iScience, Volume 26

Supplemental information

Machine learning-based risk model incorporating

tumor immune and stromal contexture predicts

cancer prognosis and immunotherapy efficacy

Li-Na He, Haifeng Li, Wei Du, Sha Fu, Linfeng luo, Tao Chen, Xuanye Zhang, Chen Chen, Yongluo Jiang, Yixing Wang, Yuhong Wang, Hui Yu, Yixin Zhou, Zuan Lin, Yuanyuan Zhao, Yan Huang, Hongyun Zhao, Wenfeng Fang, Yunpeng Yang, Li Zhang, and Shaodong Hong

Supplemental information

This file includes Figures S1-S12, Table S1, and Table S9-S14.

Tables S2-S8 that accompany this paper are provided in xlsx format.

Figure S1. Identification of cut-points of immune scores and stromal scores and their association with prognosis, related to Figure 2

Figure S2. Univariate analyses of candidate mode genes and identification of cut-point for TMErisk, related to Figure 2 and Figure 3

Figure S3. Distribution of TMErisk according to different baseline clinicopathological characteristics in the TCGA-LUSC cohort, related to Figure 2

Figure S4. Forest plot for hazard ratios of OS in the TCGA-LUSC training cohort according to TMErisk and other baseline clinicopathological characteristics, related to Figure 2

Figure S5. Association of TMErisk with chemokine genes and immune-related genes, related to Figure 4 **Figure S6.** Pearson correlation analysis between selected immune-related genes and TGM2 (A) as well as C4BPA (B), related to Figure 4

Figure S7. Pearson correlation analysis between selected immune-related genes and C11orf96 (A) as well as PLAAT4 (B), related to Figure 4

Figure S8. Pearson correlation analysis between selected immune-related genes and PNCK (A) as well as KLF5 (B), related to Figure 4

Figure S9. Mutation status between the low- and high-TMErisk groups, related to STAR Methods

Figure S10. Prediction of drug sensitivity for the low- and high-TMErisk groups, related to Figure 5 and Figure 6

Figure S11. Validation of predictive value of TMErisk for immunotherapy in GEO datasets, related to Figure 5

Figure S12. Forest plot for hazard ratios of OS and PFS in the POPLAR, IMvigor210 and ORIENT-11 cohorts according to TMErisk and other baseline clinicopathological characteristics, related to Figure 6 and Figure 7

Table S1. Baseline patient characteristics in the TCGA cohort, related to STAR Methods

Table S9. Infiltration levels of immune and stromal cells in the low- and high-TMErisk groups, related to Figure 4

Table S10. Univariate Cox analyses for OS and PFS in the immunochemotherapy group from ORIENT-11, related to Figure 7

Table S11. Univariate Cox analyses for OS and PFS in the chemotherapy group from ORIENT-11, relatedto Figure 7

 Table S12.
 Baseline patient characteristics in the POPLAR and IMvigor210 cohorts, related to STAR

 Methods

Table S13. Baseline patient characteristics in the ORIENT-11 cohort, related to STAR Methods

Table S14. Cut-points used in each dataset, related to STAR Methods



Figure S1. Identification of cut-points of immune scores and stromal scores and their association with prognosis, related to Figure 2. (A-B) Optimal cut-offs for immune score (A) and stromal score (B) identified by the maximally selected rank statistics. Scatters in the bottom panel indicate corresponding standardized log-rank statistics of every gene expression cut-point. The vertical dashed line indicates the optimal cut-point. The upper panel presents the density distribution histogram of the low- and high-immune/stromal score groups. (C-D) Kaplan–Meier survival analysis based on immune score (C) and stromal score (D) strata.



Figure S2. Univariate analyses of candidate mode genes and identification of cut-point for TMErisk, related to Figure 2 and Figure 3. (A) Forest plot for hazard ratios of OS in the TCGA-LUSC training cohort according to 15 candidate mode genes. (B) Optimal cut-off for TMErisk identified by the maximally selected rank statistics. Scatters in the bottom panel indicate corresponding standardized log-rank statistics of every gene expression cut-point. The vertical dashed line indicates the optimal cut-point. The upper panel presents the density distribution histogram of the low- and high-TMErisk groups. OS, overall survival; TCGA, the cancer genome atlas; LUSC, lung squamous cell carcinoma.



Figure S3. Distribution of TMErisk according to different baseline clinicopathological characteristics in the TCGA-LUSC cohort, related to Figure 2. Statistical difference between two groups was tested by Wilcoxon rank sum test, and Kruskal-Wallis test was used for comparison among three or more groups. *P < 0.05; ns, not significant; ECOG, Eastern Cooperative Oncology Group; PS, performance status.

A	Baseline characteristics		Hazard ratio (95% CI)	<i>P</i> -value
	TMErisk (High vs. Low)	⊢ −−−−1	2.877 (2.089-3.961)	<0.001
	Age, years (≥68 vs. <68)	H=-1	1.132 (0.861-1.488)	0.374
	ECOG PS (≥2 vs. <2)	⊢ →	3.306 (1.351-6.823)	0.007
	T stage			
	T2 vs. T1	H	1.244 (0.878-1.764)	0.220
	T3 vs. T1	⊢ •−−−+	1.817 (1.161–2.844)	0.009
	T4 vs. T1	·	2.323 (1.248-4.326)	0.008
	N stage			
	N1 vs. N0	⊢ ⊷i	1.071 (0.782-1.466)	0.669
	N2 vs. N0	H	1.320 (0.831-2.094)	0.239
	N3 vs. N0	·	2.511 (0.618-10.205)	0.198
	M stage (M1 vs. M0)	⊢ →	3.483 (1.534-7.911)	0.003
	AJCC pTNM stage			
	Stage II vs. Stage I	H	1.137 (0.827-1.565)	0.429
	Stage III vs. Stage I		1.553 (1.087–2.218)	0.015
	Stage IV vs. Stage I	·	3.331 (1.348-8.228)	0.009
	Smoking status (Yes vs. No)) • • • • •	0.589 (0.260-1.334)	0.205
	Gender (Male vs. Female)	H=1	1.196 (0.868-1.647)	0.275
	C).2 11.522.533.5 4.55 6	3	
в				
	Baseline characteristics	H	azard ratio (95% CI) <i>P</i> –	value
	TMErisk (High vs. Low)		3.057 (2.156-4.336) <0	0.001

Figure S4. Forest plot for hazard ratios of OS in the TCGA-LUSC training cohort according to TMErisk and other baseline clinicopathological characteristics, related to Figure 2. (A) Univariate Cox analysis for OS according to different variables. (B) Multivariate Cox analysis for OS according to different variables. OS, overall survival.

Т Т

4.5 5

٦

6

2.043 (0.895-4.664)

1.490 (1.036-2.142)

2.071 (1.293-3.317)

1.779 (0.939-3.369)

3.583 (1.549-8.287)

0.090

0.031

0.002

0.077

0.003

В

ECOG PS (≥2 vs. <2)

M stage (M1 vs. M0)

Г

Т

0.55 1.5 2 2.5 3 3.5

Т

Т

Т Т

T stage T2 vs. T1

T3 vs. T1

T4 vs. T1



Figure S5. Association of TMErisk with chemokine genes and immune-related genes, related to Figure 4. (A) Expression level of chemokine genes in the low- and high-TMErisk groups. (B) Expression level of immune-related genes in the low- and high-TMErisk groups. Differences were examined by Wilcoxon test. * P < 0.05; ** P < 0.01; *** P < 0.001; ns, not significant.



Figure S6. Pearson correlation analysis between selected immune-related genes and TGM2 (A) as well as C4BPA (B), related to Figure 4.



Figure S7. Pearson correlation analysis between selected immune-related genes and C11orf96 (A) as well as PLAAT4 (B), related to Figure 4.



Figure S8. Pearson correlation analysis between selected immune-related genes and PNCK (A) as well as KLF5 (B), related to Figure 4.



High (n=70) vs. Low (n=419)

Gene name	High TMErisk (%)	Low TMErisk (%)	OR (95% CI)	P-value
CDK6	23 (32.86)	234 (55.85)		0.000
FAM133B	23 (32.86)	232 (55.37)		0.001
LOC101927497	23 (32.86)	232 (55.37)	—	0.001
MTOR	21 (30.00)	217 (51.79)	—	0.001
PLCH1	0 (0.00)	49 (11.69)	·	0.001
BCL11A	27 (38.58)	253 (60.38)	—	0.001
MIR4432	27 (38.58)	250 (59.67)	—	0.001
MIR4432HG	27 (38.58)	250 (59.67)		0.001
ATP13A2	21 (30.00)	215 (51.31)		0.001
COL6A3	19 (27.14)	201 (47.97)		0.001
IGSF3	24 (34.26)	232 (55.37)	—	0.001
CASZ1	21 (30.00)	212 (50.60)	—	0.002
KIF17	21 (30.00)	212 (50.60)	—	0.002
MMEL1	21 (30.00)	212 (50.60)		0.002
PRDM2	21 (30.00)	212 (50.60)	—	0.002
			0 0.3 0.5 1	1.2 1.5

2		
High	(n=70) vs	low(n=419)

Gene name	High TMErisk (%)	Low TMErisk (%)	OR (95% CI)
OR7C2	5 (7.14)	3 (0.72)	
RNF10	5 (7.14)	3 (0.72)	
SLC35F1	5 (7.14)	4 (0.95)	·
VGLL1	5 (7.14)	5 (1.19)	
JMJD1C	7 (10.00)	12 (2.86)	·
DDC	5 (7.14)	6 (1.43)	· · ·
HEG1	5 (7.14)	6 (1.43)	· · ·
CDKL5	6 (8.57)	9 (2.15)	· · · · · · · · · · · · · · · · · · ·
ATAD2	7 (10.00)	13 (3.10)	
MYO15A	7 (10.00)	15 (3.58)	
IRF6	5 (7.14)	8 (1.91)	
ITPR1	5 (7.14)	8 (1.91)	
KIAA1524	5 (7.14)	8 (1.91)	
VWA3A	5 (7.14)	8 (1.91)	
	4 (5 71)	5 (1 19)	



Figure S9. Mutation status between the low- and high-TMErisk groups, related to STAR Methods. (A-B) Comparison of gene mutation frequencies between the low- and high-TMErisk groups. Forest plots show the top 15 genes mutated more frequently in the low-TMErisk (A) and high-TMErisk (B) group. (C) The top 20 mutated genes and distribution of mutation types in the low-TME risk group. (D) The top 20 mutated genes and distribution of mutation types in the high-TME risk group. (E) The lollipop plot shows the mutation types and mutation sites of TP53 according to TMErisk strata. OR, odds ratio.

В



Figure S10. Prediction of drug sensitivity for the low- and high-TMErisk groups, related to Figure 5 and Figure 6. Prediction was generated based on GDSC and CTRP cell line databases. Statistical difference was examined by Wilcoxon test. CTRP, Cancer Therapeutics Response Portal; GDSC, Genomics of Drug Sensibility in Cancer.



Figure S11. Validation of predictive value of TMErisk in GEO datasets, related to Figure 5. (A) Kaplan–Meier analysis for progression-free survival based on TMErisk strata in the GSE135222 cohort. (B) Kaplan–Meier analysis for overall survival based on TMErisk strata in the GSE78220 cohort. (C) Association of TMErisk with respond to immunotherapy in the GSE78220 cohort. BOR, best overall response; R, responders, patients with BOR of complete response or partial response; NR, non-responders, patients with BOR of stable disease or progressive disease.

A	Baseline characteristics	POPLAR Atezolizumab	HR (95% CI)	<i>P</i> -value	B Baseline characteristics OS POPLAR Docetaxel	HR (95% CI)
	TMErisk (high vs. low)		4 53 (2 18-9 44)	<0.001	TMErisk (high vs. low)	4.07 (2.13-7.79)
	Age, vears (≥62 vs. <62)		0.88 (0.51-1.53)	0.658	Age, years (≥62 vs. <62)	1.35 (0.80-2.30)
	Gender (male vs. female)		1.06 (0.59-1.93)	0.840	Gender (male vs. female)	1.28 (0.76-2.17)
	TMB_mute/Mb			0.040	TMB, muts/Mb	
	<10 vs >10		1 22 (0 57-2 63)	0.611	<10 vs. ≥10	1.55 (0.78-3.07)
	NA vs. ≥10		1.86 (1.00-3.44)	0.050	NA vs. ≥10	1.70 (0.95-3.05)
	PD-I 1 TPS %			0.000	PD-L1 TPS, %	
	<1 vs >50		2 91 (0 90-9 48)	0.075	<1 vs. ≥50	0.77 (0.30-1.96)
	1-49 vs >50		2,75 (0,75-10,07)	0.127	1-49 vs. ≥50	1.03 (0.48-2.22)
	PES		2.10 (0.10 10.01)	0.121	PFS	. ,
	TMErisk (high vs. low)		2,73 (1,60-4,67)	<0.001	TMErisk (high vs. low)	1.89 (1.02-3.50)
	Ane vers (>62 vs $<$ 62)		0.65 (0.41-1.04)	0.070	Age, years (≥62 vs. <62)	0.79 (0.48-1.29)
	Gender (male vs. female)		0.82 (0.50-1.36)	0.822	Gender (male vs. female)	1.01 (0.62-1.65)
	TMB muts/Mb		0.02 (0.00-1.00)	0.022	TMB. muts/Mb	
	<10 vs >10		1 08 (0 57-2 05)	0.800	<10 vs ≥10	1.07 (0.56-2.01)
	NA vo. 210		1.83 (1.09-3.08)	0.009		1.69 (0.97-2.93)
			1.03 (1.03-3.00)	0.022	PD-I 1 TPS %	,
	PD-L1 TPS, %		2 10 (0 00 4 84)	0.054	<1 vs >50	0.61 (0.30-1.26)
	1-49 vg >50		2.19 (0.99-4.04)	0.054	1-49 vs ≥50	0.86 (0.41-1.80)
			2.04 (0.05-4.90)	0.110		
_	0.2	11.522.533.5 4.55 6			0.2 11.522.533.5 4.55	6
С	08				D	
	03				PFS PFS	
I	Baseline characteristics	IMvigor210 Atezolizumab	HR (95% CI)	<i>P</i> -value	 PFS Baseline characteristics IMvigor210 Atezolizuma 	ıb HR (95% CI)
ا ۲	Baseline characteristics FMErisk (high vs. low)	IMvigor210 Atezolizumab	HR (95% CI) 3.41 (2.17-5.36)	<i>P</i> -value <0.001	PFS Baseline characteristics IMvigor210 Atezolizuma TMErisk (high vs. low)	HR (95% CI) 2.46 (1.73-3.50)
T (Baseline characteristics IMErisk (high vs. low) Gender (male vs. female)	IMvigor210 Atezolizumab	HR (95% CI) 3.41 (2.17-5.36) 0.87 (0.58-1.30)	P-value <0.001 0.497	PFS Baseline characteristics IMvigor210 Atezolizuma TMErisk (high vs. low) Gender (male vs. female)	HR (95% CI) 2.46 (1.73-3.50) 0.81 (0.57-1.15)
 	Baseline characteristics IMErisk (high vs. low) Gender (male vs. female)	IMvigor210 Atezolizumab	HR (95% CI) 3.41 (2.17-5.36) 0.87 (0.58-1.30)	<i>P</i> -value <0.001 0.497	 PFS Baseline characteristics TMErisk (high vs. low) Gender (male vs. female) TMB, muts/Mb 	HR (95% Cl) 2.46 (1.73-3.50) 0.81 (0.57-1.15)
 -	Baseline characteristics IMErisk (high vs. low) Gender (male vs. female) IMB, muts/Mb <10 vs. ≥10	IMvigor210 Atezolizumab	HR (95% CI) 3.41 (2.17-5.36) 0.87 (0.58-1.30) 1.48 (0.72-3.06)	<i>P</i> -value <0.001 0.497 0.258	 PFS Baseline characteristics TMErisk (high vs. low) Gender (male vs. female) TMB, muts/Mb <10 vs. ≥10 IMvigor210 Atezolizuma 	 HR (95% CI) 2.46 (1.73-3.50) 0.81 (0.57-1.15) 1.41 (0.72-2.79)
	Baseline characteristics IMErisk (high vs. low) Gender (male vs. female) IMB, muts/Mb <10 vs. ≥10 NA vs. ≥10	IMvigor210 Atezolizumab	HR (95% Cl) 3.41 (2.17-5.36) 0.87 (0.58-1.30) 1.48 (0.72-3.06) 1.08 (0.75-1.56)	<i>P</i> -value <0.001 0.497 0.258 0.678	 PFS Baseline characteristics TMErisk (high vs. low) Gender (male vs. female) TMB, muts/Mb <10 vs. ≥10 NA vs. ≥10 	 HR (95% CI) 2.46 (1.73-3.50) 0.81 (0.57-1.15) 1.41 (0.72-2.79) 1.07 (0.78-1.48)
	Baseline characteristics IMErisk (high vs. low) Gender (male vs. female) IMB, muts/Mb <10 vs. ≥10 NA vs. ≥10 PD-L1 TPS, %	IMvigor210 Atezolizumab	HR (95% Cl) 3.41 (2.17-5.36) 0.87 (0.58-1.30) 1.48 (0.72-3.06) 1.08 (0.75-1.56)	P-value <0.001 0.497 0.258 0.678	 PFS Baseline characteristics TMErisk (high vs. low) Gender (male vs. female) TMB, muts/Mb <10 vs. ≥10 NA vs. ≥10 PD-L1 TPS, % 	 HR (95% Cl) 2.46 (1.73-3.50) 0.81 (0.57-1.15) 1.41 (0.72-2.79) 1.07 (0.78-1.48)
	Baseline characteristics IMErisk (high vs. low) Gender (male vs. female) TMB, muts/Mb <10 vs. ≥10 NA vs. ≥10 PD-L1 TPS, % <1 vs. ≥50	IMvigor210 Atezolizumab	HR (95% Cl) 3.41 (2.17-5.36) 0.87 (0.58-1.30) 1.48 (0.72-3.06) 1.08 (0.75-1.56) 1.58 (0.65-3.87)	P-value <0.001 0.497 0.258 0.678 0.317	 PFS Baseline characteristics IMvigor210 Atezolizuma TMErisk (high vs. low) Gender (male vs. female) TMB, muts/Mb <10 vs. ≥10 NA vs. ≥10 PD-L1 TPS, % <1 vs. ≥50 	 HR (95% Cl) 2.46 (1.73-3.50) 0.81 (0.57-1.15) 1.41 (0.72-2.79) 1.07 (0.78-1.48) 1.46 (0.68-3.12)
 	Baseline characteristics IMErisk (high vs. low) Gender (male vs. female) TMB, muts/Mb <10 vs. ≥10 NA vs. ≥10 PD-L1 TPS, % <1 vs. ≥50 1-49 vs. ≥50	IMvigor210 Atezolizumab	HR (95% Cl) 3.41 (2.17-5.36) 0.87 (0.58-1.30) 1.48 (0.72-3.06) 1.08 (0.75-1.56) 1.58 (0.65-3.87) 1.50 (0.57-3.93)	<i>P</i> -value <0.001 0.497 0.258 0.678 0.317 0.415	PFS Baseline characteristics TMErisk (high vs. low) Gender (male vs. female) TMB, muts/Mb <10 vs. ≥10 PD-L1 TPS, % <1 vs. ≥50 1-49 vs. ≥50 IMvigor210 Atezolizuma	 HR (95% Cl) 2.46 (1.73-3.50) 0.81 (0.57-1.15) 1.41 (0.72-2.79) 1.07 (0.78-1.48) 1.46 (0.68-3.12) 1.44 (0.63-3.29)
	Baseline characteristics IMErisk (high vs. low) Gender (male vs. female) ↓ IMB, muts/Mb <10 vs. ≥10 NA vs. ≥10 PD-L1 TPS, % <1 vs. ≥50 1-49 vs. ≥50 0.2	IMvigor210 Atezolizumab	HR (95% Cl) 3.41 (2.17-5.36) 0.87 (0.58-1.30) 1.48 (0.72-3.06) 1.08 (0.75-1.56) 1.58 (0.65-3.87) 1.50 (0.57-3.93)	<i>P</i> -value <0.001 0.497 0.258 0.678 0.317 0.415	PFS Baseline characteristics IMvigor210 Atezolizuma TMErisk (high vs. low) Imvigor210 Atezolizuma Gender (male vs. female) Imvigor210 Atezolizuma TMB, muts/Mb Imvigor210 Atezolizuma <10 vs. ≥10	hR (95% Cl) 2.46 (1.73-3.50) 0.81 (0.57-1.15) 1.41 (0.72-2.79) 1.07 (0.78-1.48) 1.46 (0.68-3.12) 1.44 (0.63-3.29) 16
E	Baseline characteristics IMErisk (high vs. low) Gender (male vs. female) TMB, muts/Mb <10 vs. ≥10 NA vs. ≥10 PD-L1 TPS, % <1 vs. ≥50 1-49 vs. ≥50 0.2	IMvigor210 Atezolizumab	HR (95% Cl) 3.41 (2.17-5.36) 0.87 (0.58-1.30) 1.48 (0.72-3.06) 1.08 (0.75-1.56) 1.58 (0.65-3.87) 1.50 (0.57-3.93)	<i>P</i> -value <0.001 0.497 0.258 0.678 0.317 0.415	 PFS Baseline characteristics IMvigor210 Atezolizuma TMErisk (high vs. low) Gender (male vs. female) TMB, muts/Mb <10 vs. ≥10 PD-L1 TPS, % <1 vs. ≥50 1-49 vs. ≥50 E 	hk HR (95% Cl) 2.46 (1.73-3.50) 0.81 (0.57-1.15) 1.41 (0.72-2.79) 1.07 (0.78-1.48) 1.46 (0.68-3.12) 1.44 (0.63-3.29)
E Bas	Baseline characteristics IMErisk (high vs. low) Gender (male vs. female) IMB, muts/Mb <10 vs. ≥10 NA vs. ≥10 PD-L1 TPS, % <1 vs. ≥50 1-49 vs. ≥50 C2 Seline characteristics	IMvigor210 Atezolizumab	HR (95% Cl) 3.41 (2.17-5.36) 0.87 (0.58-1.30) 1.48 (0.72-3.06) 1.08 (0.75-1.56) 1.58 (0.65-3.87) 1.50 (0.57-3.93) HR (95% Cl)	P-value <0.001 0.497 0.258 0.678 0.317 0.415 P-value	 PFS Baseline characteristics IMvigor210 Atezolizuma TMErisk (high vs. low) Gender (male vs. female) TMB, muts/Mb <10 vs. ≥10 PD-L1 TPS, % <1 vs. ≥50 1-49 vs. ≥50 F Baseline characteristics ORIENT-11 Chemo 	 HR (95% Cl) 2.46 (1.73-3.50) 0.81 (0.57-1.15) 1.41 (0.72-2.79) 1.07 (0.78-1.48) 1.46 (0.68-3.12) 1.44 (0.63-3.29) HR (95% Cl)
E Bas	Baseline characteristics FMErisk (high vs. low) Gender (male vs. female) TMB, muts/Mb <10 vs. ≥10 NA vs. ≥10 PD-L1 TPS, % <1 vs. ≥50 1-49 vs. ≥50 C2 Seline characteristics O S	IMvigor210 Atezolizumab	HR (95% Cl) 3.41 (2.17-5.36) 0.87 (0.58-1.30) 1.48 (0.72-3.06) 1.08 (0.75-1.56) 1.58 (0.65-3.87) 1.50 (0.57-3.93) HR (95% Cl)	<i>P</i> -value <0.001 0.497 0.258 0.678 0.317 0.415 <i>P</i> -value	 PFS Baseline characteristics IMvigor210 Atezolizuma TMErisk (high vs. low) Gender (male vs. female) TMB, muts/Mb <10 vs. ≥10 PD-L1 TPS, % <1 vs. ≥50 1-49 vs. ≥50 F Baseline characteristics ORIENT-11 Chemo OS 	 HR (95% Cl) 2.46 (1.73-3.50) 0.81 (0.57-1.15) 1.41 (0.72-2.79) 1.07 (0.78-1.48) 1.46 (0.68-3.12) 1.44 (0.63-3.29) HR (95% Cl)
E Bas TM	Baseline characteristics FMErisk (high vs. low) Gender (male vs. female) , TMB, muts/Mb <10 vs. ≥10 PD-L1 TPS, % <1 vs. ≥50 1-49 vs. ≥50 C2 Beline characteristics O S Erisk (high vs. low) 0.2 DD (1 _ 2)	IMvigor210 Atezolizumab	HR (95% CI) 3.41 (2.17-5.36) 0.87 (0.58-1.30) 1.48 (0.72-3.06) 1.08 (0.75-1.56) 1.58 (0.65-3.87) 1.50 (0.57-3.93) HR (95% CI) 3.81 (2.33-6.24)	P-value <0.001	 PFS Baseline characteristics IMvigor210 Atezolizuma TMErisk (high vs. low) Gender (male vs. female) TMB, muts/Mb <10 vs. ≥10 PD-L1 TPS, % <1 vs. ≥50 1-49 vs. ≥50 F Baseline characteristics ORIENT-11 Chemo OS TMErisk (high vs. low) 	 HR (95% Cl) 2.46 (1.73-3.50) 0.81 (0.57-1.15) 1.41 (0.72-2.79) 1.07 (0.78-1.48) 1.46 (0.68-3.12) 1.44 (0.63-3.29) HR (95% Cl) 4.04 (1.84-8.87)
E Bas O: TM EC	Baseline characteristics FMErisk (high vs. low) Gender (male vs. female) TMB, muts/Mb <10 vs. ≥10 PD-L1 TPS, % <1 vs. ≥50 1-49 vs. ≥50 C Beline characteristics C S Erisk (high vs. low) OG PS (1 vs. 0)	IMvigor210 Atezolizumab	HR (95% CI) 3.41 (2.17-5.36) 0.87 (0.58-1.30) 1.48 (0.72-3.06) 1.08 (0.75-1.56) 1.58 (0.65-3.87) 1.50 (0.57-3.93) HR (95% CI) 3.81 (2.33-6.24) 2.55 (1.41-4.61)	P-value <0.001	 PFS Baseline characteristics IMvigor210 Atezolizuma TMErisk (high vs. low) Gender (male vs. female) TMB, muts/Mb <10 vs. ≥10 PD-L1 TPS, % <1 vs. ≥50 1-49 vs. ≥50 F Baseline characteristics ORIENT-11 Chemo OS TMErisk (high vs. low) PFS 	 HR (95% Cl) 2.46 (1.73-3.50) 0.81 (0.57-1.15) 1.41 (0.72-2.79) 1.07 (0.78-1.48) 1.46 (0.68-3.12) 1.44 (0.63-3.29) HR (95% Cl) 4.04 (1.84-8.87)
E Bas O: TM EC PI	Baseline characteristics FMErisk (high vs. low) Gender (male vs. female) TMB, muts/Mb <10 vs. ≥10 PD-L1 TPS, % <1 vs. ≥50 1-49 vs. ≥50 C Geline characteristics C S Erisk (high vs. low) OG PS (1 vs. 0) S S	IMvigor210 Atezolizumab	HR (95% CI) 3.41 (2.17-5.36) 0.87 (0.58-1.30) 1.48 (0.72-3.06) 1.08 (0.75-1.56) 1.58 (0.65-3.87) 1.50 (0.57-3.93) HR (95% CI) 3.81 (2.33-6.24) 2.55 (1.41-4.61)	P-value <0.001	 PFS Baseline characteristics IMvigor210 Atezolizuma TMErisk (high vs. low) Gender (male vs. female) TMB, muts/Mb <10 vs. ≥10 PD-L1 TPS, % <1 vs. ≥50 1-49 vs. ≥50 F Baseline characteristics ORIENT-11 Chemo OS TMErisk (high vs. low) PFS ECOG PS (1 vs. 0) 	 HR (95% Cl) 2.46 (1.73-3.50) 0.81 (0.57-1.15) 1.41 (0.72-2.79) 1.07 (0.78-1.48) 1.46 (0.68-3.12) 1.46 (0.63-3.29) HR (95% Cl) 4.04 (1.84-8.87) 2.00 (1.06-3.75)
E Bass O: TM EC PI TM	Baseline characteristics [MErisk (high vs. low) Gender (male vs. female) TMB, muts/Mb <10 vs. ≥10 NA vs. ≥10 PD-L1 TPS, % <1 vs. ≥50 1-49 vs. ≥50 Erisk (high vs. low) OG PS (1 vs. 0) =S Erisk (high vs. low) in motoctoric	IMvigor210 Atezolizumab	HR (95% CI) 3.41 (2.17-5.36) 0.87 (0.58-1.30) 1.48 (0.72-3.06) 1.08 (0.75-1.56) 1.58 (0.65-3.87) 1.50 (0.57-3.93) HR (95% CI) 3.81 (2.33-6.24) 2.55 (1.41-4.61) 3.01 (1.83-4.94)	P-value <0.001	 PFS Baseline characteristics IMvigor210 Atezolizuma TMErisk (high vs. low) Gender (male vs. female) TMB, muts/Mb <10 vs. ≥10 PD-L1 TPS, % <1 vs. ≥50 1-49 vs. ≥50 F Baseline characteristics ORIENT-11 Chemo OS TMErisk (high vs. low) PFS ECOG PS (1 vs. 0) Brain metastasis 	 HR (95% Cl) 2.46 (1.73-3.50) 0.81 (0.57-1.15) 1.41 (0.72-2.79) 1.07 (0.78-1.48) 1.46 (0.68-3.12) 1.46 (0.68-3.12) 1.44 (0.63-3.29) HR (95% Cl) 4.04 (1.84-8.87) 2.00 (1.06-3.75) 2.51 (1.21-5.20)
E Bass O: TM EC PI TM Bra (yet	Baseline characteristics [MErisk (high vs. low) Gender (male vs. female) TMB, muts/Mb <10 vs. ≥10 NA vs. ≥10 PD-L1 TPS, % <1 vs. ≥50 1-49 vs. ≥50 S Erisk (high vs. low) OG PS (1 vs. 0) =S Erisk (high vs. low) ain metastasis s vs. no)	IMvigor210 Atezolizumab	HR (95% CI) 3.41 (2.17-5.36) 0.87 (0.58-1.30) 1.48 (0.72-3.06) 1.08 (0.75-1.56) 1.58 (0.65-3.87) 1.50 (0.57-3.93) HR (95% CI) 3.81 (2.33-6.24) 2.55 (1.41-4.61) 3.01 (1.83-4.94) 2.12 (1.03-4.34)	 P-value <0.001 0.497 0.258 0.678 0.317 0.415 P-value <0.001 0.002 <0.001 0.041 	 PFS Baseline characteristics IMvigor210 Atezolizuma TMErisk (high vs. low) Gender (male vs. female) TMB, muts/Mb <10 vs. ≥10 PD-L1 TPS, % <1 vs. ≥50 1-49 vs. ≥50 F Baseline characteristics ORIENT-11 Chemo OS TMErisk (high vs. low) PFS ECOG PS (1 vs. 0) Brain metastasis (yes vs. no) D2 115225335 455 	 HR (95% Cl) 2.46 (1.73-3.50) 0.81 (0.57-1.15) 1.41 (0.72-2.79) 1.07 (0.78-1.48) 1.46 (0.68-3.12) 1.44 (0.63-3.29) HR (95% Cl) 4.04 (1.84-8.87) 2.00 (1.06-3.75) 2.51 (1.21-5.20)

P-value

< 0.001 0.264 0.360

0.210 0.074

0.581 0.946

0.042 0.787 0.976

0.844 0.063

0.181 0.865

> P-value <0.001 0.242

0.320 0.683

0.329 0.382

P-value

<0.001

0.031 0.013

Figure S12. Forest plot for hazard ratios of OS and PFS in the POPLAR, IMvigor210 and ORIENT-11 cohorts according to TMErisk and other baseline clinicopathological characteristics, related to Figure 6 and Figure 7. (A-B) Univariate Cox analysis for OS and PFS according to different variables in the POPLAR immunotherapy (A) and chemotherapy (B) datasets. (C-D) Univariate Cox analysis for OS (C) and PFS (D) according to different variables in the IMvigor210 cohort. (E-F) Multivariate Cox analysis for OS and PFS in the immunochemotherapy (E) and chemotherapy (F) groups from ORIENT-11 cohort.

HR, hazard ratio; CI, confidence interval; OS, overall survival; PFS, progression-free survival; TMB, tumor mutation burden; PD-L1, programmed cell death ligand 1; TPS, tumor proportion score; Chemo, chemotherapy; IO+Chemo, immunotherapy plus chemotherapy; ECOG, Eastern Cooperative Oncology Group; PS, performance status

Patient characteristics	TCGA-LUSC cohort (N = 494)
Age, years	
Median (range)	68 (39-90)
< 68, n (%)	229 (46.4)
≥ 68, n (%)	260 (52.6)
NA, n (%)	5 (1.0)
Gender, n (%)	
Male	366 (74.1)
Female	128 (25.9)
ECOG PS, n (%)	
< 2	88 (17.8)
≥2	21 (4.3)
NA	385 (77.9)
Smoking status, n (%)	
Never smoker	18 (3.6)
Current or former smoker	464 (93.9)
NA	12 (2.4)
T stage, n (%)	
Τ1	114 (23.1)
T2	287 (58.1)
ТЗ	70 (14.2)
Τ4	23 (4.6)
N stage, n (%)	
NO	316 (64.0)
N1	127 (25.7)
N2	40 (8.1)
N3	5 (1.0)
Nx	6 (1.2)
M stage, n (%)	
MO	405 (82.0)
M1	8 (1.6)
Mx	77 (15.6)
NA	4 (0.8)

Table S1. Baseline p	patient characteristics in the	TCGA-LUSC cohort,	related to STAR Methods
----------------------	--------------------------------	--------------------------	-------------------------

AJCC pTNM stage, n (%)	
Stage I	242 (49.0)
Stage II	158 (32.0)
Stage III	83 (16.8)
Stage IV	7 (1.4)
NA	4 (0.8)
TMErisk score	
Median (range)	-0.45 (-1.33-1.15)
Low risk, n (%)	421 (85.2)
High risk, n (%)	73 (14.8)

Abbreviations: ECOG, Eastern Cooperative Oncology Group; PS, performance status; AJCC, American Joint Committee on Cancer; NA, not available.

Dreatom	Enriched in low-TMErisk	Enriched in high-TMErisk group
Program	group	
CIBERSORT	Plasma cells, Monocytes,	B cells memory, T cells regulatory (Tregs),
	Dendritic cells resting, Mast	Macrophages M2, Eosinophils, Neutrophils
	cells activated	
EPIC	T cell CD8+	B cell, Cancer associated fibroblast, T cell CD4+,
		Endothelial cell
xCell	T cell CD4 ⁺ central memory,	Myeloid dendritic cell activated, T cell CD4+ effector
	T cell CD8+ naive,	memory, Class-switched memory B cell, Myeloid dendritic
	Common lymphoid progenitor,	cell, Endothelial cell, Eosinophil, Cancer associated
	T cell CD4+ Th1, T cell CD4+	fibroblast,
	Th2	Granulocyte-monocyte progenitor, Hematopoietic stem cell,
		Macrophage, Macrophages M1, Macrophages M2, Mast
		cell, Monocyte, Neutrophil, T cell NK, Tregs

Table S9. Infiltration levels of immune and stromal cells in the low- and high-TMErisk groups,related to Figure 4

	OS		PFS	
Patient characteristics -	HR (95% CI)	P value	HR (95% CI)	P value
Age, years				
> 60 vs. ≤ 60	1.45 (0.90-2.38)	0.121	1.53 (0.92-2.53)	0.101
Gender				
Male vs. Female	1.48 (0.81-2.67)	0.200	1.66 (0.89-3.09)	0.110
ECOG PS				
1 vs. 0	2.16 (1.20-3.90)	0.010	1.51 (0.87-2.63)	0.146
Smoking status				
Current/former smoker vs.	1 07 (0 65 1 75)	0 701	1 10 (0 66 1 84)	0 720
Never smoker	1.07 (0.05-1.75)	0.791	1.10 (0.00-1.04)	0.720
PD-L1 TPS, %, n (%)				
1-49 vs. < 1	0.79 (0.42-1.47)	0.448	1.06 (0.57-1.98)	0.843
≥ 50 vs. < 1	0.64 (0.36-1.13)	0.122	0.73 (0.40-1.33)	0.301
Brain metastases, n (%)				
Yes vs. No	1.67 (0.85-3.28)	0.137	2.14 (1.04-4.35)	0.036
TMErisk score				
High vs. Low	3.42 (2.10-5.56)	< 0.001	3.01 (1.83-4.94)	< 0.001

Table S10. Univariate Cox analyses for OS and PFS in the immunochemotherapy group from ORIENT-11, related to Figure 7

Abbreviations: ECOG, Eastern Cooperative Oncology Group; PS, performance status; EGFR, epidermal growth factor receptor; ALK, anaplastic lymphoma kinase; PD-L1, programmed cell death ligand 1; TPS, tumor proportion score; HR, hazard ratio; CI, confidence interval; OS, overall survival; PFS, progression-free survival.

	OS		PFS	
Patient characteristics -	HR (95% CI)	P value	HR (95% CI)	P value
Age, years				
> 60 vs. ≤ 60	0.75 (0.41-1.39)	0.365	1.06 (0.60-1.87)	0.838
Gender				
Male vs. Female	1.72 (0.79-3.73)	0.172	1.34 (0.69-2.62)	0.389
ECOG PS				
1 vs. 0	1.03 (0.54-2.00)	0.921	1.84 (0.99-3.43)	0.054
Smoking status				
Current/former smoker vs.	1 25 (0 62 2 45)	0.522	0 00 (0 54 1 81)	0.069
Never smoker	1.23 (0.03-2.43)	0.525	0.99 (0.54-1.61)	0.900
PD-L1 TPS, %, n (%)				
1-49 vs. < 1	0.77 (0.30-1.96)	0.581	1.22 (0.56-2.69)	0.617
≥ 50 vs. < 1	1.45 (0.72-2.95)	0.295	1.06 (0.57-1.98)	0.860
Brain metastases, n (%)				
Yes vs. No	0.82 (0.38-1.78)	0.613	2.24 (1.10-4.56)	0.026
TMErisk score				
High vs. Low	4.04 (1.84-8.87)	< 0.001	1.27 (0.71-2.26)	0.423

Table S11. Univariate Cox analyses for OS and PFS in the chemotherapy group from ORIENT-11,related to Figure 7

Abbreviations: ECOG, Eastern Cooperative Oncology Group; PS, performance status; EGFR, epidermal growth factor receptor; ALK, anaplastic lymphoma kinase; PD-L1, programmed cell death ligand 1; TPS, tumor proportion score; HR, hazard ratio; CI, confidence interval; OS, overall survival; PFS, progression-free survival.

Patient characteristics	POPL	IMvigor210	
_	Atezolizumab (N = 81) Docetaxel (N = 75)		(N = 208)
Age, years			
Median (range)	61 (42-82)	63 (36-80)	NA
< 62, n (%)	41 (50.6)	32 (42.7)	NA
≥ 62, n (%)	40 (49.4)	43 (57.3)	NA
Gender, n (%)			
Male	56 (69.1)	44 (58.7)	162 (77.9)
Female	25 (30.9)	31 (41.3)	46 (22.1)
TMB,muts/Mb			
< 10	15 (18.5)	16 (21.3)	10 (4.8)
≥ 10	35 (43.2)	34 (45.3)	134 (64.4)
Not available	31 (38.3)	25 (33.3)	64 (30.8)
PD-L1 TPS, %, n (%)			
< 1	55 (69.7)	37 (49.3)	164 (78.8)
1-49	17 (21.0)	27 (36.0)	35 (16.8)
≥ 50	9 (11.1)	11 (14.7)	9 (4.3)
TMErisk score			
Median (range)	0.25 (-1.53 – 1.27)	0.67 (-0.11 – 2.20)	0.46 (-1.13 – 1.42)

Table S12. Baseline patient characteristics in the POPLAR and IMvigor210 cohorts, related to STAR Methods

Abbreviations: TMB, tumor mutation burden; PD-L1, programmed cell death ligand 1; TPS, tumor proportion score.

Patient characteristics	IO+Chemo cohort (N = 113)	Chemo cohort (N = 58)		
Age, years				
Median (range)	61 (30-74)	60 (37-74)		
≤ 60, n (%)	52 (46.0)	30 (51.7)		
> 60, n (%)	61 (54.0)	28 (48.3)		
Gender, n (%)				
Male	87 (77.0)	45 (77.6)		
Female	26 (23.0)	13 (22.4)		
ECOG PS				
0	35 (31.0)	17 (29.3)		
1	78 (69.0)	41 (70.7)		
Smoking status				
Never smoker	42 (37.2)	16 (27.6)		
Current or former smoker	71 (62.8)	42 (72.4)		
Histology				
Adenocarcinoma	111 (98.2)	55 (94.8)		
Other types	2 (1.8)	3 (5.2)		
EGFR mutation status				
Positive	2 (1.8)	1 (1.7)		
Negative	111 (98.2)	57 (98.3)		
ALK translocation				
Positive	0 (0)	0 (0)		
Negative	113 (100)	58 (100)		
TNM stage,				
IIIB	7 (6.2)	4 (6.9)		
IIIC	2 (1.8)	5 (8.6)		
IV	104 (92.0)	49 (84.5)		
PD-L1 TPS, %, n (%)				
< 1	31 (27.4)	19 (32.8)		
1-49	33 (29.2)	13 (22.4)		
≥ 50	49 (43.4)	26 (44.8)		
Brain metastases, n (%)	12 (10.6)	11 (19.0)		
TMErisk score				
Median (range)	0.00 (-1.68 – 1.93)	-2.00 (-3.56 – 1.16)		

Table S13. Baseline patient characteristics in the ORIENT-11 cohort, related to STAR Methods

Abbreviations: ECOG, Eastern Cooperative Oncology Group; PS, performance status; EGFR, epidermal growth factor receptor; ALK, anaplastic lymphoma kinase; PD-L1, programmed cell death ligand 1; TPS, tumor proportion score; IO+Chemo, immunotherapy plus chemotherapy.

Table S14. Cut-points use	d in each dataset, relate	d to STAR Methods
---------------------------	---------------------------	-------------------

Datasets	C11orf96	C4BPA	PLAAT4	TGM2	KLF5	PNCK	TMErisk
TCGA-LUSC	5.952	23.190	51.288	72.246	73.092	5.914	0.27
TCGA-LUAD	19.323	16.440	15.327	93.926	63.064	0.060	-0.02
GSE81089	7.261	56.762	97.238	142.426	47.928	10.799	-0.38
GSE30219	9.348	7.942	9.022	8.105	8.346	5.949	0.00
GSE37745	9.192	10.207	8.509	8.302	10.701	6.521	-1.47
GSE157011	9.641	9.768	9.208	8.168	9.064	6.689	-0.68
GSE31210	6518.450	9749.500	3058.135	2259.950	1328.665	19.605	-1.35
GSE135222	10.570	0.390	119.120	38.630	46.380	5.750	0.67
GSE78220	24.298	0.021	NA	12.825	1.108	0.468	-2.15
OAK_LUSC_immunotherapy	3.000	0.569	4.682	5.881	8.701	2.369	-0.80
OAK_LUAD_immunotherapy	6.149	6.837	5.559	8.160	7.700	0.119	-0.34
POPLAR_immunotherapy	5.450	6.560	6.600	6.580	7.130	0.400	-0.10
IMvigor210_immunotherapy	9.425	1.535	97.605	40.615	682.025	3.905	0.06
OAK_LUSC_chemotherapy	2.480	0.380	5.383	5.550	6.460	1.590	-2.07
OAK_LUAD_chemotherapy	3.465	3.600	6.114	8.187	7.600	1.650	-0.61
POPLAR_chemotherapy	4.440	4.825	4.150	9.120	7.350	2.230	1.25
ORIENT-11_combo	105.955	89.240	55.889	409.449	161.353	2.110	0.03
ORIENT-11_chemotherapy	15.558	2.250	27.786	109.750	43.179	1.844	-2.04