Figures and figure supplements

Early prediction of clinical response to checkpoint inhibitor therapy in human solid tumors through mathematical modeling
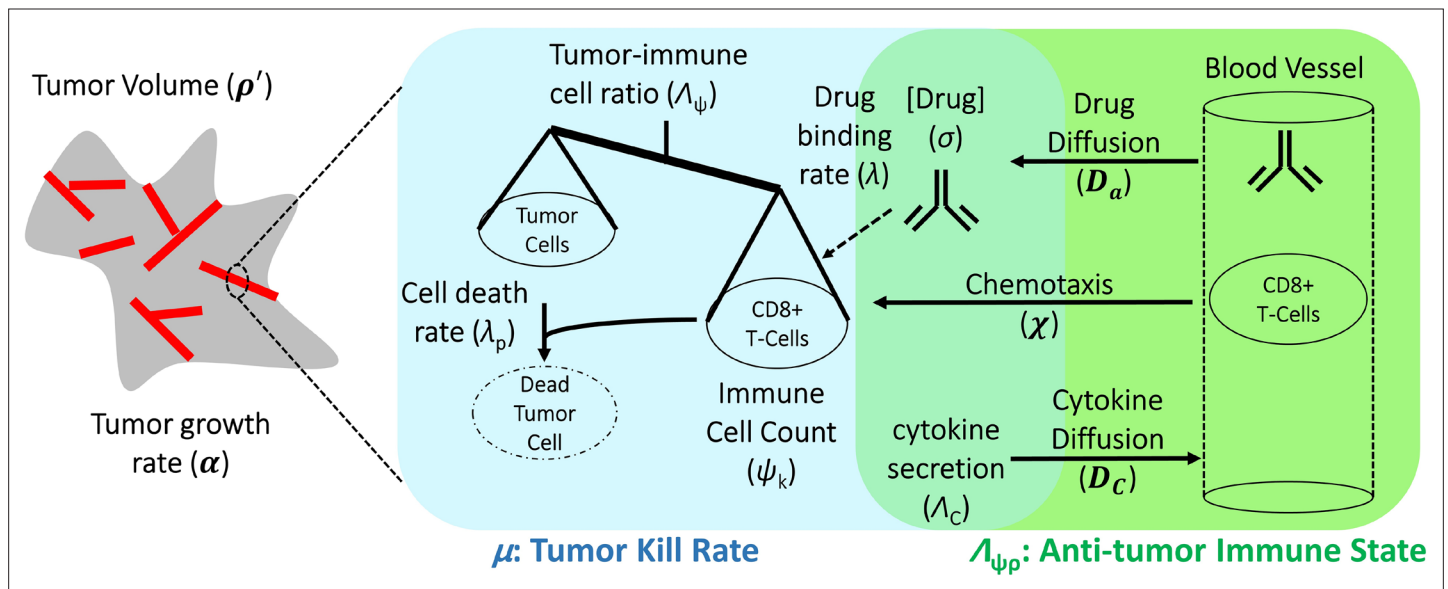
Joseph D Butner *et al*

Figure 1. Schematic representation of biological mechanisms included in the mathematical model. These processes are described by four partial differential equations, which are solved to obtain Equation (1). Briefly, the checkpoint inhibitor enters the tumor via diffusion (D_a) leading to time-dependent drug concentration (σ), which then binds to the conjugate receptor on immune cells at rate λ. Immune cells (ϕ_k) are drawn into the tumor microenvironment via cytokine-mediated chemotaxis (χ), resulting in immune checkpoint inhibitor-mediated cancer cell kill at rate λ_p. The full mathematical model derivation and its underlying assumptions are provided in a recent modeling and analysis report (**Butner et al., 2020**).
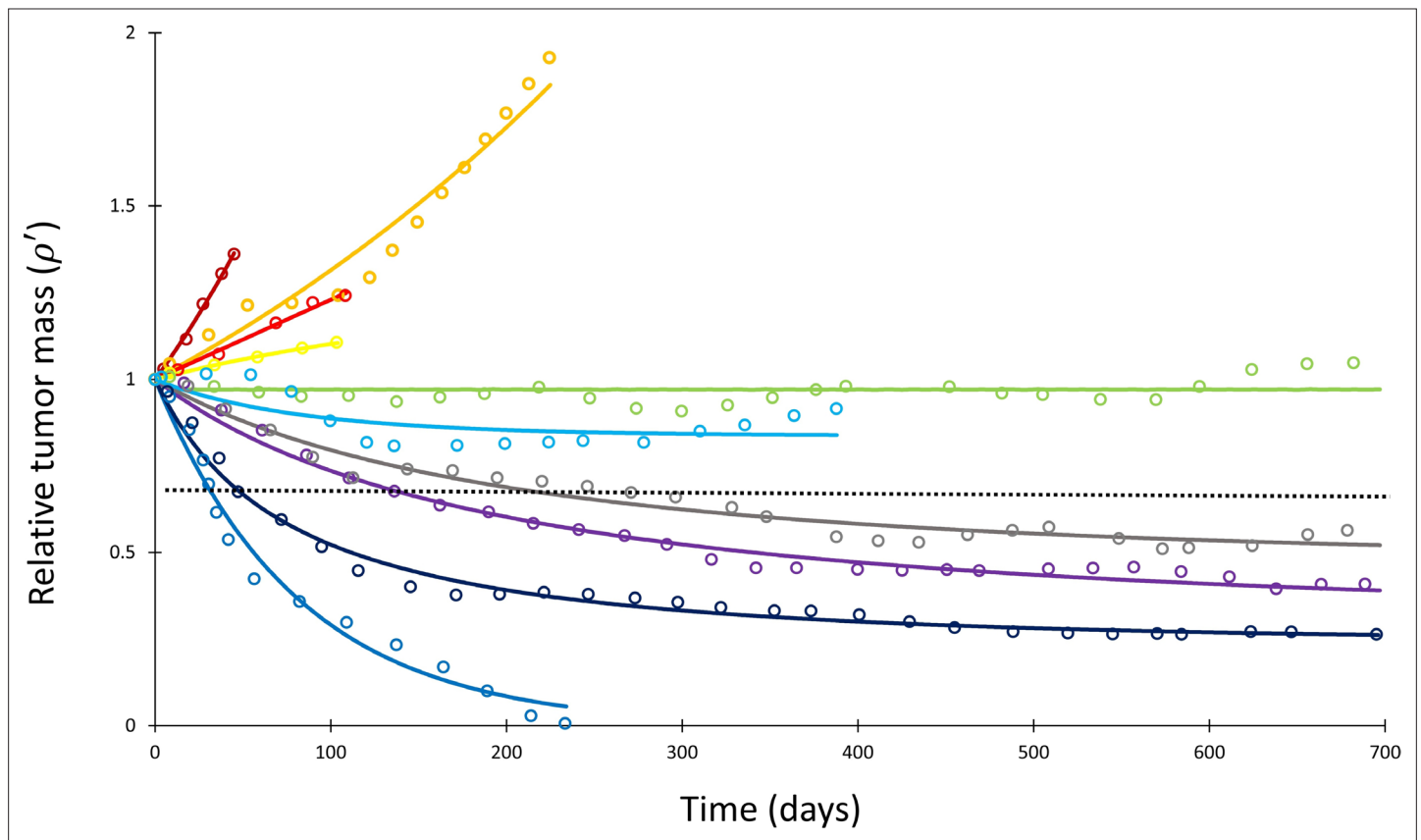
Figure 2. Mathematical model fit to individual responses to immune checkpoint inhibition. Open circles represent data points of clinical response in 10 patients extracted from *Topalian et al., 2012*, while solid lines represent best curve fits of *Equation (1)* to those data (with α^{-1} = 144 days). Each color represents a different patient. Immunotherapy was begun at $t = 0$, and tumor volume was designated as the relative change in volume from $t = 0$ (i.e., tumor volume of 1 at $t = 0$). The dashed line depicts the cutoff used for classifying patients deemed as responders (partial or complete response) versus nonresponders (stable disease or disease progression) according to the RECIST v1.1 criteria.
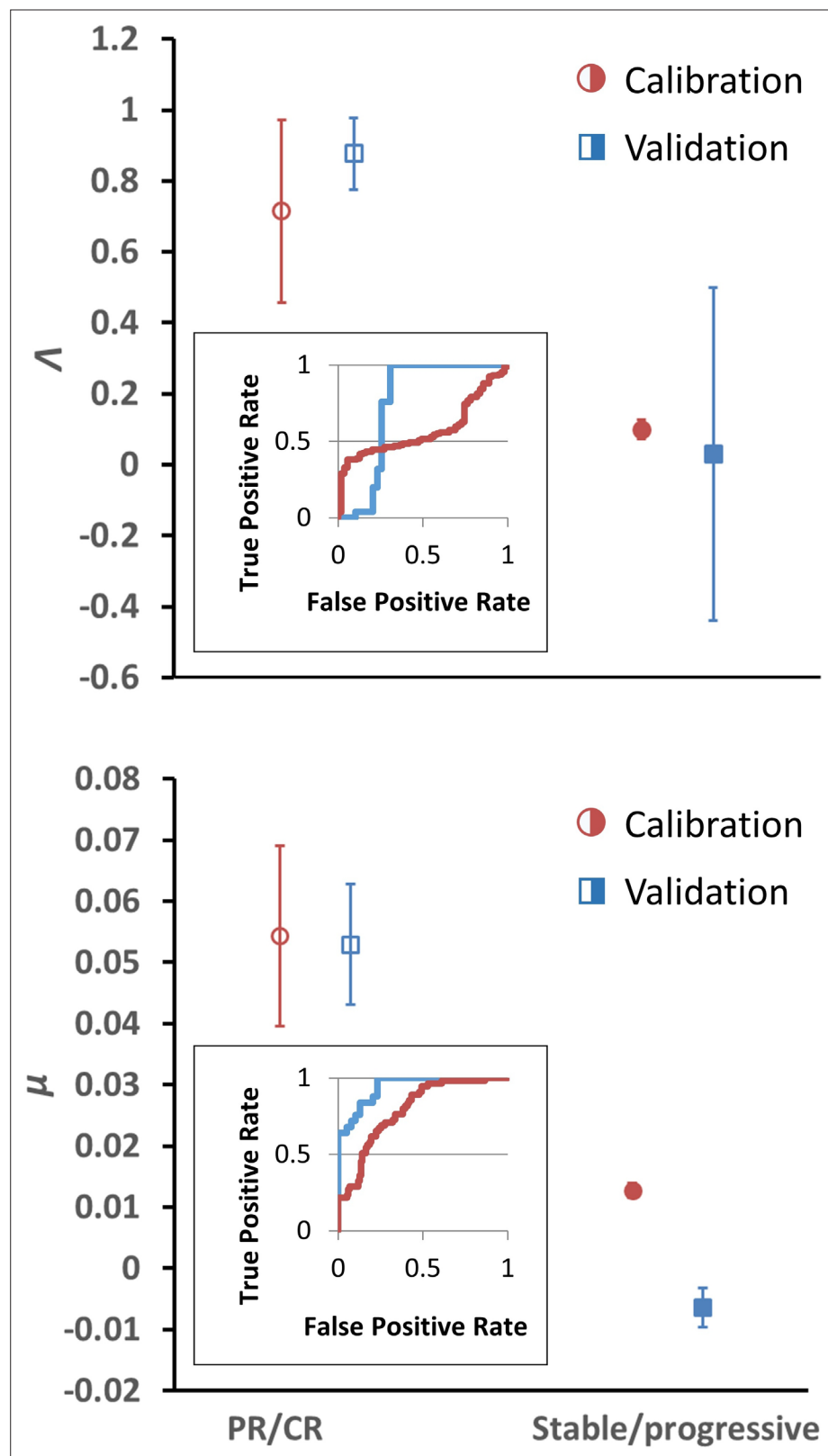
Figure 3. Depiction of average Λ and μ values in patients with response (n = 55) versus nonresponse (n = 134) in the calibration cohort (circular markers), while n = 25 patients had objective response and 39 patients demonstrated stable/progressive disease in the validation cohort (square markers) as determined by RECIST v1.1 criteria. Open markers represent the average values of patients with response, and solid markers represent

Figure 3 continued on next page

Figure 3 continued

patients with stable/progressive disease. Error bars represent the standard error of the mean (SEM). p-Values of separation between groups by Wilcoxon rank sum (two tails): Λ, p=0.119 and p<0.001 for literature (calibration) and non-small cell lung cancer (NSCLC) (validation) cohorts, respectively; μ, p<0.001 for both literature (calibration) and NSCLC (validation) cohorts. Insets: receiver-operator characteristic (ROC) curves for patient response versus model parameters for both cohorts; Λ, literature cohort: sensitivity = 0.381, specificity = 0.945, accuracy = 545; μ, literature cohort: sensitivity = 0.891, specificity = 0.567, accuracy = 0.661; Λ, NSCLC clinical cohort: sensitivity = 0.600, specificity = 0.744, accuracy = 0.688; μ, NSCLC clinical cohort: sensitivity = 0.960, specificity = 0.769, accuracy = 0.844. PR, partial response; CR, complete response. Examples of cancer drug-specific parameter values may be found in *Butner et al., 2020*.
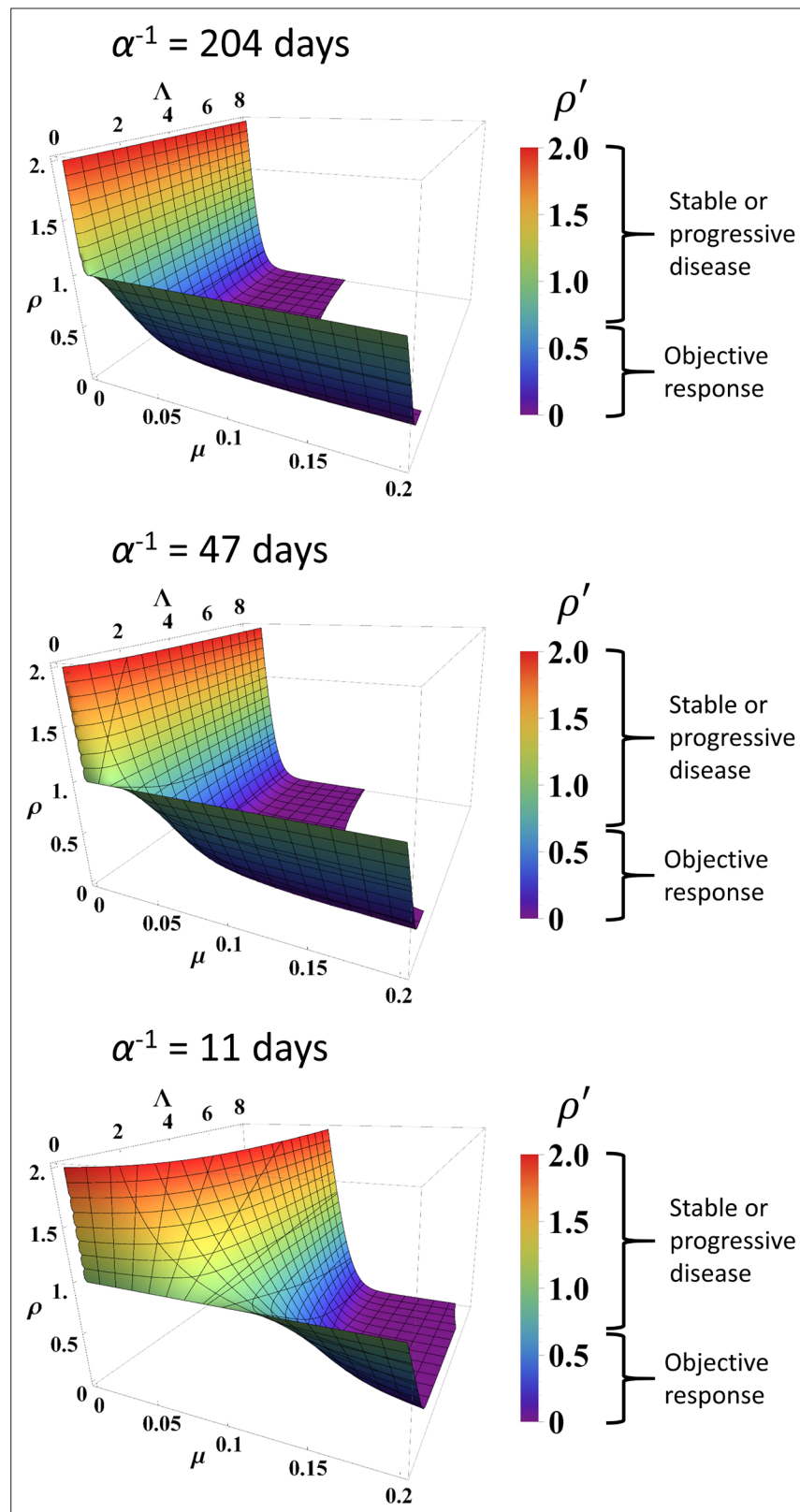
Figure 4. Simulated response to immune checkpoint inhibition at different values of α, Λ, and μ. Data are obtained from **Equation (1)**. Normalized tumor volume (ρ') was determined at $t = 200$ days. Three different α values were used that represent the minimum, average, and maximum values derived from fitting the calibration cohort, as described in the text. Λ and μ were varied continuously over their respective ranges. Colors also correspond with ρ' as per color map on the right. RECIST v1.1 criteria of response are listed to the right of the color bars.
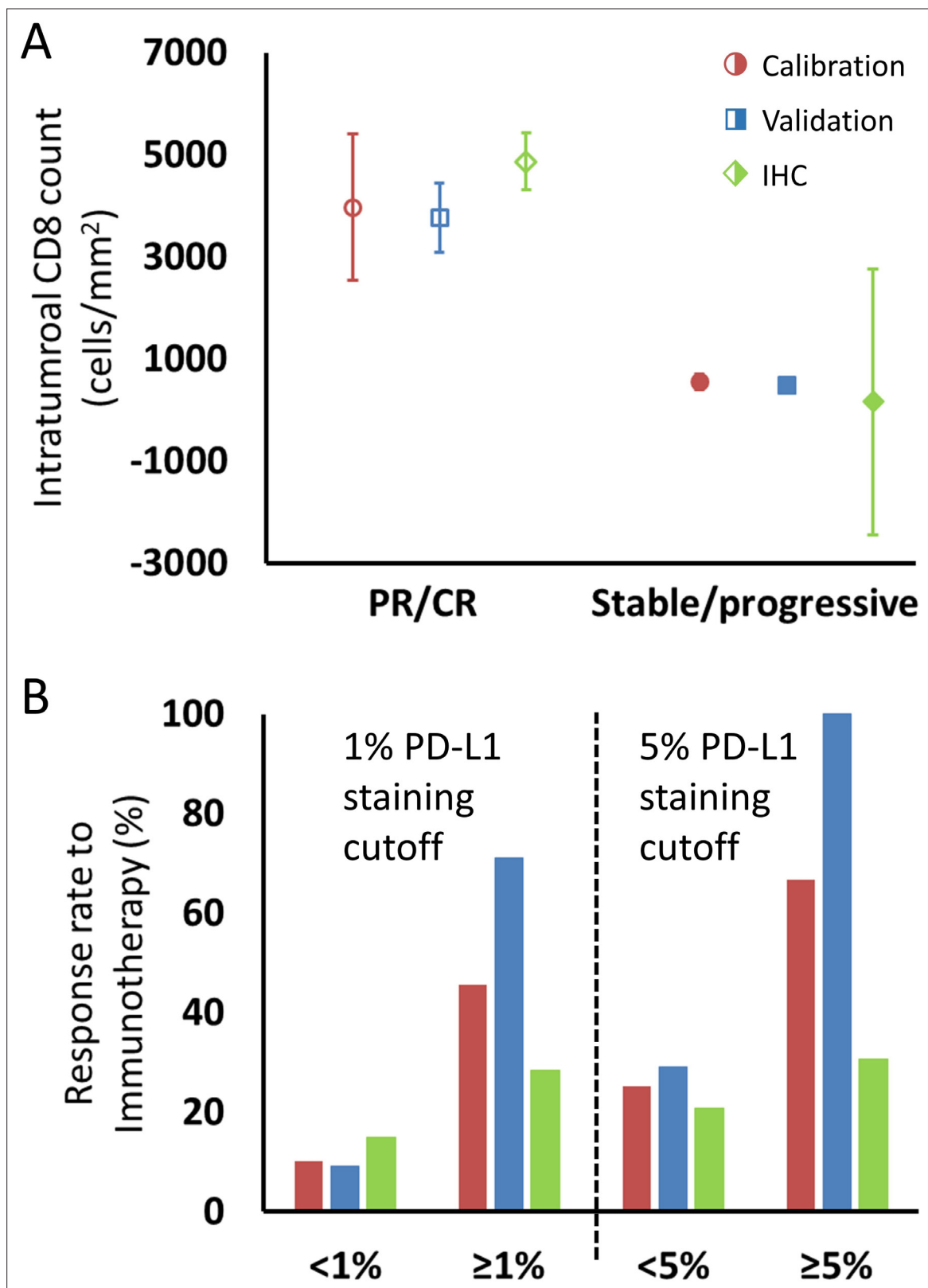
Figure 5. Comparison of intratumoral CD8+ T cell count and tumor PD-L1 staining derived from fitting the model to clinical data and values reported in the literature, as described in the text. (**A**) Model intratumoral CD8+ T cell count (circles: calibration cohort, p=0.119 [Wilcoxon, two-tail]; squares: validation cohort, p<0.001) was derived from Λ and literature CD8 intratumoral count was taken from immunohistochemical (IHC) staining in *Tumeh et al., 2014* in melanoma (diamonds; average CD8 counts including on-treatment values [n = 23]). CD8+ T cell counts from pretreatment biopsies

Figure 5 continued on next page

Figure 5 continued

only (n = 46) demonstrated mean values (± SEM) of 2632 ± 518 cells/mm^2 in patients with response to immunotherapy and 322 ± 133 cells/mm^2 in nonresponding patients, respectively. Values for CD8+ T cell counts are plotted as averages with error bars representing the standard error. (**B**) Patient response rates to immunotherapy stratified by PD-L1 staining were derived from μ from the model (calibration: red; validation: blue) and from references (*Borghaei et al., 2015*; *Robert et al., 2015b*; *Brahmer et al., 2012*; *Tumeh et al., 2014*; *Motzer et al., 2015*; *Powles et al., 2014*; *Topalian et al., 2012*; *Garon et al., 2015*; *Herbst et al., 2014*; *Kefford et al., 2014*; *Spira et al., 2015*; *Taube et al., 2014*; *Weber et al., 2015*) for the literature data (green; n = 975 for 1% cutoff, n = 1492 for 5% cutoff; see Appendix 1—table 1). Response to immune checkpoint inhibition was determined by RECIST v1.1 criteria. PR, partial response; CR, complete response.

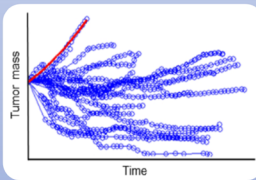**Model Calibration: Fitting the model to measured tumor data**
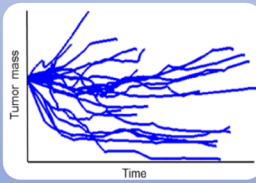
Data collection

Obtain time-course tumor measurements for individual patients

Response curves for checkpoint inhibitor therapy:
- From the literature: obtained with WebPlotDigitizer (see Table 1)
- From in-house clinical study: obtained from in-house patient records

Establish baseline pre-treatment growth rate (α)

Pre-treatment growth rate estimated as an exponential $e^{\alpha t}$

Different approaches for each cohort due to lack of pre-treatment measurements in the literature data:
- Literature data: pre-treatment data unavailable; estimated as fastest growing tumor in each study (red line)
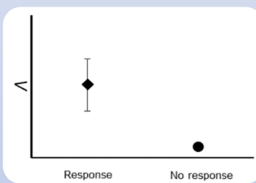- Clinical data: per-patient data were available; calculated per-patient using data before and at start of treatment
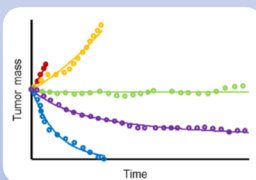
Quantify remaining parameters by fitting Equation 1 to data

Performed numerical regression to obtain best-fit values

Obtained unique per-patient values of Λ and μ
- μ: tumor cell kill rate by activated immune cells
- Λ: anti-tumor immune state
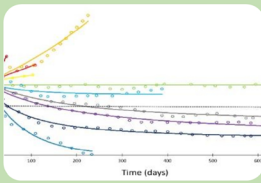
Comparing model values for response criteria

Confirm model parameters are unique for response criteria

RECIST v1.1 response criteria were combined into two categories
- Partial/complete response: last measured tumor burden <70% of tumor burden measured at therapy start
- Stable/progressive disease: last measured tumor burden >70% of tumor burden measured at therapy start

Appendix 1—figure 1. Steps for calibration of the mathematical model with clinical data. First, checkpoint inhibitor response curves were extracted from the literature. In all cases, immunotherapy treatment began at time $t = 0$. Second, a tumor-specific proliferation constant (α) was determined for each cancer type by fitting exponential function ($e^{\alpha t}$ to fastest progressing patient in each clinical trial [red line]). Third, individual patient response data were fit to **Equation (1)** by using the respective α to determine Λ and μ. Λ and μ values were then with compared in patients with partial/complete response versus patients with stable/progressive disease after immunotherapy by using the RECIST v1.1 criteria.

Appendix 1—figure 2. Model validation, sensitivity studies, and comparison of model parameters to immunohistochemical (IHC) measures. Model parameters were obtained from a second in-house patient cohort of patients with non-small cell lung cancer (NSCLC) (n = 64), which were compared to values obtained in the calibration cohort in a validation study. To study the sensitivity of the model to changes in model parameter values, key parameters were perturbed ±10% and the resultant simulated expected tumor burden was compared to measured values pre-perturbation. Tumor burden measures were also truncated, and results of truncated and full dataset model fits were compared. Lastly, the full parameter space of the model was examined. In order to compare model parameters to the underlying biology, model parameters were converted to intratumoral CD8+ lymphocyte counts (for Λ) and PD-L1 staining (for μ), which were compared to IHC measures obtained from the literature.
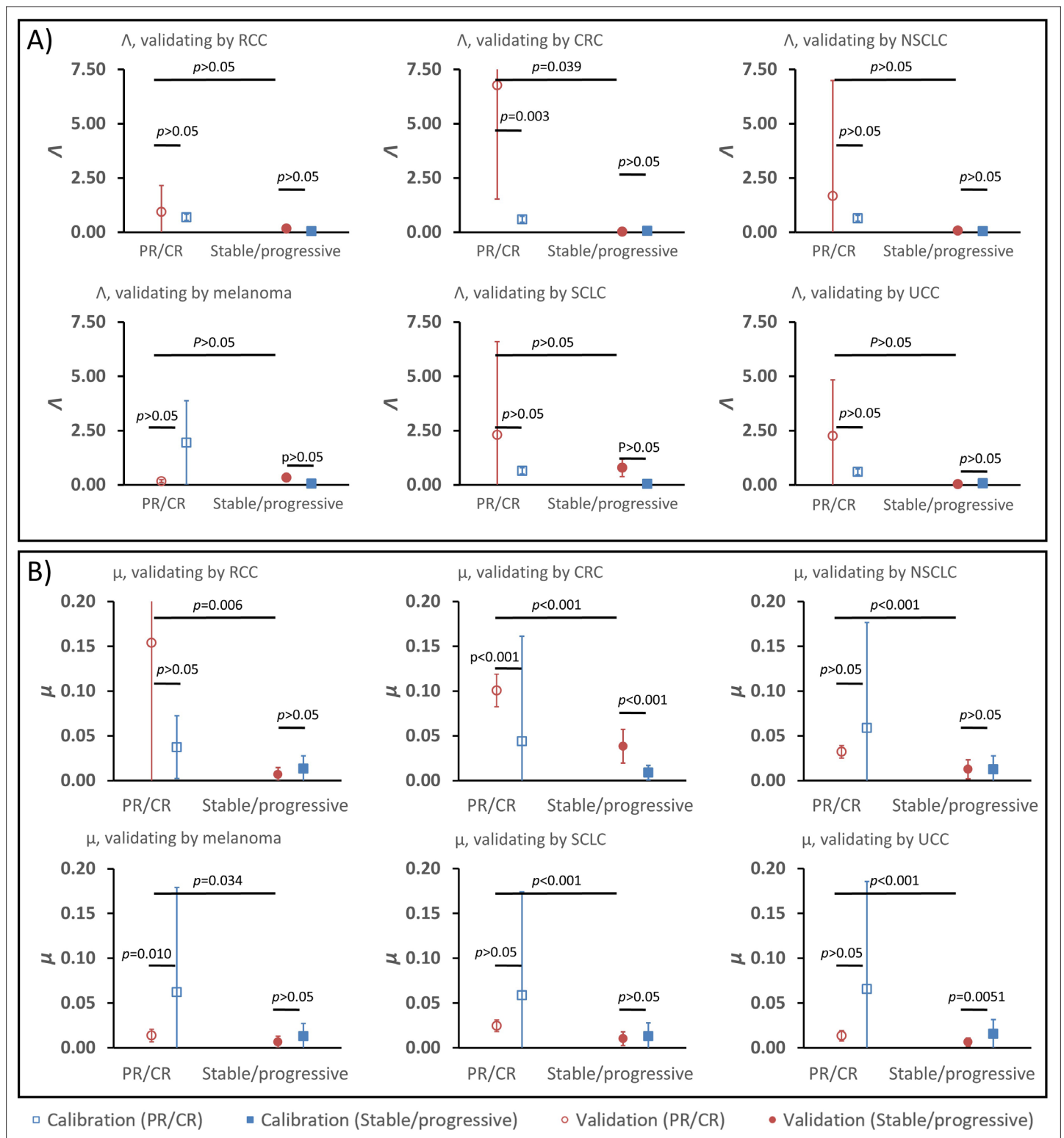
Appendix 1—figure 3. Parameter validation analysis within the calibration cohort. In order to examine the robustness of ranges for (**A**) parameter Λ and (**B**) parameter μ between partial and complete response (PR/CR) versus stable/progressive disease among different cancer types, a validation study was performed where one cancer type was removed from the calibration cohort and used as validation against the parameter ranges in the reduced calibration set obtained from *Borghaei et al., 2015*; *Antonia et al., 2015*; *Le et al., 2015*; *Motzer et al., 2015*; *Powles et al., 2014*; and *Topalian et al., 2012*. Analysis was repeated once for each cancer type, and

Appendix 1—figure 3 continued on next page

Appendix 1—figure 3 continued

results are shown as mean ± standard deviation (error bars). Parameter ranges were found to vary between individual cancer types, and with μ showing more consistent significant difference between response categories relative to Λ (these results are consistent with results shown in *Butner et al., 2020*).