Title: Development and validation of the Immune Profile Score (IPS), a novel multi-omic algorithmic assay for stratifying outcomes in a real-world cohort of advanced solid cancer patients treated with immune checkpoint inhibitors

Authors

Alia Zander[1], Rossin Erbe[1], Yan Liu[1], Ailin Jin[1], Seung Won Hyun[1], Sayantoni Mukhopadhyay[1], Ben Terdich[1], Mario Rosasco[1], Nirali Patel[1], Brett Mahon[1], Kate Sasser[1], Michelle Ting-Lin[1], Halla Nimeiri[1], Justin Guinney[1], Douglas R. Adkins[2], Matthew Zibelman[3], Kyle A. Beauchamp[1], Chithra Sangli[1], Michelle M. Stein[1], Timothy Taxter[1], Timothy A. Chan[4], Sandip Pravin Patel[5], Ezra Cohen[1,5,*]

1. Tempus AI, Inc., Chicago, IL
2. Division of Oncology, Washington University School of Medicine, St. Louis, MO
3. Department of Hematology/Oncology, Fox Chase Cancer Center, Philadelphia, PA
4. Center for Immunotherapy and Precision Immuno-Oncology, Lerner Research Institute, Cleveland Clinic Foundation, Cleveland, OH
5. Division of Hematology/Oncology, Department of Medicine, University of California San Diego, Moores Cancer Center, La Jolla, CA.
6. *corresponding author; ezra.cohen@tempus.com

Abstract

Background: Immune checkpoint inhibitors (ICIs) have transformed the oncology treatment landscape. Despite substantial improvements for some patients, the majority do not benefit from ICIs, indicating a need for predictive biomarkers to better inform treatment decisions.

Methods: A de-identified pan-cancer cohort from the Tempus multimodal real-world database was used for the development and validation of the Immune Profile Score (IPS) algorithm leveraging Tempus xT (648 gene DNA panel) and xR (RNAseq). The cohort consisted of advanced stage cancer patients treated with any ICI-containing regimen as the first or second line of therapy. The IPS model was developed utilizing a machine learning framework that includes tumor mutational burden (TMB) and 8 RNA-based biomarkers as features.

Results: IPS-High patients demonstrated significantly longer overall survival (OS) compared to IPS-Low patients (HR=0.45, 90% CI [0.40-0.52]). IPS was consistently prognostic in PD-L1 (positive/negative), TMB (high/low), microsatellite status (MSS/MSI), and regimen (ICI only/ICI + other) subgroups. In a subgroup of TMB-Low patients who received ICI-only therapy (n=323), IPS-high patients had longer survival than IPS-Low patients (HR=0.41 [0.30-0.57]). In a subgroup of MSS patients who received ICI only therapy in LOT1, IPS-High patients had longer survival than IPS-Low patients (HR=0.33 [0.24-0.45]). Additionally, IPS remained significant in multivariable models controlling for TMB, MSI, and PD-L1, with IPS HRs of 0.49 [0.42-0.56], 0.47 [0.41-0.53], and 0.45 [0.38-0.53] respectively.

In an exploratory predictive utility analysis of the subset of patients (n=345) receiving first-line (1L) chemotherapy (CT) and second-line (2L) ICI, there was no significant effect of IPS for time to next treatment on CT in L1 (HR=1.06 [0.85-1.33]). However, there was a significant

effect of IPS for OS on ICI in L2 (HR=0.63 [0.46-0.86]). A test of interaction was statistically significant ($p<0.01$).

Conclusions: Our results demonstrate that IPS is a generalizable multi-omic biomarker that can be widely utilized clinically as a prognosticator of ICI based regimens.

Background

Cancer immunotherapies, particularly immune checkpoint inhibitors (ICIs) targeting PD-[L]1 and CTLA-4, have transformed the oncology treatment landscape. This transformation has been especially notable in cases where conventional systemic therapy options were associated with poor long-term outcomes [1]. Despite substantial improvements, the majority of patients do not benefit from ICIs, emphasizing the need for predictive biomarkers to inform treatment decisions [2].

To date, identifying candidates for immunotherapy relies on myriad PD-L1 immunohistochemistry (IHC) staining criteria across cancer types in addition to pan-cancer biomarkers of microsatellite instability (MSI) status and tumor mutational burden (TMB). Although PD-L1 positivity or high TMB may suggest potential responsiveness to ICIs, there remains a clinical need to improve our ability to determine whether patients will benefit from ICI treatment given the significant number of patients who do not under current guidelines [3].

Translational research efforts have made significant strides in identifying molecular biomarkers beyond PD-L1 IHC, TMB, and MSI, which characterize various aspects of the cancer-immunity cycle that hold promise as predictive immunotherapy biomarkers [4]. Advancements in RNA profiling technologies for both fresh tissue and formalin fixed paraffin embedded tissues have been essential in enabling analysis of routine pathology samples from clinical trials. As evidenced in the comprehensive analysis from Litchfield *et al* of publically available ICI clinical trial data sets, RNA biomarkers hold significant value in complementing DNA biomarkers for characterizing ICI response across solid organ cancers [5]. However, while large-panel DNA

sequencing is commonly performed in advanced-stage cancers to guide treatment decisions, the clinical utility and routine implementation of RNA sequencing are still emerging, and consequently, RNA sequencing is less frequently available [6]. Additionally, the clinical validation of predictive biomarkers is constrained by the limited availability of large-scale multi-omic datasets that include high-quality clinical outcomes data [7]. Driven by these challenges and unmet clinical needs, we developed and validated a multi-omic, pan-solid cancer biomarker using the Tempus testing platform, incorporating both DNA and RNA analysis, to predict outcomes of ICI therapy.

Methods

Patient Cohorts

The model development and validation cohorts consist of patients from the de-identified Tempus real-world multimodal database, all of whom underwent clinical next-generation sequencing. **Figure 1** illustrates the CONSORT diagram for the validation cohort. Patients included in the study were diagnosed with stage IV cancer and received an approved ICI (**Table S1**) in 1L or 2L therapy after January 1, 2018 and before July 1, 2023 (1L) or January 1, 2024 (2L). Cancer types included in the validation and development cohorts are listed in **Tables S2** and **S3**. Patients with an ECOG score ≥ 3 were excluded. To be eligible, samples had to be collected prior to any exposure to ICI therapy, with the time between sample collection and treatment within the standard of care range. Exclusion criteria included low tumor purity (<20% for development, <30% for validation) and samples collected from cytology or lymph node biopsies due to ambiguity of anatomic location of lymph node biopsy, high expression of immune genes in the lymph node, and background noise. Eligible patients were then representatively divided

into development (n=1707) and validation (n=1600) cohorts. Further characterization of the overall validation cohort is listed in **Table 1**.

NGS-based DNA and RNA sequencing

The Tempus testing platform includes both a targeted DNA sequencing assay (xT), and an exome capture RNA sequencing assay (xR) [8–10]. The current xT assay targets 648 genes, with a panel size of 1.9MB. Prior versions of xT assay, including a 596-gene version and other DNA sequencing assays, were also utilized in the analysis (**Table S4**). TMB was calculated by dividing the number of nonsynonymous mutations by the size of the panel size (**Supplemental Methods**) [11]. The xT assay also includes probes for loci frequently unstable in tumors with mismatch repair deficiencies, allowing for the assessment of microsatellite instability (MSI) and classifies tumors into MSI-H, and MSS categories [9]. The xR assay is based on the IDT xGen Exome Research Panel v2 backbone, comprising >415K individually synthesized probes and spans a 34 Mb target region (19,433 genes) of the human genome. Tempus-specific custom spike-in probes are added to enhance target region detection in key areas like fusion and viral probes. Clinically, the xR assay is used for reporting gene fusions, alternative gene splicing, and gene expression algorithms [9,10,12,13].

PD-L1 Immunohistochemistry

PD-L1 status for each patient was determined by clinical Tempus testing or curated from pathology reports associated with external PD-L1 IHC testing performed at the referring pathology lab. PD-L1 positive and negative classification for each cancer subtype was defined

per the FDA guidelines or clinical trials (**Table S5**). For cancer types lacking established PD-L1 IHC criteria, a generalized threshold of TPS > 1 was used to define positivity; this was also generalizable across PD-L1 clones used in testing.

Model development / ML and AI methodologies

DNA and RNA features adapted to the Tempus IO platform were used as the basis for feature selection for the Tempus IPS assay. The features in the IO platform consists of a comprehensive list of DNA and RNA based IO biomarkers that have been established in the literature as associated with tumor immune biology and immunotherapy outcomes [14]. In addition to the candidate features selected from the literature, two novel gene signatures were developed by Tempus as part of this study. The first is a signature of tumor-intrinsic immune resistance derived from single-cell RNA-sequencing data, which we term the single-cell immune resistance (scIR) signature [15]. Briefly, this signature was created using a variational autoencoder to extract biological signal from a single-cell RNA-sequencing sample taken from a lung adenocarcinoma patient. The scIR signature was strongly weighted in a small population of tumor cells within a highly immune-activated tumor environment and included known pathways of immune inhibitory signaling on tumor associated macrophages. The second signature was created to capture known literature meta-analysis signals using 105 genes [16].

Using a cohort of 1707 patients treated with ICI, 1094 patients were used to select the features for the model and 613 were held out for model evaluation. This train-evaluation split was performed to create comparable cohorts, stratified on line of therapy and cancer type. To avoid

overreliance on this training set, candidate features were further evaluated in publicly available ICI data sets [5–8] using univariate Cox models. Features that did not reach $p < 0.05$ in any of these datasets were excluded from consideration. Using the remaining features, we fit a multivariate Cox proportional hazards model, stratifying by line of therapy (1L or 2L). The model was trained using 10-fold cross-validation, where balanced L1/L2 regularization was applied to remove redundant features, with cross-validation used to determine the regularization weights. The resulting model was then applied to the remaining 613 held-out patients to verify that the model performed consistently outside of the initial training data. After this assessment, the model's final feature coefficients were determined from the full 1707 patient training cohort. The IPS score was calculated as a linear combination of the coefficients and was min-max scaled to fall between 0-100. The threshold for IPS-Low was set at all patients below the 55th percentile among the full training cohort, IPS-High at greater than or equal to the 60th percentile, and the patients between the 55th and 60th percentiles form an indeterminate category.

Statistical analyses

The analyses conducted in this study were defined prospectively in a statistical analysis plan. The primary objective was to demonstrate in a pan-cancer ICI treated population that IPS-High patients had longer overall survival compared to IPS-Low patients. A stratified Cox proportional hazards model was employed for the primary endpoint of overall survival, with adjustment for treatment regimen type (ICI only vs. ICI+additional), and stratification by line of therapy (first-line vs. second-line). Risk set adjustment was applied in patients where sequencing (and

therefore study entry) occurred after the initiation of ICI [17]. The significance of the hazard ratio (HR) was evaluated using a one-sided Wald test at a 5% significance level. Consequently, the one-sided upper 90% confidence interval is provided for all survival analyses. The primary endpoint was also descriptively evaluated across several subgroups. These subgroups included PD-L1 positive and negative patients (based on available IHC data), TMB high and low (<10 mut/Mb and ≥10mut/Mb), age categories (<65 and ≥65), sex (male and female), regimens (ICI only vs. ICI+additional), and cancer types (restricted to those with at least 15 patients in both the IPS-High and IPS-Low groups). For each of these subgroups, a stratified Cox PH model (incorporating risk set adjustment) similar to the one described in the primary endpoint analysis was fit.

The prognostic utility of IPS over PD-L1 was evaluated by a likelihood ratio test that compared the full Cox model including both PD-L1 and IPS to a reduced Cox model that included PD-L1 alone (Methods - Statistical analysis). The prognostic utility of the IPS score in relation to TMB and MSI-H was assessed using a similar approach.

An exploratory analysis of the predictive utility of the IPS score was performed by combining the training and validation cohorts of patients who received chemotherapy (CT) as first line treatment and ICI as second line treatment. Patients served as their own control in this analysis, and outcomes were evaluated for two lines of therapy: time to next treatment (TTNT) on CT and OS on ICI. If IPS was purely prognostic, time to next treatment (as a surrogate for progression) would be anticipated to be longer in IPS-H patients than in IPS-Low patients. The

HR for TTNT of IPS-H to IPS-L would then be of a similar magnitude as the HR for OS on second line treatment with ICI. A conditional model for recurrent events was fit to the selected subset of patients. Specifically, a Cox proportional hazards model, stratified by line of therapy, was used to model the two ordered time periods: period 1 in which the patient received CT and period 2 in which the patient received ICI. A Wald test p-value of less than 0.05 for the interaction between IPS and line of treatment would indicate a significant difference in the hazard ratios between the two time periods.

Statement of Ethics

This study was conducted in accordance with HIPAA regulations where applicable, and IRB exempt determinations (Advarra Pro00076072, Pro00072742).

Data Availability

Deidentified data used in the research were collected in a real-world health care setting and subject to controlled access for privacy and proprietary reasons. When possible, derived data supporting the findings of this study have been made available within the paper and its Supplementary Figures and Tables.

Results

IPS model development and feature characterization

To develop a biomarker that robustly stratifies outcomes in pan-cancer, solid tumor, metastatic ICI-treated patients, we randomly divided the Tempus ICI cohort into a 1,707 development

patient cohort and held out 1,600 patients for clinical validation (**Tables S6** and **S7**). The development cohort was further subdivided into 1,094 patients for feature selection and model training and 613 were reserved for initial model evaluation. Potential features included in the model were drawn from a comprehensive set of RNA and DNA biomarkers that had been previously implicated in tumor-immune biology or associated with IO-related outcomes. We also considered two novel gene signatures developed as a part of this study that characterize expression patterns of tumor-intrinsic immune resistance (see "Model Development", Methods).

Candidate model features were initially selected using a combination of biological plausibility, association with rwOS in publicly available ICI studies, and favorable analytical properties. [5–8]. These candidate biomarkers were included in a preliminary multivariate Cox model, stratified by line of therapy feature weights were determined using the combined development and evaluation cohorts (n=1,707) and included the following features: TMB [18], expression of *CD74* [19], *CD274* [20], *CD276* [21], *CXCL9* [22], *IDO1* [23], *PDCD1LG2* [20], *SPP1* [22], *TNFRSF5* [24], scIR signature [15], the meta-analysis literature signature, and a gMDSC signature [25] (**Figure 2**). The IPS-L and IPS-H thresholds were set as the 55th and 60th percentile of the full training cohort respectively. Patients that fell between the 55th and 60th percentile thresholds were classified as indeterminate and excluded from further analysis.

Patient characterization of validation cohort

The validation cohort was comprised of 1600 patients with stage IV cancer: median (IQR) age of 65.0 (58.0-73.0) years, 40% female (n=645), 1,114 (70%) were treated at community-based hospital or medical practices, and 1,043 (65%) were smokers, 1,016 patients (64%) were White (**Table 1, Table S8**). The majority of patients in the study were *de novo* stage IV at the time of diagnosis (1,219 [76%]). There were 16 cancer types included in the validation study (**Table S2**). The most common cancer was NSCLC (647 patients [40.4%]), followed by GEJ (171 [11%]), urothelial (137 [9%]), RCC (131 [8%]) and HNSCC (125 [8%]) (**Table S9**). Of note, the following cancer subtype roll-ups were used for NSCLC (lung adenocarcinoma - 371 [23%], lung squamous carcinoma - 155 [9.7%], and NSCLC-NOS - 121 [7.6%]), gastro-esophageal (gastroesophageal adenocarcinoma - 147 [9.2%], gastroesophageal squamous cell carcinoma - 24 [1.5%]) (**Table S9**). The highest rates of IPS-H were observed in colorectal cancer (27 [59%]), melanoma (56 [55%]), and RCC (69 [53%]) subcohorts (**Table 1**). Consistent with current standards of care, 91% of the colorectal cancer patients were MSI-H (**Table S10**). The lowest rates of IPS-H were observed in GEJ (26 [15%]), urothelial (36 [26%]) and HNSCC (35 [28%]). PD-L1 IHC results were available on 1,132 patients (PD-L1 positive - [637], PD-L1 negative - [495]), the vast majority of cases were stained with PD-L1 22c3 (1,145) (**Table S11**). Notably, a higher proportion of IPS-H patients were PD-L1 positive (250 [43%]) versus PD-L1 negative (149 [26%]). TMB data were available on all patients in the study, and a higher proportion (783 [83%]) of IPS-L patients were TMB-L versus TMB-H (160 [17%]).

Patients were treated with one of ten FDA-approved ICIs (**Table S1**). The majority of patients received ICI therapy as part of the first line (1,326 [83%]) versus the second line (274 [17%]).

Treatment patterns with ICI were generally consistent with established standards of care. Of the ICI regimen types, ICI+chemotherapy (869 [54%]) was the most common, followed by ICI monotherapy (381 [24%]) and ICI doublet (153 [9.6%]). Notable cancer types and regimens include NSCLC (ICI mono - (92 [14%]), ICI doublet (30 [4.6%]) ICI+chemo - (525 [81%]), melanoma (ICI mono - (56 [55%]), ICI doublet - (40 [39%])), and RCC (ICI doublet - (53 [40%]), ICI+other (66 [50%]) (**Table S12**). Of the patients receiving ICI+other, the "other" consisted mainly of tyrosine kinase inhibitors (78 [4.8%]) of which the majority was used in RCC patients (ICI+TKI - [66]), and ICI with a biologic such as anti-VEGF in hepatocellular carcinoma (Biologic + ICI [26]) and anti-EGFR in GEJ (Biologic + Chemo + ICI - [30]) (**Table S13**).

The median follow-up time was 21.2 months (IPS-H) or 18.9 months (IPS-L); follow-up time was calculated from reverse Kaplan Meier (**Table S14**).

Clinical validation of IPS as a pan-cancer ICI biomarker

A multivariate CoxPH controlling for regimen (ICI monotherapy or ICI in combination with other therapies), and stratified by line of therapy (1L or 2L), was used to assess the prognostic association of IPS-H and IPS-L labels with patient outcomes. OS was demonstrated to be significantly longer in patients with tumors classified as IPS-H vs IPS-L (HR=0.45 [0.40, 0.52], p-value < 0.01) (**Fig.3**). Differences in survival between IPS-H and IPS-L were consistent across lines of therapy and regimens. The predicted OS from the CoxPH model is shown in (**Fig.3 a-b**) for the setting of ICI only in 1L or 2L. Notably, the predicted OS curves for ICI combination therapy in 1L and 2L demonstrate a similar relationship to the IPS result and predicted OS

(**Fig.S1**). In an exploratory analysis, we assessed the prognostic association of IPS as a continuous variable. The HR (for a 50 unit increase, HR=0.29 [0.24, 0.35]) was similar in magnitude and direction to the categorical representation of IPS.

Performance of IPS in clinical and biomarker subgroups

The prognostic association of IPS was also evaluated in clinical and biomarker subgroups. Patients whose tumors were classified as IPS-H had significantly longer OS than IPS-L tumors across all subgroups. Notably, significant associations were maintained across molecular biomarker subgroups of TMB-H, TMB-L, PD-L1+, PD-L1-, and MSS as well as clinical subgroups of presence/absence of brain or liver metastasis **(Fig.4).** HR subgroup estimates were similar in direction and magnitude to that of the overall estimate. Among the cancer subgroups evaluated, RCC (0.34 [0.20, 0.59]), HNSCC (0.38 [0.22, 0.67]), NSCLC (0.42 [0.34, 0.52]), and melanoma (0.47 [0.27, 0.82]) had the largest effects while the smallest effects for the IPS score were observed in GEJ, HCC, breast and CRC. An exploratory subgroup analysis was performed in RCC and melanoma to evaluate IPS in patients receiving ICI-doublet regimens which are enriched in those disease groups, with melanoma showing (HR=0.56 [0.25-1.23]) and RCC (HR=0.25 [0.10-0.63]) (**Fig.S2**; **Fig.S3**).

IPS complements TMB, PD-L1, and MSI in forecasting responses to ICI

To demonstrate the complementary and added value that IPS score has in relation to the clinically established biomarkers of TMB, PD-L1 IHC, and MSI, we compared the full model including both IPS score and the biomarker of interest to a reduced model of either TMB, PD-L1

IHC, or MSI without IPS (see Methods). In this analysis, we observed a significant association of IPS over TMB, PD-L1 IHC, and MSI ($p < 0.001$) which

The predicted OS curves for these biomarker subgroups, categorized by IPS status, are presented in patients treated with ICI monotherapy in 1L (**Fig.5 b-d**). Similar predicted OS curves for treatment conditions and lines of therapy are shown in the supplementary data (**Fig.S4**-**S7**). Given the size and clinical significance of the NSCLC cohort, these results are also broken out for NSCLC by PD-L1 status. The predicted OS curve is shown for combination therapy in 1L along with similar predicted OS curves for monotherapy and 2L treated patients (**Fig.5e**; **Figs.S7**-**S9**). HRs and 90% CI for the most relevant curves shown in the predicted OS plots are listed in **Fig.5f**.

Exploratory evaluation of predictive utility for IPS

In an exploratory analysis to test the potential predictive utility of IPS, we examined a combined cohort of training and validation patients that had been exposed to non-ICI and ICI therapies in 1L and 2L respectively. While IPS was not associated with TTNT on CT in 1L (HR=1.06 [0.85, 1.33]; **Fig.6a**), it was significantly associated with OS in patients receiving 2L ICI (HR=0.63 [0.46, 0.86]; **Fig.6b**). An interaction test between the two lines was significant ($p<0.01$) indicating that the HR for 2L ICI between IPS-H and IPS-L is significantly different from the HR for 1L CT.

To further evaluate prognostication of IPS in non-ICI treated patients as a means to understand predictive utility, an exploratory analysis was performed in stage IV patients from The Cancer

Genome Atlas (TCGA, N=722, patient selection criteria is described in Supplemental Methods). The TCGA enrollment period was prior to the approval and usage of ICI therapies thus ensuring a representative non-ICI comparator cohort that also had DNA and RNA sequencing available to generate IPS scores. There was a significant association of IPS with OS in this cohort (HR=0.75 [0.56-0.99]), however the hazard ratio was attenuated relative to the IPS validation cohort (**Fig.S10**).

Tumor distribution and IPS prevalence in an expanded pan-cancer cohort

To characterize IPS prevalence more generally including in cancer types without approved ICI indications, we examined the distribution of IPS-H and IPS-L in an expanded pan-cancer cohort of patients sequenced at Tempus. In the entire cohort encompassing 25 different cancer types, prevalence of IPS-H was 28.64%. Lung adenocarcinoma, RCC, and melanoma had IPS-H prevalence greater than 50% (**Fig. S11**). On the opposite side of the spectrum, GI neuroendocrine cancer, cholangiocarcinoma, CRC, gynecologic sarcomas, and PDAC all had IPS-H prevalence of less than 20%. Of note, lung squamous cell carcinoma had a prevalence of 25.59% and NSCLC-NOS had a prevalence of 45.75% indicating a likely high proportion of lung adenocarcinomas in the NOS group of patients. To further characterize how IPS may identify ICI responders outside of current cancer type or pan-cancer biomarker ICI approvals, we calculated the proportion of patients who are IPS-H and TMB-L (14.1%) after excluding cancer types with an ICI approval or tumors that were MSI-H (**Table S15**). We also generated a more granular cancer subtype type visualization of IPS status in relation to TMB status (**Fig. S12**).

Discussion/Conclusions

Leveraging the Tempus xT / xR assays and the IO platform along with real-world data from ICI treated patients, we developed and validated the multi-omic IPS algorithm in a prospectively designed retrospective study using a real-world cohort of advanced solid organ cancer patients treated with an ICI containing regimen in the first or second line of therapy. Using a prespecified statistical analysis plan, IPS was validated as a generalizable pan-cancer prognostic biomarker demonstrating that IPS-High patients have significantly longer OS than IPS-Low patients. Additionally, the validation demonstrated that IPS-High patients had significantly longer OS compared to IPS-Low patients across relevant ICI biomarker subgroups, including PD-L1 status, TMB levels, and microsatellite stability. Specifically, in TMB-low patients receiving ICI-only therapy, and microsatellite-stable (MSS) patients treated with ICI in their first line of therapy, IPS-High patients showed substantially longer survival than their IPS-Low counterparts. Notably, IPS retained its prognostic significance in multivariable models, even when controlling for TMB, MSI status, and PD-L1 expression. Overall, these analyses demonstrate the clinical value of IPS in assessing the potential benefits of ICI regimens beyond current standard of care biomarkers. Finally, a post-hoc exploratory analysis into the predictive capabilities of IPS was performed with patients who received chemotherapy in the first line of therapy and ICI in the second line. IPS did not predict time to the next treatment following chemotherapy, however, IPS was a significant predictor of OS when patients were subsequently treated with ICI.

Our study results build upon the growing body of evidence supporting that multi-omic biomarkers developed using machine learning / artificial intelligence methodologies, high-

throughput commercial NGS assays, and real-world clinical data can provide insights into tumor/immune biology and clinical outcomes. The current treatment paradigm of approved immunotherapies therapies in addition to the vast number of clinical trials (including ALCHEMIST, OptimICE-PCR, EQUATE, PET-Stop trials) utilizing immunotherapies with a diverse range of mechanisms and targets highlights opportunities and unmet clinical needs for patient selection using multi-omic biomarkers [26].

In the current treatment paradigm of stage IV solid organ cancers, there are opportunities for biomarkers to help inform clinical management for approved ICI regimens in indications with equipoise between regimens or indications that lack biomarkers for patient selection. This opportunity is perhaps most apparent in NSCLC where patients of all PD-L1 levels are approved for ICI+chemo while in tumors with PD-L1 IHC high (TPS > 50) patients can receive ICI+chemo or ICI monotherapy [27]. A significant focus of clinical research has therefore focused on further sub-stratification of PD-L1 IHC. Aguilar *et al.* showed in an RWD retrospective analysis that patients with TPS scores greater than 90 have significantly better outcomes than patients with TPS between 50 and 89, which may be informative for ICI monotherapy patient selection [28]. In our exploratory analysis of NSCLC patients receiving ICI monotherapy and subgrouped by PD-L1 IHC levels, patients with IPS high tumors in all PD-L1 IHC subgroups were observed to have longer OS than patients with IPS low tumors. This finding may represent the importance of CD274 (PD-L1) gene expression as a continuous feature in the IPS model. The analysis is notably limited by small sample sizes but generally highlights the potential of IPS to capture tumor immune biomarker granularity and precision. Currently the INSIGNA study which is a large

randomized control trial in NSCLC has aims focused on elucidating the optimal clinical management for these patients [29].

Regarding the potential for IPS to inform new treatment indication strategies, **Figure 6** highlights the cancer-specific IPS-High/low prevalence in an expanded cohort that includes diseases that currently lack ICI indications. Of note, MSS colorectal cancer and pancreatic cancer were shown to have among the lowest IPS-High rates which tracks with the historically limited response rates seen in ICI monotherapy trials for those cancer types [30–32]. IPS therefore could have value in identifying the rare potential ICI responders in these or similarly challenging diseases. Additionally, we showed the proportion of patients who are IPS-H/TMB-L/MSS in cancer subtypes that currently lack an ICI approval, indicating the potential pan-cancer role of IPS for identifying ICI responders in stage IV patients that currently lack an approved use. Prospective clinical trial designs utilizing IPS stratification or selection may be considered in the future [33]. However, perhaps even more impactful than development of monotherapy ICI studies is the potential application of multi-omic biomarkers such as IPS to inform patient selection for novel ICI combinatorial strategies and the next generation of immunotherapy modalities such as T-cell/NK-cell engagers and RNA cancer vaccines. These novel applications may require modified versions of IPS along with additional biomarkers that characterize the cancer-immunity cycle relevant to a specific combinatorial strategy [4]. Lastly, as ICI based regimens move into neoadjuvant and adjuvant settings, there are significant opportunities for patient selection strategies to reduce ICI exposure in patients unlikely to respond and therefore reducing the number of overall adverse events.

Limitations of this study reflect the real-world, retrospective nature of the validation cohort. While our study inclusion and exclusion criteria attempted to control for confounding variables and non-standard care scenarios, additional biases may be unaccounted for. Regarding our attempts to characterize the predictive nature of IPS, Tempus clinical testing and our subsequent clinical-molecular data set was generated predominantly in the post ICI-era. Therefore in our predictive analysis with Tempus sequenced patients we did not have the ability to perform case-control matching with patients who received non-ICI regimens prior to the approvals. We attempted to address this with an analysis of stage IV patients who did not receive ICI, collected from TCGA. Among these patients, we observed a significant difference in OS between patients classified as IPS-High versus IPS-Low. This result suggests that the IPS model has generalized prognostic utility, as would be biologically expected given the known prognostic association of immune infiltration in tumors [34] which the model is intended to capture. However, given the attenuated hazard ratio we observed in these non-ICI treated patients in TCGA versus the ICI treated patients in the study cohort, IPS appears to have additionally predictive utility. Also of note, the proportion of patients in each cancer type and biomarker subgroup is representative of clinical testing at Tempus which expectedly results in disproportionately sized cancer subgroups in the development and validation cohorts representative of cancer prevalence and NGS testing frequency. The variability of cohort size across cancer types therefore limits the ability to comprehensively evaluate the heterogeneity of IPS performance across cancer types. Additionally, the IPS model did not include clinical and lab features that have been demonstrated to add prognostic utility in combination with

molecular markers such as TMB as evidenced by Chowell *et al,* which could be considered for future model iterations [35].

In summary, we demonstrated in a large RWD clinical validation study that IPS is a generalizable multi-omic biomarker that can be widely utilized clinically as a prognosticator of ICI based regimens. Importantly, IPS-High may identify patients within subgroups (TMB-L, MSS, PD-L1 negative) who benefit from ICI beyond what is predicted by existing biomarkers. Future prospective predictive utility studies are planned for evaluating the numerous clinical applications of IPS.

Acknowledgments

We thank Dana DeSantis, Ph.D. of the Tempus Scientific Communications team for facilitating manuscript writing. Parts of this work were presented at SITC 2024.

Competing Interests

A.Z., R.E., Y.L., A.J., S.W.H, S.M., B.T., M.R, N.P., B.M., K.S., M.T-L., H.N., J.G, K.B, C.S., M.M.S., T.T., and E.C. are employees of and stockholders in Tempus AI, Inc. a for-profit company.

T.A.C. is a co-founder of Gritstone Oncology and holds equity. T.A.C. holds equity in An2H. T.A.C. acknowledges grant funding from Bristol-Myers Squibb, AstraZeneca, Illumina, Pfizer, An2H, and Eisai. T.A.C. has served as an advisor for Bristol-Myers, MedImmune, Squibb, Illumina, Tempus, Eisai, AstraZeneca, and An2H. T.A.C. holds ownership of intellectual property

on using tumor mutation burden to predict immunotherapy response, with pending patent, which has been licensed.

S.P. receives scientific advisory income from: Amgen, AstraZeneca, BeiGene, Bristol-Myers Squibb, Eli Lilly, Jazz, Genentech, Illumina , Merck, Pfizer, Zai Labs. S.P.'s university receives research funding from: Amgen, AstraZeneca, A2bio, Bristol-Myers Squibb, Eli Lilly, Fate Therapeutics, Gilead, Iovance, Merck, Pfizer, and Roche/Genentech.

M.Z. reports Institutionally-directed research funding from BMS and Exelixis; Advisory income from Pfizer, Exelixis, Janssen, Merck; and Honoraria from Adicet Bio and Arcus Bio.

D.R.A. reports consulting or scientific advisory board support from Adlai Nortye, Boehringer Ingelheim, Cue Biopharma, Exelixis, Genmab, Inhibrx, Immunitas, Sanofi, Kura Oncology, Merck, Merck KGaA, Merus, Natco Pharma, Purple Biotech, Regeneron, Seagen, and TargImmune Therapeutics; travel support from Natco Pharma; leadership role on the NCCN practice guidelines and the Barnes Jewish Hospital Pharmacy and Therapeutics committee; research support from Tempus, Pfizer, Eli Lilly, Merck, Celgene/BMS, Novartis, AstraZeneca, Blueprint Medicine, Kura Oncology, Cue Biopharma, Cofactor Genomics, Debiopharm International, Inhibrx, ISA Therapeutics, Gilead Sciences, BeiGene, Roche, Vaccinex, Hookipa Biotech, Adlai Nortye USA, Epizyme, BioAtla, Boehringer Ingelheim, Calliditas Therapeutics, Genmab, Natco Pharma, Tizona Therapeutics, Erasca, Alentix, Seagen, Coherus, Takeda, Xilio, GSK, Johnson & Johnson, and Immunotep.

Funding

All work was funded by Tempus AI, Inc. a for-profit company.

References

1 Herbst RS, Baas P, Kim D-W, *et al.* Pembrolizumab versus docetaxel for previously treated, PD-L1-positive, advanced non-small-cell lung cancer (KEYNOTE-010): a randomised controlled trial. *Lancet*. 2016;387:1540–50.

2 Zhao B, Zhao H, Zhao J. Efficacy of PD-1/PD-L1 blockade monotherapy in clinical trials. *Ther Adv Med Oncol*. 2020;12:1758835920937612.

3 Kim MS, Prasad V. Pembrolizumab for all. *J Cancer Res Clin Oncol*. 2023;149:1357–60.

4 Chen DS, Mellman I. Oncology meets immunology: the cancer-immunity cycle. *Immunity*. 2013;39:1–10.

5 Litchfield K, Reading JL, Puttick C, *et al.* Meta-analysis of tumor- and T cell-intrinsic mechanisms of sensitization to checkpoint inhibition. *Cell*. 2021;184:596–614.e14.

6 Lin HM, Wu Y, Yin Y, *et al.* Real-world ALK testing trends in patients with advanced non–small-cell lung cancer in the United States. *Clin Lung Cancer*. 2023;24:e39–49.

7 Anti-PD1 response prediction dream challenge. DREAM Challenges. https://dreamchallenges.org/anti-pd1-response-prediction-dream-challenge/ (accessed 12 September 2024)

8 Beaubier N, Bontrager M, Huether R, *et al.* Integrated genomic profiling expands clinical options for patients with cancer. *Nat Biotechnol*. 2019;37:1351–60.

9 Beaubier N, Tell R, Lau D, *et al.* Clinical validation of the tempus xT next-generation targeted oncology sequencing assay. *Oncotarget*. 2019;10:2384–96.

10 Wenric S, Davison JM, Guittar J, *et al.* Real-world data validation of the PurIST pancreatic ductal adenocarcinoma gene expression classifier and its prognostic implications. medRxiv. 2023;2023.02.23.23286356.

11 Aggarwal C, Ben-Shachar R, Gao Y, *et al.* Assessment of tumor mutational burden and outcomes in patients with diverse advanced cancers treated with immunotherapy. *JAMA Netw Open*. 2023;6:e2311181.

12 Michuda J, Breschi A, Kapilivsky J, *et al.* Validation of a Transcriptome-Based Assay for Classifying Cancers of Unknown Primary Origin. *Mol Diagn Ther*. 2023;27:499–511.

13 Leibowitz BD, Dougherty BV, Bell JSK, *et al.* Validation of genomic and transcriptomic models of homologous recombination deficiency in a real-world pan-cancer cohort. *BMC Cancer*. 2022;22:587.

14 Jain P, Stein MM, Fields P, *et al.* 163 A multi-modal, pan-cancer atlas of tumor-immune states across primary and metastatic disease using a large, real-world database. *Regular and Young Investigator Award Abstracts*. BMJ Publishing Group Ltd 2023.

15 Erbe R, Stein MM, Rand TA, *et al.* Abstract 2281: A tumor-intrinsic signature involving immunosuppression via MIF-CD74 signaling is associated with overall survival in ICT-treated lung adenocarcinoma. *Cancer Res*. 2024;84:2281–2281.

16 Bareche Y, Kelly D, Abbas-Aghababazadeh F, *et al.* Leveraging big data of immune checkpoint blockade response identifies novel potential targets. *Ann Oncol*. 2022;33:1304–17.

17 Tsai W-Y, Jewell NP, Wang M-C. A note on the product-limit estimator under right censoring and left truncation. *Biometrika*. 1987;74:883.

18 Chan TA, Yarchoan M, Jaffee E, *et al.* Development of tumor mutation burden as an immunotherapy biomarker: utility for the oncology clinic. *Ann Oncol*. 2019;30:44–56.

19 Wang J, Li X, Xiao G, *et al.* CD74 is associated with inflamed tumor immune microenvironment and predicts responsiveness to PD-1/CTLA-4 bispecific antibody in patients with solid tumors. *Cancer Immunol Immunother*. 2024;73:36.

20 Zhou X, Wang Y, Zheng J, *et al.* Integrative study reveals the prognostic and immunotherapeutic value of CD274 and PDCD1LG2 in pan-cancer. *Front Genet*. 2022;13:990301.

21 Zhou W-T, Jin W-L. B7-H3/CD276: An emerging cancer immunotherapy. *Front Immunol*. 2021;12:701006.

22 Bill R, Wirapati P, Messemaker M, *et al.* CXCL9:SPP1 macrophage polarity identifies a network of cellular programs that control human cancers. *Science*. 2023;381:515–24.

23 Tang K, Wu Y-H, Song Y, *et al.* Indoleamine 2,3-dioxygenase 1 (IDO1) inhibitors in clinical trials for cancer immunotherapy. *J Hematol Oncol*. 2021;14:68.

24 Beatty GL, Li Y, Long KB. Cancer immunotherapy: activating innate and adaptive immunity through CD40 agonists. *Expert Rev Anticancer Ther*. 2017;17:175–86.

25 Cristescu R, Nebozhyn M, Zhang C, *et al.* Transcriptomic determinants of response to pembrolizumab monotherapy across solid tumor types. *Clin Cancer Res*. 2022;28:1680–9.

26 Upadhaya S, Hubbard-Lucey VM, Yu JX. Immuno-oncology drug development forges on despite COVID-19. *Nat Rev Drug Discov*. 2020;19:751–2.

27 Merck & Co., Inc. KEYTRUDA (pembrolizumab) [package insert]. U.S. Food and Drug Administration website .

28 Aguilar EJ, Ricciuti B, Gainor JF, *et al.* Outcomes to first-line pembrolizumab in patients with non-small-cell lung cancer and very high PD-L1 expression. *Ann Oncol*. 2019;30:1653–9.

29 Hossein Borghaei, ECOG-ACRIN Cancer Research Group. Testing the Timing of Pembrolizumab Alone or With Chemotherapy as First Line Treatment and Maintenance in Non-small Cell Lung Cancer. ClinicalTrials.gov. https://clinicaltrials.gov/study/NCT03793179 (accessed 24 September 2024)

30 Sahin IH, Ciombor KK, Diaz LA, *et al.* Immunotherapy for microsatellite stable colorectal cancers: Challenges and novel therapeutic avenues. *Am Soc Clin Oncol Educ Book*. 2022;42:1–12.

31 Marabelle A, Le DT, Ascierto PA, *et al.* Efficacy of pembrolizumab in patients with noncolorectal high microsatellite instability/mismatch repair-deficient cancer: Results from the phase II KEYNOTE-158 study. *J Clin Oncol*. 2020;38:1–10.

32 Reardon DA, Brandes AA, Omuro A, *et al.* Effect of nivolumab vs bevacizumab in patients with recurrent glioblastoma: The CheckMate 143 phase 3 randomized clinical trial. *JAMA Oncol*. 2020;6:1003–10.

33 Marabelle A, Fakih M, Lopez J, *et al.* Association of tumour mutational burden with outcomes in patients with advanced solid tumours treated with pembrolizumab: prospective biomarker analysis of the multicohort, open-label, phase 2 KEYNOTE-158 study. *Lancet Oncol*. 2020;21:1353–65.

34 Fridman WH, Zitvogel L, Sautès-Fridman C, *et al.* The immune contexture in cancer prognosis and treatment. *Nat Rev Clin Oncol*. 2017;14:717–34.

35 Chowell D, Yoo S-K, Valero C, *et al.* Improved prediction of immune checkpoint blockade efficacy across multiple cancer types. *Nat Biotechnol*. 2022;40:499–506.

36 Grossman RL, Heath AP, Ferretti V, *et al.* Toward a shared vision for cancer genomic data. *N Engl J Med*. 2016;375:1109–12.

37 Cerami E, Gao J, Dogrusoz U, *et al.* The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov*. 2012;2:401–4.

Figures

Figure 1. Cohort funnel

Figure 1

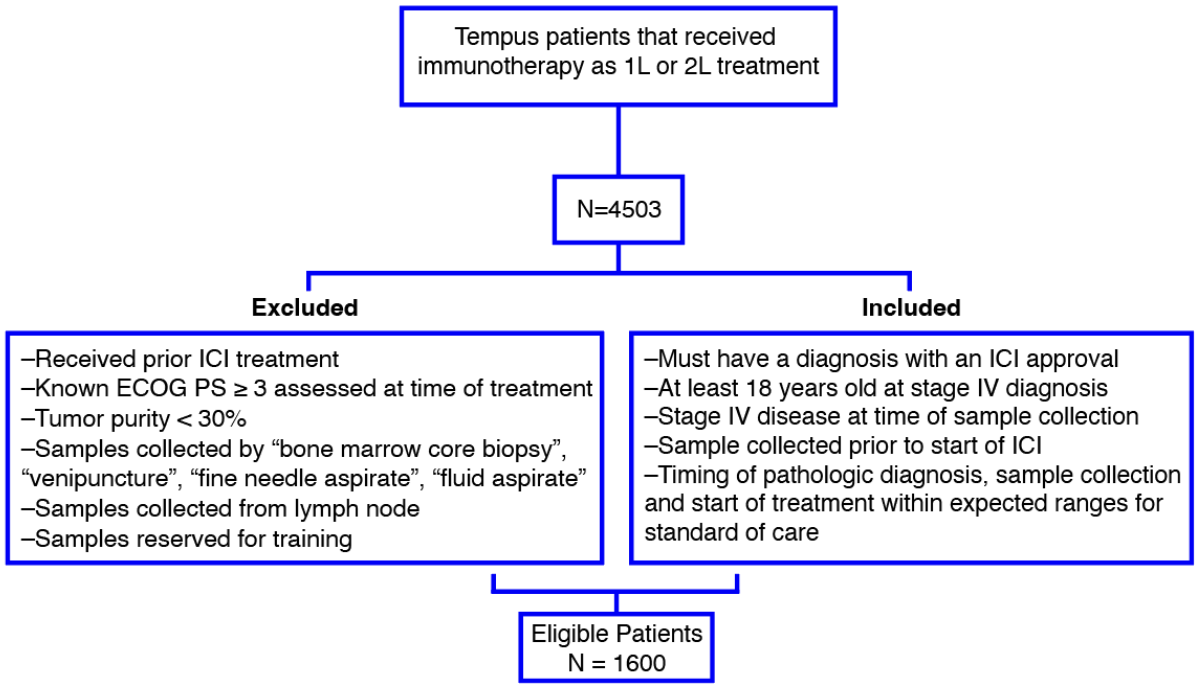


Figure 1. Inclusion/exclusion criteria for the study cohort.

Figure 2. IPS Model Features

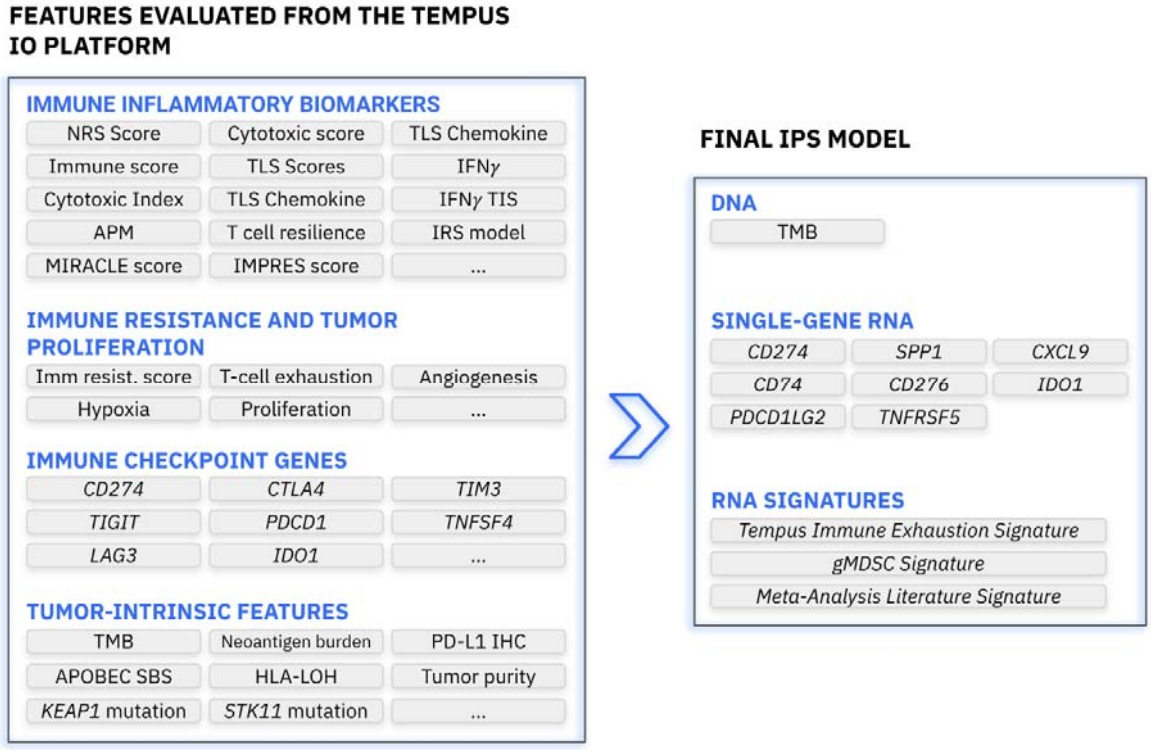
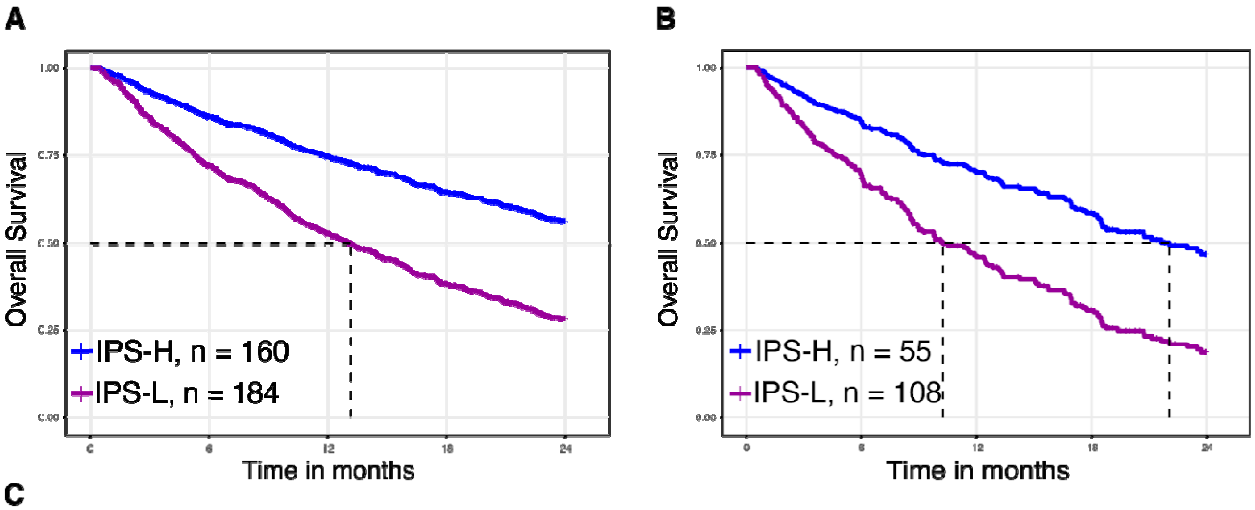


Figure 2. The Tempus IO Platform was leveraged for developing the IPS model. Various machine learning (ML) techniques were implemented to reduce the feature space. The final IPS model includes 11 RNA-based features and TMB.

Figure 3. OS is significantly higher in IPS-H vs. IPS-L

Figure 3



C

	1L mono	2L mono	1L combo	2L combo
IPS-H	>24 (>24, >24)	22.1 (18.3, >24)	>24 (23.0, >24)	20.75 (16.95, >24)
IPS-L	13.1 (10.8, 15.6)	10.2 (8.56, 13.1)	12.2 (10.8, 13.8)	9.77 (8.23, 12.6)

Figure 3. The HR was 0.45 (0.40, 0.52), $p < 0.01$. Predicted OS from the CoxPH model for a) 1L monotherapy and b) 2L monotherapy patients. Predicted survival for 1L and 2L combination therapy patients are similar to above. c) The median OS and 95% confidence interval for IPS-H and IPS-L groups for each line of therapy/treatment group combination.

Figure 4.

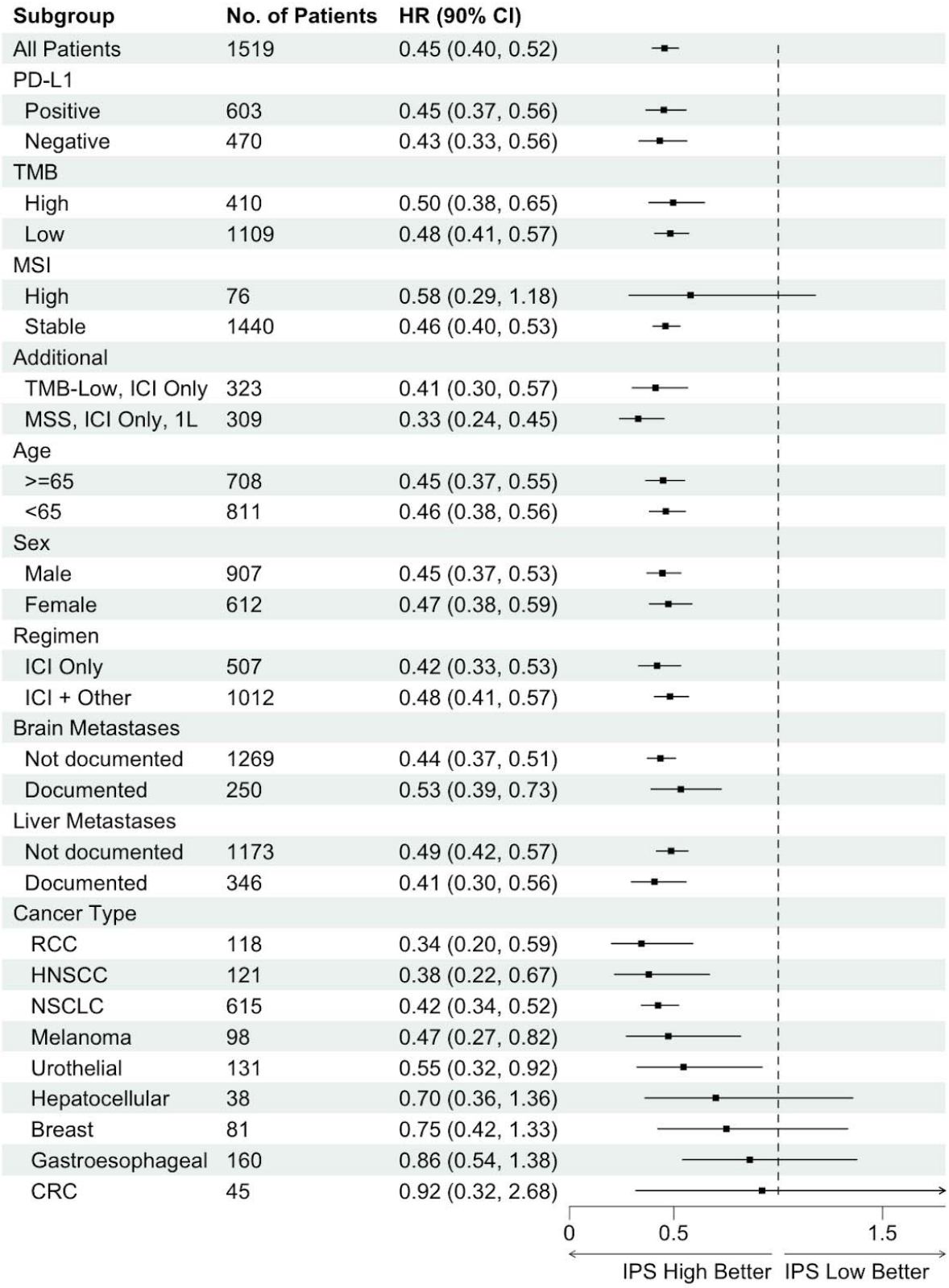
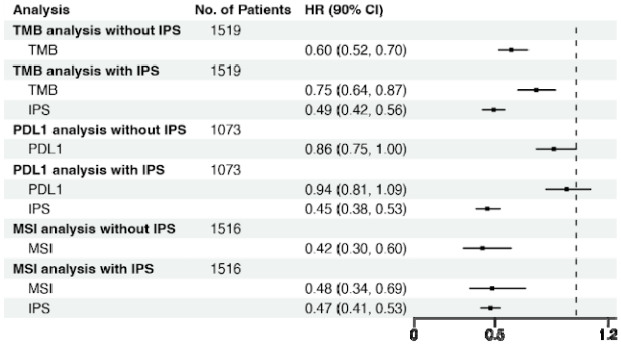


Figure 4. Forest plot showing IPS-H vs. IPS-L hazard ratios and confidence intervals across demographics and clinically relevant subgroups. Subgroups may have <1519 patients due to availability of data.

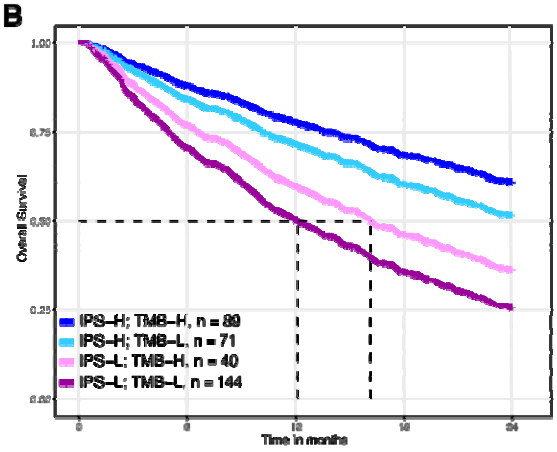
Figure 5.

Figure 5

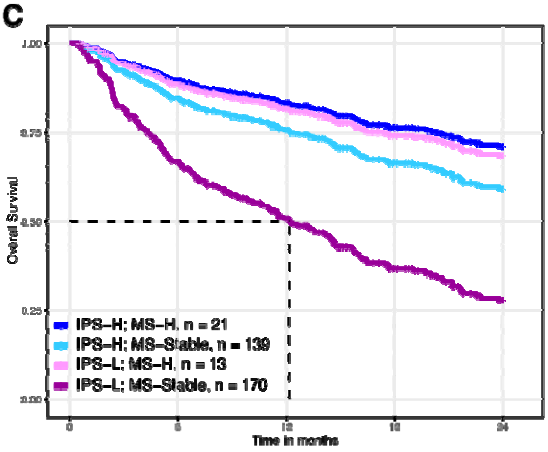
A



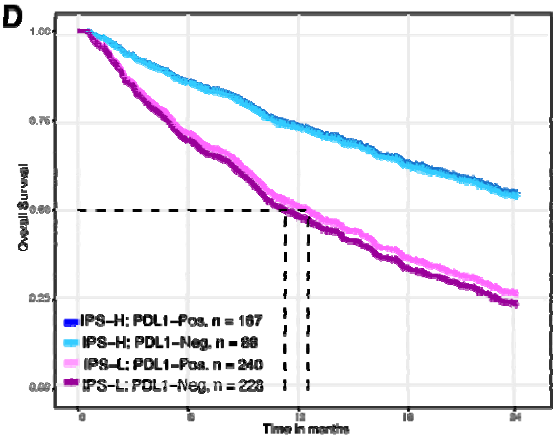
B



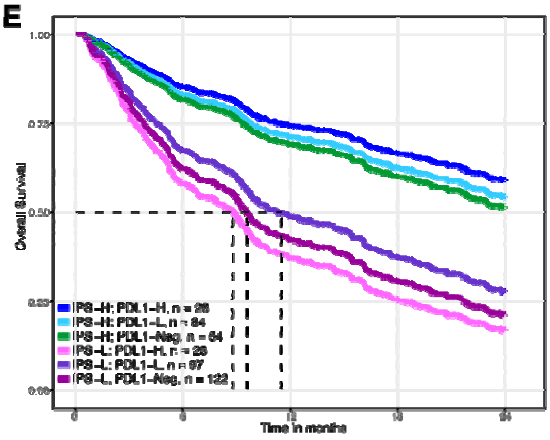
C



D



E



F

Comparisons	HR
B. IPS-H/TMB-L to IPS-L/TMB-L	0.49 (0.41, 0.58)
C. IPS-H/MSS to IPS-L/MSS	0.41 (0.32, 0.53)
D. IPS-H/PDL1- to IPS-L/PDL1-	0.43 (0.33, 0.56)
E. IPS-H/PDL1- to IPS-L/PDL1- (within NSCLC)	0.43 (0.30, 0.63)

Figure 5. A. Forest plot showing univariate (UV) HRs for TMB, PD-L1, MSI and multivariate (MV) HRs that include IPS. A likelihood ratio test between the UV and MV models was significant ($p<0.01$) for all three biomarkers, indicating that IPS has significant prognostic utility beyond TMB, MSI, and PD-L1. Plots b-e show predicted OS from a model stratified by line of therapy and fit on IPS, treatment group, and the MV model with the listed biomarker: **B.** TMB pan-cancer, **C.** MSI pan-cancer, **D.** PD-L1 pan-cancer and **E.** PD-L1 in NSCLC patients. The predicted OS curves represent patients treated with monotherapy in 1L for TMB and MSI (**B-C**), and combination therapy in 1L for PD-L1 and NSCLC (**D-E**). **F.** HR and 90% CI for the most relevant curves shown in the predicted OS plots in (**B-E**).

Figure 6. Predictive utility for IPS

Figure 6

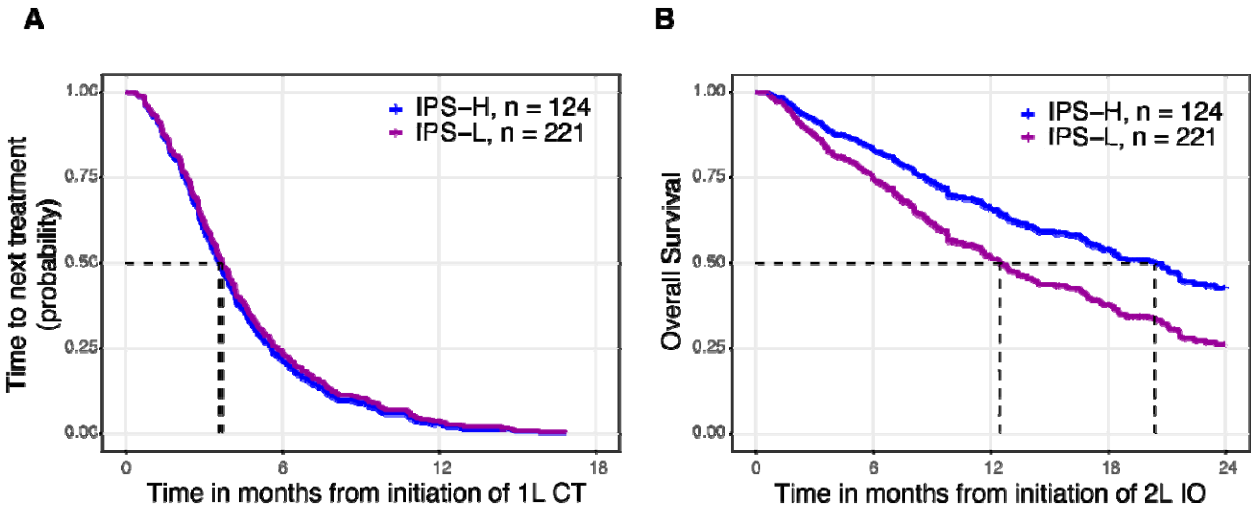


Figure 6. A. Predicted TTNT for 1L CT with no significant effect for IPS (HR = 1.06 (0.85, 1.33)). **B.** Predicted OS for 2L ICI shows that IPS does have a significant effect (HR = 0.63 (0.46, 0.86)). Interaction test $p < 0.01$, indicating that the HR in 2L ICI is significantly different from HR in 1L CT.

Table 1. Cohort Characteristics

Characteristics	Overall N=1600	IPS-High N=576	IPS-Low N=943	Indeterminate N=81
Age				
Mean (SD)	64.6 (11.8)	64.9 (12.1)	64.5 (11.6)	64.2 (11.7)
Median	65.0	65.0	65.0	66.0
[IQR]	[58.0, 73.0]	[59.0, 74.0]	[58.0, 73.0]	[56.0, 74.0]
Min / Max	20.0 / 89.0	27.0 / 89.0	20.0 / 88.0	34.0 / 83.0
Sex				
Female	645 (40%)	252 (44%)	360 (38%)	33 (41%)
Male	955 (60%)	324 (56%)	583 (62%)	48 (59%)
Line of therapy				
1	1,326 (83%)	482 (84%)	774 (82%)	70 (86%)
2	274 (17%)	94 (16%)	169 (18%)	11 (14%)
Treatment regimen				
IO Only	534 (33%)	215 (37%)	292 (31%)	27 (33%)
IO mono	381 (24%)	146 (25%)	219 (23%)	16 (20%)
IO doublet	153 (9.6%)	69 (12%)	73 (7.7%)	11 (14%)
IO + Other	1,066 (67%)	361 (63%)	651 (69%)	54 (67%)
IO+Chemo	869 (54%)	283 (49%)	542 (57%)	44 (54%)
IO+Chemo+Other	72 (4.5%)	15 (2.6%)	55 (5.8%)	2 (2.5%)
IO+Other	125 (7.8%)	63 (11%)	54 (5.7%)	8 (9.9%)
ECOG				
0	334 (21%)	137 (24%)	186 (20%)	11 (14%)
1	473 (30%)	168 (29%)	279 (30%)	26 (32%)
2	140 (9%)	45 (8%)	93 (10%)	2 (2%)
Unknown/Missing	653 (40%)	226 (39%)	385 (40%)	42 (52%)
Stage at primary dx				
Stage I	47 (3%)	22 (4%)	23 (2%)	2 (2%)
Stage II	68 (4%)	25 (4%)	41 (4%)	2 (2%)
Stage III	94 (6%)	33 (6%)	58 (6%)	3 (4%)
Stage IV	1,219 (76%)	430 (75%)	721 (76%)	68 (84%)
Unknown/Missing	172 (11%)	66 (11%)	100 (11%)	6 (7%)
Cancer type*				
Breast	86 (5%)	40 (47%)	41 (48%)	5 (6%)
CRC	46 (3%)	27 (59%)	18 (39%)	1 (2%)

????Gastroesophageal	171 (11%)	26 (15%)	134 (78%)	11(6%)
????Hepatocellular	40 (2%)	16 (40%)	22 (55%)	2 (5%)
????HNSCC	125 (8%)	35 (28%)	86 (69%)	4 (3%)
????Melanoma	102 (6%)	56 (55%)	42 (41%)	4 (4%)
????NSCLC	647 (40%)	248 (38%)	367 (57%)	32 (5%)
????RCC	131 (8%)	69 (53%)	49 (37%)	13 (10%)
????Urothelial	137 (9%)	36 (26%)	95 (69%)	6 (4%)
????Other	115 (7%)	23 (20%)	89 (77%)	3 (3%)
Brain metastases documented	265 (17%)	107 (19%)	143 (15%)	15 (19%)
Liver metastases documented	362 (23%)	94 (16%)	252 (27%)	16 (20%)
PD-L1 by IHC				
? ???Negative	495 (31%)	149 (26%)	321 (34%)	25 (31%)
?? ??Positive	637 (40%)	250 (43%)	353 (37%)	34 (42%)
? ???Unknown/Missing	468 (29%)	177 (31%)	269 (29%)	22 (27%)
TMB				
??? ?High	430 (27%)	250 (43%)	160 (17%)	20 (25%)
??? ?Low	1,170 (73%)	326 (57%)	783 (83%)	61 (75%)
MSI				
High	80 (5.0%)	45 (7.8%)	31 (3.3%)	4 (4.9%)
Stable	1517 (95%)	531 (92%)	909 (96%)	77 (95%)
Undetermined	3 (0.2%)	0 (0%)	3 (0.3%)	0 (0%)

*Overall category sums down column; IPS-H, IPS-L, and Indeterminate sum across rows