Prediction of pan-solid tumor pembrolizumab benefit by integrating tumor mutation and gene expression profiling

Nickolay Khazanov

Strata Oncology

Melissa Shreve

Strata Oncology   https://orcid.org/0000-0002-5983-4216

Laura Lamb

Strata Oncology

Daniel Hovelson

Strata Oncology

Marc Matrana

Ochsner Cancer Institute

Mark Burkard

University of Wisconsin–Madison

Eddy Yang

University of Alabama at Birmingham

William Edenfield

Prisma Health Greenville Memorial Hospital

Claire Dees

Lineberger Comprehensive Cancer Center

Adedayo Onitilo

ancer Care and Research Center, Marshfield Clinic Research Institute

Michael Thompson

Aurora Cancer Care, Advocate Aurora Health   https://orcid.org/0000-0001-7930-4610

Gary Buchschacher

Kaiser Permanente Southern California

Alan Miller

SCL Health-CO

Alexander Menter

Kaiser Permanente Colorado

Benjamin Parsons

Gundersen Health System

Timothy Wassenaar

Waukesha Memorial Hospital

Leon Hwang

Kaiser Permanente of the Mid-Atlantic States

J. Suga

Kaiser Permanente - Northern California

Robert Siegel

Bon Secours St. Franci

William Irvin, Jr.

Bon Secours St. Franci

Suresh Nair

Lehigh Valley Health Network

Jennifer Slim

MultiCare Regional Cancer Center

Kat Kwiatkowski

Strata Oncology

Khalis Mitchell

Strata Oncology

Tina Hu-Seliger

Strata Oncology

Stephanie Drewery

Strata Oncology

Andrew Fischer

Strata Oncology

Jennifer Hipp

Strata Oncology

Travis Reeder

Strata Oncology

Hana Vakil

Strata Oncology

Bryan Johnson

Strata Oncology

Daniel Rhodes

Strata Oncology

Scott Tomlins (✉ scott.tomlins@strataoncology.com)

Strata Oncology   https://orcid.org/0000-0001-8661-9821

Article

Prediction of pan-solid tumor pembrolizumab benefit by integrating tumor mutation burden and gene expression profiling

Nickolay Khazanov[1], Melissa J. Shreve[1], Laura E. Lamb[1], Daniel H. Hovelson[1], Marc R. Matrana [2], Mark E. Burkard[3], Eddy Shih-Hsin Yang[4], William Jeffery Edenfield[5], E. Claire Dees[6], Adedayo A. Onitilo[7], Michael Thompson[8], Gary L. Buchschacher[9], Alan M. Miller[10], Alexander Menter[11], Benjamin Parsons[12], Timothy Wassenaar[13], Leon C. Hwang[14], J. Marie Suga[15], Robert Siegel[16], William Irvin, Jr.[16], Suresh Nair[17], Jennifer N. Slim[18], Kat Kwiatkowski[1], Khalis Mitchell[1], Tina Hu-Seliger[1], Stephanie Drewery[1], Andrew Fischer[1], Jennifer Hipp[1], Travis Reeder[1], Hana Vakil[1], D. Bryan Johnson[1], Daniel R. Rhodes[1*], Scott A. Tomlins[1*]

[1] Strata Oncology, Ann Arbor, MI; [2] Ochsner Cancer Institute, New Orleans, LA; [3] University of Wisconsin Carbone Cancer Center, Madison, WI; [4] O'Neal Comprehensive Cancer Center, University of Alabama at Birmingham School of Medicine, Birmingham, AL; [5] Prisma Health Greenville Memorial Hospital, Greenville, SC; [6] UNC Lineberger Comprehensive Cancer Center, Chapel Hill NC; [7] Cancer Care and Research Center, Marshfield Clinic Research Institute, Marshfield, WI; [8] Aurora Cancer Care, Advocate Aurora Health, Milwaukee WI; [9] Kaiser Permanente Southern California, Los Angeles, CA; [10] SCL Health-CO, Broomfield, CO; [11] Kaiser Permanente Colorado, Lone Tree, CO; [12] Gundersen Health System, La Crosse, WI; [13] UW Health Cancer Center at ProHealth Care, Waukesha, WI; [14] Kaiser Permanente of the Mid-Atlantic States, Rockville, MD; [15] Kaiser Permanente - Northern California, Oakland, CA; [16] Bon Secours St. Francis, Midlothian, VA; [17] Lehigh Valley Health Network, Allentown, PA; [18] MultiCare Regional Cancer Center - Tacoma, WA

Corresponding Author:
*Daniel R. Rhodes Ph.D.
Strata Oncology
8192 Jackson Road, Suite A
Ann Arbor, MI 48103 USA
Phone: 734-527-0992
email: daniel.rhodes@strataoncology.com

*Scott A. Tomlins, M.D., Ph.D.
Strata Oncology
8192 Jackson Road, Suite A
Ann Arbor, MI 48103 USA
Phone: 734-527-0992
email: scott.tomlins@strataoncology.com

Manuscript word count: 6,816

Running Title: Immune Response Score predicts pembrolizumab benefit

Keywords: immune response, oncology, pembrolizumab, biomarkers, pan tumor, tumor mutation burden, gene expression signature

Abstract

Pembrolizumab is approved in many advanced solid tumor types, however predictive biomarkers and the proportion of pembrolizumab-benefiting patients vary. Biomarkers beyond PD-L1 immunohistochemistry, microsatellite instability (MSI) status, and tumor mutation burden (TMB) may improve benefit prediction. Here, leveraging treatment data (time to next treatment [TTNT]) and comprehensive genomic and quantitative transcriptomic profiling on routine tumor tissue from 708 patients (24 tumor types) collected in an ongoing observational trial (NCT03061305), we report a multivariate, integrative predictor of pan-solid tumor pembrolizumab benefit. The Immune Response Score (IRS) model, which includes TMB and quantitative *PD-1*, *PD-L2*, *ADAM12* and *CD4* RNA expression, was confirmed as predictive through comparison of pembrolizumab TTNT with previous chemotherapy TTNT in a subset of 166 patients treated with both. Applying IRS to the entire NCT03061305 cohort (n=25,770 patients), 13.2-30.7% of patients (2.2-9.6% of patients outside of pembrolizumab approved tumor types [including TMB-High and MSI-High]) are predicted to benefit substantially from pembrolizumab. Hence, if prospectively validated, the IRS model may improve pembrolizumab benefit prediction across approved tumor types including patients outside of currently approved indications.

INTRODUCTION

Checkpoint inhibitors (CPIs) have transformed cancer care with anti-PD-1 and anti-PD-L1 monoclonal antibodies approved for use in multiple tumor types and pan tumor indications (microsatellite instability high/mismatch repair deficient [MSI-H/dMMR] and tumor mutation burden [TMB] ≥ 10 mutations/megabase [Muts/Mb])[1-3]. Improved biomarkers capable of predicting CPI benefit have the potential to expand CPIs to additional patient populations outside of currently approved indications, and to focus their application more effectively on likely responsive patients when alternative therapies exist. PD-L1 immunohistochemistry (IHC) is required for treatment in many tumor types and serves as a companion diagnostic biomarker; although, antibodies, staining platforms, PD-L1 expressing cells included in scoring algorithms, and cutoffs vary across tumor types[4-14]. In addition, high TMB predicts CPI response across multiple tumor types, although TMB determination approaches vary across studies and tests, and only a fraction of TMB high (TMB-H) patients benefit[15-24]. For example, in the KEYNOTE-158 study of 9 tumor types leading to pan-solid tumor approval of second-line pembrolizumab in patients with TMB ≥ 10 Muts/Mb by the FoundationOne companion diagnostic (CDx) device, objective responses were observed in 37%, 13%, and 6% of patients with TMB ≥13 Muts/Mb, ≥10 and <13 Muts/Mb, and <10 Muts/Mb, respectively[25,26].

Numerous translational studies have demonstrated that PD-L1 expression, TMB (with clonal TMB showing increased predictive ability vs. TMB methods including all somatic mutations), and other immune related gene expression markers focusing on the tumor microenvironment (TME) are independent predictors of response [15,27-39]; however, a single, integrative, clinically applicable and validated test for treatment selection across solid tumors is lacking. Herein, leveraging pembrolizumab real-world data (RWD) for treatment and comprehensive genomic and quantitative transcriptomic profiling (CGqTP) data from the Strata Trial (NCT03061305)—an observational clinical trial evaluating the impact of molecular profiling on patients with advanced solid tumors — we report the development of an integrated clinical Immune Response Score (IRS) that predicts pembrolizumab response across solid

26   tumors from small, real-world, formalin-fixed paraffin-embedded (FFPE) tumor tissue specimens.

27   Although prospective clinical validation is required, these results demonstrate that integrated CGqTP may

28   be able to increase the clinical benefit of the CPI pembrolizumab in patients with advanced solid tumors.

29

METHODS

Cohort

The Strata Clinical Molecular Database (SCMD) contains deidentified subject, molecular profiling, treatment, and survival data captured from the Strata Trial (NCT03061305), a 500,000-patient observational study for patients with advanced solid tumors. The Strata Trial has been reviewed and approved by Advarra Institutional Review Board (IRB; IRB Pro00019183) prior to study start. At enrolling health care systems, all adult patients with unresectable or metastatic solid tumors and available FFPE tumor tissue were eligible. Although the protocol allowed enrollment of patients with rare early-stage tumors, some analyses herein were restricted to patients with advanced (clinical stage III or IV) disease as indicated at the time of enrollment, or unstaged tumor types. Prior antineoplastic therapy, including start and stop dates, were collected for trial participants at the time of study entry. Antineoplastic therapy data and survival status were prospectively collected for 3 years from the time of enrollment and/or informed consent. Post-hoc power analysis was not performed to determine the sample size. A case series analysis was performed herein focusing on the development of an integrative CGqTP based pembrolizumab predictor, an exploratory aim of the trial. Patients in the SCMD tested by a version of StrataNGS assessing TMB (see **Biomarker Data** below) with parallel gene expression testing data from 25 January 2017 to 04 May 2021 were potentially eligible for analysis using a data cutoff of 19 May 2021.

To generate an integrative CGqTP based pembrolizumab predictor, patients in the SCMD with valid StrataNGS derived TMB and gene expression data (including meeting the minimum 20% tumor content requirement) and had greater than 1 month on pembrolizumab were identified as eligible. Real-world time to next treatment (TTNT) was defined as the time in months from the initiation of a therapy to the date of commencement of the next line of therapy (or date of death). Patients without an event (i.e., new therapy start or death) were censored at their last date of medical history record update. Patients treated with either pembrolizumab monotherapy or combination pembrolizumab plus chemotherapy were included.

55   Samples collected after the start date of pembrolizumab were excluded. Source data verification in the

56   Strata Trial was performed for high-risk data fields such as demographics and treatment history per an

57   approved Trial Monitoring Plan. Data completeness, consistency, and quality assurance checks were

58   performed across the Strata electronic data capture (EDC) system per an approved Data Management

59   Plan. Additional details on the Strata Trial experience and Strata molecular profiling have been

60   described[40-42].

61   **Biomarker Data**

62   Multiplex PCR-based comprehensive genomic profiling (PCR-CGP), including TMB assessment, was

63   performed on FFPE solid tumor tissue using StrataNGS (Strata Oncology, Ann Arbor, MI). The current

64   version of StrataNGS is a 437 gene laboratory-developed test (LDT) for FFPE tumor tissue samples

65   performed on co-isolated DNA and RNA, which has been validated on over 1,900 FFPE tumor samples,

66   and is covered for Medicare beneficiaries[41]. While earlier StrataNGS versions were also used during the

67   study period, all had similar performance for the TMB assessment (and MSI) used herein[42]. In parallel,

68   immune gene expression was determined by analytically and clinically validated multiplex PCR-based

69   quantitative transcriptomic profiling via an investigational test performed on the same co-isolated RNA as

70   described[40]; different versions of this quantitative transcriptomic profiling test have been run in parallel

71   with StrataNGS (assessing 26, 46 and currently 103 expression targets), however only quantification of

72   RNA from the 46 target version was used herein. One or more exon-spanning PCR amplicons were

73   selected for each target gene and multiple housekeeping genes were included, with 3 pan-cancer stable

74   housekeeping genes used for clinical testing. Multiplex RNAseq was performed using Ampliseq after

75   reverse transcription followed by Ion Torrent-based next-generation sequencing. Expression target

76   transcripts were measured in normalized reads per million, whereby raw expression target read counts

77   were normalized by a factor that results in the median housekeeping gene expression value matching the

78   same gene's standard reads per million in a reference FFPE normal cell line sample (GM24149) run in

79 parallel with all clinically tested samples.[40] Formal analytical validation of the multiplex PCR-based

80 quantitative transcriptomic panel and the integrated StrataIO model will be described separately.

81 **Data Analysis**

82 TTNT across groups and treatments were visualized using the Kaplan Meier method with the log-rank

83 test used to test TTNT curve differences. Correlation between TTNT and overall survival (OS) was

84 calculated using Spearman's p among patients with both a documented death event and at least two lines

85 of therapy. Throughout this study, TMB-H was defined as $\geq$10 Muts/Mb by StrataNGS, given the

86 previous validation of TMB by StrataNGS and high concordance with TMB estimates from

87 FoundationOne tissue testing[41]. All analyses were performed in python.

88 **Strata Clinical Molecular Database (SCMD) Validity Analysis**

89 Analyses to assess the clinical validity of the SCMD included an analysis of TTNT in first line, stage III

90 and IV non-small cell lung cancer (NSCLC) adenocarcinoma stratified by the presence and absence of

91 standard- of-care (SOC), actionable alterations in *EGFR* (excluding the recent SOC actionable exon 20

92 insertions), *ALK,* and *ROS1*, as well as general TTNT in non-pembrolizumab treated patients receiving at

93 least two lines of antineoplastic therapy stratified by therapy class. (e.g., chemotherapy + chemotherapy,

94 hormonal therapy + chemotherapy, small molecule inhibitors + chemotherapy). For analysis of NSCLC

95 adenocarcinoma SOC alterations, the presence of both the genomic alteration and treatment history with

96 one of the FDA-approved targeted therapy for the alteration was considered as SOC treatment.

97 **Real-World Progression Free Survival (rwPFS) as Measured by Time To Next Treatment**

98 **(TTNT)**

99 Real-world progression free survival was measured by time to next treatment (TTNT) for patients within

100 the SCMD database. Patient treatment history was standardized as described below, to ensure TTNT

101 calculations were performed appropriately for each treatment type. Medications were classified into

102 antineoplastic or non-antineoplastic treatments, and chemotherapy medications were defined as a subset

103   of the antineoplastic treatments. Non-antineoplastic treatments were excluded from TTNT calculations;

104   patient treatment records with invalid or non-informative dates were also excluded (e.g., no start date or

105   start date in the future). Consecutive single-dose treatments were combined into a course of treatment

106   with a single start and end date. Since chemotherapy medications are frequently administered together,

107   any chemotherapy treatment(s) with a temporal overlap of 80% or more were merged into a single

108   treatment record. Lines of therapy were defined when a different therapy was started at, or after, the end

109   of another therapy (with the former therapy not being administered at any time after the latter). For

110   example, multiple starts and stops of the same therapy with a different therapy in between did not

111   delineate different lines of treatment.

112   To determine TTNT, an effective end date was defined for each course of treatment as either a) date of

113   record if treatment is ongoing, b) date of death if patient died while on treatment, or c) the latest available

114   end date. Furthermore, a likely progression event for the end of treatment was identified if either a) the

115   patient died during treatment or b) the patient started another antineoplastic treatment at the end of the

116   current treatment. TTNT was calculated as the difference, in months, between the start date and effective

117   end date of the treatment. Only records with 1 month or more anti-neoplastic TTNT were retained.

118   For analysis of pembrolizumab monotherapy vs combination therapy, a pembrolizumab course of

119   treatment was classified as combination therapy if there was 10% or more temporal overlap between the

120   pembrolizumab treatment and chemotherapy treatment(s).

121   **Immune Response Score (IRS) Model Development**

122   The association of TMB and 21 candidate immune and proliferation gene expression biomarkers with

123   pembrolizumab TTNT was determined using standard Cox proportional hazards regression. TMB

124   measurements were $\log_2$-transformed and gene expression measurements were $\log_2$-transformed and

125   median-centered per laboratory workflow (two PCR cycling conditions were used with the 46 gene

126   expression test) prior to analysis. For gene expression biomarkers with two independent expression

127   amplicons on the 46 gene RNA test version (*PD-1* and *PD-L1*), results were averaged prior to inclusion in

128   the analysis (e.g. *PD-1* composite). For multivariate model building, backward stepwise regression was

129   used, first including all variables in the model, then selectively removing the least significant variables so

130   long as the overall model significance improved. This approach was performed several times on a sub-

131   sampling and showed that most significant models [as measured by Akaike information criterion (AIC)]

132   contained no more than half-dozen factors. We therefore performed a brute-force search of all

133   combinations of expression targets to find the most significant factors. To minimize the risk of

134   overfitting, a final 5-factor model was selected. Individual patient IRS were derived from the Cox model

135   as:

136    $IRS = 4.03 * \exp(0.29 * TMB + 0.15 * PD\text{-}1 + 0.14 * PD\text{-}L2 - 0.14 * CD4 - 0.07 * ADAM12)$

137   where an IRS of 10 is equal to the median hazard rate observed in the dataset, values greater than 10

138   represent decreased hazard (i.e., more benefit from pembrolizumab) and values less than 10 represent

139   increased hazard (i.e., less benefit from pembrolizumab). We assigned patients to one of three IRS groups

140   to compare patient outcomes (i.e., Low (L) < 8.5, 8.5 ≤ Intermediate (I) < 11.9, and High (H) ≥ 11.9)

141   through dividing the dataset into 8 equal IRS bins and combining bins based on overlapping TTNT

142   curves. Cox proportional hazards models were utilized to examine the interaction between

143   pembrolizumab vs. prior chemotherapy TTNT within the same patient and IRS as a continuous variable.

144   The likelihood ratio test for interaction compared the reduced model, which excluded the IRS by

145   treatment interaction, with the competing full model, which included the IRS by treatment interaction.

146

147

148 **RESULTS**

149 **Clinical Molecular Data**

150 The Strata Trial (NCT03061305) is an observational clinical trial evaluating the impact of tumor

151 molecular profiling for patients with advanced solid tumors. De-identified demographic, clinical and

152 molecular data from patients in the Strata Trial is maintained in the Strata Clinical Molecular Database

153 (SCMD). With a data-cutoff of 19 May 2021, the SCMD contains clinical and molecular data from

154 39,252 unique patients with stage III or IV solid tumors (from 28 tumor types) enrolled from 25 United

155 States health care systems who had routine FFPE tumor tissue molecularly profiled by the StrataNGS

156 CGP test[41,42] with 7,978 patients (from 28 tumor types) having treatment data from at least one

157 antineoplastic agent.

158 For all SCMD patients, antineoplastic treatment start and stop dates (for all prior therapies and up to 3

159 years after Strata trial enrollment) were obtained from automated electronic health record queries or

160 manual entry; data was updated regularly by submitting institutions, and date of death was obtained

161 similarly. Real-world TTNT was determined directly from treatment start and stop dates for each line of

162 therapy. Among the 7,978 patients, the median follow-up from start of first treatment and Strata trial

163 enrollment was 11 months [interquartile range (IQR) 4-24 months] and 5 months (IQR 1-11 months),

164 respectively. The median number of total therapies and lines of therapy per patient was 2 (both IQR 1-3),

165 with a median of 1 total therapy and 1 line of therapy (both IQR 1-2) after Strata trial enrollment. As

166 expected, in patients who had at least two lines of therapy, median TTNT was shorter with each

167 subsequent line of therapy (**Figure S1**).

168 Given the substantial proportion of patients in the SCMD with NSCLC and extensive previous

169 characterization of molecular subtypes and associated therapies, we leveraged the NSCLC cohort to

170 assess the validity of using the SCMD to support this study. Of the 7,978 total patients, 1,173 (14.7%) had

171 NSCLC, with a median age at enrollment of 65 years (IQR 60-73), 51.8% were women, 56.4% were

172   white, and 8.0% had NSCLC squamous cell carcinoma, similar to data reported for patients in the initial

173   report from the Flatiron/Foundation Medicine clinical molecular database [17]. At enrollment, of the 1,173

174   NSCLC patients in the SCMD, 24.8% and 75.2% had stage III and IV cancer, respectively. Of note, in

175   139 patients with NSCLC adenocarcinoma harboring standard of care (SOC) alterations in *EGFR*, *ALK* or

176   *ROS1*, 84% received at least one matched targeted therapy, while only 68 of 935 (7.2%) patients without

177   SOC *EGFR*, *ALK* or *ROS1* alterations received one or more of these targeted therapies. These treatment

178   results contrast with the initial report of the Flatiron/Foundation Medicine clinical molecular database,

179   where only 480 of 737 (65%) patients with NCCN-driver alterations in *EGFR* and *ALK* rearrangements

180   received targeted therapy after advanced NSCLC diagnosis, while 26% of EGFR inhibitor treatment was

181   in patients without an *EGFR* alteration[17]. In part this may be due to the contemporary nature of our series,

182   as for example, 57 of 81 (70%) patients with *EGFR* SOC alterations in the SCMD treated with EGFR

183   TKI received osimertinib (or an osimertinib containing combination regimen) as their first line of EGFR

184   TKI therapy, and 26 of 28 (93%) patients with *ALK* SOC alterations treated with ALK TKIs received

185   alectinib or brigatinib as their first line of ALK TKI therapy. Additional details and NSCLC analyses

186   supporting the validity of TTNT and the SCMD are shown in **Figure S2**.

187   **Biomarkers of Pembrolizumab Benefit Analysis**

188   We have previously demonstrated that molecular alteration frequency in the first ~30,000 patients

189   enrolled in the Strata Trial[42] was similar to that observed in the Memorial Sloan Kettering single

190   institution pan-cancer profiling effort, MSK-IMPACT[43], supporting the generalizability of the SCMD.

191   Herein, to assess general associations and develop an integrative CGqTP tumor-agnostic tumor

192   pembrolizumab predictive biomarker, we first limited results to the 5,233 patients in the SCMD who met

193   the following criteria: TMB measurements from StrataNGS testing (including meeting the overall 20%

194   tumor content requirement), immune gene expression quantification from an investigative multiplex PCR

195   based transcriptomic profiling test, and treatment for at least one month with at least one antineoplastic

196   agent. Of these 5,223 patients, 708 (13.5%) were treated with pembrolizumab. As shown in **Figure 1a**,

197   this cohort was comprised of patients with 24 tumor types, with NSCLC accounting for 293 (41.4%).

198   Real-world TTNT was inferred for each patient as the time from starting pembrolizumab to the time of

199   stopping pembrolizumab and starting a new therapy or death. To establish the appropriateness of TTNT

200   for studying pembrolizumab treatment outcomes, TTNT was compared to overall survival (OS). As

201   shown in **Figure S3a**, the overall correlation (Spearman ρ= 0.61) was impacted by two outliers, one of

202   which was a patient with metastatic melanoma who was briefly treated with pembrolizumab, then

203   ipilimumab + nivolumab, prior to an extended course with imatinib (the patient harbored two VUS in

204   *KIT;* **FigureS3a blue box**), while the other was a patient with metastatic NSCLC harboring an *EML4-*

205   *ALK* fusion by StrataNGS testing who was briefly treated with pembrolizumab and chemotherapy before

206   prolonged treatment with crizotinib and lorlatinib (**FigureS3a red box**); excluding these two patients,

207   TTNT and OS were more strongly correlated (Spearman ρ = 0.75). Lastly, to confirm the validity of ≥10

208   Muts/Mb from StrataNGS testing to define TMB-H, we demonstrated that TMB-H patients (n = 208) had

209   significantly longer pembrolizumab TTNT vs. TMB-L patients (n = 500; median TTNT >24 months vs.

210   10.3, log rank p<0.0001; **Figure S3b**).

211   To identify potential biomarkers of pembrolizumab benefit, we first considered 21 candidate immune and

212   proliferation gene expression biomarkers assessed across clinical RNA tests run in parallel with the

213   StrataNGS CGP test (which generates TMB). Importantly, target gene expression (in normalized reads

214   per million [nRPM]), were highly correlated from independent amplicons targeting different exons of *PD-*

215   *L1* (n=25,769 samples, concordance correlation coefficient =0.83) and *PD-1* (n=25,769 samples,

216   concordance correlation coefficient =0.78). Likewise, expression profiles of these 21genes across 27

217   directly comparable tumor types were highly correlated between 8,424 TCGA tumors and 18,062 Strata

218   RNA component profiled tumors (median Spearman ρ =0.871 for all candidate genes; **Table S1**). We

219   therefore assessed the association of pembrolizumab TTNT with StrataNGS derived TMB and the 21

220   candidate immune gene expression biomarkers. As shown in **Table 1**, significant (p<0.01) univariate

221   predictors included TMB (HR = 0.77; p<0.0001), composite *PD-L1* expression (HR = 0.90; p=0.0007),

222   composite *PD-1* expression (HR = 0.90; *p*=0.0011), *PD-L2* expression (HR = 0.91; *p*=0.005) and

223   *TNFRSF9* expression (HR = 0.92; *p*=0.007).

224   To develop an integrative pembrolizumab benefit predictor, we performed a multi-part process. First, we

225   performed backwards stepwise regression to fit a multivariate Cox proportional hazards model, iterating

226   100 times on randomly selected two-thirds of the dataset. TMB, *PD-1* and *PD-L2* were the three most

227   frequently included variables across the 100 models. Hence, we then used a brute-force approach, locking

228   in these three variables and using backwards stepwise regression adding in 1, 2 or 3 of the remaining

229   candidate variables, iterating 100 times on randomly selected 2/3rds of the dataset, and evaluating the

230   Akaike's Information Criteria (AIC) and Bayesian Information Criterion (BIC) of the trained model on

231   the held-out 1/3 of the dataset. As additional components increased the BIC with minimal decrease in

232   AIC, we chose a five variable multivariate model including TMB, *PD-1*, *PD-L2*, *CD4,* and *ADAM12*. We

233   therefore entered all five variables into model selection 100 times on randomly selected two-thirds of the

234   dataset and determined the model coefficients, confirming that that model coefficients developed in the

235   full 708 patient cohort were stable, and these coefficients were used in the final integrative model (**Figure**

236   **S4a**). Multivariate analysis on only the final five variable set confirmed that all five biomarkers were

237   independent predictors of pembrolizumab treatment outcome (**Table 1**). Notably, *PD-L1* was not included

238   in the final model and forced addition to the five variable model had essentially no impact on

239   performance, even when trained on the full 708 patient cohort (**Figure S4b**). As shown in **Figure S4c**,

240   TMB was minimally correlated with all final model gene expression biomarkers (Spearman ρ = -0.106

241   [*CD4*] to -0.0.015 [*PD-1]*), while correlation of individual gene expression biomarkers ranged from ρ =

242   0.217 (*ADAM12* vs. *PD-1*) to ρ = 0.675 (*PD-L2* vs. *CD4*).

243   **Integrative Immune Response Score (IRS) Development**

244   To evaluate the potential of the multivariate model to predict pembrolizumab treatment outcome in

245   patients, we derived individual Immunotherapy Response Scores (IRS) from the final five variable model,

246   assigned the 708 patients to one of three IRS groups based on potential clinical utility (see Methods; IRS-

247   High [-H; n=266], IRS-Intermediate [-I; n=176] and IRS-Low [-L; n=266]; IRS-H associated with

248   greatest benefit of pembrolizumab), and compared group outcomes. Kaplan Meier analysis of

249   pembrolizumab TTNT showed that treatment outcome varied widely across groups, with median TTNT

250   ranging from > 24 months in IRS-H, to 17.5 months in IRS-I, to 7 months in IRS-L (log-rank test

251   p<0.0001 IRS-H vs. -L; **Figure 1b**). To compare to OS, Kaplan Meier analysis was also performed with

252   respect to time-to-death, censoring patients with respect to latest date of follow-up, producing similarly

253   significant difference in outcome for IRS-H vs. -L (n=707, log-rank test p<0.0001; **Figure S5a**).

254

255   To establish the IRS model as predictive and not prognostic, we evaluated a chemotherapy comparator

256   cohort, consisting of the most recent previous chemotherapy line from the 166 of 708 (23%)

257   pembrolizumab patients with documented chemotherapy treatment prior to monotherapy pembrolizumab.

258   While chemotherapy median TTNT was similar across all three IRS groups at 7.0-8.2 months (**Figure**

259   **1c**), pembrolizumab had significantly longer TTNT than chemotherapy in IRS-H (median TTNT >24

260   months vs. 7.1 months; log-rank *p* value <0.001) and IRS-I (median TTNT 13.5 months vs. 8.2 months,

261   log-rank *p* value 0.02), but no significant difference was observed for pembrolizumab vs. chemotherapy

262   TTNT in IRS-L (median TTNT 5.8 vs. 7.1 months; log-rank p value = 0.65). The test for interaction

263   between pembrolizumab vs. previous chemotherapy treatment and continuous IRS was significant

264   (likelihood ratio test for interaction p<0.005), confirming the predictive nature of the IRS biomarker.

265   Through a similar analysis using continuous TMB (instead of the IRS model), TMB alone was also

266   confirmed as a predictive biomarker of pembrolizumab TTNT (likelihood ratio test for interaction

267   p<0.005). However, IRS (expression component + TMB) had significantly greater predictive ability than

268   TMB alone (likelihood ratio test between models, p=0.04). As in the overall cohort, TTNT and OS

269   comparisons were similar between IRS-H vs. -L patients in this predictive analysis (in addition to a cohort

270   of 201 patients considering pembrolizumab monotherapy or combination therapy vs. previous

271   chemotherapy), although IRS-I did not appear to be a separate group in these comparisons (**Figure S5b-**

272   **d**). Taken together, these results demonstrate the predictive nature of the IRS model for pembrolizumab

273   benefit prediction vs. an internal chemotherapy comparator and highlight the benefit of combining CGP

274   biomarkers (TMB) with parallel quantitative gene expression of the tumor and TME to improve

275   performance.

276   **Robustness of IRS model to potential confounding factors**

277   We next evaluated potential factors that could confound the utility of IRS [i.e., tumor type, therapy type

278   (monotherapy vs combination therapy) and TMB status]. First, we compared pembrolizumab TTNT in

279   the 293 patients with NSCLC (41.4%) to 415 patients with other tumor types (48.6%) across IRS groups

280   and found no significant differences (median TTNT >24 vs. >24, log-rank p= 0.13 for IRS-H; median

281   TTNT > 24 vs. 13.5, log-rank p=0.19 for IRS-I; median TTNT 7.2 vs. 6.1, log-rank p=0.10 for IRS-L;

282   **Figure 2a**). Then we compared pembrolizumab TTNT in the 481 patients treated with pembrolizumab

283   monotherapy (67.9%) to 227 patients treated with pembrolizumab plus chemotherapy (combination

284   therapy) (32.1%), and found no significant difference in any IRS risk group (median TTNT >24 vs. >24,

285   log-rank p= 0.87 for IRS-H; median TTNT > 17.5 vs. 15.6, log-rank p=0.77 for IRS-I; median TTNT 6.4

286   vs. 7.2, log-rank p=0.20 for IRS-L; **Figure 2b).** Lastly, given the pan-tumor approval of pembrolizumab

287   in TMB-H patients, if the IRS risk groups were exactly overlapping with TMB status, the IRS would have

288   no clinical utility. Therefore, we examined the predictive effect of IRS groups among the 208 (29.4%)

289   TMB-H patients and 500 (70.6%) TMB-L patients. While 167 (80%), 28 (13%) and 13 (6%) of the 208

290   TMB-H patients were IRS-H, -I, and -L, respectively, the TMB-H/IRS-H group still had significantly

291   longer pembrolizumab TTNT compared to TMB-H/IRS-L (median TTNT >24 months vs. 10.4 months;

292   log rank p-value = 0.001, **Figure 3c**), demonstrating the added predictive value of immune gene

293   expression, even among TMB-H patients. Among the 500 TMB-L patients, 99 (20%), 148 (30%) and 253

294   (50%) were IRS-H, -I, and -L, respectively. IRS robustly stratified pembrolizumab TTNT, with median

295   pembrolizumab TTNT of 20.8 months, and 7 months in IRS-H, and -L groups, respectively (log rank p-

296   value = <0.001, **Figure 3c**), demonstrating that TMB alone is insufficient for maximizing the prediction

297 of pembrolizumab benefit. Together, these results demonstrate that the IRS biomarker is robust to tumor

298 type, pembrolizumab monotherapy vs. combination chemotherapy treatment, and TMB status.

299

300 **Stability of IRS across temporal sample collection variability prior to CPI treatment**

301 Tissue based TMB has recently been shown to be stable for nearly all patients with advanced cancer

302 through whole genome sequencing of sequential tissue samples,[44] however less is known about the

303 stability of an integrative CGqTP model predicting pembrolizumab benefit. Hence, we first confirmed

304 that in the 426 total patients treated with both chemotherapy (regardless of whether chemotherapy was

305 pre- or post- pembrolizumab treatment) and pembrolizumab, the timing of sample collection (pre-

306 chemotherapy and pembrolizumab vs. post-chemotherapy but prior to pembrolizumab) did not

307 significantly impact median pembrolizumab TTNT across IRS groups (**Figure S6a**). Next, we directly

308 assessed IRS stability across patients in the SCMD with sequentially tested tissue samples. As analyses

309 presented thus far were limited to the most recently tested sample per patient (if testing had been

310 performed more than once) and to patients who started pembrolizumab after the collection date of the

311 included sample, we therefore identified 69 total patients in the SCMD who 1) had valid IRS scores from

312 two specimens with different collection dates, 2) were confirmed to be of clonal origin as part of routine

313 StrataNGS clinical testing, and 3) did not have CPI therapy starting between the collection dates of the

314 samples. As shown in **Figure S6b**, the integrative IRS model scores were highly correlated (Pearson

315 r=0.75, respectively) in paired specimens, and only two (3%) patients (n=2) moved from the IRS-H to -L

316 (or vice versa), supporting the stability of the IRS across temporal sampling in the absence of checkpoint

317 inhibitor therapy. Lastly, we assessed the performance of IRS in 84 patients who otherwise would have

318 been included in the 708 total patient discovery cohort described above, but had their sample collected

319 after starting pembrolizumab. Hypothesizing that CGP testing in this clinical scenario would usually be

320 performed as the patient was progressing on pembrolizumab, we predicted that IRS would be minimally

321 predictive of pembrolizumab TTNT. In these 84 patients, continuous IRS was not predictive of

322 pembrolizumab TTNT (p=0.61), with median pembrolizumab TTNT of 15.9 vs. 15.6 months in IRS-H

323 vs. -L, log rank p=0.92, **Figure S6c**). Together, these results support the stability and validity of IRS in

324 sequential tumor tissue samples collected prior to CPI treatment.

325 **Pan Solid Tumor Distribution of IRS groups**

326 Although future studies are required to prospectively validate IRS performance for routine clinical use,

327 we sought to leverage IRS distributions across tumor types (and pan-cancer biomarkers) in the entire

328 SCMD to understand the potential impact of IRS both within and outside of currently approved

329 pembrolizumab indications. Thus, we determined IRS for the 25,770 patients in the SCMD

330 (NCT03061305) with valid TMB and gene expression data, with 13.2%, 17.5%, and 69.3% of all patients

331 classified as IRS-H, -I, and -L, respectively (**Figure 3a**). Pembrolizumab approved tumor types (without

332 consideration of PD-L1 IHC status) had a substantially higher proportion (22.1% vs. 7.1%) of IRS-H

333 patients than non-pembrolizumab approved tumor types, as well as a higher proportion (23.6% vs. 13.4%)

334 of IRS-I group patients (**Figure 3b**). Tumor types with the highest proportion of IRS-H group patients

335 include several known to be highly responsive to CPIs, including melanoma, non-melanoma skin cancer,

336 NSCLC, lung small cell carcinoma (Lung – Other), and bladder (urothelial) tumors (**Figure 3c**). We next

337 examined the pan-solid tumor distribution of IRS groups by TMB status, given the pan-tumor approval of

338 pembrolizumab in TMB-H tumors. Whereas 91% of TMB-H patients were also IRS-H or -I, 22.0% of all

339 patients in the SCMD were IRS-H or -I and TMB-L (vs. 8.7% IRS-H or -I and TMB-H), demonstrating

340 that while TMB-H identifies most IRS-I/-H patients, IRS identifies a larger set of TMB-L patients

341 predicted to benefit from pembrolizumab (**Figure 3d**). Finally, to estimate the overall proportion of

342 patients with solid tumors who might benefit from pembrolizumab outside of currently approved tumor

343 types and biomarkers, we stratified the SCMD population by all pembrolizumab approved indications

344 (pembrolizumab approved tumor types, TMB-H, or MSI-H as approved). As shown in **Figure 3e**, if

345 prospectively validated, an additional 2.2-9.6% of patients (2.2% IRS-H and 7.4% IRS-I) with solid

346 tumors outside of currently approved indications are predicted to have substantial benefit from

347 pembrolizumab.

348 **Discussion**

349 Leveraging a robust clinical molecular database from the StrataTrial (NCT03061305), we developed a

350 highly significant, integrative, multivariate Immune Response Score (IRS) model that combined TMB

351 and quantitative immune gene expression to predict real-world pembrolizumab treatment outcomes in 708

352 patients from 28 solid tumor types. IRS model inputs were generated from simultaneously performed,

353 clinically validated, multiplex PCR based DNA and RNA NGS (StrataNGS CGP and a separate RNA

354 panel for quantitative gene expression)[40-42]. These assays were performed on co-isolated DNA and RNA

355 and share the same key sample input requirements defined from over 30,000 consecutively received FFPE

356 tumor samples for CGP testing: $\geq$20% tumor content and 2mm^2 tumor surface area (from 10 x 5um FFPE

357 sections)[41,42]. Of note, only 38.3% of samples in the development cohort included herein, and 38.8% of

358 the 25,770 total patients in the SCMD used to assess IRS distribution, met the minimum tumor surface

359 area requirements ($\geq$25mm^2) of FoundationOne CDx[45], the FDA approved companion diagnostic device

360 to identify TMB-H tumors.

361 The IRS model predicts an individual patient's likelihood of benefit with pembrolizumab therapy.

362 Patients were grouped into three categories (IRS-H, IRS-I, and IRS-L) based on pembrolizumab TTNT as

363 a measure for potential clinical utility, with IRS-H patients having median pembrolizumab TTNT > 24

364 months, while IRS-L patients had median pembrolizumab TTNT of 7 months. Critically, in the subset of

365 166 patients treated with pembrolizumab monotherapy who had prior chemotherapy treatment, we

366 confirmed the predictive nature of the IRS model, as IRS-H patients had significantly longer TTNT on

367 pembrolizumab vs. their immediately preceding chemotherapy treatment (median TTNT >24 months vs.

368 7.1 months), whereas IRS-L patients did not (median TTNT 5.8 vs. 7.1), with a significant test for

369 interaction between continuous IRS and pembrolizumab vs. chemotherapy. Notably, the association of

370   IRS with pembrolizumab TTNT was stable when stratified by NSCLC vs. other tumor types,

371   pembrolizumab monotherapy vs. combination therapy, TMB-H vs. TMB-L tumors, and pre- vs. post-

372   chemotherapy sample collection, suggesting that the model captures universal biological features of

373   pembrolizumab benefit. When applied to all 25,770 patients in the SCMD where IRS could be generated,

374   IRS-H was more frequent in tumor types known to derive benefit from CPI, but occurred in subsets of

375   nearly every tumor type. Outside of approved pembrolizumab tumor type indications, including TMB-H

376   and MSI-H pan cancer, 2.2% patients in the SCMD were IRS-High, representing a conservative estimate

377   as many approved indications have PD-L1 IHC requirements. Hence, if subsequently validated in our

378   ongoing studies, the improved predictive clinical utility of the integrative IRS model would be

379   demonstrated to more accurately identify patients across tumor types with a high probability of response

380   to pembrolizumab over single biomarker approaches.

381   Intriguingly, an additional 7.4% of patients were IRS-I outside of tumor types with FDA-approved

382   pembrolizumab indications (also excluding MSI-H and TMB-H). Although these patients had

383   significantly longer monotherapy pembrolizumab TTNT vs. their previous chemotherapy (13.5 vs. 8.2

384   months) and the proportion of IRS-I patients in selected tumor types is consistent versus observed

385   response rates (ORR) in early phase pembrolizumab (e.g.,13% of SCMD patients with pancreatic cancer

386   are IRS-I versus an ORR of 10% in non-biomarker selected patients treated with pembrolizumab

387   monotherapy in the KEYNOTE-028 trial[36]), tumor specific trials may be needed to demonstrate a clear

388   benefit of pembrolizumab in this subset of patients.

389   As the IRS was developed from a single integrative clinical platform using co-isolated DNA and RNA to

390   generate TMB and highly quantitative gene expression assessment of the tumor and TME from over 700

391   patients across 24 tumor types, the IRS model holds several potentially interesting biological insights.

392   First, TMB, *PD-1* expression, and *PD-L2* expression were each independent predictors of pembrolizumab

393   benefit, indicating a multiplicative predictive effect across these biomarkers representing increased

394   antigenicity (TMB), the direct target of pembrolizumab (*PD-1*), and one of the two PD-1 interacting

395 ligands (*PD-L2*). Notably, although *PD-L1* expression was predictive of pembrolizumab TTNT on

396 univariate analysis, it was not an independent significant predictor identified through our multi-step

397 multivariate model developed process. In an exploratory analysis, when *PD-L1* RNA was added into the

398 IRS model and trained on the entire cohort, essentially no change in performance was observed,

399 consistent with its predictive ability being captured by the other model components. While PD-L1

400 evaluation by IHC is the current FDA-approved biomarker to predict pembrolizumab (or other CPI)

401 benefit either individually or in models[45-47], expression varies by antibody clone and nearly all studies

402 show at least some responsive PD-L1–IHC low/negative patients , suggesting that other PD-1 ligands

403 beyond PD-L1, such as PD-L2, may be relevant for predicting clinical response[30,48-51]. Consistent with

404 this observation, in head and neck squamous cell carcinoma, PD-L2 expression by IHC predicted

405 pembrolizumab response and progression free survival independent of PD-L1 IHC status[46]. *CD4* and

406 *ADAM12* were both negative predictors of pembrolizumab TTNT in numerous model training iterations

407 and were significant in the final five variable IRS model. Although both effector CD8+ and CD4+ T cells

408 have been shown to express PD-1[47], *CD8A* (which encodes CD8 and was included in our 21 candidate

409 genes), was more predictive of CPI benefit in a recent metanalysis of whole transcriptome data than either

410 *PD-L1* expression or the T cell inflamed gene expression signature[15], and hence the inclusion of *CD4* as a

411 negative predictive factor in the final IRS model likely reflects at least in part the ratio of effector (CD8+)

412 to helper/regulatory CD4+ T cells in the TME.

413 Although less is known about the direct role of *ADAM12* in CPI response, it is highly expressed by cancer

414 associated fibroblasts CAFs—as shown through single cell sequencing studies and bulk tumor profiling—

415 as a driver of feed forward TGF-β signaling, has been shown to act as a T cell co-stimulatory molecule

416 expressed on some regulatory T cells, and has been identified in a signature of negative response to ICI in

417 melanoma[52-57]. Of note, in colorectal cancer, where single cell sequencing demonstrated high *ADAM12*

418 expression in CAFs[58], as well as urothelial carcinoma, TGF-β signaling from CAFs has been shown to

419 drive T cell exclusion, a hallmark of low response to ICI[59-63]. Taken together, these results support

420   additional investigation into a potential mechanistic role for *ADAM12* in ICI resistance, as well as

421   demonstrate the complementary nature of the integrative biomarkers in the IRS model, which integrates

422   measurement of tumor neo-antigenicity (TMB), with quantification of key tumor and TME biomarkers.

423   Current FDA-approved CPI biomarkers include PD-L1 IHC, TMB and MSI-H (although the latter

424   indication was initially approved without a companion diagnostic biomarker), however these biomarkers

425   have several practical challenges for clinical use including variations in assay parameters, platforms, and

426   predictive thresholds[4,64-67]. For example, although there are multiple tissue TMB assays commercially

427   available (LDTs, FDA cleared devices, and a single FDA approved device), TMB testing typically has a

428   large tissue requirement, which is frequently not feasible in patients with advanced cancers, and such

429   approaches do not allow for parallel assessment of gene expression biomarkers. Thus, there is a need for

430   optimized CPI biomarkers with improved predictive utility that can be developed into a scalable clinical

431   test that is applicable to nearly all cancer patients, including those with limited tumor tissue available. As

432   described herein, IRS addresses these needs by 1) co-isolating DNA and RNA to measure multiple classes

433   of biomarkers from the same tissue sample; 2) optimizing pembrolizumab treatment benefit prediction via

434   explainable model development using highly quantitative gene expression data in a large pan-tumor

435   cohort with real world treatment data; and 3) utilizing a clinically validated and scalable platform

436   developed for real world FFPE samples with minimal tumor size (2 mm^2 tumor surface area)[41,42]. In

437   comparison, only 38.8% of the 25,770 patients in the SCMD used to assess IRS distribution met the

438   minimum tumor surface area requirements ($\geq$25 mm^2) for FoundationOne CDx, the FDA approved

439   companion diagnostic device to identify TMB-H tumors[45], suggesting that the majority of real-world

440   patients with advanced solid tumors have insufficient tumor samples to determine TMB.

441   Our analysis has several potential limitations. First, the real-world dataset was biased toward tumor types

442   for which pembrolizumab is FDA-approved or treatment was selected based on other biomarker results,

443   and thus, as expected, was enriched for patients benefiting from pembrolizumab. Indeed, the proportion of

444   patients in the IRS-High group was much higher in the pembrolizumab treatment cohort (38%) than the

445   broader tumor profiling dataset (13.2%). Second, the TTNT endpoint likely includes some patients who

446   stopped treatment due to treatment toxicity or switching therapy to a more appropriate regimen based on

447   molecular results (as described above) and not disease progression, although this likely represents a

448   minority of events. Additionally, only advanced patients were eligible for the Strata Trial, but adjuvant

449   therapy cannot be directly excluded using our treatment data collection approach, with this limitation

450   being relevant for our pembrolizumab vs. prior chemotherapy analysis; however, this would tend to bias

451   against pembrolizumab TTNT being longer than adjuvant chemotherapy in patients with significant delay

452   between adjuvant chemotherapy and development of metastatic disease (and pembrolizumab treatment).

453   Additionally, although it is unclear if our model is applicable to other PD-1 monoclonal antibodies, PD-

454   L1 monoclonal antibodies, and/or combined PD-1/PD-L1 and CTLA4 antibody therapy, we focused on

455   pembrolizumab herein given the large amount of treatment data in our cohort across tumor types.

456   Likewise, future studies will also investigate whether inclusion of single gene-based DNA biomarkers

457   identified as potentially predictive in one or more tumor types (e.g. *STK11*, *PBRM1*, *ARID1A*, *CDNK2A*[68-

458   75] or additional immune related genes assessed on the current expanded quantitative expression panel run

459   in parallel with StrataNGS testing can improve the performance of the IRS model. Limited PD-L1 IHC

460   data was available for subjects in the SCMD, and hence we are not able to directly compare performance

461   of IRS and PD-L1 IHC for predicting pembrolizumab benefit; additionally, this limitation also biases

462   against the overall proportion of patients outside of currently approved indications predicted to benefit

463   from pembrolizumab by IRS, as herein we considered all patients in approved tumor types to be in an

464   approved indication, although in many tumor types only a minority of patients are approved for

465   pembrolizumab treatment based on PD-L1 IHC cutoffs. Notably, we chose to use standard multivariate

466   regression with a minimum number of variables versus other approaches that have included a larger

467   number of immune related genes[27,34,39] or used more advanced machine learning approaches[76] to leverage

468   the highly quantitative nature of CGqTP and minimize the risk of overfitting, as our model was trained

469   and characterized on the same dataset. Importantly, while we were able to use an internal chemotherapy

470   control cohort to determine the predictive nature of the biomarker, additional, pre-specified validation of

471   the IRS model in an independent cohort will be required to establish the clinical utility of IRS groups for

472   predicting pembrolizumab benefit.

473   In summary, after demonstrating the face validity a clinical molecular database containing treatment data

474   and molecular profiling from a large observational trial of patients with advanced cancer, we report the

475   development of a biologically rational, integrative CGqTP based model of pembrolizumab benefit that is

476   robust to tumor type, TMB status, pembrolizumab monotherapy vs. combination therapy treatment, and

477   pre-pembrolizumab sample collection timing. Importantly, the IRS biomarker was developed from a

478   single clinically validated NGS platform capable of simultaneously performing comprehensive genomic

479   profiling (required for TMB but also for therapy options outside of CPI) and in parallel precise

480   quantification of tumor- and TME-relevant gene expression, providing a clear diagnostic pathway for

481   potential clinical application. IRS has potential application for both refining the use of pembrolizumab in

482   tumor types for which immunotherapy is indicated and therapeutic choice is present (as well as

483   monotherapy pembrolizumab vs. combination therapy as in NSCLC), as well as for guiding

484   pembrolizumab treatment decisions for patients outside of indicated tumor types. Most notably, IRS-High

485   patients treated with pembrolizumab had not reached median TTNT after 24 months (compared to a

486   median TTNT of 7.1 months on their prior line of chemotherapy), suggesting that this population may

487   benefit similarly to the TMB-H population identified in KEYNOTE-158 (29% overall response rate; 66%

488   of responders having a duration of response ≥ 24 months). Herein, across the entire SCMD, 2.2% of

489   patients were IRS-H/TMB-L/not-MSI-H and outside of approved pembrolizumab approved tumor types,

490   and an additional 7.4% were IRS-I. Hence, if further validated in additional cohorts, the IRS model has

491   the potential to markedly expand the benefit of pembrolizumab across solid tumors, addressing one of the

492   most important challenges in precision oncology.

493

494

ACKNOWLEDGEMENTS

AUTHOR CONTRIBUTIONS

Conception and design: Kat Kwiatkowski, D. Bryan Johnson, Daniel R. Rhodes, Scott A. Tomlins,

Administrative support: Laura E. Lamb, Melissa J. Shreve

Provision of study materials or patients: Marc R. Matrana, Mark E. Burkard, Eddy Shih-Hsin Yang, William Jeffery Edenfield, E. Claire Dees, Adedayo A. Onitilo, Michael Thompson, Gary L. Buchschacher, Alan M. Miller, Alexander Menter, Benjamin Parsons, Timothy Wassenaar, Leon C. Hwang, J. Marie Suga, Robert Siegel, William Irvin, Jr., Suresh Nair, Jennifer N. Slim

Collection and assembly of data: Marc R. Matrana, Mark E. Burkard, Eddy Shih-Hsin Yang, William Jeffery Edenfield, E. Claire Dees, Adedayo A. Onitilo, Michael Thompson, Gary L. Buchschacher, Alan M. Miller, Alexander Menter, Benjamin Parsons, Timothy Wassenaar, Leon C. Hwang, J. Marie Suga, Robert Siegel, William Irvin, Jr., Suresh Nair, Jennifer N. Slim, Kat Kwiatkowski, Khalis Mitchell, Stephanie Drewery, Andrew Fischer, Jennifer Hipp, Travis Reeder, Hana Vakil

Data analysis and interpretation: Nickolay Khazanov, Daniel H. Hovelson, Tina Hu-Seliger, D. Bryan Johnson, Daniel R. Rhodes, Scott A. Tomlins

Manuscript writing: All authors

Final approval of manuscript: All authors

Accountable for all aspects of the work: All authors

DATA AVAILABILITY

Relevant data supportive of the figures, tables, and results of the paper are found in the Supplemental Materials. In the interest of protecting patient privacy and in accordance with applicable data sharing agreements and/or patient informed consent forms, the authors are restricted from making raw patient-level data publicly available. However, interested parties may contact the authors at BD@strataoncology.com to request access for research purposes, and such requests will be handled on a case-by-case basis.

CODE AVAILABILITY

530   The IRS model algorithm is provided in the Methods.

531   **COMPETING INTERESTS**

560   **REFERENCES**

561   1   Sharma, P. *et al.* The Next Decade of Immune Checkpoint Therapy. *Cancer Discov* **11**, 838-857,
562   doi:10.1158/2159-8290.CD-20-1680 (2021).
563   2   Chamoto, K., Hatae, R. & Honjo, T. Current issues and perspectives in PD-1 blockade cancer
564   immunotherapy. *Int J Clin Oncol* **25**, 790-800, doi:10.1007/s10147-019-01588-7 (2020).
565   3   Ribas, A. & Wolchok, J. D. Cancer immunotherapy using checkpoint blockade. *Science* **359**,
566   1350-1355, doi:10.1126/science.aar4060 (2018).
567   4   Doroshow, D. B. *et al.* PD-L1 as a biomarker of response to immune-checkpoint inhibitors. *Nat*
568   *Rev Clin Oncol* **18**, 345-362, doi:10.1038/s41571-021-00473-5 (2021).
569   5   Gavrielatou, N., Shafi, S., Gaule, P. & Rimm, D. L. PD-L1 Expression Scoring: Non-
570   Interchangeable, Non-Interpretable, Neither, or Both. *J Natl Cancer Inst*,
571   doi:10.1093/jnci/djab109 (2021).
572   6   Rimm, D. L. *et al.* A Prospective, Multi-institutional, Pathologist-Based Assessment of 4
573   Immunohistochemistry Assays for PD-L1 Expression in Non-Small Cell Lung Cancer. *JAMA*
574   *Oncol* **3**, 1051-1058, doi:10.1001/jamaoncol.2017.0013 (2017).
575   7   Salgado, R. *et al.* How current assay approval policies are leading to unintended imprecision
576   medicine. *Lancet Oncol* **21**, 1399-1401, doi:10.1016/S1470-2045(20)30592-1 (2020).
577   8   Herbst, R. S. *et al.* Atezolizumab for First-Line Treatment of PD-L1-Selected Patients with
578   NSCLC. *N Engl J Med* **383**, 1328-1339, doi:10.1056/NEJMoa1917346 (2020).
579   9   Abdul Karim, L., Wang, P., Chahine, J. & Kallakury, B. Harmonization of PD-L1
580   Immunohistochemistry Assays for Lung Cancer: A Working Progress. *J Thorac Oncol* **12**, e45,
581   doi:10.1016/j.jtho.2016.12.022 (2017).
582   10   Dolled-Filhart, M. *et al.* Development of a Prototype Immunohistochemistry Assay to Measure
583   Programmed Death Ligand-1 Expression in Tumor Tissue. *Arch Pathol Lab Med* **140**, 1259-
584   1266, doi:10.5858/arpa.2015-0544-OA (2016).
585   11   Dolled-Filhart, M. *et al.* Development of a Companion Diagnostic for Pembrolizumab in Non-
586   Small Cell Lung Cancer Using Immunohistochemistry for Programmed Death Ligand-1. *Arch*
587   *Pathol Lab Med* **140**, 1243-1249, doi:10.5858/arpa.2015-0542-OA (2016).
588   12   Hirsch, F. R. *et al.* PD-L1 Immunohistochemistry Assays for Lung Cancer: Results from Phase 1
589   of the Blueprint PD-L1 IHC Assay Comparison Project. *J Thorac Oncol* **12**, 208-222,
590   doi:10.1016/j.jtho.2016.11.2228 (2017).
591   13   Velcheti, V. *et al.* Real-world PD-L1 testing and distribution of PD-L1 tumor expression by
592   immunohistochemistry assay type among patients with metastatic non-small cell lung cancer in
593   the United States. *PLoS One* **13**, e0206370, doi:10.1371/journal.pone.0206370 (2018).
594   14   Grant, M. J., Herbst, R. S. & Goldberg, S. B. Selecting the optimal immunotherapy regimen in
595   driver-negative metastatic NSCLC. *Nat Rev Clin Oncol* **18**, 625-644, doi:10.1038/s41571-021-
596   00520-1 (2021).
597   15   Litchfield, K. *et al.* Meta-analysis of tumor- and T cell-intrinsic mechanisms of sensitization to
598   checkpoint inhibition. *Cell* **184**, 596-614 e514, doi:10.1016/j.cell.2021.01.002 (2021).
599   16   Rizvi, H. *et al.* Molecular Determinants of Response to Anti-Programmed Cell Death (PD)-1 and
600   Anti-Programmed Death-Ligand 1 (PD-L1) Blockade in Patients With Non-Small-Cell Lung
601   Cancer Profiled With Targeted Next-Generation Sequencing. *J Clin Oncol* **36**, 633-641,
602   doi:10.1200/JCO.2017.75.3384 (2018).
603   17   Singal, G. *et al.* Association of Patient Characteristics and Tumor Genomics With Clinical
604   Outcomes Among Patients With Non-Small Cell Lung Cancer Using a Clinicogenomic Database.
605   *JAMA* **321**, 1391-1399, doi:10.1001/jama.2019.3241 (2019).
606   18   Rizvi, N. A. *et al.* Cancer immunology. Mutational landscape determines sensitivity to PD-1
607   blockade in non-small cell lung cancer. *Science* **348**, 124-128, doi:10.1126/science.aaa1348
608   (2015).

609   19   McGranahan, N. *et al.* Clonal neoantigens elicit T cell immunoreactivity and sensitivity to
610   immune checkpoint blockade. *Science* **351**, 1463-1469, doi:10.1126/science.aaf1490 (2016).
611   20   Van Allen, E. M. *et al.* Genomic correlates of response to CTLA-4 blockade in metastatic
612   melanoma. *Science* **350**, 207-211, doi:10.1126/science.aad0095 (2015).
613   21   Samstein, R. M. *et al.* Tumor mutational load predicts survival after immunotherapy across
614   multiple cancer types. *Nat Genet* **51**, 202-206, doi:10.1038/s41588-018-0312-8 (2019).
615   22   Vega, D. M. *et al.* Aligning tumor mutational burden (TMB) quantification across diagnostic
616   platforms: phase II of the Friends of Cancer Research TMB Harmonization Project. *Ann Oncol*,
617   doi:10.1016/j.annonc.2021.09.016 (2021).
618   23   Steuer, C. E. & Ramalingam, S. S. Tumor Mutation Burden: Leading Immunotherapy to the Era
619   of Precision Medicine? *J Clin Oncol* **36**, 631-632, doi:10.1200/JCO.2017.76.8770 (2018).
620   24   Anagnostou, V. *et al.* Multimodal genomic features predict outcome of immune checkpoint
621   blockade in non-small-cell lung cancer. *Nat Cancer* **1**, 99-111, doi:10.1038/s43018-019-0008-8
622   (2020).
623   25   Marcus, L. *et al.* FDA Approval Summary: Pembrolizumab for the Treatment of Tumor
624   Mutational Burden-High Solid Tumors. *Clin Cancer Res* **27**, 4685-4689, doi:10.1158/1078-
625   0432.CCR-21-0327 (2021).
626   26   Marabelle, A. *et al.* Association of tumour mutational burden with outcomes in patients with
627   advanced solid tumours treated with pembrolizumab: prospective biomarker analysis of the
628   multicohort, open-label, phase 2 KEYNOTE-158 study. *Lancet Oncol* **21**, 1353-1365,
629   doi:10.1016/S1470-2045(20)30445-9 (2020).
630   27   Cristescu, R. *et al.* Pan-tumor genomic biomarkers for PD-1 checkpoint blockade-based
631   immunotherapy. *Science* **362**, doi:10.1126/science.aar3593 (2018).
632   28   Lee, J. S. & Ruppin, E. Multiomics Prediction of Response Rates to Therapies to Inhibit
633   Programmed Cell Death 1 and Programmed Cell Death 1 Ligand 1. *JAMA Oncol* **5**, 1614-1618,
634   doi:10.1001/jamaoncol.2019.2311 (2019).
635   29   Rolfo, C. *et al.* Liquid Biopsy for Advanced NSCLC: A Consensus Statement From the
636   International Association for the Study of Lung Cancer. *J Thorac Oncol* **16**, 1647-1662,
637   doi:10.1016/j.jtho.2021.06.017 (2021).
638   30   Herbst, R. S. *et al.* Predictive correlates of response to the anti-PD-L1 antibody MPDL3280A in
639   cancer patients. *Nature* **515**, 563-567, doi:10.1038/nature14011 (2014).
640   31   Taube, J. M. *et al.* Colocalization of inflammatory response with B7-h1 expression in human
641   melanocytic lesions supports an adaptive resistance mechanism of immune escape. *Sci Transl*
642   *Med* **4**, 127ra137, doi:10.1126/scitranslmed.3003689 (2012).
643   32   Chen, D. S. & Mellman, I. Elements of cancer immunity and the cancer-immune set point. *Nature*
644   **541**, 321-330, doi:10.1038/nature21349 (2017).
645   33   Sanmamed, M. F. & Chen, L. A Paradigm Shift in Cancer Immunotherapy: From Enhancement
646   to Normalization. *Cell* **175**, 313-326, doi:10.1016/j.cell.2018.09.035 (2018).
647   34   Ayers, M. *et al.* IFN-gamma-related mRNA profile predicts clinical response to PD-1 blockade. *J*
648   *Clin Invest* **127**, 2930-2940, doi:10.1172/JCI91190 (2017).
649   35   Fountzilas, E., Kurzrock, R., Hiep Vo, H. & Tsimberidou, A. M. Wedding of Molecular
650   Alterations and Immune Checkpoint Blockade: Genomics as a Matchmaker. *J Natl Cancer Inst*,
651   doi:10.1093/jnci/djab067 (2021).
652   36   Ott, P. A. *et al.* T-Cell-Inflamed Gene-Expression Profile, Programmed Death Ligand 1
653   Expression, and Tumor Mutational Burden Predict Efficacy in Patients Treated With
654   Pembrolizumab Across 20 Cancers: KEYNOTE-028. *J Clin Oncol* **37**, 318-327,
655   doi:10.1200/JCO.2018.78.2276 (2019).
656   37   Lu, S. *et al.* Comparison of Biomarker Modalities for Predicting Response to PD-1/PD-L1
657   Checkpoint Blockade: A Systematic Review and Meta-analysis. *JAMA Oncol* **5**, 1195-1204,
658   doi:10.1001/jamaoncol.2019.1549 (2019).

659 38 Zhang, Z. *et al.* RNF2 ablation reprograms the tumor-immune microenvironment and stimulates
660   durable NK and CD4+ T-cell-dependent antitumor immunity. *Nature Cancer* **2**, 1018-1038,
661   doi:10.1038/s43018-021-00263-z (2021).
662 39 Cristescu, R. *et al.* Transcriptomic Determinants of Response to Pembrolizumab Monotherapy
663   Across Solid Tumor Types. *Clin Cancer Res*, doi:10.1158/1078-0432.CCR-21-3329 (2021).
664 40 Harms, K. L. *et al.* Virus-positive Merkel Cell Carcinoma Is an Independent Prognostic Group
665   with Distinct Predictive Biomarkers. *Clin Cancer Res* **27**, 2494-2504, doi:10.1158/1078-
666   0432.CCR-20-0864 (2021).
667 41 Tomlins, S. A. *et al.* Development and Validation of StrataNGS, a Multiplex PCR,
668   Semiconductor Sequencing-Based Comprehensive Genomic Profiling Test. *J Mol Diagn* **23**,
669   1515-1533, doi:10.1016/j.jmoldx.2021.08.005 (2021).
670 42 Tomlins, S. A. *et al.* Real-World Performance of a Comprehensive Genomic Profiling Test
671   Optimized for Small Tumor Samples. *JCO Precis Oncol* **5**, doi:10.1200/PO.20.00472 (2021).
672 43 Zehir, A. *et al.* Mutational landscape of metastatic cancer revealed from prospective clinical
673   sequencing of 10,000 patients. *Nat Med* **23**, 703-713, doi:10.1038/nm.4333 (2017).
674 44 van de Haar, J. *et al.* Limited evolution of the actionable metastatic cancer genome under
675   therapeutic pressure. *Nat Med* **27**, 1553-1563, doi:10.1038/s41591-021-01448-w (2021).
676 45 Medicine, F.   (2021).
677 46 Yearley, J. H. *et al.* PD-L2 Expression in Human Tumors: Relevance to Anti-PD-1 Therapy in
678   Cancer. *Clin Cancer Res* **23**, 3158-3167, doi:10.1158/1078-0432.CCR-16-1761 (2017).
679 47 Giraldo, N. A. *et al.* Multidimensional, quantitative assessment of PD-1/PD-L1 expression in
680   patients with Merkel cell carcinoma and association with response to pembrolizumab. *J
681   Immunother Cancer* **6**, 99, doi:10.1186/s40425-018-0404-0 (2018).
682 48 Garon, E. B. *et al.* Pembrolizumab for the treatment of non-small-cell lung cancer. *N Engl J Med*
683   **372**, 2018-2028, doi:10.1056/NEJMoa1501824 (2015).
684 49 Robert, C. *et al.* Pembrolizumab versus Ipilimumab in Advanced Melanoma. *N Engl J Med* **372**,
685   2521-2532, doi:10.1056/NEJMoa1503093 (2015).
686 50 Tumeh, P. C. *et al.* PD-1 blockade induces responses by inhibiting adaptive immune resistance.
687   *Nature* **515**, 568-571, doi:10.1038/nature13954 (2014).
688 51 Mahoney, K. M. & Atkins, M. B. Prognostic and predictive markers for the new
689   immunotherapies. *Oncology (Williston Park)* **28 Suppl 3**, 39-48 (2014).
690 52 Yu, C. *et al.* A five-gene signature is a prognostic biomarker in pan-cancer and related with
691   immunologically associated extracellular matrix. *Cancer Med* **10**, 4629-4643,
692   doi:10.1002/cam4.3986 (2021).
693 53 Liu, Y. *et al.* ADAM12 is a costimulatory molecule that determines Th1 cell fate and mediates
694   tissue inflammation. *Cell Mol Immunol* **18**, 1904-1919, doi:10.1038/s41423-020-0486-8 (2021).
695 54 Cheon, D. J. *et al.* ADAM12 is a prognostic factor associated with an aggressive molecular
696   subtype of high-grade serous ovarian carcinoma. *Carcinogenesis* **36**, 739-747,
697   doi:10.1093/carcin/bgv059 (2015).
698 55 Cui, C. *et al.* Ratio of the interferon-gamma signature to the immunosuppression signature
699   predicts anti-PD-1 therapy response in melanoma. *NPJ Genom Med* **6**, 7, doi:10.1038/s41525-
700   021-00169-w (2021).
701 56 Ruff, M. *et al.* The Disintegrin and Metalloprotease ADAM12 Is Associated with TGF-beta-
702   Induced Epithelial to Mesenchymal Transition. *PLoS One* **10**, e0139179,
703   doi:10.1371/journal.pone.0139179 (2015).
704 57 Atfi, A. *et al.* The disintegrin and metalloproteinase ADAM12 contributes to TGF-beta signaling
705   through interaction with the type II receptor. *J Cell Biol* **178**, 201-208,
706   doi:10.1083/jcb.200612046 (2007).
707 58 Pelka, K. *et al.* Spatially organized multicellular immune hubs in human colorectal cancer. *Cell*
708   **184**, 4734-4752 e4720, doi:10.1016/j.cell.2021.08.003 (2021).

709   59   Mariathasan, S. *et al.* TGFbeta attenuates tumour response to PD-L1 blockade by contributing to
710   exclusion of T cells. *Nature* **554**, 544-548, doi:10.1038/nature25501 (2018).
711   60   Tauriello, D. V. F. *et al.* TGFbeta drives immune evasion in genetically reconstituted colon
712   cancer metastasis. *Nature* **554**, 538-543, doi:10.1038/nature25492 (2018).
713   61   Jiang, P. *et al.* Signatures of T cell dysfunction and exclusion predict cancer immunotherapy
714   response. *Nat Med* **24**, 1550-1558, doi:10.1038/s41591-018-0136-1 (2018).
715   62   Bonaventura, P. *et al.* Cold Tumors: A Therapeutic Challenge for Immunotherapy. *Front*
716   *Immunol* **10**, 168, doi:10.3389/fimmu.2019.00168 (2019).
717   63   Gajewski, T. F., Schreiber, H. & Fu, Y. X. Innate and adaptive immune cells in the tumor
718   microenvironment. *Nat Immunol* **14**, 1014-1022, doi:10.1038/ni.2703 (2013).
719   64   Cottrell, T. R. & Taube, J. M. PD-L1 and Emerging Biomarkers in Immune Checkpoint Blockade
720   Therapy. *Cancer J* **24**, 41-46, doi:10.1097/PPO.0000000000000301 (2018).
721   65   Kerr, K. M. The PD-L1 Immunohistochemistry Biomarker: Two Steps Forward, One Step Back?
722   *J Thorac Oncol* **13**, 291-294, doi:10.1016/j.jtho.2018.01.020 (2018).
723   66   Sholl, L. M. *et al.* The Promises and Challenges of Tumor Mutation Burden as an
724   Immunotherapy Biomarker: A Perspective from the International Association for the Study of
725   Lung Cancer Pathology Committee. *J Thorac Oncol* **15**, 1409-1424,
726   doi:10.1016/j.jtho.2020.05.019 (2020).
727   67   Addeo, A., Friedlaender, A., Banna, G. L. & Weiss, G. J. TMB or not TMB as a biomarker: That
728   is the question. *Crit Rev Oncol Hematol* **163**, 103374, doi:10.1016/j.critrevonc.2021.103374
729   (2021).
730   68   Han, G. *et al.* 9p21 loss confers a cold tumor immune microenvironment and primary resistance
731   to immune checkpoint therapy. *Nat Commun* **12**, 5606, doi:10.1038/s41467-021-25894-9 (2021).
732   69   Sholl, L. M. Biomarkers of response to checkpoint inhibitors beyond PD-L1 in lung cancer. *Mod*
733   *Pathol*, doi:10.1038/s41379-021-00932-5 (2021).
734   70   Di Federico, A., De Giglio, A., Parisi, C. & Gelsomino, F. STK11/LKB1 and KEAP1 mutations
735   in non-small cell lung cancer: Prognostic rather than predictive? *Eur J Cancer* **157**, 108-113,
736   doi:10.1016/j.ejca.2021.08.011 (2021).
737   71   Skoulidis, F. *et al.* STK11/LKB1 Mutations and PD-1 Inhibitor Resistance in KRAS-Mutant
738   Lung Adenocarcinoma. *Cancer Discov* **8**, 822-835, doi:10.1158/2159-8290.CD-18-0099 (2018).
739   72   Shen, J. *et al.* ARID1A deficiency promotes mutability and potentiates therapeutic antitumor
740   immunity unleashed by immune checkpoint blockade. *Nat Med* **24**, 556-562, doi:10.1038/s41591-
741   018-0012-z (2018).
742   73   Miao, D. *et al.* Genomic correlates of response to immune checkpoint therapies in clear cell renal
743   cell carcinoma. *Science* **359**, 801-806, doi:10.1126/science.aan5951 (2018).
744   74   Braun, D. A. *et al.* Clinical Validation of PBRM1 Alterations as a Marker of Immune Checkpoint
745   Inhibitor Response in Renal Cell Carcinoma. *JAMA Oncol* **5**, 1631-1633,
746   doi:10.1001/jamaoncol.2019.3158 (2019).
747   75   Liu, X. D. *et al.* PBRM1 loss defines a nonimmunogenic tumor phenotype associated with
748   checkpoint inhibitor resistance in renal carcinoma. *Nat Commun* **11**, 2135, doi:10.1038/s41467-
749   020-15959-6 (2020).
750   76   Banchereau, R. *et al.* Molecular determinants of response to PD-L1 blockade across tumor types.
751   *Nat Commun* **12**, 3969, doi:10.1038/s41467-021-24112-w (2021).

752

753

FIGURE LEGENDS

755 **Figure 1. An integrative Immune Response Score (IRS) model predicts real-world pembrolizumab**

756 **treatment outcome across solid tumors.** a) The tumor type distribution of 708 patients in the StrataTrial

757 (NCT03061305) clinical molecular database used to develop the IRS model, an integrative algorithm

758 predicting pembrolizumab benefit. Included patients were treated with pembrolizumab and had available

759 tumor mutation burden (from comprehensive genomic profiling) and in-parallel quantitative gene

760 expression data of immune relevant biomarkers from clinical testing of routine tumor tissue. **b)** Real-

761 world pembrolizumab progression-free survival, stratified by IRS group. After IRS model development,

762 patients were binned into three groups based on predicted pembrolizumab benefit: IRS-Low (grey line),

763 IRS-Intermediate (inter., light blue line), and IRS-High (dark blue line). Real world progression free

764 survival was determined using time to next therapy (TTNT; see Methods). **c)** Confirmation of the

765 predictive nature of the IRS model. Real-world progression-free survival on monotherapy pembrolizumab

766 (Pembro; purple) vs. immediately preceding chemotherapy (Chemo; line green) stratified by IRS group

767 was compared for the applicable subset of 166 patients. The interaction test for continuous IRS and

768 pembrolizumab vs. chemotherapy treatment was significant (likelihood ratio test for interaction $p<0.005$).
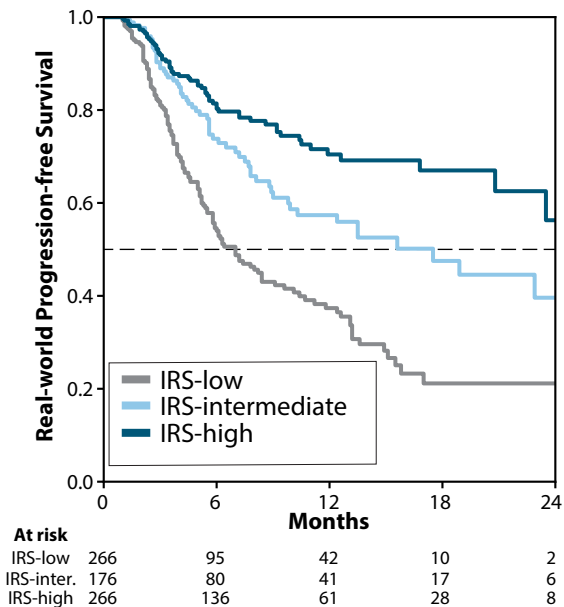
Figure 1
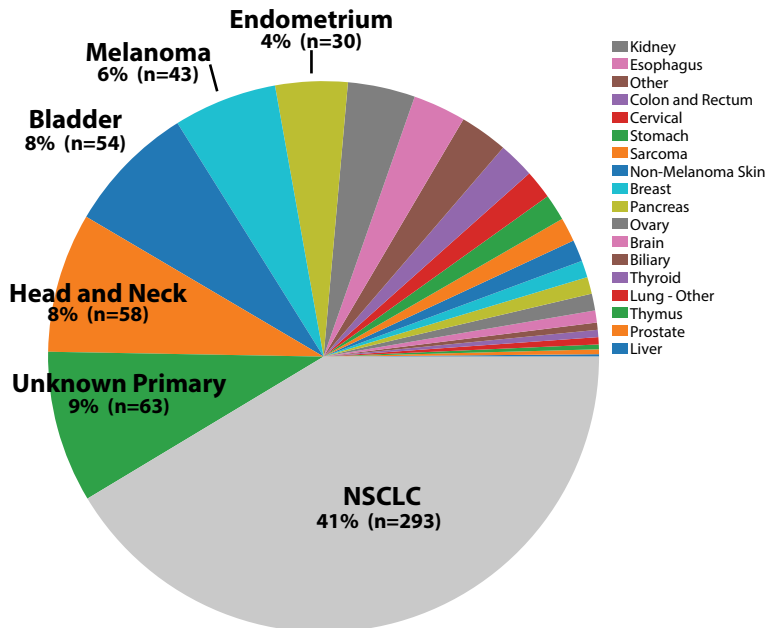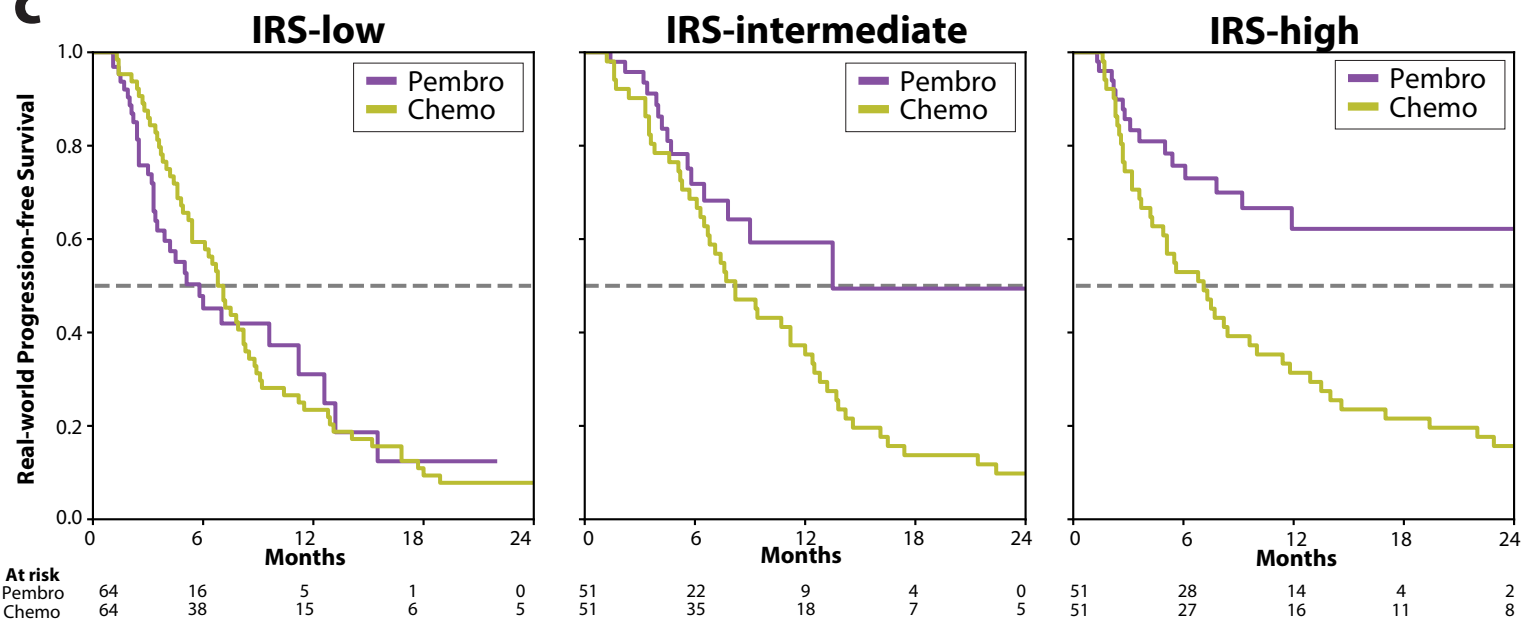
769 **Figure 2. Robustness of the IRS model to potential confounding factors. a)** Real-world progression-

770 free survival on pembrolizumab in patients with non-small cell lung cancer (NSCLC, blue) versus other

771 tumor types (Other, lime green), stratified by IRS group. **b)** Real-world progression-free survival on

772 pembrolizumab in patients treated with pembrolizumab monotherapy (mono, blue) versus pembrolizumab

773 + chemotherapy combination (combo, lime green), stratified by IRS group. **c)** Real-world progression-

774 free survival on pembrolizumab in TMB-low and TMB-high patients, stratified by IRS groups (IRS-low

775 [grey], IRS-intermediate [light blue], or IRS-high [dark blue line]). TMB, tumor mutation burden.
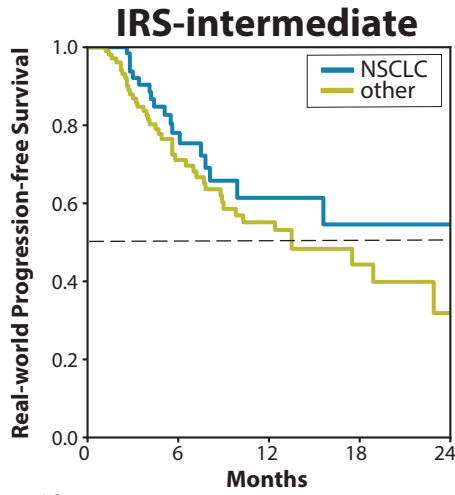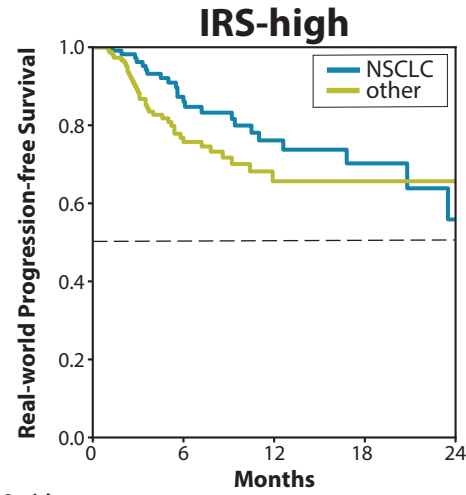
776

Figure 2

a

IRS-low
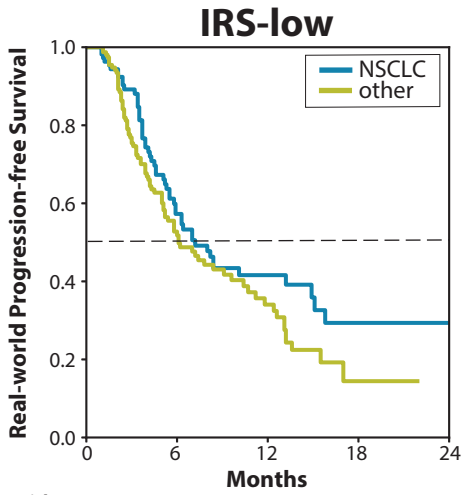
At risk				
NSCLC	108	44	21	7	2
Other	158	51	21	3	0

IRS-intermediate

At risk				
NSCLC	70	29	13	6	3
Other	106	51	28	11	3

IRS-high

At risk				
NSCLC	115	66	35	20	7
Other	151	70	26	8	1

b
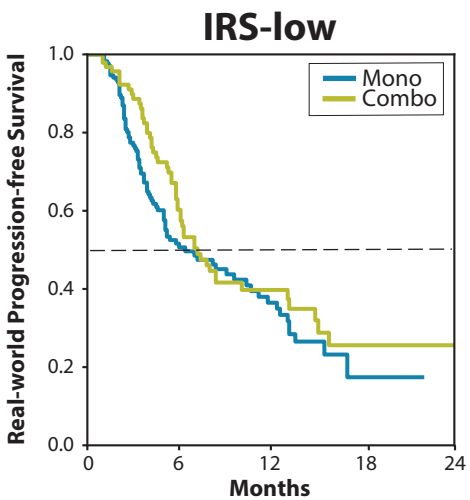
IRS-low

At risk				
Mono	172	52	24	3	0
Combo	94	43	18	7	2
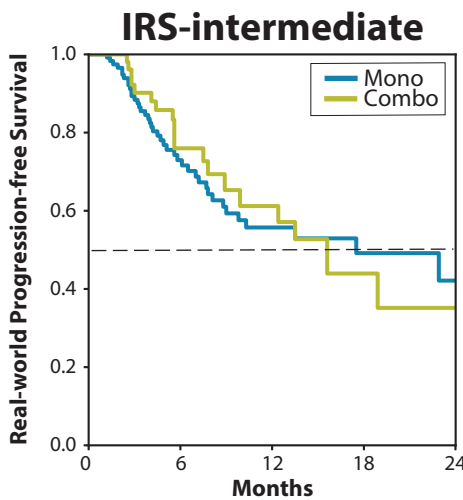
IRS-intermediate

At risk				
Mono	120	54	26	12	4
Combo	56	26	15	5	2
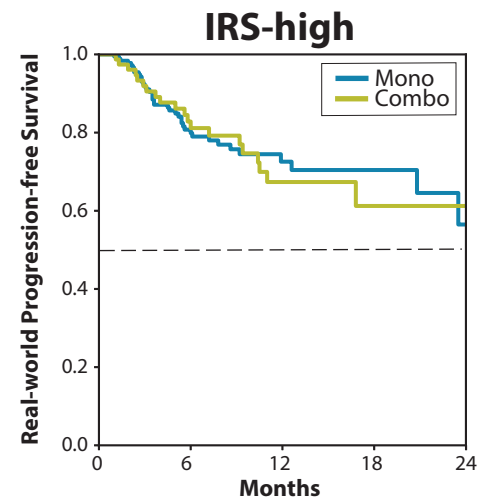
IRS-high

At risk				
Mono	189	90	37	19	6
Combo	77	46	24	9	2

c

TMB-Low
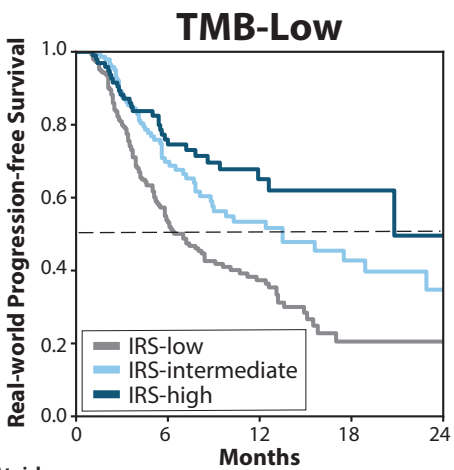
At risk				
IRS-low	253	87	39	9	2
IRS-intermediate	148	65	32	15	5
IRS-high	99	54	24	10	1

TMB-High
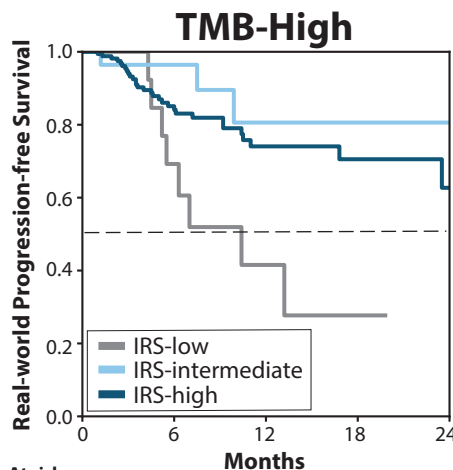
At risk				
IRS-low	13	8	3	1	0
IRS-intermediate	28	15	9	2	1
IRS-high	167	82	37	18	7

33

Figure 3. Distribution of IRS scores across the Strata Clinical Molecular Database (SCMD) to

778 **assess potential clinical utility. a)** IRS groups were determined for all 25,770 patients in the SCMD with

779 valid TMB and gene expression data. The distribution by IRS group (IRS-low [grey], IRS-intermediate

780 [light blue], and IRS-high [dark blue]) is shown. **b**) IRS distribution by pembrolizumab approved vs. not

781 not-approved tumor types. **c)** Stratification of **b**) by tumor type. **d)** IRS distribution by TMB-high vs.

782 TMB-low. **e)** IRS distribution by any pembrolizumab approved indication (approved tumor type, MSI-H

783 *or* TMB-H as approved) vs. not approved indication. Pembro, Pembrolizumab, TMB, Tumor Mutation

784 Burden; CRC, colorectal cancer; CUP, cancer of unknown primary; Lung- Other, lung small cell

785 carcinoma; NSCLC, non-small cell lung cancer; NMSC, nonmelanoma skin cancer; Small Intes., small

786 intestine. Numbers in subpanels may not add to the totals in **a)** due to rounding.

Figure 3

a

Total

25,770 subjects in SCMD

| 13.2 | 17.5 | 69.3 |

b

Pembro Not-Approved | Pembro Approved

Tumor Type

4.2

| | 8.0 | 47.3 | 8.9 | 9.5 | 22.0 |

c

d

Low | High

1.9 0.8

TMB

| 6.3 | 15.7 | 68.5 | 6.8 | |

e

Not Approved | Approved

All Pembro Indications

2.2

| 7.4 | 47.0 | 11.0 | 10.0 | 22.3 |

Legend: ■ IRS-high ■ IRS-intermediate ■ IRS-low
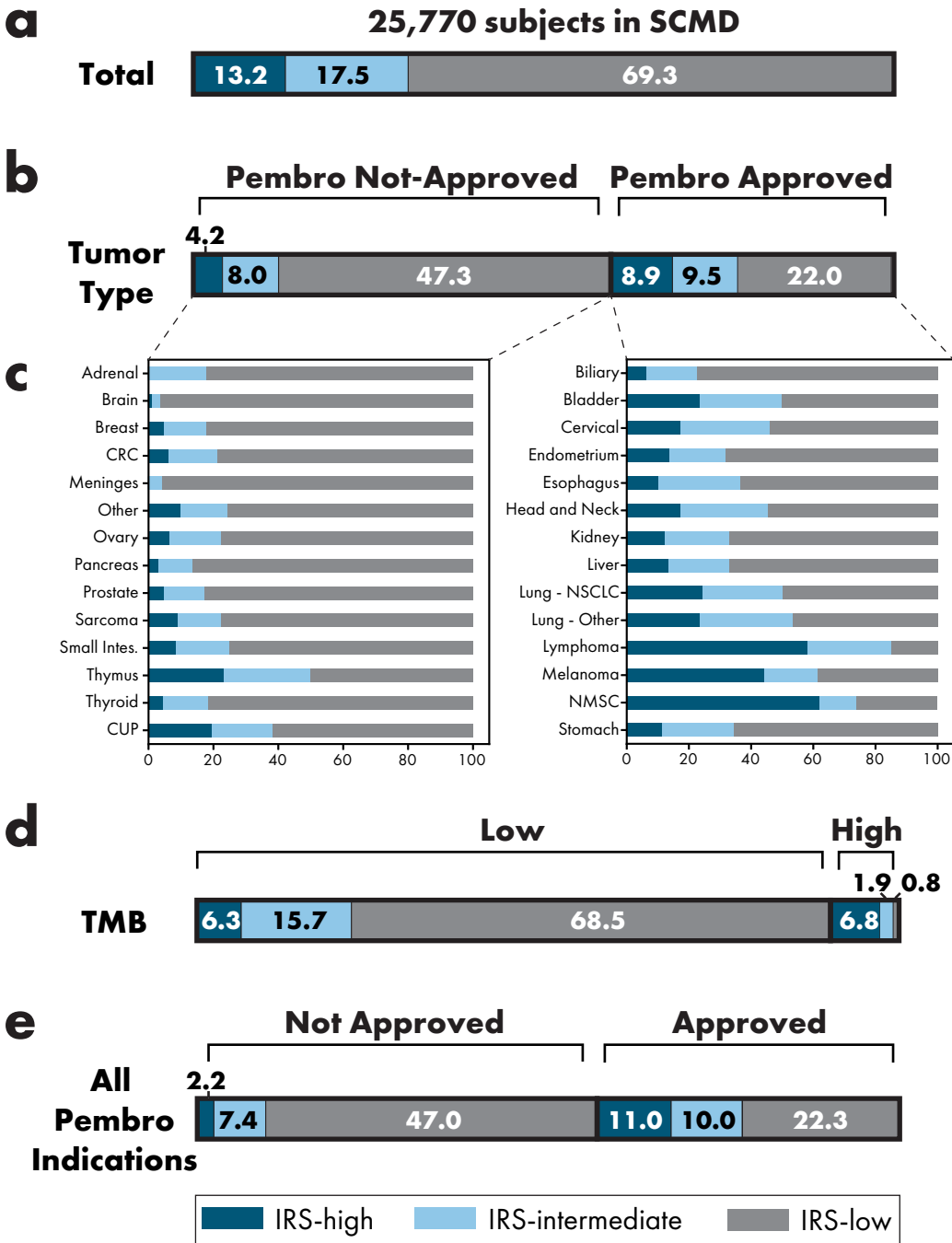
Table 1. Univariate and multivariate associations of comprehensive genomic and quantitative transcriptomic profiling derived candidate biomarkers and real world pembrolizumab progression free survival in 708 patients

Biomarker	Univariate		Multivariate (IRS model)	
	HR (95% CI)	p	HR	p
TMB	0.77 (0.70 – 28.02)	3.7E-09	0.75 (0.69 – 0.82)	5.8E-10
CD274 (*PD-L1*) Composite	0.90 (0.85 – 10.47)	7.1E-04		
PDCD1 (*PD-1*) Composite	0.89 (0.84 – 9.83)	1.1E-03	0.86 (0.79 – 0.94)	5.0E-04
PDCD1LG2 (*PD-L2*)	0.91 (0.85 – 7.77)	4.6E-03	0.87 (0.79 – 0.95)	1.3E-03
TNFRSF9	0.92 (0.87 – 7.07)	7.4E-03		
IDO1	0.95 (0.92 – 5.42)	0.023		
*UBE2C**	0.91 (0.83 – 5.09)	0.029		
TIGIT	0.95 (0.90 – 3.51)	0.088		
LAG3	0.94 (0.88 – 3.19)	0.109		
CD8A	0.95 (0.90 – 3.17)	0.111		
IFNG	0.98 (0.95 – 2.51)	0.176		
TCF7	0.95 (0.87 – 2.35)	0.196		
CTLA4	0.97 (0.92 – 2.10)	0.233		
GZMA	0.97 (0.91 – 1.57)	0.337		
*TOP2A**	0.96 (0.87 – 1.49)	0.356		
VTCN1	0.99 (0.96 – 1.28)	0.413		
FOXP3	0.98 (0.92 – 1.19)	0.439		
ADAM12	1.02 (0.95 – 0.69)	0.620	1.08 (1.01 – 1.15)	0.03
CD4	1.02 (0.93 – 0.63)	0.646	1.15 (1.01 – 1.31)	0.04
AXL	0.98 (0.90 – 0.52)	0.699		
HAVCR2	1.00 (0.93 – 0.11)	0.929		
TNFRSF4	1.00 (0.93 – 0.08)	0.946		

For each biomarker, the hazard ratio (with 95% confidence interval bounds) and log-likelihood p-value are shown. TMB ($\log_2$) was from StrataNGS CGP testing; the remaining biomarkers were target gene expression from in-parallel quantitative transcriptomic profiling by Multiplex PCR-based RNA sequencing. The multivariate analysis was performed using the final five component Immune Response Score (IRS) model. Candidate proliferation markers are indicated by *. For *PD-L1* and *PD-1*, two independent target amplicons were assessed for each gene; normalized target gene expression was averaged from the independent amplicons (per gene) to yield a composite result.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- IRSSuppfinal.pdf