scientific reports

Check for updates

OPEN

Clinically interpretable nomogram combining body composition and clinicopathological features for one year survival prediction in advanced solid tumors

Giulia Bruschi[1,12], Francesco Paoloni[2,12], Federica Pecci[2,3,12]✉, Elisabetta Tola[4], Valeria Cognigni[2], Tommaso Galassi[2], Alessandra Borgheresi[4], Luca Cantini[5], Luca Santamaria[2], Mariangela Gualtieri[2], Valentina Lunerti[2], Natalia Chiodi[6], Veronica Agostinelli[2], Marzia Di Pietro Paolo[2], Agnese Sbrollini[1], Andrea Agostini[4], Giulia Mentrasti[2], Salvatore Ficarra[7,8], Giulia Mazzaschi[3,9], Alessandro Parisi[2], Riccardo Giampieri[2], Kamal S. Saini[5,10], Sebastiano Buti[3,9], Marcello Tiseo[3,9], Arianna Vignini[11], Andrea Giovagnoni[4], Laura Burattini[1,12] & Rossana Berardi[2,12]✉

Immune checkpoint inhibitors (ICIs) have improved outcomes for patients with solid tumors, but reliable predictors of overall survival (OS) are limited. This retrospective study of 146 advanced solid tumor patients treated with ICIs aims to provide a nomogram to predict 1-year (1y) OS integrating body composition (BC) parameters with standard clinicopathological (CP) features. A two-stage approach was implemented: first random survival forest models were trained and tested to evaluate the prognostic value of (a) CP features alone, (b) BC metrics alone or as newly introduced BC scores, and (c) their combination. The best predictive performance (average cumulative AUC of 0.73 in test set) was achieved by combining 12 CP features with the BC score comprising intramuscolar adipose tissue content, visceral fat area index, and the visceral-to-subcutaneous fat area index ratio. Finally, a nomogram was developed with this feature set, offering a tool for personalized risk stratification and treatment planning (mean absolute error in calibration curve of 0.03 and overall AUC of 0.76). Integrating BC parameters with CP features substantially enhances 1y OS prediction in patients receiving ICIs.

Keywords Cancer, Machine learning, Body composition, Nomogram, Overall survival prediction, Immunotherapy

In recent years, immune checkpoint inhibitors (ICIs) have revolutionized the clinical practice of solid tumors, leading to a significant survival benefit across various malignancies. However, ICIs are effective only in a relatively small and unpredictable portion of patients. On the other hand, a fraction of patients does not benefit from ICI treatment but instead experience immune-related adverse events (irAEs)[1]. Therefore, precise patient

[1]Department of Information Engineering, Marche Polytechnic University, Ancona, Italy. [2]Marche Polytechnic University, Clinical Oncology, Azienda Ospedaliero Universitaria delle Marche, Ancona, Italy. [3]Department of Medicine and Surgery, University of Parma, Parma, Italy. [4]Department of Clinical, Special, and Dental Sciences (DISCO), Department of Radiological Sciences, Marche Polytechnic University, Azienda Ospedaliero Universitaria delle Marche, Ancona, Italy. [5]Fortrea Inc., Durham, NC, USA. [6]Department of Clinical and Molecular Sciences, Marche Polytechnic University, Ancona, Italy. [7]Division of Population Sciences, Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, USA. [8]Medical School, Harvard University, Cambridge, USA. [9]Medical Oncology Unit, Ospedale di Parma, Parma, Italy. [10]Cambridge University Hospitals NHS Foundation Trust, Addenbrooke's Hospital, Cambridge, UK. [11]Department of Clinical, Special, and Dental Sciences (DISCO), Department of Clinical Sciences, Section of Biochemistry, Biology and Physics, Marche Polytechnic University, Ancona, Italy. [12]G. Bruschi, F. Paoloni, F. Pecci: These authors contributed equally to this work. L. Burattini, R. Berardi: These authors contributed equally as senior authors of the manuscript. ✉email: peccifede91@gmail.com; r.berardi@staff.univpm.it

selection, with the optimization of biomarkers that are predictive or prognostic for patients treated with ICIs remains an important clinical need[2]. While validated biomarkers, such as programmed cell death ligand 1 (PD-L1), tumor mutational burden (TMB), and DNA mismatch repair (MMR) are currently used in clinical practice, their predictive value can vary significantly, depending on tumor type and treatment setting[3–5]. Notably, these biomarkers are primarily based on cancer type and on tumor-infiltrating lymphocytes, not considering other factors external to the tumor. Indeed, immune activity against tumor is shaped by a variety of factors extrinsic to the tumor microenvironment, including concomitant medications, infections, vaccinations, diet, and lifestyle[6,7]. Among these factors, body mass index (BMI) has historically been widely explored as potential predictor of clinical outcomes in patients treated with ICIs[8–10]. Several studies have suggested that patients with higher BMI have better clinical outcomes compared to those with normal BMI, a phenomenon known as "obesity paradox"[11]. Despite this evidence, mostly based on retrospective data, this relation between a higher BMI and improved clinical outcomes in patients treated with ICIs is still debatable. In fact, while the BMI is easy to compute, it does not consider the body composition (BC) in terms of fat and muscle mass. BC can be evaluated using imaging modalities, which are already routinely used to assess the disease status. Computed tomography (CT) is highly reliable for evaluating and quantifying adipose and muscle tissue present in specific compartments. Changes in BC, as assessed through CT during the immunotherapy-based treatment, are associated with variations in clinical outcomes[12].

Recently, the use of machine learning (ML) models that integrate multiomics with clinical data[13–15] has enhanced the identification of factors predicting response to ICIs. Despite the widespread use of ML models to predict ICI outcomes and patient survival, no large-scale study, to the best of the authors' knowledge, has integrated both clinicopathologic (CP) features and BC parameters to predict overall survival (OS) in patients receiving ICIs. Furthermore, no study has comprehensively assessed the combined predictive potential of BC parameters. The aim of this study is to develop a potentially clinically applicable and interpretable nomogram for predicting 1-year (1y) OS in patients receiving ICIs. To achieve this, CP features and BC parameters were integrated to more accurately characterize patient responses. Using a two-stage approach, first the predictive performance of CP features and BC parameters, both individually and combined into novel BC scores, were assessed. Then, the best feature set was identified and used to construct the nomogram for 1y OS prediction.

Results

Patient characteristics

The study CONSORT diagram is shown in Supplementary Fig. S1. A total of 146 patients with advanced solid tumors, treated with ICIs either alone or in combination with chemotherapy or tyrosine kinase inhibitor (TKI), were enrolled, with a median follow-up of 26.1 months (95% CI 23.2-29.6). CP features and comorbidities of the retrospective cohort are summarized in Table 1. Baseline BC parameters, assessed via CT scan across the entire retrospective cohort, were summarized and analyzed by sex. As expected, notable differences in BC emerged between the sexes (Supplementary Table S1). Skeletal muscle index (SMI), psoas muscle index (PMI), visceral fat area index (VFAI), and the visceral-to-subcutaneous fat area ratio (VF/SF) were significantly higher in males compared to females, whereas intramuscular adipose tissue content (IMAC) and subcutaneous fat area index (SFAI) were significantly lower in males. Due to the role of the Controlling Nutritional Status (CONUT) score in assessing the risk of malnutrition, BC parameters and BMI were evaluated in patients stratified by CONUT score into two groups: low risk (CONUT score ≤ 4) and high risk (CONUT score > 4) (Supplementary Table S2). No statistically significant differences were observed between the low-risk and high-risk groups. Comorbidities were collected at the baseline of ICIs, including cardiovascular disease (CD) history, diabetes mellitus (DM) and hypertension (HTA). In terms of differences for the BC parameters according to comorbidities, the 40 patients with history of CD showed lower muscular content in terms of lean psoas muscle area index (LPMAI) (Supplementary Table S3). Additionally, fat tissue features such as VFAI, total fat area index (TFAI), and VF/SF were significantly higher in patients with DM at the baseline of ICI treatment (Supplementary Table S4). Patients with HTA at the baseline of ICI showed significantly higher values of VF/SF, TFAI, and VFAI and lower values of LPMAI (Supplementary Table S5).

Two-stage procedure

Analysis of the predictive performance of CP variables for OS prediction with random survival forest (RSF) revealed that the CP feature set including sex, tumor type, treatment group, smoking history, DM, HTA, BMI, neutrophil-to-lymphocyte ratio (NLR), and the CONUT score, achieved the highest predictive accuracy. This feature set achieved a validation average cumulative dynamic area under the curve (avg cdAUC) of 0.75 ± 0.11 and a test avg cdAUC of 0.61 ± 0.08 across bootstraps. The OS predictive ability of each baseline BC parameter, evaluated alone, is reported in Supplementary Table S6 with VF/SF identified as the best predictor with test avg cdAUC of 0.69 ± 0.07. The generated BC scores, along with RSF performance for predicting OS, are provided in Supplementary Table S7. Table 2 reports the performance of RSF models trained and tested using CP features in combination with BC scores. The optimal feature combination (FC), named FC3, includes a BC score derived from three BC parameters (BC3 score), IMAC, VFAI, and VF/SF, together with 12 CP features: sex, tumor type, treatment group, treatment line, Eastern Cooperative Oncology Group (ECOG), number of metastatic sites, smoking history, CD, DM, HTA, NLR, and CONUT score. FC3 achieved the best performance in terms of OS prediction with an avg cdAUC of 0.84 ± 0.08 and 0.73 ± 0.07 respectively in the validation and test. The feature ranking is depicted in Fig. 1. The top three features contributing the most to OS prediction are NLR, BC3 score, and treatment group. The implemented optimization approach identified 46.08 $\frac{cm^2}{m^2}$ as the optimal threshold (Thr) for the BC3 score, enabling the stratification and categorization of patients into two clinically interpretable risk groups. The stability of this Thr is illustrated in Supplementary Fig. S2, which shows that 44%

Feature		Overall (N=146)	Feature		Overall (N=146)	Feature		Overall (N=146)
Age			Tumor type			Statins assumption		
	Median [Min, Max]	71.0 [38.0, 88.0]		NSCLC	87 (59.6%)		No	96 (65.8%)
Age classification				RCC	23 (15.8%)		Yes	50 (34.2%)
	$<$ 70 years	64 (43.8%)		Melanoma	12 (8.2%)	Neutrophil-to-Lymphocyte Ratio		
	$\geq$ 70 years	82 (56.2%)		Urothelial carcinoma	10 (6.8%)		Median [Min, Max]	4.22 [1.10, 28.0]
Sex				Head and neck carcinoma	5 (3.4%)	CONUT score		
	Male	107 (73.3%)		SCLC	4 (2.7%)		Low ($\leq$ 4)	101 (69.0%)
	Female	39 (26.7%)		HCC	3 (2.1%)		High ($>$ 4)	45 (31.0%)
Smoking history				Breast carcinoma	1 (0.7%)	BMI		
	Current/former	97 (66.4%)		Squamous carcinoma	1 (0.7%)		Median [Min, Max]	24.8 [16.9, 36.9]
	Never	49 (33.6%)	Line of treatment			Treatment group		
ECOG				First	115 (78.8%)		ICI monotherapy	101 (69.2%)
	(0,1)	139 (95.2%)		Second	24 (16.4%)		ICI combination with chemotherapy or TKI or targeted therapy	45 (30.8%)
	(2,3)	7 (4.8%)		Third	6 (4.1%)	Liver metastasis		
Cardiovascular disease				Fourth	1 (0.7%)		No	129 (88.4%)
	No	106 (72.6%)	Type of ICI				Yes	17 (11.6%)
	Yes	40 (27.4%)		Pembrolizumab	96 (65.8%)	Brain metastasis		
Hypertension				Nivolumab	25 (17.1%)		No	123 (84.2%)
	No	66 (45.2%)		Atezolizumab	15 (10.2%)		Yes	23 (15.8%)
	Yes	80 (54.8%)		Cemiplimab	4 (2.7%)	Bone metastasis		
Diabetes Mellitus				Avelumab	3 (2.1%)		No	102 (69.9%)
	No	111 (76.0%)		Durvalumab	2 (1.4%)		Yes	44 (30.1%)
	Yes	35 (24.0%)		Nivolumab + Ipilimumab	1 (0.7%)			

Table 1. Clinicopathologic features and comorbidities of the cohort. ECOG: Eastern Cooperative Oncology Group, CONUT: controlling nutritional status, ICI: immune checkpoint inhibitors, TKI: tyrosine kinase inhibitor, BMI: body mass index, NSCLC: non-small cell lung cancer, RCC: renal cell carcinoma, SCLC: small cell lung cancer, HCC: hepatocellular carcinoma

of all identified bootstrap-derived Thr values fell within one standard deviation of the selected value. The poor risk group composed of patients with BC3 score $\geq$ Thr has a median OS of 5.0 months, while the good risk group, with BC3 score < Thr, has a median OS of 35.8 months (hazard ratio (HR) adjusted for age, sex, tumor type, treatment line, number of metastatic sites, treatment group: 3.2, 95% confidence interval (CI) $[1.8 - 5.8$]; p-value$(p) < 0.001$). The distribution of CP features and BC3 score in the two groups is reported in Table 3. Analysis of the individual BC parameters within the BC3 score across the two risk groups demonstrated elevated values in the poor-risk group for all parameters, with statistically significant differences observed for both VFAI and the VF/SF ratio. The survival curves comparing the two risk groups are depicted in Fig. 2A with $p < 0.01$. In Fig. 2B, the forest plot presents the results of the multivariable Cox proportional hazards regression analysis on the entire cohort. The analysis reveals statistically significant associations between the risk groups and the BC3 score, tumor type, treatment line, and number of metastatic sites, distinguishing poor-risk from good-risk patients. The clinical significance of the BC3 score was further evaluated in a subset of 87 patients with advanced non-small cell lung cancer (NSCLC) undergoing ICI-based therapy. A statistically significant difference in median OS was observed between the poor-risk and good-risk groups, with medians of 8.0 and 22.6 months, respectively (HR adjusted for PD-L1 status, tumor type, sex, treatment line, number of metastatic sites, and treatment group: 2.7; 95% CI, $[1.4 - 5.2]$; $p = 0.003$) (Supplementary Fig. S3.A). The forest plot in Supplementary Fig. S3.B highlights statistically significant differences in BC3 score and the number of metastatic sites between the poor-risk and good-risk groups. The nomogram is depicted in Fig. 3 and the calibration plot of the Cox model used to construct the nomogram is represented in Supplementary Fig. S4 and reports a mean absolute error of 0.03. The predictive performance of the nomogram, evaluated through bootstrap-based internal validation, yielded an overall AUC of 0.76 ± 0.07. Stratified results by tumor type, treatment line, and treatment group are presented in Supplementary Fig. S5. For tumor type, only the NSCLC cohort was assessed, given its larger sample size, showing an AUC of 0.76 with minimal variability across bootstraps (0.06). Regarding treatment line, first-line therapy achieved a higher AUC compared to second-line therapy (0.81 vs. 0.66), with considerably lower variability (0.05 vs. 0.13). For treatment group, combination therapy outperformed monotherapy (AUC 0.85 vs. 0.71), but with higher variability (0.11 vs. 0.06). An example of nomogram score

Trial	Selected Features	Validation	Test
		Avg cdAUC ± std	Avg cdAUC ± std
Baseline CP	'Sex', 'Tumor Type', 'Treatment Group', 'Smoke', 'DM', 'HTA', 'NLR', 'BMI', 'CONUT'	0.75 ± 0.11	0.61 ± 0.08
FC1	'Sex', 'Tumor Type', 'Treatment Group', 'Line', 'ECOG', 'N Metastasis', 'Age', 'Smoke', 'CD', 'DM', 'HTA', 'CONUT', 'VF/SF'	0.82 ± 0.09	0.71 ± 0.05
FC2	'Sex', 'Tumor Type', 'Treatment Group', 'Smoke', 'CD', 'HTA', 'NLR', 'BMI', 'CONUT', 'BC2 score'	0.81 ± 0.09	0.67 ± 0.08
FC3	'Sex', 'Tumor Type', 'Treatment Group', 'Line', 'ECOG', 'N Metastasis', 'Smoke', 'CD', 'DM', 'HTA', 'NLR', 'CONUT', 'BC3 score'	**0.84 ± 0.08**	**0.73 ± 0.07**
FC4	'Tumor Type', 'Treatment Group', 'N Metastasis', 'NLR', 'BMI', 'CONUT', 'BC4 score'	0.82 ± 0.10	0.70 ± 0.06
FC5	'Sex', 'Tumor Type', 'Treatment Group', 'Line', 'ECOG', 'N Metastasis', 'Smoke', 'DM', 'HTA', 'NLR', 'BMI', 'CONUT', 'BC5 score'	0.82 ± 0.10	0.65 ± 0.09
FC6	'Sex', 'Tumor Type', 'Treatment Group', 'Line', 'ECOG', 'Smoke', 'CD', 'DM', 'HTA', 'NLR', 'BMI', 'CONUT', 'BC6 score'	0.76 ± 0.10	0.62 ± 0.09
FC7	'Sex', 'Tumor Type', 'Treatment Group', 'Line', 'ECOG', 'N Metastasis', 'Smoke', 'CD', 'DM', 'HTA', 'NLR', 'BMI', 'CONUT', 'BC7 score'	0.84 ± 0.09	0.68 ± 0.07
FC8	'Sex', 'Tumor Type', 'Line', 'ECOG', 'N Metastasis', 'Age', 'Smoke', 'DM', 'Stati', 'CONUT', 'BC8 score'	0.81 ± 0.10	0.66 ± 0.07

Table 2. Clinicopathologic features and comorbidities of the cohort. The best model performance is highlighted in bold. FC: feature combination, DM: diabetes mellitus, HTA: hypertension, smoke: smoking history, NLR: neutrophil-to-lymphocyte ratio, BMI: body mass index, CONUT: controlling nutritional status, ECOG: Eastern Cooperative Oncology Group, VF/SF: Visceral to subcutaneous fat area index ratio, N Metastasis: number of metastatic sites, Line: treatment line, Stati: statins assumption, BC score: body composition score.

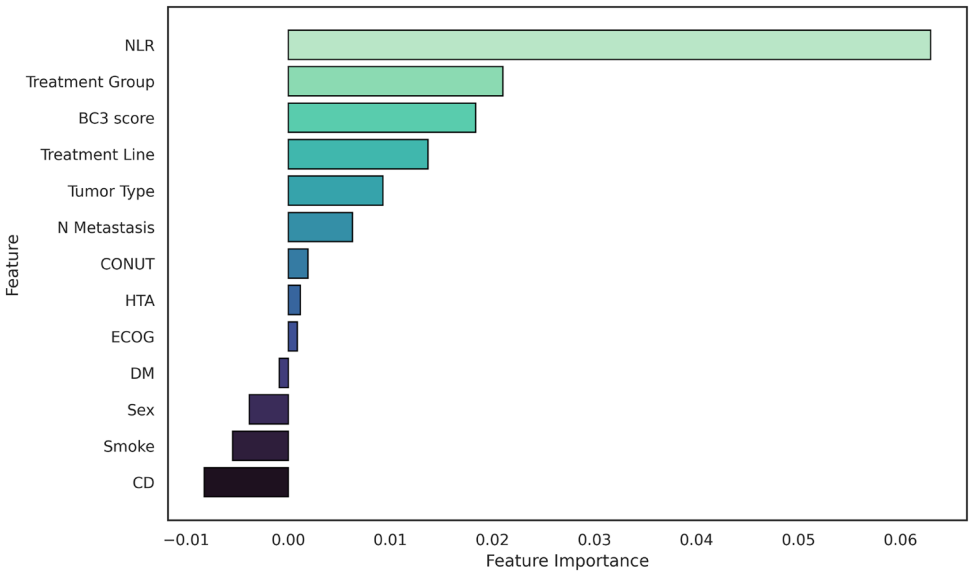

Figure 1. Average feature importance across splits. NLR: neutrophil-to-lymphocyte ratio, BC3 score: body composition score incorporating three BC parameters, N Metastasis: number of metastatic sites, smoke: smoking history, HTA: hypertension, CD: cardiovascular disease, CONUT: controlling nutritional status, DM: diabetes mellitus, ECOG: eastern cooperative oncology group.

calculation is shown for a patient with OS >12 months (Supplementary Fig. S6.A) and for a patient with OS <12 months (Supplementary Fig. S6.B).

Discussion

The aim of this study was to provide a potentially clinically applicable and interpretable nomogram that integrates both CP features and BC parameters to predict 1y OS in patients treated with ICIs.

The optimal set of CP features, identified by permutation importance feature selection algorithm, includes sex, tumor type, treatment group, smoking history, DM, HTA, NLR, BMI, and CONUT. The RSF model trained and tested with these features exhibits poor test performance with a significant drop in Avg cdAUC from validation

Feature		Poor risk (N=21)	Good risk (N=125)	P-value	Feature		Poor risk (N=21)	Good risk (N=125)	P-value
BC3 score	Score	53.9 [48.6;58.8]	24.2 [18.1;31.7]	**<0.01**	ECOG				
	IMAC	-0.16 [-0.26;-0.08]	-0.22 [-0.32;-0.08]	0.06		(0,1)	21 (100.0%)	118 (94.4%)	0.58
	VFAI	137.74 [124.02;149.98]	58.82 [37.73;82.76]	**<0.01**		(2,3)	0 (0%)	7 (5.6%)	
	VF/SF	2.54 [1.96;2.75]	1.27 [0.88;1.62]	**<0.01**					
NLR		4.5 [3.2;6.2]	4.0 [2.6;5.6]	0.39	DM	No	12 (57.1%)	99 (79.2%)	**0.05**
						Yes	9 (42.9%)	26 (20.8%)	
Sex	Male	21 (100%)	86 (68.8%)	**0.01**					
	Female	0 (0%)	39 (31.2%)		CD	No	13 (61.9%)	93 (74.4%)	0.49
						Yes	8 (38.1%)	32 (25.6%)	
Tumor Type	NSCLC	14 (66.6%)	73 (58.4%)	0.28					
	Melanoma	1 (4.8%)	11 (8.8%)		HTA	No	6 (28.6%)	60 (48.0%)	0.16
	Urothelial	2 (9.4%)	8 (6.4%)			Yes	15 (71.4%)	65 (52.0%)	
	HCC	0 (0%)	3 (2.4%)						
	RCC	1 (4.8%)	22 (17.6%)		N metastasis	0-2	14 (66.7%)	96 (76.8%)	0.47
	Squamous carcinoma	1 (4.8%)	0 (0%)			≥ 3	7 (33.3%)	29 (23.2%)	
	Breast carcinoma	0 (0%)	1 (0.8%)						
	SCLC	1 (4.8%)	3 (2.4%)		CONUT	Low (≤4)	13 (61.9%)	88 (70.0%)	0.62
	Head-neck carcinoma	1 (4.8%)	4 (3.2%)			High (>4)	8 (38.1%)	37 (30.0%)	
Line of treatment									
	First	16 (76.2%)	99 (79.2%)	0.96	Smoking history	Never	6 (28.6%)	43 (34.4%)	0.78
	Second	4 (19.0%)	20 (16.0%)			Current/Former	15 (71.4%)	82 (65.6%)	
	Third	1 (4.8%)	5 (4.0%)						
	Fourth	0 (0%)	1 (0.8%)		Treatment Group	Mono	17 (81.0%)	84 (67.2%)	0.31
						Combo	4 (19.0%)	41 (32.8%)	

Table 3. characteristics of patients stratified by risk group (good vs poor).Statistically significant differences (p-value ≤ 0.05) between groups are highlighted in bold. IMAC: intramuscolar adipose tissue content, VFAI: visceral fat area index, VF/SF: visceral to subcutaneous fat area index, NSCLC: non-small cell lung cancer, RCC: renal cell carcinoma, HCC: hepatocellular carcinoma, SCLC: small cell lung cancer, DM: diabetes mellitus, HTA: hypertension, CD: cardiovascular disease, NLR: neutrophil-to-lymphocyte ratio, CONUT: controlling nutritional status, ECOG: Eastern Cooperative Oncology Group, N Metastasis: number of metastatic sites, BC score: body composition score.

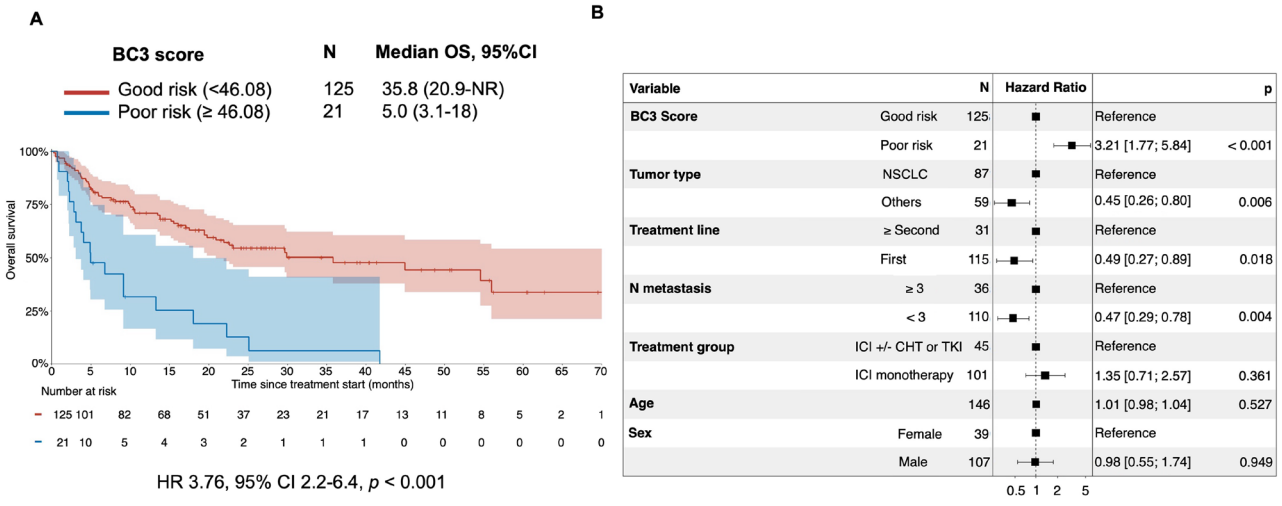

Figure 2. **A** Survival curves of whole cohort (N=146) stratified by BC3 score threshold. **B** Forest plot illustrating the results of multivariable Cox proportional hazards regression analysis in the whole cohort. BC3 Score: body composition score incorporating three BC parameters, NSCLC: non-small cell lung cancer, TKI: tyrosine kinase inhibitor, N metastasis: number of metastatic sites, CHT: chemotherapy, ICI: immune checkpoint inhibitor

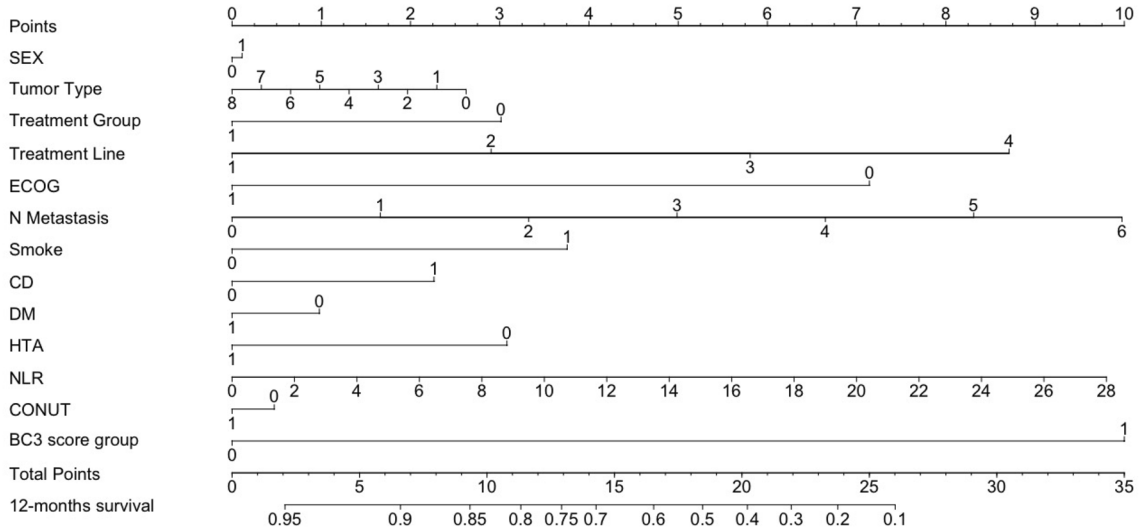

Figure 3. Nomogram with the selected variables identified in best combination (FC3) for the survival probability calculation at 1 year time horizon. NLR: neutrophil-to-lymphocyte ratio, BC3 score: body composition score incorporating three BC parameters, N Metastasis: number of metastatic sites, smoke: smoking history, HTA: hypertension, CD: cardiovascular disease, CONUT: controlling nutritional status, DM: diabetes mellitus, ECOG: eastern cooperative oncology group.

to test set. This drop indicates overfitting, highlighting the model's limited ability to generalize to unseen data. The best-performing BC parameter for OS prediction was VF/SF (0.69 ± 0.07). The RSF model trained with VF/SF demonstrated consistent performance across both the validation and test sets. However, the relatively low value of avg cdAUC in the test set indicates that VF/SF alone offers limited predictive power for OS. Analysis of the performance of RSF models trained with BC scores revealed that VF/SF and LPMAI alone outperformed all other BC-derived parameters in predicting OS. Notably, training the RSF model with BC scores did not enhance its predictive performance for OS compared to using VF/SF and LPMAI alone. However, a significant improvement in the RSF model's predictive accuracy was observed when CP features were combined with one of the newly introduced BC score, with FC3 being the best combination. The RSF model trained with FC3 achieved the highest validation scores for avg cdAUC (0.84 ± 0.08), indicating good discriminatory power in validation. While its test performance declined (avg cdAUC: 0.73 ± 0.07;), this drop is consistent with typical overfitting challenges in ML. Notably, even with this decline, the RSF model trained with FC3 achieved the best overall test performance among all models. In summary, the RSF model trained with FC3 emerged as the most promising framework, demonstrating robust validation performance and reasonable generalizability. Its superior predictive accuracy over RSF models trained solely with CP features or BC parameters, either alone or combined in a score, highlights the importance of integrating CP features with BC parameters (combined in a newly introduced score) for predicting OS in ICI-treated patients. FC3 was composed of sex, tumor type, treatment group, treatment line, ECOG, number of metastatic sites, smoking history, CD, DM, HTA, NLR, CONUT, and BC3 score. The calibration curve for the Cox proportional hazards model used to create the nomogram with the FC3 feature set, demonstrated good agreement between predicted and observed 1y OS probabilities. This suggests that the model's risk estimates are well-calibrated and can be reliably used for individual risk assessment in terms of 1y OS probability (mean absolute error of 0.03). Specifically, the calibration plot showed that the predicted survival probabilities closely followed the reference line, demonstrating that the nomogram derived from the Cox model generates accurate 1y OS predictions, thereby supporting its potential clinical utility for individualized prognostic assessment.

From a clinical point of view, the importance of the features selected as input to RSF in FC3 has also been confirmed by the work of Chowell *et al.*[16]. In this study, an ensemble random forest classifier was used to predict response to ICIs using a comprehensive pan-cancer dataset. The model called RF16, was fed with 16 clinically relevant features spanning genomic, molecular, clinical, and demographic characteristics. Specifically, RF16 included: tumor mutational burden (TMB), fraction of copy number alterations (FCNA), HLA-I evolutionary divergence (HED), HLA-I loss of heterozygosity (LOH), microsatellite instability (MSI) status, BMI, sex, NLR, tumor stage, specific ICI agent, and patient age. Additionally, cancer type, prior chemotherapy exposure, and hematologic parameters (albumin, platelets, and hemoglobin levels) were included in the model. Feature importance analysis identified TMB as the most significant predictor of treatment response, with prior chemotherapy, albumin levels, NLR, and age also demonstrating substantial predictive value. RF16 was trained on 80% of the dataset and tested on the remaining 20%, exhibiting strong discriminatory performance, with AUC values of 0.85 in the training cohort and 0.79 in the test cohort, indicating robust predictive capability across diverse patient populations. Comparing our findings, the performance achieved by our model yielded an avg cdAUC of 0.84 and 0.73 in the training and test set, respectively. Regarding the feature contribution, our model also confirmed the importance of NLR, identified as the most important feature. The importance of

tumor type, treatment line, and NLR in predicting OS was further confirmed by Chang *et al.*[15]. In this study, the authors employed logistic regression to predict ICI response in a large cohort encompassing 18 solid tumor types. Their model (LLR6), incorporating six clinically relevant features ranked by importance (TMB, prior systemic therapy, albumin levels, NLR, age, and tumor type), demonstrated good discriminative performance with an AUC of 0.75 in the test cohort. Interestingly, the predictive importance of some variables selected in our work has been confirmed also in a recent study evaluating the predictive role of blood tests and clinical characteristics in a pan-cancer cohort[16]. NLR, albumin, and BMI were among the most important features for the prediction of clinical outcomes, achieving an AUC(t) of 0.76 for 1y OS. The prognostic role of the BC parameters included in the BC3 score is also strongly supported by existing literature. Specifically, the BC3 score incorporates three well-validated metrics, VF/SF, IMAC, and VFAI, each weighted by optimized coefficients derived from the optimization analysis. The prominence of VF/SF as the most influential feature with higher weight, along with its elevated value in the poor-risk group, may underscore the distinct roles of visceral and subcutaneous adiposity in immune and metabolic regulation[17] and the association of visceral adiposity with systemic pro-inflammatory effects and metabolic dysfunctions[18]. Additionally, our findings about the negative prognostic role of VF/SF is in line with evidence showed in lung cancer literature[19]. The prognostic significance of the BC score is further demonstrated by the effective risk stratification across the study population. Through systematic Thr analysis, confirmed by the bootstrap distribution of optimal cutoff values, where 44% of all identified Thr values fell within one standard deviation of the selected value (46.08 $\frac{cm^2}{m^2}$), an optimal cut-point was identified that effectively stratified patients into two distinct prognostic groups, with significantly different survival outcomes ($p \leq 0.001$). This robust discrimination, consistent across the overall population and NSCLC subgroup, underscores the critical role of BC parameters in predicting ICI treatment outcomes. Interestingly, these results in terms of BC parameters have a clinical equivalent. Specifically: (1) CD and DM were among the most predictive comorbidities selected in FC3, and (2) a1 higher percentage of patients had CD and DM in the poor-risk group compared to the good-risk group, as stratified by the BC3 Thr, although this difference was statistically significant only for DM. In terms of endpoints, the choice of OS as the clinical endpoint evaluated in the model was based on the weak relationship between OS and surrogate endpoints in clinical trials enrolling patients treated with ICIs[20].

The results indicate that incorporating CP and BC variables improves OS prediction, consistent with the findings of Gu *et al.*[21] and the features selected in the optimal combination (FC3) closely align with established literature. Moreover, the consistent selection of tumor type and treatment groups by the feature selection algorithm, together with the fact that treatment line is included in almost all feature combinations, highlights that the model inherently accounts for the heterogeneity of the cohort. In addition, the feature importance ranking (Fig. 2) from the model trained and tested on the FC3 combination further confirms that tumor type, treatment line, and treatment group are among the most influential predictors. Specifically, the consistency between FC3 and previously identified predictive biomarkers reinforces the biological and clinical relevance of our approach while supporting its robustness in identifying clinically meaningful predictors of ICI response. Furthermore, by applying the "exposome" framework[7], which emphasizes the role of external determinants in clinical outcomes beyond intrinsic tumor characteristics on disease outcomes, our analysis demonstrates the critical role of extrinsic determinants, including comorbidities, nutritional-inflammatory biomarkers, and radiologic BC parameters in predicting OS for ICI-treated patients. This approach extends beyond conventional tumor-intrinsic features, providing a more comprehensive prognostic model that reflects the complex interplay between host factors and therapeutic response. This was achieved by integrating clinical and laboratory-based nutritional-inflammatory parameters, which are routinely evaluated in clinical practice, with BC parameters derived from CT scans, combined in an innovative score. The strength of this work lies in the development of a clinically interpretable nomogram that predicts 1y OS using readily available parameters. By integrating routine clinical and laboratory-based nutritional-inflammatory markers with CT-derived BC parameters, we created a practical tool that have prognostic performance comparable to more complex models, such as the ones reported above, while also exhibiting similar overfitting degree, despite being trained on a smaller cohort. The nomogram demonstrated good performance on the test set, across bootstraps yielding an AUC of 0.76 ± 0.06, and similar performances in the stratified analyses. In particular, the NSCLC subgroup showed good predictive accuracy. Performance was consistent in the NSCLC cohort, the largest subgroup. The higher accuracy in firs line versus second line therapy (0.81 vs. 0.66) suggests stronger predictive value earlier in the disease course, while reduced performance in later lines may reflect greater clinical heterogeneity. Similarly, the superior results in combination versus monotherapy (0.85 vs. 0.71) highlight the model's potential in more intensive treatment settings, though increased variability indicates a need for larger validation cohorts. Overall, these findings support the nomogram's clinical promise, but emphasize the importance of external validation and further refinement before routine use. Specifically, this nomogram offers two key clinical advantages: immediate feasibility, as it utilizes only parameters routinely assessed (standard blood tests and diagnostic CT scans); and broader accessibility, as it eliminates dependence on TMB, a biomarker requiring specialized genomic testing not routinely available in most practice settings. Therefore, by eliminating the need for additional molecular profiling, this approach significantly lowers implementation barriers while maintaining prognostic accuracy comparable to more complex models. Furthermore, this work introduces an innovative BC score, which significantly improves the OS predictive performance. This score not only strengthens the model's prognostic performance but also underscores the importance of integrating BC metrics into predictive frameworks for cancer outcomes. Such approach allows to embrace the complexity of cancer patients treated with ICIs in a multimodal setting. To the best of our knowledge, this is the first work that integrates BC parameters and CP features into a user-friendly nomogram capable of predicting 1y OS with high accuracy. However, this study also presents some limitations. First, this is a single-center, retrospective study including a heterogeneous cohort of

patients with various solid tumor types, treatment groups, and lines, and with a relatively limited sample size. These characteristics increase the risk of overfitting in the RSF models, particularly when evaluating combinations of features rather than individual BC parameters or scores. When a model relies on a single variable, the risk of overfitting is inherently lower due to reduced model complexity. This suggests that the overfitting observed in RSF models trained and tested on individual BC parameters or scores is more likely due to the different distribution of data in train and test set[22]. Furthermore, the small number of patients and events within each subgroup precluded performing stratified analyses. Despite these limitations, RSF was primarily used to identify the most informative combination of features for constructing the nomogram. A second limitation of this study is the absence of an external validation cohort to fully assess the clinical applicability and generalizability of the proposed nomogram. Third, we did not consider the lines of therapy in a curative setting, such as any adjuvant and neoadjuvant systemic therapies[23]. Despite these limitations, the results obtained are promising, highlighting the nomogram's ability to capture significant patterns and relationships within the data. Future efforts will focus on expanding the training cohort to include a larger number of patients and on collecting an external validation cohort from additional centers. This strategy aims to improve the model's performance, robustness, and specificity, facilitating a more in-depth understanding of its efficacy across various histological types. To further improve scalability and reproducibility, future works will integrate emerging automated segmentation tools[12] for automatic extraction of BC parameters from routine CT imaging. Future prospective studies are warranted to validate the added value of incorporating BC parameters alongside CP characteristics for more comprehensive prediction of clinical outcomes in cancer patients treated with ICIs.

Conclusions

This study presents a potentially clinically applicable nomogram that integrates routinely available CP features with a novel BC score to predict 1y OS in ICI-treated cancer patients. Although the small sample size, and the lack of external validation cohort, our findings not only demonstrate the efficacy of this approach but also highlight the prognostic importance of incorporating BC parameters with conventional CP features. Future efforts aim to expand the cohort.

Methods

Patients

We retrospectively enrolled patients with a diagnosis of advanced solid tumor treated with ICIs-based treatment (Table 1) at Polytechnic University of Marche – Azienda Ospedaliera Universitaria of Marche (Italy). Patients treated with at least one cycle of ICIs, alone or in combination with chemotherapy or with TKI, were included in the study. CP data, including ECOG, smoking history, number of metastatic sites, were collected from medical records by trained research staff. Moreover, the baseline nutritional-inflammatory status was evaluated by NLR and CONUT score (assessed as shown in Supplementary Table S8). The past medical history was reviewed to collect the CD, defined as any type of disease affecting the heart or blood vessel according to National Cancer Institute and American Heart Association, DM, HTA, statin use at ICI's baseline, and BMI. Ethical approval to conduct this study was obtained by the respective local ethical committees on human experimentation of each participating center, after previous approval by the coordinating center ("Comitato Etico Territoriale- C.E.T.," Reference Number 19/792). The procedures and data collection were performed according to the Declaration of Helsinki and to the Good Clinical Practice. Informed consent was obtained from the patients, in accordance with national law. The categorization of variables is reported in Supplementary Table S9).

Clinical outcomes

The response to ICIs was assessed by investigators according to Response Evaluation Criteria in Solid Tumors version 1.1 (RECIST 1.1). We evaluated the OS and PFS based, respectively, on the interval from the date of ICI's based treatment start until the date of death/last follow up or until the date of progression/death/last follow-up. Patients who did not experience progression or death, were censored at the date of last radiological evaluation; for the overall survival, the patients alive at the data lock were censored at the date of last follow-up.

Body composition assessment

The BC was assessed on CT considering an unenhanced, axial image at the level of L3. CT images were processed with a dedicated software (Syngo.via, Siemens Healthineers, Forcheim) by two radiologists in consensus. The analysis of body fat and muscles was performed with a specific tool using semiautomatic and threshold segmentation (Supplementary Fig. S7). Specifically, the default threshold values used were as following: skeletal muscle -50 to +142 Hounsfield Unit (HU), and for adipose tissue -150 to -50 HU. The BC parameters extracted are SMI, PMI, IMAC, lean psoas muscle area (LPMA), SBI, visceral fat area (VFA), subcutaneous fat area (SFA), VF/SF and total fat area (TFA). The equations used to compute the BC parameters, along with their distribution in the train and test sets, are reported in Supplementary Table S10.

Statistical analysis

Categorical and continuous variables were summarized descriptively using percentages and medians. The Wilcoxon-Rank Sum test and Kruskal-Wallis test were used to examine differences between continuous variables, and Fisher's exact test was used to compare associations between categorical variables. Linear correlations were evaluated using Spearman's test. p values were based on a two-sided hypothesis with confidence intervals at the 95% level, and statistical significance was set at $p \leq 0.05$. Log-rank tests were used to test for differences in event-time distributions, and Cox proportional hazards models were fitted to obtain estimates of HR in

univariate models. The multivariable Cox proportional hazards regression analysis was performed to adjust for potential confounders. The hazard HR and CI 95% were reported for each covariate.

Two-stage procedure for overall survival prediction

The proposed method, named two-stage procedure for OS prediction (Supplementary Fig. S8), follows a two-stage procedure. The aim of the first stage is to assess the RSF model perfomance trained and tested with CP features (Supplementary Fig. S9.A), BC parameters, either alone or combined into a newly introduced BC score (Supplementary Fig. S9.B) and their combination (Supplementary Fig. S9.C) to predict OS. Finally, the best identified FC is passed to the second stage whose aim is to create a nomogram.

First stage

In the first stage, the RSF model was used to assess the baseline OS predictive performance of CP features and BC parameters alone, selected using a permutation importance algorithm for feature selection. The model was trained and tested using a bootstrap approach. The 80% of the data was used to train and validate the model (80% training and 20% validation) while the remaining 20% for testing. Simultaneously, $(m-1)$ BC scores were generated by exploring all possible combinations, ranging from 1 to $(m-1)$ BC parameters, being m the total number of BC parameters. This process was carried out using an innovative optimization procedure which identifies the weights assigned to each BC parameter by integrating a genetic algorithm (GA) with survival ROC analysis (sROC) for 1y OS prediction. In detail, GA was employed to iteratively optimize the weights assigned to BC parameters, evolving them toward the optimal solution by maximizing the AUC of sROC for 1y OS prediction. Then, the predictive performance of all possible FC were systematically evaluated using the RSF model coupled with the feature selection algorithm that takes in input both CP features and the generated BC scores. This comprehensive assessment employed $m-1$ trials (named as $(FC1, ..., FC[m-1])$, where the index denotes the number of BC parameters incorporated in each score). Through subsequent feature importance analysis, the most prognostically significant variables influencing OS predictions were assessed. This approach not only enabled objective comparison of different FC configurations to determine the optimal integration of CP features and BC scores, but also revealed an optimal set of feature for the OS prediction. The best identified FC was then passed to the second stage.

Second stage

A computational approach was employed to identify the optimal Thr for the BC score, aiming to stratify patients into poor-risk and good-risk groups based on OS. Multiple candidate thresholds were systematically evaluated, and for each, the survival differences between the resulting groups were assessed using the log-rank test. The threshold yielding the greatest statistical separation, i.e. the lowest p-value, was selected as the optimal cutoff. Survival curves were estimated using the Kaplan-Meier method, with statistical significance determined via the log-rank test. To address potential overfitting and the risk of circularity from deriving the Thr within the training dataset, bootstrap resampling was used. The cutoff optimization procedure was repeated across 200 bootstrap iterations on resampled training sets, generating a distribution of optimal thresholds. The stability and variability of these thresholds are shown in Supplementary Fig. S2. Each risk group was then described by $50^{th}[25^{th}; 75^{th}]$ percentiles for continuous features and by class frequency for categorical features. Statistical significance (p) was evaluated with the statistical tests reported in Statistical Analysis section. This approach enabled the identification of an optimal Thr for the selected BC score, facilitating its clinical interpretation. Therefore, the BC score was categorized (BC score group) as follows: BC score group = 0 if BC score < Thr otherwise BC score group = 1 if BC score $\geq$ Thr. A further analysis was conducted to validate the findings of the Thr for BC score within the lung cancer cohort, specifically focusing on the NSCLC subgroup, as it represents the largest subset in the collected dataset. This analysis aimed to evaluate how the Thr for BC score stratifies NSCLC patients into two distinct risk groups (good-risk and poor-risk groups) and to assess whether these groups demonstrate significant differences in OS. Next, the selected FC from the first stage, together with the BC score group, were given in input to the Cox model to develop the nomogram. The nomogram was developed to visually represent the multivariable Cox proportional hazards model, enabling individualized prediction of 1y OS probability. Finally, the nomogram's predictive accuracy was assessed using the calibration curve that compares predicted survival probabilities against observed outcomes.

Two-stage procedure implementation and evaluation metrics

The proposed two-stage procedure was developed in Python, using the Pro version of the Google Colab cloud service and in RStudio (4.4.2). All the RSF models were trained and tested using a bootstrap approach with the number of bootstraps set to 200. The predictive performance of RSF models was evaluated using the avg cdAUC ± std across bootstraps which provides a time-dependent assessment of the model's performance, measuring its ability to predict survival across different time points. To evaluate the model's generalization capabilities, performance metrics were assessed on both the validation and test sets. The Cox model used for constructing the nomogram was evaluated with calibration curve at 1y landmark time. The predictive performance of the nomogram was evaluated using bootstrap-based validation (200 iterations). In each iteration, the nomogram score was calculated for the test set. Discrimination was quantified using the 12 months AUC, both overall and stratified by tumor histology, treatment group and treatment line.

Data Availability

The data in this study are not publicly available but may be available from the corresponding author upon reasonable request.

Received: 20 June 2025; Accepted: 22 January 2026
Published online: 12 March 2026

References

1. Suijkerbuijk, K. P. M., Van Eijs, M. J. M., Van Wijk, F. & Eggermont, A. M. M. Clinical and translational attributes of immune-related adverse events. *Nature Cancer.* **5**(4):557-571. doi: 10.1038/s43018-024-00730-3 (2024).
2. Morad, G., Helmink, B. A., Sharma, P. & Wargo, J. A. Hallmarks of response, resistance, and toxicity to immune checkpoint blockade. *Cell.* **184**(21):5309-5337. doi: 10.1016/j.cell.2022.01.008 (2021).
3. Catalano, M. et al. Immunotherapy-related biomarkers: Confirmations and uncertainties. *Crit. Rev. Oncol. Hematol.* 2023 Dec;192:104135. doi: 10.1016/j.critrevonc.2023.104135.
4. Wang, S. L. & Chan, T. A. Navigating established and emerging biomarkers for immune checkpoint inhibitor therapy. *Cancer Cell.* **43**(4): 641–664. doi: 10.1016/j.ccell.2025.03.006 (2025).
5. Holder, A. M. et al. Defining clinically useful biomarkers of immune checkpoint inhibitors in solid tumours. *Nature Rev. Cancer*, 498–512. doi: 10.1038/s41568-024- (2024).
6. Lysaght, J. & Conroy, M. J. The multifactorial effect of obesity on the effectiveness and outcomes of cancer therapies. *Nat. Rev. Endocrinol.* **20**(12):701-714. doi: 10.1038/s41574-024-01032-5 (2024).
7. Pizzutilo, E. G. et al. Immune checkpoint inhibitors and the exposome: Host-extrinsic factors determine response, survival, and toxicity. *Cancer Res.* **83**(14):2283-2296. doi: 10.1158/0008-5472.CAN-23-0161 (2023).
8. Cortellini, A. et al. A multicenter study of body mass index in cancer patients treated with anti-PD-1/PD-L1 immune checkpoint inhibitors: when overweight becomes favorable. *J. Immunother. Cancer.* 2019 Feb 27; 7(1):57. doi: 10.1186/s40425-019-0527-y.
9. Kichenadasse, G. et al. Association between body mass index and overall survival with immune checkpoint inhibitor therapy for advanced non-small cell lung cancer. *JAMA Oncol.* **6**(4):512-518. doi: 10.1001/jamaoncol.2019.5241 (2020).
10. Ma, W. et al. Impact of baseline body composition on prognostic outcomes in urological malignancies treated with immunotherapy: a pooled analysis of 10 retrospective studies. *BMC Cancer.* **24**(1):830. doi: 10.1186/s12885-024-12638-3 (2024).
11. Alden, S. L. et al. Pan-tumor analysis to investigate the obesity paradox in immune checkpoint blockade. *J. Immuno Therapy Cancer.* 2025 Jan 19;13(1):e009734. doi: 10.1136/jitc-2024-009734.
12. Chaunzwa, T. L. et al. Body composition in advanced non-small cell lung cancer treated with immunotherapy. *JAMA Oncol.* **10**(6):773-783. doi: 10.1001/jamaoncol.2024.1120 (2024).
13. Addala, V. et al. Computational immunogenomic approaches to predict response to cancer immunotherapies. *Nat. Rev. Clin. Oncol.* **21**(1):28-46. doi: 10.1038/s41571-023-00830-6 (2024).
14. Li, Y., Wu, X., Fang, D. & Luo, Y. Informing immunotherapy with multi-omics driven machine learning. *NPJ Digit. Med.* **7**, 67 (2024). doi: 10.1038/s41746-024-01043-6 (2024).
15. Chang, T. G. et al. LORIS robustly predicts patient outcomes with immune checkpoint blockade therapy using common clinical, pathologic and genomic features. *Nat. Cancer.* **5**(8):1158-1175. doi: 10.1038/s43018-024-00772-7 (2024).
16. Chowell, D. et al. Improved prediction of immune checkpoint blockade efficacy across multiple cancer types. *Nat. Biotechnol.* **40**(4):499-506. doi: 10.1038/s41587-021-01070-8 (2022).
17. Baggerman, M. R. et al. Computed tomography reference values for visceral obesity and increased metabolic risk in a caucasian cohort. *Clin. Nutr. ESPEN.* **48**:408-413. doi: 10.1016/j.clnesp.2022.01.009 (2022).
18. Yoo, S. K. et al. Prediction of checkpoint inhibitor immunotherapy efficacy for cancer using routine blood tests and clinical data. *Nat. Med.* **31**(3):869-880. doi: 10.1038/s41591-024-03398-5 (2025).
19. Ibrahim, M. M. Subcutaneous and visceral adipose tissue: structural and functional differences. *Obes. Rev.* **11**(1):11-8. doi: 10.1111/j.1467-789X.2009.00623.x (2010).
20. Nattenmüller, J. et al. Prognostic impact of CT-quantified muscle and fat distribution before and after first-line-chemotherapy in lung cancer patients. *PLoS One.* **12**(1):e0169136. doi: 10.1371/journal.pone.0169136 (2017).
21. Gu, H. et al. Improving surgical risk prediction through integrating automated body composition analysis: a retrospective trial on colectomy surgery. arXiv preprint arXiv:2506.11996https://doi.org/10.48550/arXiv.2506.11996 (2025).
22. Strobl, C., Boulesteix, A. L., Zeileis, A. & Hothorn, T. Bias in random forest variable importance measures: Illustrations, sources and a solution. *BMC Bioinformatics*. **8**:25. doi: 10.1186/1471-2105-8-25 (2007).
23. Saini, K. S. & Twelves, C. Determining lines of therapy in patients with solid cancers: A proposed new systematic and comprehensive framework. *Br. J. Cancer.* **125**(2):155-163. doi: 10.1038/s41416-021-01319-8 (2021).

Author contributions

All authors have read and approved the manuscript. Individual contributions are listed below: G.B. conceptualization, methodology, machine learning model training, testing and validation, manuscript writing. F.P. & F.P. conceptualization, data curation, methodology, analysis, manuscript writing, E.T., T.G., L.S. & M.G. data curation, critical revision of the article, A.B., L.C. & A.A. methodology and critical revision of the article, V.C., V.L. N.C., V.A., M.D.P.P., A.S., G.M., S.F., G.M., A.P., R.G., K.S.S., S.B., M.T., A.V. & A.G., critical revision of the article, L.B. & R.B. critical revision of the article, supervision.

Funding

There are no funds to declare in support for this work.

Declarations

Competing interests

The authors declare no competing interests.

Ethics and consent to participate

The study was approved by the institutional review board ("Comitato Etico Regionale delle Marche - C.E.R.M.," Reference Number 19/792) and was performed in accordance with the Declaration of Helsinki and in accordance with Good Clinical Practice. Informed consent was obtained from the patients, in accordance with national law. For this study, due to its retrospective nature, (Comitato Etico Regionale delle Marche - C.E.R.M.) waived the need of obtaining informed consent.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-026-37510-1.

Correspondence and requests for materials should be addressed to F.P. or R.B.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.