Supplementary Online Content

Li B, Cui Y, Diehn M, Li R. Development and validation of an individualized immune prognostic signature in early-stage nonsquamous non–small cell lung cancer. *JAMA Oncol*. Published online July 6, 2017. doi:10.1001/jamaoncol.2017.1609

This supplementary material has been provided by the authors to give readers additional information about their work.

eMethods. Data Preprocessing, Assessment of Robustness, and Comparison With Commercialized Lung Biomarker

Data preprocessing

All microarray data and corresponding clinical information were collected from Gene Expression Omnibus (https://www.ncbi.nlm.nih.gov/geo/). All affymetrix microarrays were normalized with MAS5.0 method (R package *affy*, v1.50.0). For Illumina datasets, Model-Based Background Correction (MBCB) processed data were downloaded. To improve reproducibility of data measured by different platforms, raw median intensity values of tumor samples in each Agilent datasets were extracted by R package *agilp* (v3.4.0).[1] Normalized TCGA RNA-Seq data and clinical information were downloaded from Cancer Genomics Browser (https://genome-cancer.ucsc.edu/).[2] Entrez IDs were used to represent genes across different platform. Probe sets correspond to multiple Entrez IDs were removed. If multiple probe sets correspond to the same Entrez ID, the one with the highest mean signal was selected as the expression level of the corresponding gene.[3] Gene symbols in TCGA dataset were also converted to Entrez IDs.

Assessing robustness of the 25 IRGPs and variability of prediction for different biomarkers

To evaluate the robustness of the selected 25 IRGPs, feature selection and model building were repeated 1000 times in different randomization of the meta-cohort. The same criteria described in Methods were used to find prognostic IRGPs and build IRGP-based models in each meta-training cohort. IRGPs that appeared in at least one prognostic model were considered. The frequencies of selected IRGPs were used to estimate a background distribution. The frequencies of the 25 IRGPs consisting of the IRGPI were examined and compared with the background distribution with Wilcoxon rank-sum test.

The C-index of IRGP-based prognostic biomarker, Cell Cycle Progression (CCP) score and the 14-gene biomarker was also calculated in 1000 randomization of the meta-testing cohort to estimate performance variations of each biomarker. Implementations of CCP score and 14-gene biomarker were described below. The standard deviation of C-index was used to measure the variability of survival prediction for different biomarkers.

Comparison with commercialized lung biomarker:

The 14-gene biomarker consisting of 11 prognostic genes and 3 reference genes and the CCP biomarker consisting of 31 CCP genes were implemented.[4,5] Briefly, for 14-gene biomarker, gene expression values of 11 prognostic genes were normalized by those of 3 reference genes and risk scores for each patient were calculated based on coefficients proposed by the original paper.[4] The construction of CCP biomarker was based on the pipeline designed for microarray data.[5] In short, Affymetrix datasets were normalized with Robust Multi-array Average (RMA)[6] method. For Illumina and Agilent datasets, GEO normalized datasets were used. The average expression value of 31 genes were used as the CCP score.[5]

The performance of commercialized mPS[7], which is an integration of CCP score with pathological staging was compared with our immune-clinical composite score ICPI in terms of C-index in the validation datasets.

eResults. Analysis and Robustness of IRGPs

Analysis of IRGPs with constant orderings

To further illustrate the possible causes of constant ordering, we categorized the total 126615 IRGPs with constant orderings into three groups based on measurement platforms and datasets. First, ~23% (29200) of IRGPs were consistently constant (all 0 or 1) in datasets only measured by a single platform (Affymetrix, Agilent or Illumina) and ~9% (11712) of IRGPs were consistently constant within one platform, but constant in the opposite direction in other platforms. These types of IRGPs, in our opinion, were caused by platform bias. Second, 58% (73805) IRGPs were consistently constant in datasets measured by at least two different platforms, which might be caused by biological preferential transcription. Third, the remaining 10% (11898) IRGPs were inconsistently constant (1 for some datasets and 0 for others) between different datasets of the same microarray platform. The reasons for this may be various, and one possibility could be batch effects.

Robustness of 25 IRGPs in IRGPI

There were 1588 IRGPs that appeared in at least 1 model. Among the top 10 IRGPs with the highest frequencies, 5 of them were included in our 25 IRGPs. The smallest frequency of our 25 IRGPs ranked among top 25 percentile of that for all 1588 IRGPs. Wilcoxon rank-sum test showed significantly higher frequencies for our 25 IRGPs compared with the background distribution of frequencies for all 1588 IRGPs ($P<1.7\times10^{-14}$). Hence, our IRGPI biomarker was relatively robust to different randomization.

eReferences

1. Stafford P, Brun M. Three methods for optimization of cross-laboratory and cross-platform microarray expression data. *Nucleic Acids Res.* 2007;35(10):e72.
2. Sanborn JZ, Benz SC, Craft B, et al. The UCSC Cancer Genomics Browser: update 2011. *Nucleic Acids Res.* 2011;39(Database issue):D951-959.
3. Miller JA, Cai C, Langfelder P, et al. Strategies for aggregating gene expression data: the collapseRows R function. *BMC Bioinformatics.* 2011;12:322.
4. Kratz JR, He J, Van Den Eeden SK, et al. A practical molecular assay to predict survival in resected non-squamous, non-small-cell lung cancer: development and international validation studies. *The Lancet.* 2012;379(9818):823-832.
5. Wistuba, II, Behrens C, Lombardi F, et al. Validation of a proliferation-based expression signature as prognostic marker in early stage lung adenocarcinoma. *Clin Cancer Res.* 2013;19(22):6261-6271.
6. Irizarry RA, Bolstad BM, Collin F, Cope LM, Hobbs B, Speed TP. Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Res.* 2003;31(4):e15.
7. Bueno R, Hughes E, Wagner S, et al. Validation of a molecular and pathological model for five-year mortality risk in patients with early stage lung adenocarcinoma. *Journal of thoracic oncology : official publication of the International Association for the Study of Lung Cancer.* 2015;10(1):67-73.

eTable 1. Details About the Data Sets Used in This Study

	Name/Accession No.	Platform	No. of Nonsquamous NSCLC
Independent validation datasets	TCGA lung adenocarcinoma (TCGA)	Illumina HiSeq	439
	Director's Challenge Consortium (DCC)	Affymetrix Human Genome U133A Array	323
	GSE30219	Affymetrix Human Genome U133 Plus 2.0 Array	207
Meta (training/testing) dataset	GSE31210	Affymetrix Human Genome U133 Plus 2.0 Array	204
	GSE41271	Illumina HumanWG-6 v3.0 expression beadchip	192
	GSE50081	Affymetrix Human Genome U133 Plus 2.0 Array	138
	GSE13213	Agilent-014850 Whole Human Genome Microarray 4x44K G4112F	117
	GSE26939	Agilent-UNC-custom-4X44K	115
	GSE83227	Affymetrix Human Genome U95 Version 2 Array	112
	GSE11969	Agilent Homo sapiens 21.6K custom array	105
	GSE42127	Illumina HumanWG-6 v3.0 expression beadchip	94
	GSE68571	Affymetrix Human Full Length HuGeneFL Array	86
	GSE19188	Affymetrix Human Genome U133 Plus 2.0 Array	58
	GSE3141	Affymetrix Human Genome U133 Plus 2.0 Array	58
	GSE37745	Affymetrix Human Genome U133 Plus 2.0 Array	47
	GSE10245	Affymetrix Human Genome U133 Plus 2.0 Array	40
	GSE14814	Affymetrix Human Genome U133A Array	36
	GSE31547	Affymetrix Human Genome U133A Array	30
	GSE31546	Affymetrix Human Genome U133 Plus 2.0 Array	13

eTable 2. Clinical and Pathologic Features of Patients in Meta-training, Meta-testing, and Independent Validation Cohorts

	Meta-training	Meta-testing	*P* Value*	DCC	GSE30219	TCGA
No. of samples	729	716		323	207	439
Median age in years (range)	63 (35 – 86)	64 (30 - 90)	.32	66 (33 - 87)	61 (15 - 84)	67 (38 - 88)
Female (%)	364 (50)	321 (45)	.11	158 (49)	37 (18)	231 (53)
Male (%)	322 (44)	340 (47)		165 (51)	170 (82)	208 (47)
Stage IA (%)	125 (17)	118 (16)		108 (33)	98 (47)	123 (28)
Stage IB (%)	155 (21)	157 (22)		116 (36)	32 (15)	116 (26)
Stage IA/B (%)[a]	128 (18)	109 (15)		0	0	6 (1)
Stage IIA (%)	22 (3)	18 (3)		17 (5)	9 (4)	44 (10)
Stage IIB (%)	65 (9)	55 (8)	.59	41 (13)	22 (11)	55 (13)
Stage IIA/B (%)[b]	22 (3)	28 (4)		0	0	1 (0)
Stage IIIA (%)	42 (6)	63 (9)		33 (10)	22 (11)	62 (14)
Stage IIIB (%)	25 (3)	17 (2)		7 (2)	16 (8)	10 (2)
Stage IIIA/B (%)[c]	10 (1)	9 (1)		0	0	0
Stage IV (%)	8 (1)	6 (1)		0	3 (1)	21 (5)
Positive smoking history (%)	397 (54)	393 (55)	1.00	206 (64)	NA	371 (85)
Non-smoker (%)	151 (21)	148 (21)		32 (10)	NA	61 (14)
Median Follow-up in months	47.6	43.85	.29	48	55	14.47
No. of death (%)	291 (40)	303 (42)		155 (48)	134 (65)	118 (27)

Abbreviations: DCC, Director's Challenge Consortium; TCGA, TCGA lung adenocarcinoma dataset;
[a] annotated as stage I patients only
[b] annotated as stage II patients only
[c] annotated as stage III patients only
NA represents information was unavailable
* The difference between meta-testing and meta-training dataset was calculated in terms of clinical pathologic factors. Specifically, age was compared with Wilcoxon rank-sum test; gender, stage and smoking history was compared with chi-squared test; follow-up difference was assessed with log-rank test.

eTable 3. Model Information About IRGPI

IRG 1	Full name	Immune processes	IRG 2	Full name	Immune processes	Coefficient
ESM1	endothelial cell-specific molecule 1	Cytokine	*AGTR1*	angiotensin II receptor, type 1	Cytokine receptor	-0.151291958
CX3CR1	chemokine (C-X3-C motif) receptor 1	Cytokine receptor	*FCGR2B*	Fc fragment of IgG, low affinity IIb, receptor (CD32)	BCR signaling	0.10001878
ADRB2	adrenergic, beta-2-, receptor, surface	Cytokine receptor	*HTR3A*	5-hydroxytryptamine (serotonin) receptor 3A	Cytokine receptor	0.014243079
ADRB2	adrenergic, beta-2-, receptor, surface	Cytokine receptor	*INHBB*	inhibin, beta B	Cytokine	0.074206935
EDN3	endothelin 3	Cytokine	*CD1B*	CD1b molecule	Antigen processing and presentation	-0.035977941
EDNRB	endothelin receptor type B	Cytokine receptor	*INHBB*	inhibin, beta B	Cytokine	0.06319796
ANGPT1	angiopoietin 1	Cytokine receptor	*IL1B*	interleukin 1, beta	Cytokine	0.028027231
GZMB	granzyme B	Natural Killer cell cytotoxicity	*IL6R*	interleukin 6 receptor	Cytokine receptor	-0.047644839
HSPA6	heat shock 70kDa protein 6 (HSP70B')	Antigen processing and presentation	*CD1C*	CD1c molecule	Antigen processing and presentation	-0.025903083
IL7	interleukin 7	Cytokine	*MIA*	melanoma inhibitory activity	Cytokine	0.028908034
INHBB	inhibin, beta B	Cytokine	*CX3CL1*	chemokine (C-X3-C motif) ligand 1	Cytokine	-0.00418581
NR3C2	nuclear receptor subfamily 3, group C, member 2	Cytokine receptor	*PIK3CA*	phosphoinositide-3-kinase, catalytic, alpha polypeptide	Multiple	0.030567093
NFKBIB	nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, beta	Multiple	*PTGER2*	prostaglandin E receptor 2 (subtype EP2), 53kDa	Cytokine receptor	-0.013838628
NRAS	neuroblastoma RAS viral (v-ras) oncogene homolog	Multiple	*BMPR2*	bone morphogenetic protein receptor, type II (serine/threonine kinase)	Cytokine receptor	-0.104349443
PSMD2	proteasome (prosome, macropain) 26S subunit, non-ATPase, 2	Antigen processing and presentation	*TGFBR2*	transforming growth factor, beta receptor II (70/80kDa)	Cytokine receptor	-0.006260886
RORA	RAR-related orphan receptor A	Cytokine receptor	*TGFA*	transforming growth factor, alpha	Cytokine	0.063627166
RORA	RAR-related orphan receptor A	Cytokine receptor	*RABEP1*	rabaptin, RAB GTPase binding effector protein 1	Cytokine	0.035503811
CCL7	chemokine (C-C motif) ligand 7	Cytokine	*CCL17*	chemokine (C-C motif) ligand 17	Cytokine	-0.068004041
CCL13	chemokine (C-C motif) ligand 13	Cytokine	*IL32*	interleukin 32	Cytokine	0.066796769
SECTM1	secreted and transmembrane 1	Cytokine	*VAV1*	vav 1 guanine nucleotide exchange factor	Multiple	-0.052319774
BTK	Bruton agammaglobulinemia	BCR signaling	*IL1R2*	interleukin 1 receptor, type II	Cytokine receptor	0.097609448

	tyrosine kinase					
C5	complement component 5	Cytokine	*ZAP70*	zeta-chain (TCR) associated protein kinase 70kDa	Multiple	0.046488127
C5	complement component 5	Cytokine	*MIA*	melanoma inhibitory activity	Cytokine	0.039305453
IL1R2	interleukin 1 receptor, type II	Cytokine receptor	*CD1C*	CD1c molecule	Antigen processing and presentation	-0.092033309
MIA	melanoma inhibitory activity	Cytokine	*CD1B*	CD1b molecule	Antigen processing and presentation	-0.185391168

eTable 4. Univariate and Multivariate Analyses of Prognostic Factors in Meta-training, Meta-testing, and Independent Validation Data Sets

Datasets		Univariate		Multivariate	
	Variable	HR (95% CI)	*P* Value[a]	HR (95% CI)	*P* Value[b]
Meta-training	Age	1.04 (1.02 – 1.05)	2.68×10^{-7}	1.03 (1.00 – 1.06)	.055
	Gender	1.59 (1.24 - 2.02)	.00017	1.33 (0.72 - 2.44)	.37
	Smoking	2.08 (1.46 - 2.96)	3.27×10^{-5}	1.09 (0.47 - 2.55)	.84
	Stage	1.39 (1.29 – 1.50)	$<2\times10^{-16}$	1.24 (1.05 - 1.47)	.010
	Grade	2.55 (1.73 - 3.76)	9.45×10^{-7}	1.82 (1.12 - 2.95)	.016
	Immune risk	4.38 (3.37 - 5.70)	$<2\times10^{-16}$	4.86 (2.28 - 10.34)	4.10×10^{-5}
Meta-testing	Age	1.02 (1.01 – 1.03)	.0044	1.02 (1.00 – 1.03)	.022
	Gender	1.38 (1.09 – 1.75)	.0075	1.14 (0.82 - 1.58)	.44
	Smoking	1.57 (1.14 - 2.17)	.0050	1.19 (0.81 - 1.77)	.38
	Stage	1.45 (1.35 – 1.57)	$<2\times10^{-16}$	1.46 (1.34 - 1.58)	$< 2\times10^{-16}$
	Grade	1.34 (0.99 – 1.81)	.056	-	-
	Immune risk	2.21 (1.75 – 2.79)	7.47×10^{-12}	1.72 (1.26 - 2.33)	.00054
DCC	Age	1.03 (1.01 – 1.05)	.00054	1.04 (1.02 – 1.05)	.00011
	Gender	1.41 (1.02 – 1.95)	.035	1.07 (0.77 - 1.49)	.68
	Smoking	1.31 (0.70 – 2.45)	.39	-	-
	Stage	1.54 (1.39 - 1.70)	$<2\times10^{-16}$	1.52 (1.37 - 1.69)	5.66×10^{-15}
	Grade	1.03 (0.81 - 1.31)	.81	-	-
	Immune risk	2.00 (1.41 - 2.84)	8.54×10^{-5}	1.75 (1.22 - 2.50)	.0023
GSE30219	Age	1.04 (1.02 – 1.06)	6.48×10^{-6}	1.03 (1.01 – 1.05)	.00050
	Gender	2.00 (1.19 - 3.37)	.0081	1.49 (0.88 – 2.54)	.14
	Smoking	NA	NA	-	-
	Stage	1.28 (1.17 - 1.40)	1.59×10^{-8}	1.19 (1.08 - 1.31)	.00048
	Grade	NA	NA	-	-
	Immune risk	3.72 (2.43 - 5.70)	1.10×10^{-10}	2.36 (1.47 – 3.79)	.00038
TCGA	Age	1.02 (1.00 – 1.04)	.073	-	-
	Gender	1.13 (0.79 - 1.62)	.51	-	-
	Smoking	0.85 (0.49 - 1.46)	.54	-	-
	Stage	1.32 (1.20 - 1.45)	1.90×10^{-9}	1.30 (1.18 - 1.43)	1.40×10^{-7}
	Grade	NA	NA	-	-
	Immune risk	2.15 (1.46 - 3.16)	7.28×10^{-5}	1.85 (1.26 - 2.73)	.0019

Abbreviations: DCC, Director's Challenge Consortium; TCGA, TCGA adenocarcinoma dataset;
a P value was calculated by log-rank test
b P value was calculated with wald-test of Cox-proportional hazard regression model
- represents variables were not included in multivariate Cox regression
NA represents variables were not publicly available
Age, stage, grade was coded as continuous variable. Specifically, stage was coded as IA=1, IA/IB=1.5, IB=2, IIA=3, IIA/B=3.5, IIB=4 etc. Grade was coded as well differentiated=0, moderate differentiated=1, poorly differentiated=2. The risk factors of gender, smoking and immune risk are male, smoker and high risk based on IRGPI.

eTable 5. Biological Processes Overrepresented by Genes Consisting of IRGPI

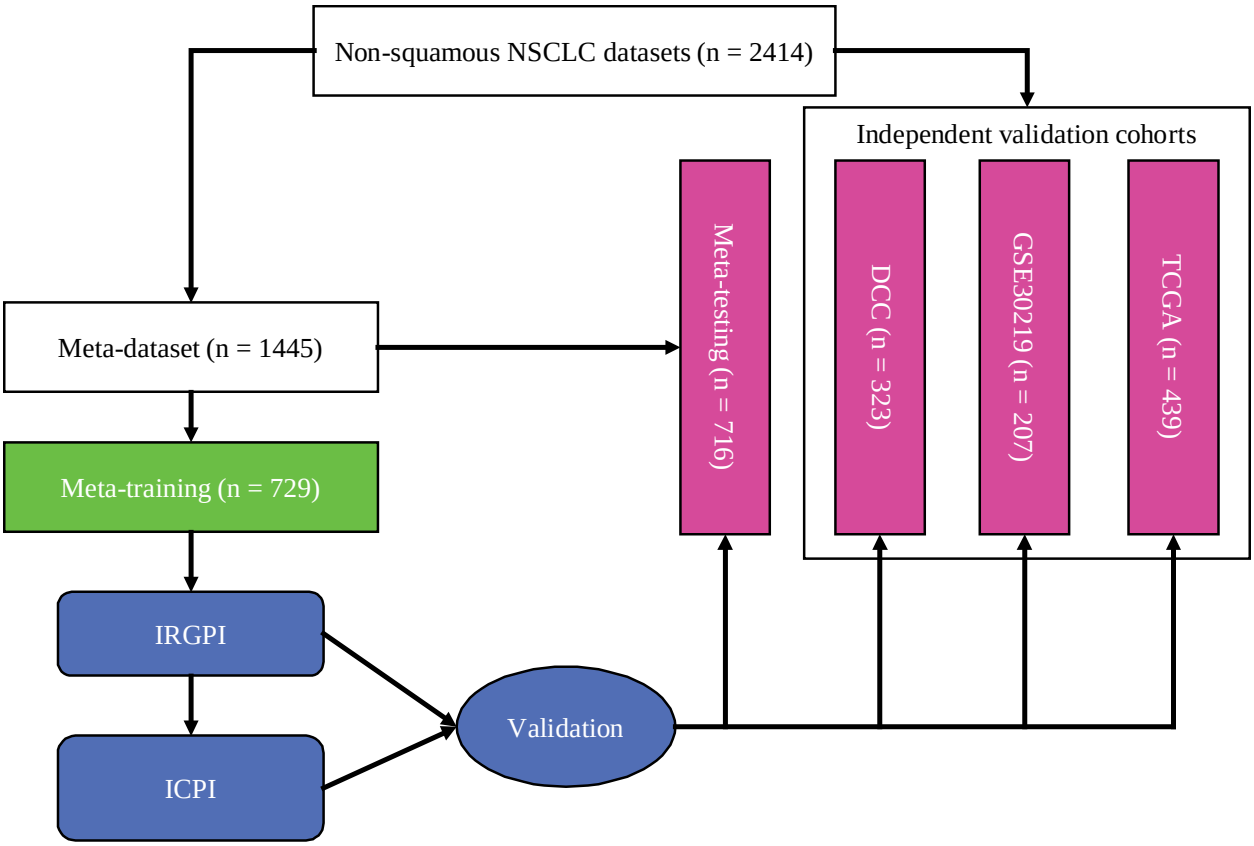
GO ID	BP name	*P* Value
GO:0030593	neutrophil chemotaxis	.018
GO:2001240	negative regulation of extrinsic apoptotic signaling pathway in absence of ligand	.042
GO:0010468	regulation of gene expression	.053
GO:0002548	monocyte chemotaxis	.063
GO:0048246	macrophage chemotaxis	.090
GO:0071347	cellular response to interleukin-1	.095

eTable 6. RMS Time Ratio Between Low- and High-Risk Groups Based on IRGPI or ICPI in Different Data Sets

RMS time (months)	IRGPI			ICPI		
	low risk (95% CI)	high risk (95% CI)	ratio (95% CI)	low risk (95% CI)	high risk (95% CI)	ratio (95% CI)
Meta-training	99.40 (95.05 - 103.75)	59.88 (53.89 - 65.87)	1.66 (1.49 - 1.85)	101.51 (97.31 - 105.72)	58.47 (52.62 - 64.32)	1.74 (1.56 - 1.94)
Meta-testing	88.88 (83.78 - 93.98)	63.37 (57.14 - 69.60)	1.40 (1.25 - 1.57)	95.76 (90.88 - 100.64)	57.68 (51.89 - 63.47)	1.66 (1.48 - 1.86)
DCC	84.87 (76.73 - 93.00)	64.75 (57.83 - 71.68)	1.31 (1.14 - 1.51)	91.00 (83.62 - 98.38)	59.03 (52.10 - 65.96)	1.54 (1.34 - 1.78)
GSE30219	96.32 (87.12 - 105.52)	52.44 (44.49 - 60.39)	1.84 (1.54 - 2.20)	91.71 (82.74 - 100.69)	48.93 (40.70 - 57.15)	1.88 (1.54 - 2.28)
TCGA	74.77 (63.47 - 86.08)	50.60 (41.51 - 59.68)	1.48 (1.17 - 1.87)	83.90 (72.03 - 95.77)	47.65 (39.67 - 55.63)	1.76 (1.41 - 2.19)

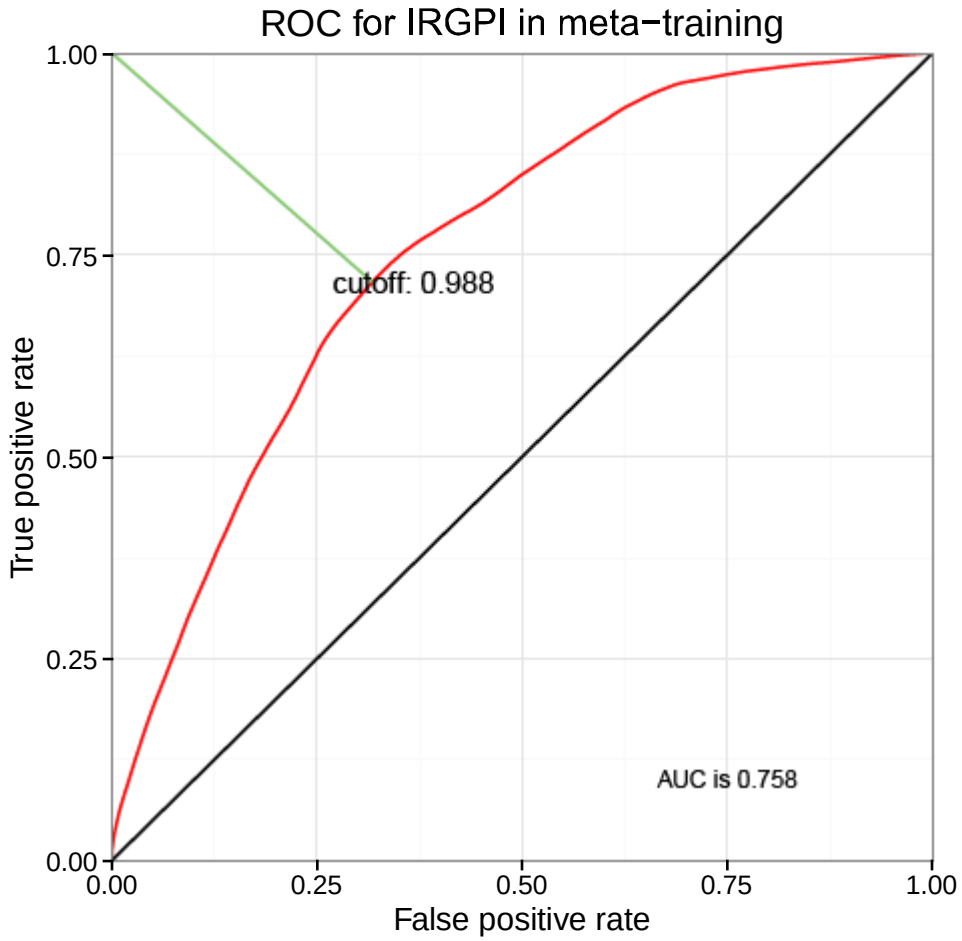
eFigure 1. Overview of the Construction and Validation of Immune and Composite Immune/Clinical Signatures

Nineteen datasets were included in this study. The three largest individual datasets were used for independent validation, while the remaining 16 datasets were merged to a meta-dataset. The meta-dataset was randomly divided into meta-training and meta-testing datasets. The meta-training dataset was used to build an immune prognostic signature (IRGPI). Age, stage and IRGPI were used to construct the composite immune clinical prognostic signature (ICPI). Both IRGPI and ICPI were evaluated on the meta-testing, Director's Challenge Consortium (DCC), GSE30219 and TCGA lung adenocarcinoma (TCGA) datasets.



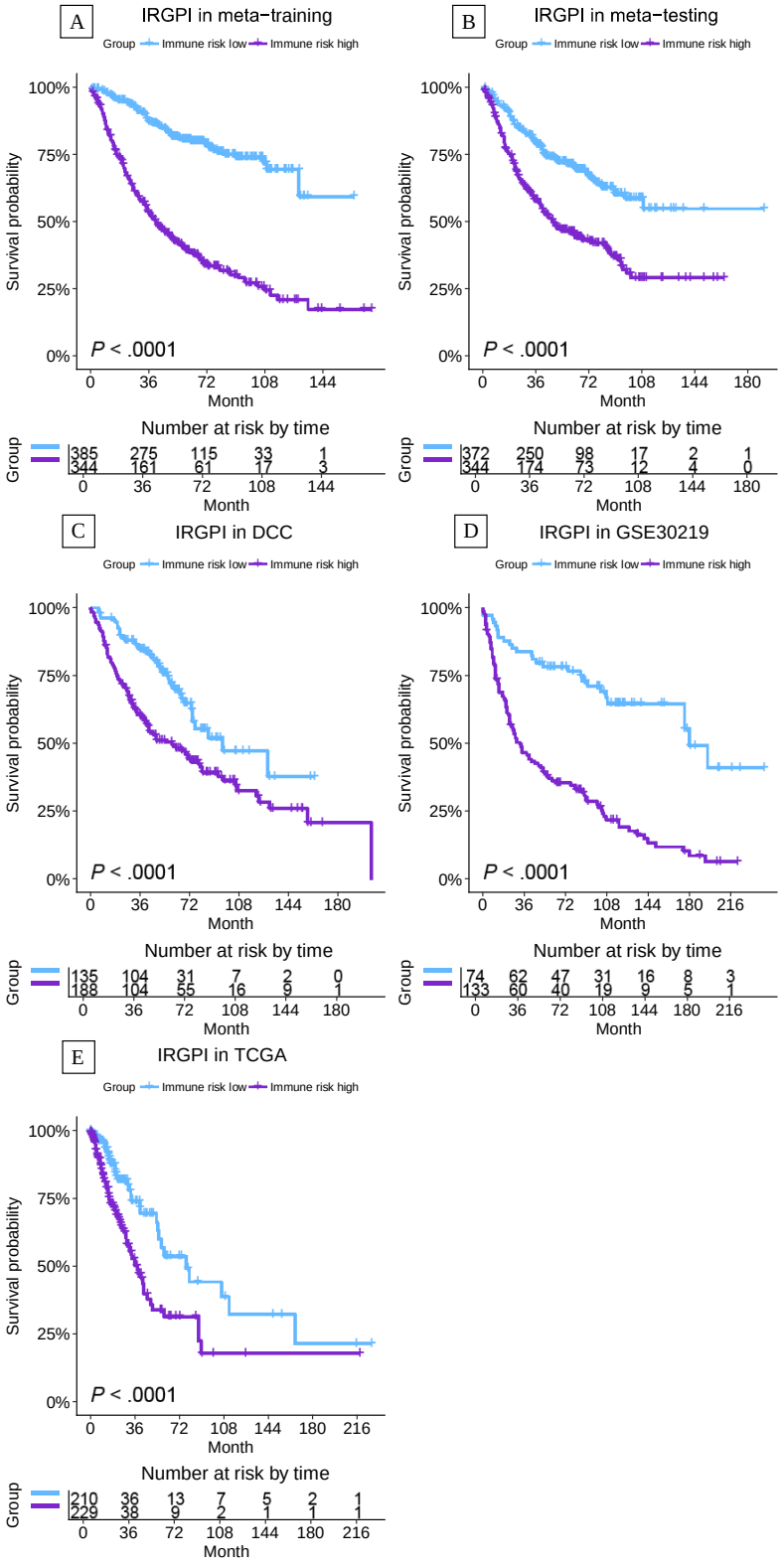
eFigure 2. Time-Dependent ROC Curve for IRGPI in the Meta-training Data Set at 5 Years

Green line indicates the shortest distance between the receiver operating characteristic (ROC) curve and the point represents 100 percent true positive rate and 0 percent false positive rate. Intersection of green line with ROC curve corresponds to the IRGPI score of 0.988 which was used as cutoff for IRGPI to stratify patients into low or high immune risk group.



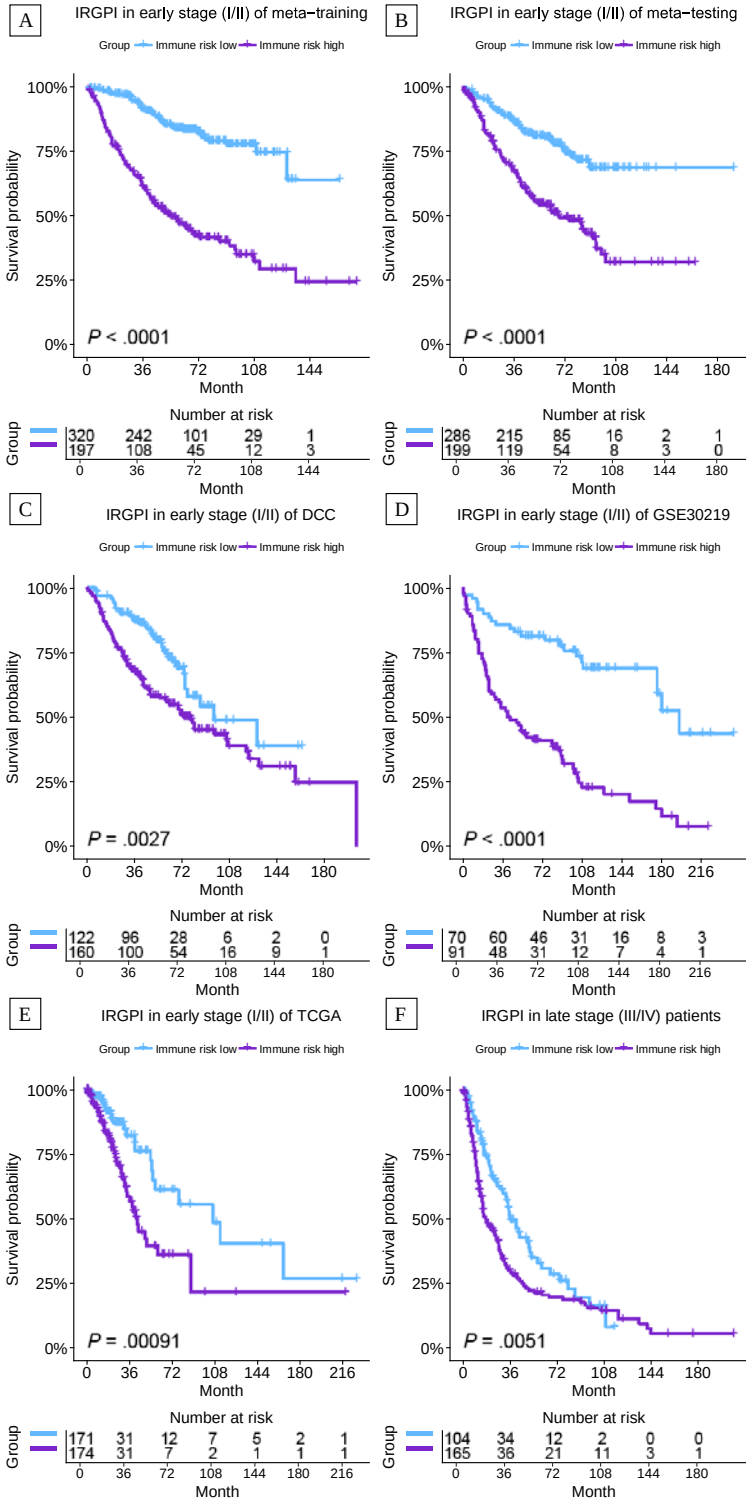
eFigure 3. Kaplan-Meier Curve of Overall Survival for Patients With Different IRGPI Risks

All patients in meta-training (A), meta-testing (B) and 3 independent validation datasets (C-E) were stratified into low or high immune risk groups based on IRGPI score.



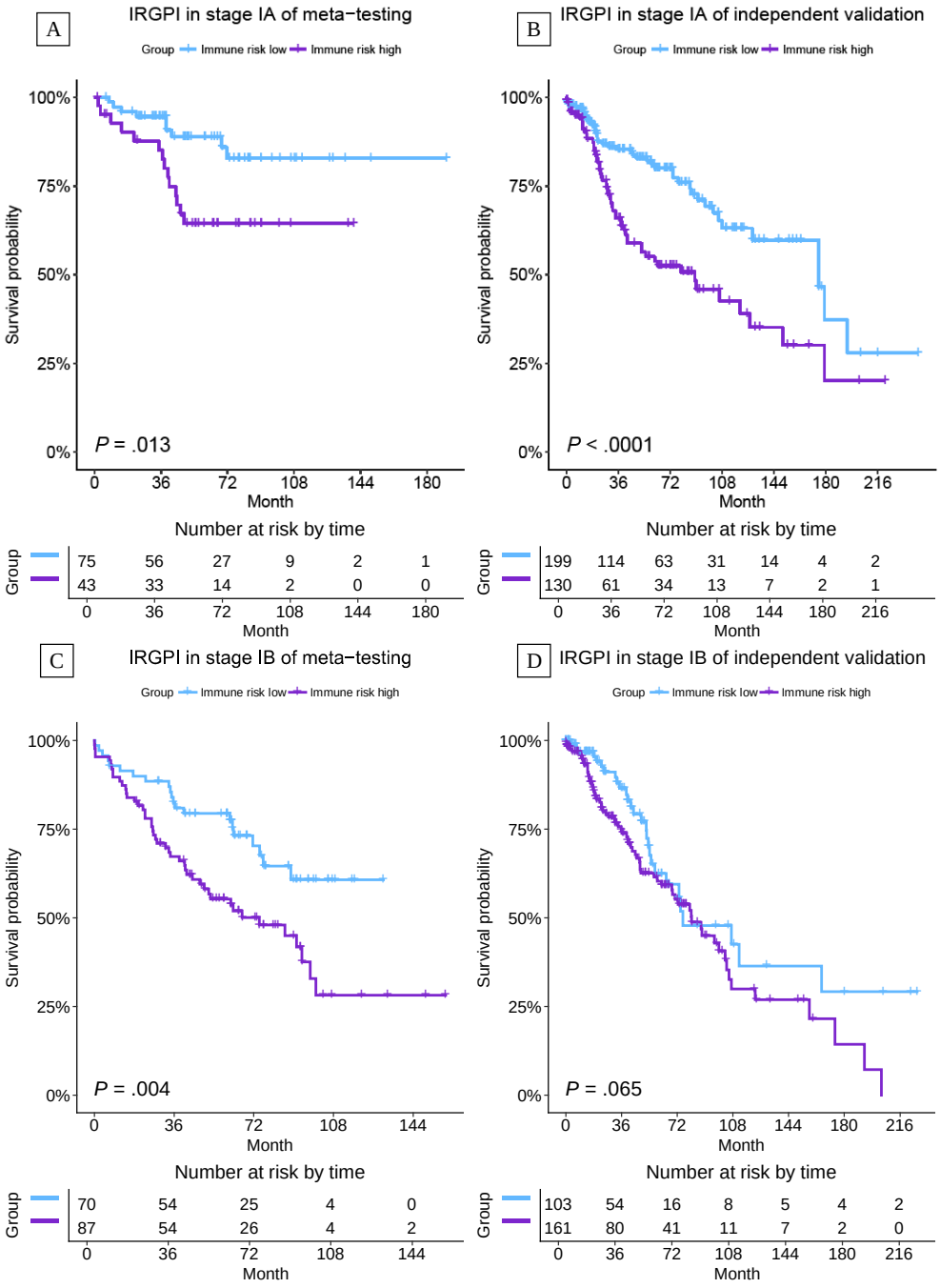
eFigure 4. Kaplan-Meier Curve of Overall Survival for Early- and Late-Stage Patients With Different IRGPI Risks

Early stage (stage I or II) patients in meta-training (A), meta-testing (B) and 3 independent validation datasets (C-E) were stratified into low or high immune risk groups based on IRGPI score. (F) Late stage (stage III or IV) patients of meta-testing, DCC, GSE30219 and TCGA were stratified into low or high immune risk groups based on IRGPI score.



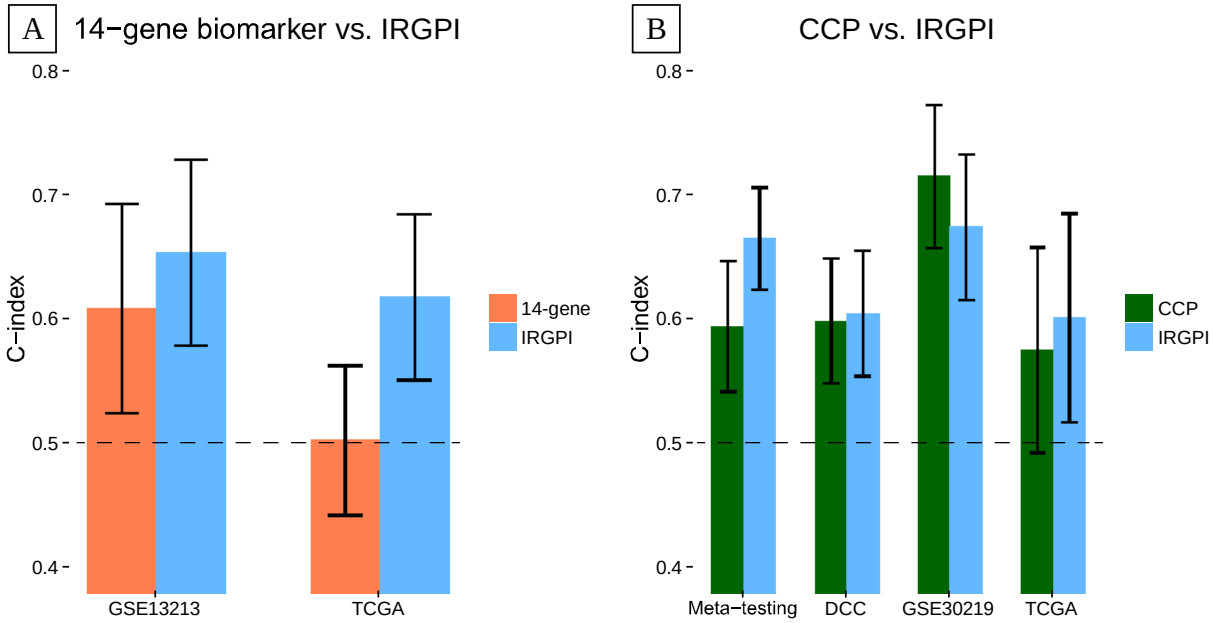
eFigure 5. Kaplan-Meier Curve of Overall Survival for Stage IA and IB Patients With Different IRGPI Risks

A: survival curve of stage IA patients of meta-testing dataset; B: survival curve of all stage IA patients in 3 independent validation datasets; C: survival curve of stage IB patients of meta-testing dataset; D: survival curve of all stage IB patients in 3 independent validation datasets.



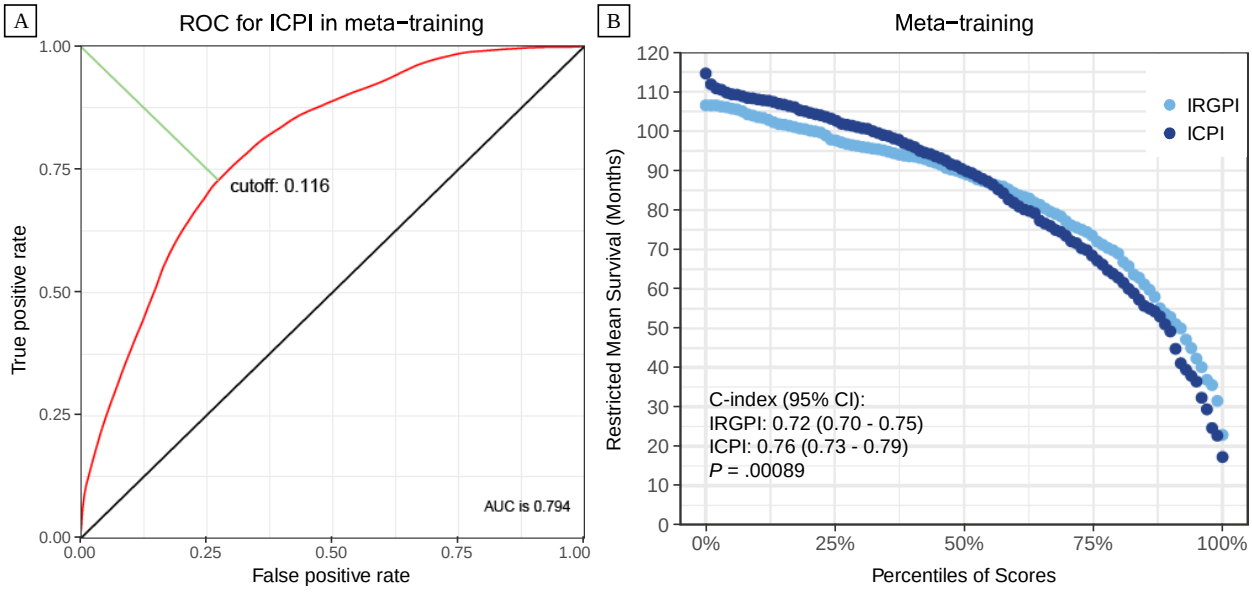
eFigure 6. C-index Comparison Between IRGPI and 2 Existing Biomarkers

A: comparison of C-index between 14-gene biomarker and IRGPI in stage I to III patients in GSE13213 and TCGA dataset. B: comparison of C-index between CCP score and IRGPI in early stage patients in meta-testing, DCC, GSE30219 and TCGA datasets. Whisker represents confident interval of estimated corresponding C-index. Dashed horizontal line indicates random prediction (C-index = 0.5).



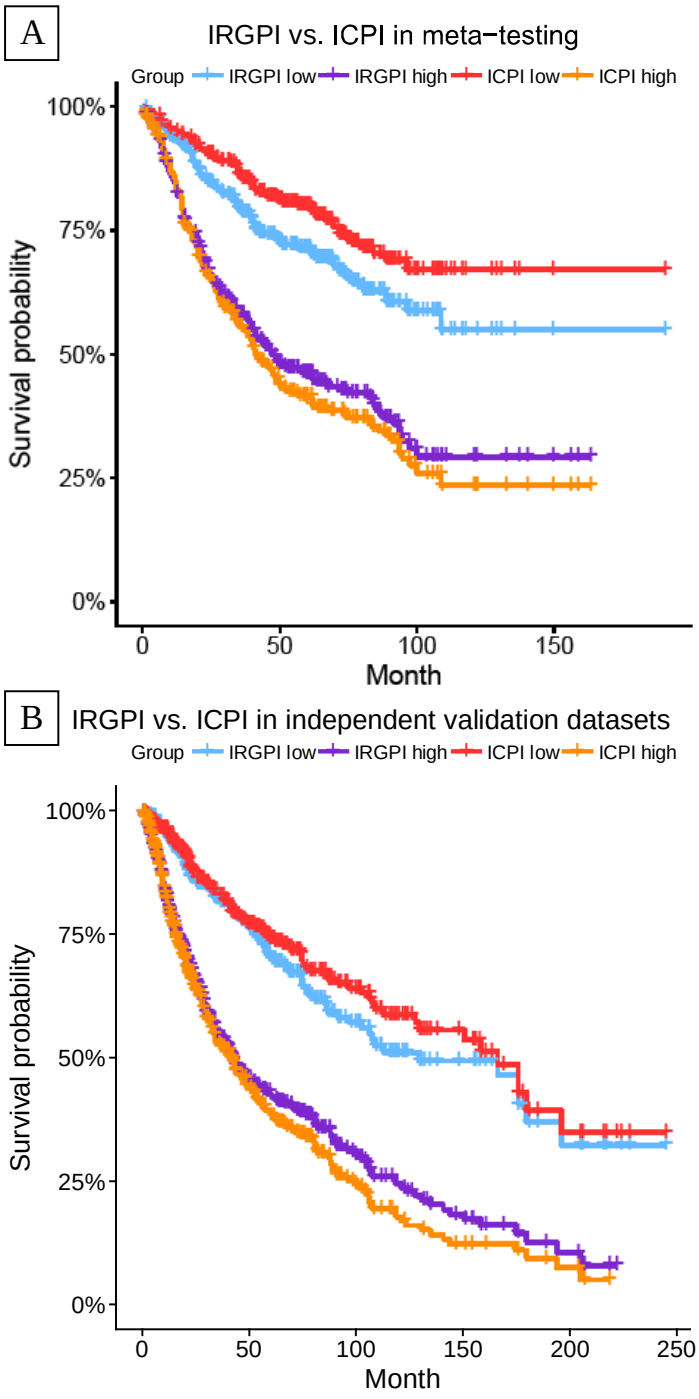
eFigure 7. Time-Dependent ROC for ICPI and RMS Curve for ICPI and IRGPI in Meta-training Data Set

(A) Green line indicates the shortest distance between the receiver operating characteristic (ROC) curve and the point represents 100 percent true positive rate and 0 percent false positive rate. Intersection of green line with ROC curve corresponds to the ICPI score of 0.116 which was used as cutoff for ICPI to stratify patients into low or high risk groups. (B) RMS (restricted mean survival) curve of continuous ICPI and IRGPI score in meta-training dataset



eFigure 8. Kaplan-Meier Curves for Overall Survival of All Patients Stratified by the IRGPI and the ICPI

Patients were stratified with both IRGPI and ICPI in all patients of meta-testing (A) and all patients of 3 independent validation datasets (B). The composite index had higher separation than immune index alone between high versus low risk groups in both cases.



eFigure 9. C-index Comparison Between ICPI and mPS Score in Validation Data Sets

Comparison of C-index between ICPI and mPS score in early stage patients in meta-testing, DCC, GSE30219 and TCGA datasets. Dashed horizontal line indicates random prediction (C-index = 0.5).