease characteristics, treatment responses, or patient outcomes. Initial studies demonstrated the potential of radiomics in characterizing the abundance of CD8+tumor-infiltrating lymphocytes and predicting treatment outcomes.^{8–11} However, these studies have shown limited reproducibility across different cancer types and have lacked external validation and signature explainability.^{12–14}

To address these challenges, we propose a novel approach—integrating radiomics with T cell-inflamed GEP to develop an ICI response signature. Our objectives encompass: (1) developing and validating a robust radiomic signature for the T cell-inflamed tumor micro-environment using machine learning, (2) studying the biological relevance of identified radiomic signature in relation to TIME, (3) analyzing spatiotemporal evolution of CT-TIME, and (4) exploring the clinical utility of the radiomic signature as a non-invasive tool for guiding personalized immunotherapy decisions, allowing for extended analysis compared with current techniques based on tissue sampling.

We anticipate that our radiogenomics signature holds the potential to advance personalized digital health solutions. CT-TIME not only serves as a novel imaging biomarker to assess the inflamed TIME but is also able to predict immunotherapy responses. This facilitates patient-centric care and contributes to tailored treatment strategies, ultimately improving clinical outcomes for patients with advanced cancer undergoing ICI treatment.

METHODS

Patient data

This study involved 428 patients (1360 evaluated tumors) from four independent patient cohorts (figure 1). Cohorts 1 was used for signature development and model training, cohort 2 for external signature validation, cohort 3 for biological associations and cohort 4 to analyze clinical utility of the signature. Cohorts 1, 3 and 4 consisted of patients with contrast-enhanced CT acquired no more than 4 weeks before the start of the immunotherapy (baseline) and before tissue sampling. Cohorts 1, 3 and 4 included patients treated with ICIs at the Vall d'Hebron Hospital in Barcelona, with data retrieved from digital clinical records. Cohort 3 was from an ongoing prospective study (PREDICT), while others were collected retrospectively. Cohort 2 data were obtained from the radiogenomics the Cancer Imaging Archive (TCIA) repository,¹⁵ including presurgical intervention patient CT images. Online supplemental S1 provides details on relevant patient characteristics.

Gene expression analysis

For the training cohort (cohort 1), NanoString computed gene expressions were analyzed as in Frigola *et al.*^{3,16} Clinical and immune characteristics of evaluated patients are presented in online supplemental S2. Univariate and multivariate regression analysis of immune variables with clinical benefit at 5 months were conducted. Only the T-cell-inflamed GEP, reflecting the TIME's activity and state rather than the abundance of specific immune cell types, achieved statistically significant associations with clinical outcome (online supplemental S3 and S4).

For cohort 2, RNA sequencing and data processing were performed as in Bakr *et al.*¹⁵ and summarized in online supplemental S5. T-cell-inflamed GEP scores from cohort 2 were harmonized by adjusting normal distribution parameters to match cohort 1. A GEP status (inflamed/uninflamed) was computed using median T-cell-inflamed GEP score and was used as the prediction target to guide development of the radiomics signature in subsequent sections.

Radiomic analysis

A robust radiomics methodology and computational pipeline were implemented in adherence to best-practice procedures. The Radiomics Quality Score (RQS)¹⁷ assessed the study quality online supplemental S6, and guidelines from the Image Biomarker Standardization Initiative¹⁸ for reporting imaging protocols and feature extraction were followed.

Contrast-enhanced CT images were identified, downloaded as Digital Imaging and Communication in Medicine (DICOM), anonymized and converted to NIFTI format using DICOM for Quantitative Imaging library.¹⁹ Image acquisition and reconstruction parameters are summarized in table 1.

All measurable lesions according to the Response Evaluation Criteria 1.1²⁰ were segmented using semiautomatic delineation tools from 3D Slicer V.4.11²¹ by a radiologist with more than 10 years of experience in oncological imaging (RP-L).

A total of 107 radiomic features describing lesion location, size, shape, first-order and high-order texture features were extracted using the PyRadiomics package.²² Images were processed in the same way across studied cohorts. Details on image processing and feature extraction parameters used are provided in table 2. Additionally, two aggregated features of total tumor volume and total surface were computed by considering all segmented tumors in a patient. Thus, a total of 109 features were analyzed per tumor online supplemental S7. Feature selection was carried out in three steps to find robust, non-redundant and informative radiomic features. Details are provided in online supplemental S8 and S9. Groups of highly correlated radiomic features were identified via hierarchical clustering and reduced to a single archetypal feature per cluster.

As detailed in table 2, 74% of images from cohort 2 were reconstructed using a sharp imaging kernel. To

Table 1 Imaging protocols across studied cohorts

Parameter	Cohort 1	Cohort 2	Cohort 3	Cohort 4
Number of images	41	49	38	421
Tube potential (kVp)	120	120	120	120
Tube current (mA), median (IQR)	419 (311–466)	341 (231–507)	395 (306–492)	319 (230–550)
Slice Thickness (mm), median (IQR)	2 (2–3)	1.25 (1.25–1.25)	2 (2)	2.5 (2–3)
Convolution Kernel, n (%)				
Soft	41 (100%)	10 (26%)	19 (100%)	421 (100%)
Sharp	–	39 (74%)	–	–

minimize the differences from imaging protocols, we applied feature harmonization using ComBat correction, as previously studied in Orlhac and Eertink^{23,24} and detailed in online supplemental S10.

Radiomic T-cell-inflamed signature development and validation

Elastic-net penalized logistic regression, based on 10 times repeated five-fold cross-validation, was applied to the training set (Cohort 1) to perform hyperparameter tuning and radiomic feature selection, using GEP status as an outcome. The receiver operating characteristic (ROC) curve served as the performance metric and was used to select an optimal model. The tuned hyperparameters included α (balancing L1 and L2 regularization)

and λ (regularization strength). The CT-TIME signature comprised robust, relevant and non-redundant radiomic features.

The radiomic signature was used to train various classification models, encompassing a generalized linear model (glm) based on logistic regression, K-nearest neighbors (knn), feed-forward neural networks model (nnet), random forest (rf), regression trees (rpart) and support vector machine (svmRadial).

Internal validation was conducted through repeated fivefold cross-validation. For each fold, the dataset was divided into five subsets, the model was trained on four subsets and validated on the fifth, computing performance metrics such as AUC, sensitivity, and specificity.

Table 2 Image processing and radiomic feature extraction parameters

Image processing	
Software	PyRadiomics V.3.0.1, installed in Python V.3.7.10
Bounding box	Defined by the segmentation, extended by default padding distance.
Resampled voxel spacing (mm)	1×1×1
Image interpolation method	B-spline
Intensity rounding	None
ROI interpolation method	Nearest neighbor
Re-segmentation	Intensity range (–100; 300)
Feature computation	
Kernel radius	1 mm
Discretization (fixed bin size)	25 HU
Image filter	None
maskedKernel	True (only voxels in the kernel that were also segmented in the ROI were used for calculation)
Initvalue	NaN (voxels outside ROI were considered as transparent)
Distance weighting for GLCM, GLRLM, NGTDM	No weighting
GLCM Symmetry	Symmetric
GLCM distance, GLSZM linkage distance, GLDZM linkage distance, NGTDM distance	Chebyshev distance $\delta=1$
NGTDM coarseness	Coarseness parameter $\alpha=0$
GLCM, gray-level co-occurrence matrix; GLSZM, gray-level size zone matrix; NGTDM, neighboring gray tone difference matrix; ROI, region of interest.	

The AUC values from all the folds were then aggregated to calculate the mean and SD, providing a comprehensive assessment of the model's performance. Furthermore, external validation of the models was carried out using cohort 2 data. The selection of the best-performing model was based on the AUC metric.

Both signature training and validation were implemented using glmnet and caret packages. It is essential to emphasise that these models, trained with GEP status as the outcome, will produce a CT-TIME score ranging from 0 to 1, representing the probability of the tumor belonging to the T cell inflamed group. This scoring system provides a quantitative measure of the likelihood of the tumor immune responsiveness.

The CT-TIME scoring system, functioning as a dichotomous variable, categorizes scores surpassing 0.5 as indicative of an immune-inflamed CT-TIME status, while scores below 0.5 are classified as uninflamed CT-TIME status. This dichotomy provides a clear and clinically relevant distinction, facilitating the identification of tumors with heightened immune responsiveness.

Exploration of biological correlates of CT-TIME score in a prospective cohort

A prospective cohort (cohort 3) of patients with matched baseline CT images and tumor biopsies was used to understand the biological associations of the developed model. Biopsies were processed using next-generation immunohistochemistry (NGI), to characterize different aspects of the tumor and its microenvironment. NGI consisted of a multiplex immune panel that provided co-localized CD8, CD3, CD163 and FOXP3 markers as well as Pan-CK for tumor region delineation and Ki67 for proliferation analyses. Furthermore, standard IHC single-plex staining was used to quantify PDL1, PD1 and CD31 expressions. Stained sections were digitized on a NanoZoomer 2.0-HT slide scanner (Hamamatsu Photonics), and digital images processed using the Visiopharm Image analysis software,²⁵ resulting in marker expression scores expressed in densities (cell/mm²).

Spatiotemporal analysis of TIME using CT-TIME

CT-TIME was comprehensively studied in cohort 4, comprising 319 patients with advanced pan-solid tumors, with over half presenting multiple tumors. Aggregating per-tumor into per-patient CT-TIME scores using mean, median, minimum, and maximum values allowed the classification of patients based on their aggregated scores. Those exceeding 0.5 were categorized as having an immune-inflamed CT-TIME status, while others were classified as uninflamed.

Additionally, CT-TIME status was longitudinally examined in a subset of 51 patients (153 tumors) with baseline, first follow-up and best response images acquired during immunotherapy treatment. A quantitative analysis of these transitions, including the number of cases, was conducted and visualized using the Sankey diagram to

provide a comprehensive understanding of the dynamics within the TIME.

Clinical applications of CT-TIME

Our investigation explored two distinct CT-TIME applications: (1) enhancing the monitoring of immunotherapy patients and its potential role in response assessment and (2) predicting the response to immunotherapy based on the evaluation of baseline CT-TIME scores.

To explore the predictive capability of aggregated scores at baseline, Cox proportional hazard models were fitted for progression-free survival (PFS). Kaplan-Meier analysis, using median CT-TIME status, was performed both in the entire cohort and in a subgroup focusing on patients with lung cancer.

As a benchmark, CT-TIME was compared against PDL1 status and tumor burden (volume) at baseline. This comparative analysis was performed in a subcohort, comprising patients with available PDL1 status (n=42), providing valuable insights into the performance of CT-TIME relative to established biomarkers in the field.

Statistical analysis and code availability

AUC and 95% CI were determined from the ROC curve, and the p value was assessed by using the Mann-Whitney U test to assess the model performance. A p value of 0.05 or lower was considered as statistically significant. Calibration was assessed using the Hosmer-Lemeshow goodness-of-fit test. Statistical analyses were performed by using R software, version 4.2.0.²⁶ To ensure the reproducibility of the signature, we made our code available on GitHub: (<https://github.com/kingaber/CT-TIME-public.git>).

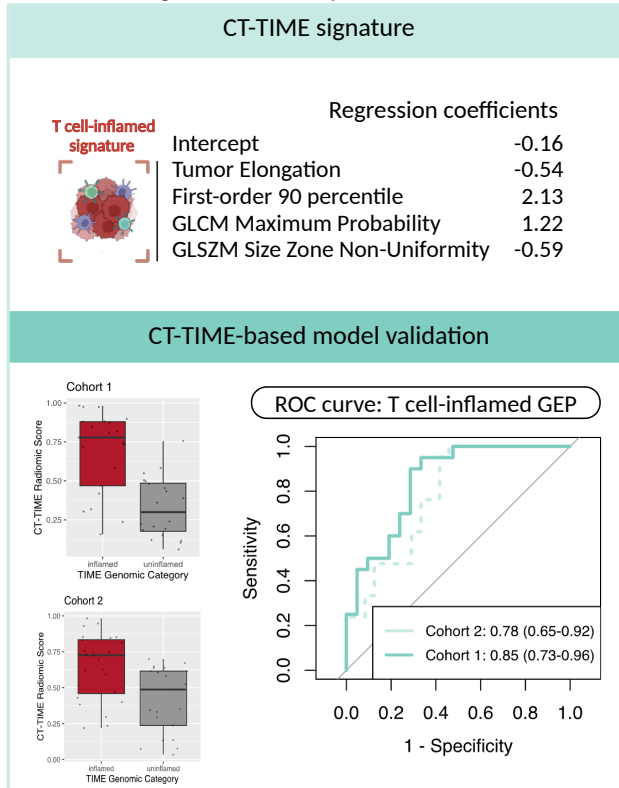
RESULTS

Radiomic T-cell-inflamed signature development and validation

Out of 109 radiomic features, 34 were identified as robust to different image perturbations (see online supplemental S9). The repeated cross-validation process revealed consistent and reliable performance metrics across different folds, attesting to the stability of the selected features and the effectiveness of the elastic net regularization. Optimal values for λ (0.012) and α (0.55) were identified. 13 radiomic features were selected, and 4 distinctive and non-redundant archetypal features were used in the signature: elongation, first-order 90th percentile, gray-level co-occurrence matrix (GLCM) maximum probability, and gray-level size zone (GLSZM) non-uniformity. Univariate and multivariate analyses of the signature features, with the T cell inflamed GEP score dichotomized to classify tumors into inflamed and uninflamed groups, are reported in online supplemental S11.

The radiomic signature features were used to train different classification models, as summarized in online supplemental S12. Following cross-validation and external validation, a generalized linear model with logistic regression was selected as the final CT-TIME-based model,

A CT-TIME signature development and validation



B Biological interpretability of CT-TIME

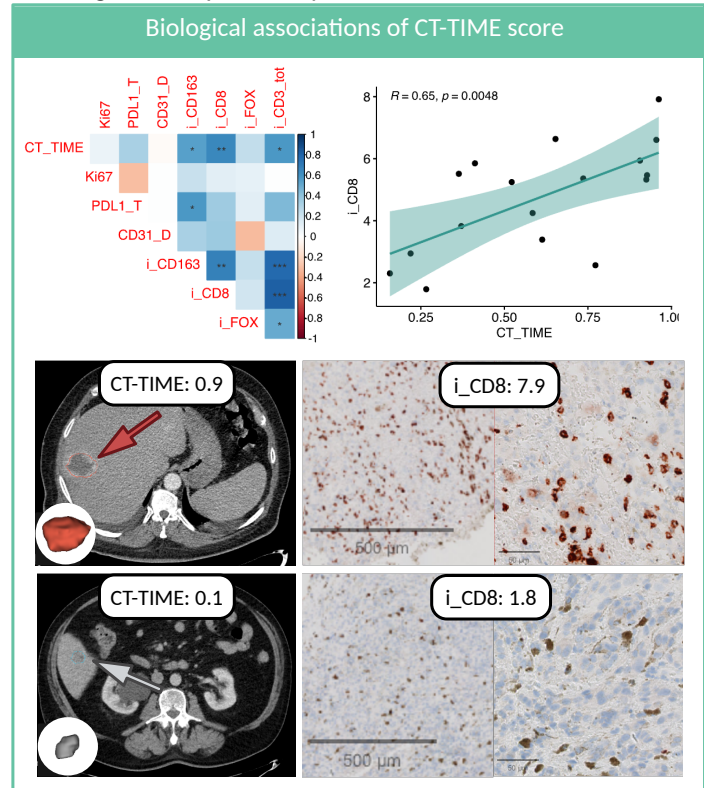


Figure 2 Development and evaluation of the radiomic T-cell-inflamed signature (CT-TIME). (A) Upper panel: Radiomic features, and their corresponding regression coefficients used in computing the CT-TIME signature. Lower panel: ROC curve depicting CT-TIME model training performance (cohort 1) and external validation results (cohort 2). (B) Correlation of the signature with immunohistochemistry staining, along with examples illustrating high and low signature scores and intratumoral CD8 expression. GEP, gene-expression profile; ROC, receiver operating characteristic; TIME, tumor immune microenvironment.

providing CT-TIME scores from 0 to 1 for each tumor, representing the probability of belonging to the inflamed group. This model yielded an AUC (95% CI) of 0.85 (0.73 to 0.96) and 0.78 (0.56 to 1) in training and internal validation, respectively. The Hosmer and Lemeshow goodness-of-fit test indicated good model calibration based on the χ^2 statistics of 5.2 and $p=0.74$. Observed and expected values were similar across the groups, further confirming well model calibration. Additionally, the CT-TIME model was validated in an external and independent cohort, demonstrating an AUC of 0.78 (0.65 to 0.92), as illustrated in (figure 2A).

Exploration of biological correlates of CT-TIME score in a prospective cohort

The biological correlates of CT-TIME score were investigated within a prospective cohort of patients with matched baseline CT images and tumor biopsies assessed using immunohistochemistry. Various markers were examined to characterize different aspects of the tumor and its microenvironment, as illustrated in (figure 2B). A significant correlation between CT-TIME score and intratumoral CD8 staining ($R=0.65, p=0.005$), a marker of the cytotoxic T cell population, was identified. Additionally, correlations were observed with CD163 ($R=0.54, p=0.02$),

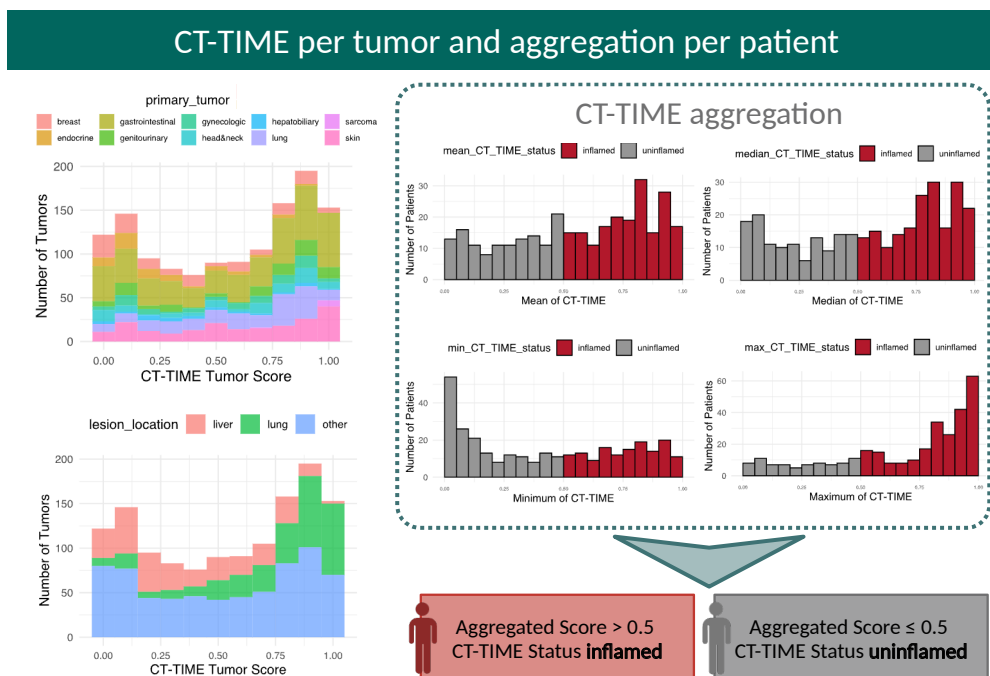
a marker of tumor-associated macrophages and the CD3 marker of the overall presence of T cells ($R=0.57, p=0.02$). Four selected radiomic features were also explored individually. We observed a correlation between GLSZM size zone non-uniformity and CD31 expression, which is indicative of vascular differentiation ($R=0.5, p=0.034$). Additionally, the first order 90th percentile displayed a negative correlation with FOXP3 expression, a marker associated with regulatory T cells ($R=-0.49, p=0.038$), as shown in online supplemental S13.

Spatiotemporal analysis of TIME using CT-TIME

CT-TIME score was computed in a cohort of 319 patients with 1314 advanced pan-solid tumors segmented. In a subset of 153 patients, additional CT images acquired at first follow-up and at best response were further analyzed. The examination of interpatient and inpatient (intra-tumor) heterogeneity of the CT-TIME score extracted from baseline images underscored the dynamic nature of immune responsiveness across diverse tumors.

Patients with multiple tumors showed moderate to high inpatient heterogeneity in CT-TIME in more than 67% of cases as shown in online supplemental S14. Interestingly, the CT-TIME score exhibited non-specificity to primary tumors, but a distinct pattern emerged, with

A Distribution of CT-TIME score at baseline



B CT-TIME score changes with time

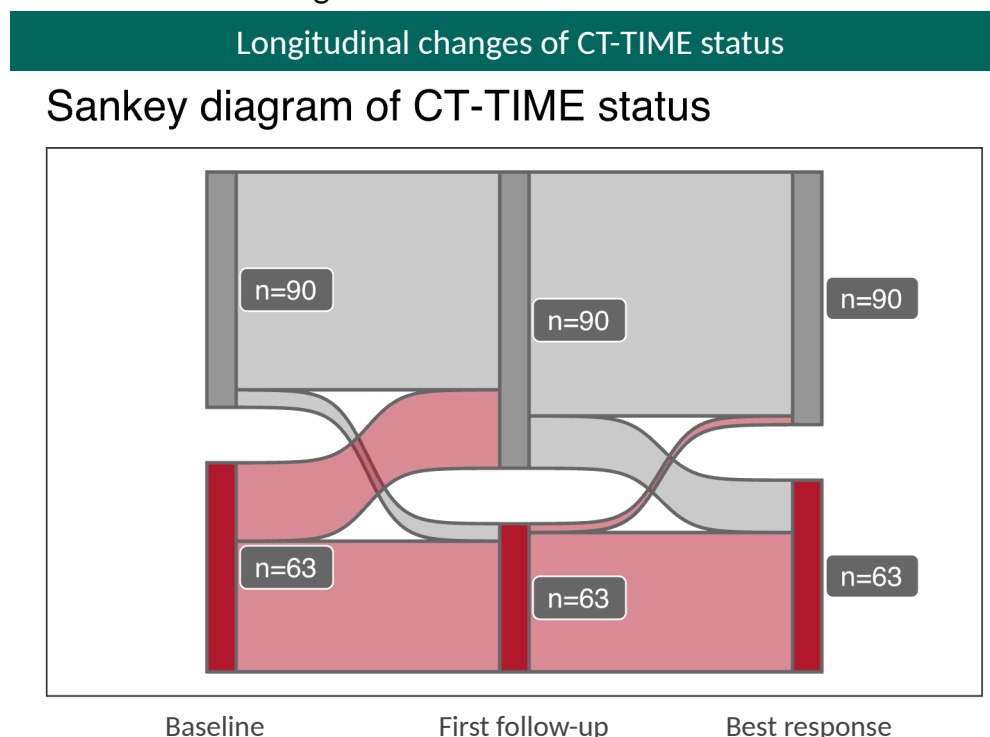


Figure 3 Spatiotemporal analysis of CT-TIME in cohort 4. (A) Exploration of CT-TIME score heterogeneity across tumor types and locations. Aggregation of per-tumor CT-TIME scores to the patient level and translation to CT-TIME status. (B) Longitudinal change in CT-TIME status. TIME, tumor immune microenvironment.

lung tumors demonstrating a higher score and liver tumors showing a lower score. This observation suggests that lung tumors may possess a more immune-inflamed microenvironment (figure 3A).

To further characterize the CT-TIME score, per-tumor scores were aggregated using mean, median, minimum, and maximum values. Patients with aggregated scores exceeding 0.5 were classified as having an

immune-inflamed CT-TIME status. The distribution of aggregated scores varied based on the aggregation method, as depicted in [figure 3A](#).

A dynamic fluctuation in CT-TIME status during immunotherapy emerged, depicted in [figure 3B](#). Analysis of baseline and first follow-up images revealed transitions from inflamed to uninflamed in 18% (9/51) of patients, with only 4% (2/51) changing from uninflamed to inflamed. From the first follow-up to best response, transitions from inflamed to uninflamed occurred in 2% (1/51) of cases, and 12% (6/51) changed from uninflamed to inflamed CT-TIME status.

Clinical applications of CT-TIME

The clinical utility of the CT-TIME, revealing notable insights into its potential applications

Given the acknowledged limitations of current response monitoring, particularly in the context of iRECIST unconfirmed progression disease, the CT-TIME score emerged as a compelling complementary tool. When computed at the first follow-up and indicating an uninflamed status, it could serve as an early indicator of confirmed progression. In a subset of progressing patients at the first follow-up, significant disparities in PFS were observed between inflamed and uninflamed groups ($p=0.047$), as depicted in [figure 4A](#).

Exploring the predictive capability of aggregated scores at baseline, Cox proportional hazard models were employed for PFS. Notably, aggregated scores using the minimum value exhibited an $HR<1$ with a CI that did not cross one in both the pan-cancer cohort and the lung subcohort (see [figure 4B](#)).

Survival analysis further reinforced the clinical significance of CT-TIME status derived from the minimum CT-TIME score, elucidating significant group separation through log-rank tests ($p=0.0054$ for pan-cancer and $p=0.00073$ for lung; [figure 4B](#)).

Lastly, within a subcohort enriched with available PD-L1 status, benchmarking against PD-L1 status and tumor burden (volume) at baseline revealed CT-TIME as the superior predictor, demonstrating significance ($p=0.018$) in Kaplan-Meier curves ([figure 4C](#)). This comprehensive investigation underscores the promising clinical utility of CT-TIME in dynamically assessing the immune microenvironment and predicting immunotherapy response across diverse advanced pan-cancer scenarios.

DISCUSSION AND CONCLUSION

In this study, we successfully developed and validated a robust radiomics signature for predicting the T cell-inflamed tumor microenvironment (CT-TIME) using advanced machine learning techniques. Notably, our signature included a single shape feature (elongation) while omitting features related to tumor volume. Due to substantial heterogeneity in previously published signatures,¹² it is difficult to compare our signature with the literature. However, some overlap could be observed with

other radiomics studies related to TIME: elongation was one of the selected features by Ligerio *et al.*²⁷ first-order 90 percentile was selected by Katsoulakis *et al.*²⁸ and GLSZM SZN was used by Li *et al.*²⁹ Nevertheless, we did not find GLCM maximum probability in previous publications, the GLCM family was frequently used. Importantly, our study's assessed quality (RQS=83.33%) substantially exceeded the mean RQS of previously published works (mean RQS=33.3%, range 0%–61.1%).¹²

The biological relevance of the identified radiomic features comprising our signature was explored in relation to the TIME. The CT-TIME signature demonstrated correlations with CD8, consistent with findings from prior studies.^{8 12} Notably, we also observed correlations with CD163, a marker associated with macrophages, and CD3 lymphocytes. Furthermore, our study revealed a novel correlation between signature radiomic features and CD31 (a vasculature marker) and FOXP3, suggesting potential regulatory T cell (Treg) activity. The negative correlation with FOXP3 implies that CT-TIME has the potential to identify a more inflamed microenvironment characterized by heightened T-cell activity. The association with CD31 expression suggests potential links between specific radiomic features and vascularity within the tumor microenvironment. Although these findings are promising, establishing causal relationships will require further investigation through functional studies or animal models.

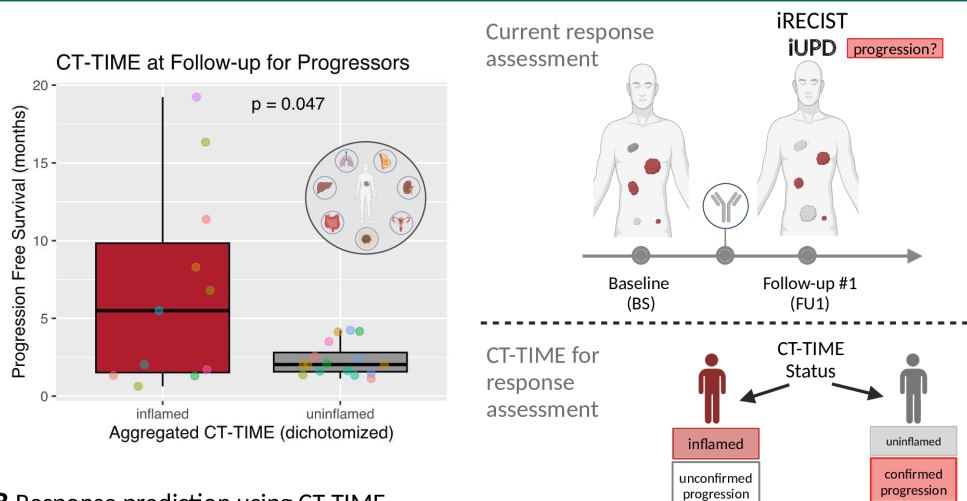
We demonstrated the utility of radiomics as a potent tool to study the spatiotemporal landscape of a patient's immune tumor microenvironment. Patients with multiple tumors exhibited substantial inpatient heterogeneity in TIME, emphasising the need for tools capable of capturing TIME across different patient locations beyond current techniques relying on tissue sampling. This aligns with other studies investigating TIME using RNA expression or histology.^{30 31} Interestingly, the CT-TIME score was independent of the primary tumor type, but it was higher in lung tumors than in liver tumors. Lung metastases were previously reported as highly immunogenic in several transcriptomic studies.^{32 33}

In the longitudinal analysis of CT-TIME, we observed transitions between inflamed and uninflamed statuses, and vice versa. Patients initially classified as having unconfirmed progression by iRECIST at the first follow-up exhibited an uninflamed CT-TIME status at the same follow-up, followed by rapid progression. This observation suggests the potential involvement of a resistance mechanism to immunotherapy treatment. CT-TIME proves to be a valuable complementary tool for guiding response assessments in clinical trials and facilitating the optimal submission of samples for laboratory analysis of immuno-resistant tumors. Moreover, early detection of progression enables the exploration of different treatment options, such as therapies that transform cold tumors into hot ones.^{34 35}

Patients with uninflamed CT-TIME status at baseline were more likely to progress, pointing to CT-TIME

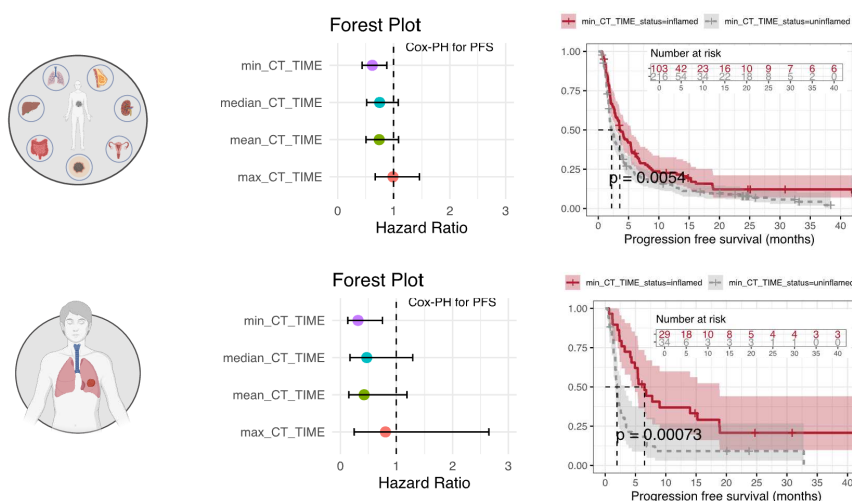
A CT-TIME score changes with time

Aggregated CT-TIME at different timepoints



B Response prediction using CT-TIME

Aggregated CT-TIME at baseline



C Comparison of CT-TIME to established biomarkers

Benchmark

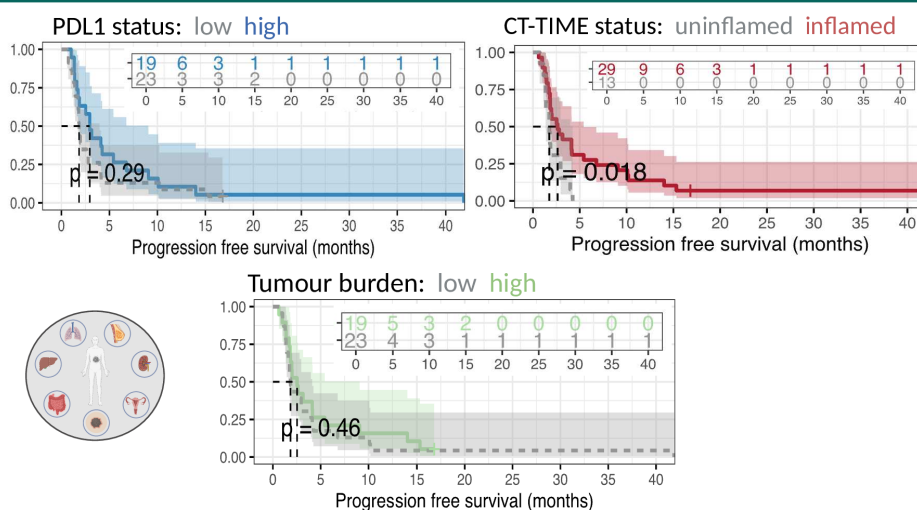


Figure 4 Clinical utility of CT-TIME signature in cohort 4. (A) Aggregated CT-TIME status for response assessment. (B) CT-TIME status as a predictor of immunotherapy response. (C) Comparative analysis of CT-TIME status with PDL1 status and tumor burden. TIME, tumor immune microenvironment.

potential for predicting immunotherapy response. We found that the signature performance was better in lung cancer patients than across diverse advanced pan-cancer scenarios. Considering the evolving landscape of immunotherapy biomarkers, a combination of CT-derived signatures (eg, CT-TIME+CT-TMB + CT-PDL1) promises accurate immunotherapy response prediction, offering broad accessibility and spatio-temporal insights for oncologists.

While our study explores the clinical utility of the CT-TIME signature and highlights potential advantages, we acknowledge inherent limitations. The need for fast and automatic radiomics analysis requires the implementation of automatic segmentation tools, an aspect beyond the scope of this work. However, recent advancements in AI, particularly automatic segmentation models, are anticipated to enhance workflow implementation.³⁶ Feature harmonization is crucial for generalizing radiomics pipelines, and standardizing reference datasets across studies could alleviate inconsistencies. Additionally, using an immune score derived from tumor samples for CT-TIME estimate could serve as a quality check, offering a quick cross-check alongside predictions. Ongoing efforts to define reproducible immune assays are crucial for this purpose.^{37,38} Proposed CT-TIME signature was successfully developed and validated in lung cancer cohorts (cohort 1 and 2). While we demonstrated its clinical application in a pan-cancer cohort, the restricted number of cases per cancer type group precludes drawing definitive conclusions, although the results are promising.

Further investigation into the relationship between radiomic features and underlying biological mechanisms, particularly through the use of functional imaging techniques and animal models, has the potential to provide deeper insights into how radiomic features correlate with specific biological processes. This avenue of research holds promise for advancing the applicability of radiomic analyses in various clinical settings. In conclusion, our study introduces an innovative radiomics T-cell-inflamed signature, providing biological insights, spatiotemporal evaluation, and clinical utility exploration. By advancing the understanding of the TIME and its radiomics features, we aim to catalyze further research and discussions, ultimately leading to the clinical implementation of radiomics tools and improved personalized treatment strategies for patients undergoing ICI therapy.

X Kinga Bernatowicz @KingaBarcelona, Joan Frigola @Frigola_Joan, Francesco Grussu @fragrussu and Raquel Perez-Lopez @raqperezlopez

Contributors KB: conceptualisation, data curation, formal analysis, investigation, methodology, validation, visualisation, writing—original draft, review and editing. RA: conceptualisation, data curation, investigation, methodology, validation, writing—review and editing. OP: methodology, visualisation, writing—original draft, review and editing. JF: data curation, investigation, formal analysis, methodology, validation, writing—review. ML: methodology, software, writing—review. FG: methodology, writing—review and editing. CZ: data curation, writing—review. GS: formal analysis, software, validation, writing—review and editing. PN: supervision, resources, writing—review. RT: supervision, resources, writing—review. ME: project administration, resources, writing—review. EG: supervision, funding acquisition, project administration, resources, writing—review. EF: project administration,

funding acquisition, resources, writing—review. RP-L: conceptualisation, funding acquisition, project administration, resources, supervision, writing—review and editing and guarantor of the study. AI was used in preparation of the manuscript to revise parts of the text.

Funding KB is supported by MSCA COFUND Beatriz de Pinós Grant (2019BP/00182). FG receives the support of a fellowship from LaCaixa Foundation (ID 100010434). The fellowship code is LCF/BQ/PR22/11920010. RP-L is supported by CRIS Cancer Foundation (TALENT19-05), the FERO Foundation, the Instituto de Salud Carlos III-Investigación en Salud (PI18/01395 and PI21/01019) and the Prostate Cancer Foundation (18YOUN19). OP was funded by Fundació “LaCaixa” (ID 100010434, code LCF/BQ/DR21/11880027).

Competing interests KB, RA, ME, JF, FG, ML, OP, GS, and CZ—nothing to disclose. EF—consulting fees: Abbvie, Amgen, Astra Zeneca, Bayer, Beigene, Boehringer Ingelheim, Bristol Myers Squibb, Eli Lilly, F. Hoffmann-La Roche, Gilead, Glaxo Smith Kline, Janssen, Merck Serono, Merck Sharp & Dohme, Novartis, Peptomyc, Pfizer, Regeneron, Sanofi, Takeda, Turning Point, Daiichi Sankyo. Payment or honoraria for lectures, presentations, speakers bureaus, manuscript writing or educational events: Amgen, Astra Zeneca, Bristol Myers Squibb, Daiichi Sankyo, Eli Lilly, F. Hoffmann – La Roche, Genentech, Janssen, Medical Trends, Medscape, Merck Serono, Merck Sharp & Dohme, Peervoice, Pfizer, Sanofi, Takeda, Touch Oncology. Support for attending meetings and/or travel: Astra Zeneca, Janssen, Roche. EG: grants or contracts from any entity: Novartis, Roche, Thermo Fisher, AstraZeneca, Taiho, BeiGene, Janssen. Consulting fees: Roche, Ellipses Pharma, Boehringer Ingelheim, Janssen Global Services, Seattle Genetics, Thermo Fisher, MabDiscovery, Anaveon, F-Star Therapeutics, Hengrui, Sanofi, Incyte, Medscape. Payment or honoraria for lectures, presentations, speakers bureaus, manuscript writing or educational events: Merck Sharp & Dohme, Roche, Thermo Fisher, Novartis, SeaGen. Stock or stock options: 1 TRIALSP. PI or Co-PI—Institutional: Adaptimmune—Affimed—Amgen SA—Anaveon AG—AstraZeneca AB—Bicycletx—BioInvent International AB—Biontech SE—Biontech Small Molecules—Boehringer Ingelheim International—Catalym—Cyclacel Biopharmaceuticals—Cytovation AS—Cytomx—F.Hoffmann La Roche—F-Star Beta—Genentech—Genmab B.V.—Hifibio Therapeutics—Hutchison Medipharma—Icon—Imcheck Therapeutics—Immunocore—Incyte Corporation—Incyte Europe Sàrl—Janssen-Cilag International NV—Janssen-Cilag SA—Laboratorios Servier SL—Medimmune—Merck & Co—Merck Kgga—Novartis Farmacéutica, S.A—Peptomyc—Pfizer Slu—Relay Therapeutics—Replimmune - Ribon Therapeutics—Ryvü Therapeutics SA—Seattle Genetics—Sotio as—Sqz Biotechnologies—Symphogen A/S—Taiho Pharma Usa—T-Knife GmbH, NEXT Oncology. PN: payment or honoraria for lectures, presentations, speakers bureaus, manuscript writing or educational events: NOVARTIS, Invited Speaker. Participation on a Data Safety Monitoring Board or Advisory Board: Bayer, Advisory Board, MSD ONCOLOGY, Advisory Board. Other financial or non-financial interests: Targos Molecular Pathology GmbH, Other, Consultant. RP-L: all support for the present manuscript: Astra Zeneca PoC Grant, CRIS Cancer Foundation. Consulting fees: Roche Pharma, Steering Committee Member of the ATHECA clinical trial. Leadership or fiduciary role in another board, society, committee or advocacy group, paid or unpaid: Cancer Core Europe Imaging Task Force Leader. RT: grants or contracts from any entity: Research grants from Astrazeneca, Beigene.

Patient consent for publication All authors authorised and granted full consent to the corresponding author of the manuscript (KB) to enter into publishing agreement with BMJ on their behalf.

Ethics approval This study involves human participants and was approved by The Vall d'Hebron Institute of Oncology (VHIO) institutional review board approved this study for retrospective data [PR(AG)70/2018] and the prospective PREDICT study [PR(AG)29/2020]. Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available on reasonable request. Validation data are publicly available on TCIA platform, as stated in the manuscript. Training and biological validation data are property of the VHIO and could be available on reasonable request to the senior author.

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

terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iDs

Kinga Bernatowicz <http://orcid.org/0000-0001-9166-1709>

Ramon Amat <http://orcid.org/0000-0001-8284-2540>

Marta Ligeró <http://orcid.org/0000-0001-9824-7316>

Garazi Serna <http://orcid.org/0000-0003-2666-0982>

REFERENCES

- Pilard C, Ancion M, Delvenne P, et al. Cancer immunotherapy: it's time to better predict patients' response. *Br J Cancer* 2021;125:927–38.
- Johnson DB, Nebhan CA, Moslehi JJ, et al. Immune-checkpoint inhibitors: long-term implications of toxicity. *Nat Rev Clin Oncol* 2022;19:254–67.
- Frigola J, Navarro A, Carbonell C, et al. Molecular profiling of long-term responders to immune checkpoint inhibitors in advanced non-small cell lung cancer. *Mol Oncol* 2021;15:887–900.
- Eckstein M, Gupta S. New insights in predictive determinants of the tumor immune microenvironment for immune checkpoint inhibition: a never ending story? *Ann Transl Med* 2019;7:S135.
- Ayers M, Lunceford J, Nebozhyn M, et al. IFN- γ -related mRNA profile predicts clinical response to PD-1 blockade. *J Clin Invest* 2017;127:2930–40.
- Cristescu R, Mogg R, Ayers M, et al. Pan-tumor genomic biomarkers for PD-1 checkpoint blockade-based immunotherapy. *Science* 2018;362:eaar3593.
- Ott PA, Bang Y-J, Piha-Paul SA, et al. T-Cell-Inflamed Gene-Expression Profile, Programmed Death Ligand 1 Expression, and Tumor Mutational Burden Predict Efficacy in Patients Treated With Pembrolizumab Across 20 Cancers: KEYNOTE-028. *J Clin Oncol* 2019;37:318–27.
- Sun R, Sundahl N, Hecht M, et al. Radiomics to predict outcomes and abscopal response of patients with cancer treated with immunotherapy combined with radiotherapy using a validated signature of CD8 cells. *J Immunother Cancer* 2020;8:e001429.
- Trebeschi S, Drago SG, Birkbak NJ, et al. Predicting response to cancer immunotherapy using noninvasive radiomic biomarkers. *Ann Oncol* 2019;30:998–1004.
- Jiang Y, Wang H, Wu J, et al. Noninvasive imaging evaluation of tumor immune microenvironment to predict outcomes in gastric cancer. *Ann Oncol* 2020;31:760–8.
- Bortolotto C, Lancia A, Stelitano C, et al. Radiomics features as predictive and prognostic biomarkers in NSCLC. *Expert Rev Anticancer Ther* 2021;21:257–66.
- Ramlee S, Hulse D, Bernatowicz K, et al. Radiomic Signatures Associated with CD8⁺ Tumour-Infiltrating Lymphocytes: A Systematic Review and Quality Assessment Study. *Cancers (Basel)* 2022;14:3656.
- Spadarella G, Stanzione A, Akinci D'Antonoli T, et al. Systematic review of the radiomics quality score applications: an EuSoMI Radiomics Auditing Group Initiative. *Eur Radiol* 2023;33:1884–94.
- Zhang C, de A. F. Fonseca L, Shi Z, et al. Systematic review of radiomic biomarkers for predicting immune checkpoint inhibitor treatment outcomes. *Methods* 2021;188:61–72.
- Bakr S, Gevaert O, Echegaray S, et al. A radiogenomic dataset of non-small cell lung cancer. *Sci Data* 2018;5:180202.
- Frigola J, Carbonell C, Irazzo N, et al. High levels of chromosomal aberrations negatively associate with benefit to checkpoint inhibition in NSCLC. *J Immunother Cancer* 2022;10:e004197.
- Lambin P, Leijenaar RTH, Deist TM, et al. Radiomics: the bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol* 2017;14:749–62.
- Zwanenburg A, Vallières M, Abdalah MA, et al. The Image Biomarker Standardization Initiative: Standardized Quantitative Radiomics for High-Throughput Image-based Phenotyping. *Radiology* 2020;295:328–38.
- Herz C, Fillion-Robin J-C, Onken M, et al. dcmqi: An Open Source Library for Standardized Communication of Quantitative Image Analysis Results Using DICOM. *Cancer Res* 2017;77:e87–90.
- Eisenhauer EA, Therasse P, Bogaerts J, et al. New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1). *Eur J Cancer* 2009;45:228–47.
- Fedorov A, Beichel R, Kalpathy-Cramer J, et al. 3D Slicer as an image computing platform for the Quantitative Imaging Network. *Magn Reson Imaging* 2012;30:1323–41.
- van Griethuysen JJM, Fedorov A, Parmar C, et al. Computational Radiomics System to Decode the Radiographic Phenotype. *Cancer Res* 2017;77:e104–7.
- Orlhac F, Eertink JJ, Cottreau A-S, et al. A Guide to ComBat Harmonization of Imaging Biomarkers in Multicenter Studies. *J Nucl Med* 2022;63:172–9.
- Ligeró M, Jordi-Ollero O, Bernatowicz K, et al. Minimizing acquisition-related radiomics variability by image resampling and batch effect correction to allow for large-scale data analysis. *Eur Radiol* 2021;31:1460–70.
- Visiopharm image analysis program. n.d. Available: <https://visiopharm.com/>
- R Core Team. R: a language and environment for statistical computing. Vienna, Austria R Foundation for Statistical Computing; 2022. Available: <https://www.R-project.org/>
- Ligeró M, García-Ruiz A, Viaplana C, et al. A CT-based Radiomics Signature Is Associated with Response to Immune Checkpoint Inhibitors in Advanced Solid Tumors. *Radiology* 2021;299:109–19.
- Katsoulakis E, Yu Y, Apte AP, et al. Radiomic analysis identifies tumor subtypes associated with distinct molecular and microenvironmental factors in head and neck squamous cell carcinoma. *Oral Oncol* 2020;110:104877.
- Li J, Shi Z, Liu F, et al. XGBoost Classifier Based on Computed Tomography Radiomics for Prediction of Tumor-Infiltrating CD8⁺ T-Cells in Patients With Pancreatic Ductal Adenocarcinoma. *Front Oncol* 2021;11:671333.
- de Bruin EC, McGranahan N, Mitter R, et al. Spatial and temporal diversity in genomic instability processes defines lung cancer evolution. *Science* 2014;346:251–6.
- Tabata M, Sato Y, Kogure Y, et al. Inter- and intra-tumor heterogeneity of genetic and immune profiles in inherited renal cell carcinoma. *Cell Rep* 2023;42:112736.
- García-Mulero S, Alonso MH, Pardo J, et al. Lung metastases share common immune features regardless of primary tumor origin. *J Immunother Cancer* 2020;8:e000491.
- Remark R, Alifano M, Cremer I, et al. Characteristics and clinical impacts of the immune environments in colorectal and renal cell carcinoma lung metastases: influence of tumor origin. *Clin Cancer Res* 2013;19:4079–91.
- Liu YT, Sun ZJ. Turning cold tumors into hot tumors by improving T-cell infiltration. *Theranostics* 2021;11:5365–86.
- Zhang J, Huang D, Saw PE, et al. Turning cold tumors hot: from molecular mechanisms to clinical applications. *Trends Immunol* 2022;43:523–45.
- Mazurowski MA, Dong H, Gu H, et al. Segment anything model for medical image analysis: An experimental study. *Med Image Anal* 2023;89:102918.
- Marliot F, Pagès F, Galon J. Usefulness and robustness of Immunoscore for personalized management of cancer patients. *Oncoimmunology* 2020;9:1832324.
- Lanzi A, Pagès F, Lagorce-Pagès C, et al. The consensus immunescore: toward a new classification of colorectal cancer. *Oncoimmunology* 2020;9:1789032.