Early prediction of clinical response to checkpoint inhibitor therapy in human solid tumors through mathematical modeling

Joseph D. Butner^{1,#}, Geoffrey V. Martin^{2,#}, Zihui Wang^{1,3,#,*}, Bruna Corradetti^{4,5}, Mauro Ferrari⁴, Nestor Esnaola⁶, Caroline Chung¹, David S. Hong⁷, James Welsh¹, Naomi Hasegawa⁸, Elizabeth A. Mittendorf⁹, Steven A. Curley¹⁰, Shu-Hsia Chen^{11,12}, Ping-Ying Pan^{11,12}, Steven K. Libutti^{13,14}, Shridar Ganesan¹³, Richard L. Sidman¹⁵, Renata Pasqualini^{16,17}, Wadih Arap^{16,18}, Eugene J. Koay^{2,*}, Vittorio Cristini^{1,3}

¹Mathematics in Medicine Program, Houston Methodist Research Institute, Houston, TX 77030, USA.

²Department of Radiation Oncology, The University of Texas M.D. Anderson Cancer Center, Houston TX 77030, USA.

³Department of Imaging Physics, The University of Texas M.D. Anderson Cancer Center, Houston TX 77030, USA.

⁴Department of Nanomedicine, Houston Methodist Research Institute, Houston, TX 77030, USA

⁵Swansea University Medical School, Singleton Park, Swansea SA28PP, UK.

⁶Department of Surgery, Houston Methodist Cancer Center, Houston, TX 77030, USA.

⁷Department of Investigational Cancer Therapeutics, The University of Texas M.D. Anderson Cancer Center, Houston, TX 77030, USA.

⁸University of Texas Health Science Center (UTHealth), McGovern Medical School, Houston, TX 77030, USA.

⁹Breast Oncology Program, Dana Farber/Brigham and Women's Cancer Center, Boston, MA 02215, USA.

¹⁰Michael E. DeBakey Department of Surgery, Baylor College of Medicine, Houston, TX 77030, USA.

¹¹Immunotherapy Research Center, Houston Methodist Research Institute, Houston, TX 77030, USA.

¹²Cancer Center, Houston Methodist Research Institute, Houston, TX 77030, USA.

¹³Rutgers Cancer Institute of New Jersey, New Brunswick, NJ 08901, USA.

¹⁴Department of Surgery, Rutgers Robert Wood Johnson Medical School, New Brunswick, NJ 08901, USA.

¹⁵Department of Neurology, Harvard Medical School, Boston, MA 02115, USA.

¹⁶Rutgers Cancer Institute of New Jersey, Newark, NJ 07101, USA.

¹⁷Division of Cancer Biology, Department of Radiation Oncology, Rutgers New Jersey Medical School, Newark, NJ 07103, USA.

¹⁸Division of Hematology/Oncology, Department of Medicine, Rutgers New Jersey Medical School, Newark, NJ 07103, USA.

#J.D.B., G.V.M., and Z.W. contributed equally to this work.

*To whom correspondence may be addressed. Email: zwang@houstonmethodist.org, EKoay@mdanderson.org

Running Title: Mathematical model of cancer response to immunotherapy

Keywords: biomarkers; immunotherapy; patient stratification; PD-1; PD-L1; translational research

6,925 words, 5 figures, 2 tables, 2 supplemental figures, and 1 supplemental table

ABSTRACT

Checkpoint inhibitor therapy of cancer has led to markedly improved survival of a subset of patients in multiple solid malignant tumor types, yet the factors driving these clinical responses or lack thereof are not known. Here we present a translational mathematical model dependent on three key parameters for describing efficacy of checkpoint inhibitors in human cancer: tumor growth rate (α), tumor immune infiltration (λ), and immunotherapy-mediated amplification of anti-tumor response (μ). The model was validated by fitting it to a compiled clinical tumor response dataset ($n = 189$ patients) obtained from published anti-PD-1 and anti-PD-L1 clinical trials, and then validated on an additional validation cohort ($n = 64$ patients) obtained from our in-house clinical trials. The derived parameters λ and μ were both significantly different between responding versus non-responding patients. Of note, our model appropriately classified response in 81.4% of patients by using only tumor volume measurements and within two months of treatment initiation in a retrospective analysis. The model reliably predicted clinical response to the PD-1/PD-L1 class of checkpoint inhibitors across multiple solid malignant tumor types. Comparison of model parameters to immunohistochemical measurement of PD-L1 and CD8+ T cells confirmed robust relationships between model parameters and their underlying biology, demonstrating reliable methods to inform model parameters directly from biopsy samples, which are conveniently obtainable as early as the start of treatment. Together, these results suggest that the model parameters may serve as early and robust biomarkers of the efficacy of checkpoint inhibitor therapy on an individualized per-patient basis.

68 **SIGNIFICANCE**

69 Checkpoint inhibitors have revolutionized cancer immunotherapy, but only a subset of patients with
70 solid tumors responds clinically. The ability to predict tumor responses *a priori* or soon after starting
71 therapy would allow for personalized and timely adaptive clinical applications of checkpoint inhibitor-
72 based immunotherapy in patients. By applying a mechanistic mathematical model, we show that
73 checkpoint inhibitor therapeutic effectiveness is accurately predictable in most patients within two
74 months after treatment initiation. Our method may be implemented directly into clinical practice, as it
75 relies on standard-of-care imaging and pathology. If successful in prospective studies, this model will
76 improve selection of cancer patients for checkpoint inhibitor therapy, and perhaps for other forms of
77 humoral- or cell-based immunotherapy.

78

79

INTRODUCTION

Recent advances in the understanding of immunological pathways responsible for antibody- and/or cell-mediated destruction of tumors has led to the development of unique cancer therapeutics in recent years, leading to markedly improved survival in the setting of previously intractable metastatic melanoma, bladder, kidney, lung, and head and neck cancers, among several other human solid tumor types (1-3). In particular, one of the most successful clinical applications of immune checkpoint inhibitors include antibodies directed against the programmed death protein 1 (PD-1) pathway, which inhibits cellular immune killing of cancer cells via complementary binding of tumor expressed programmed death ligand 1 (PD-L1) to PD-1 on immune cells (4). As a still emerging yet quite compelling immunotherapy approach in contemporary cancer medicine, targeting of immune checkpoints is being extensively investigated in ongoing and upcoming clinical trials aimed at unleashing T-cell activity, augmenting immune recognition against tumor metastases, and boosting immune-memory for long-lasting clinical remission post-treatment (5, 6). While the remarkable potential for checkpoint inhibitors in treating cancer is unequivocally exciting, the combined clinical trial experience has shown that durable effective treatment outcomes occur only in a limited subset of patients (7). Alas, some cancer types are presumed to be minimally immunogenic, resulting in little or no response to this treatment strategy (8). An accumulating body of evidence has established that certain immunological features, including T-cell exhaustion and exclusion (9), senescence markers (10) such as CD27, Tim-3, CD57, and/or T-cell receptor (TCR) repertoires (11) could perhaps reflect or even predict sensitivity and resistance to checkpoint inhibitor-based cancer immunotherapy. However, early attempts at identification of specific pathological biomarkers to predict immunotherapy response have thus far shown somewhat inconsistent results, challenged further by the inherent molecular, cellular, and biophysical diversity of human tumors (12-14). Considering the sheer complexity of the biological interactions between the immune system and the tumor microenvironment, one could reason that the

introduction of more sophisticated mathematical analytic techniques would hopefully have potential to enhance the qualitative and quantitative understanding of such interactions and to improve malignant tumor treatment with checkpoint inhibitors, ultimately supporting the development of a predictive clinical tool that could either minimize or overcome this clear unmet need of cancer medicine.

In previous work, mechanistic mathematical models of immunomodulatory interventions on cancer control utilizing coupled ordinary differential equations (ODE) have already allowed the prediction of tumor response after Bacillus Calmette-Guérin (BCG) administration in superficial transitional cell carcinoma of the urinary bladder and the serum prostate-specific antigen response after prostate cancer vaccine administration with considerable accuracy (15-17). Moreover, other investigators have shown that mechanistic models of interleukin-21 (IL-21) therapy schedules based on tumor mass and antigenic properties can predict growth patterns of multiple tumor types on patients receiving personalized doses of IL-21 (18). In another notable work, mathematical modeling of neoantigen fitness based on cancer population evolution and antigen recognition potential was able to predict tumor response and patient survival following treatment with checkpoint blockade therapy, as validated in both melanoma patients treated with anti-CTLA4 and in lung cancer patients treated with anti-PD-1 therapies (19). Finally, modeling of chemotherapy and targeted drug therapy by our group with similar mathematical techniques was able to accurately reproduce entire dose-response curves of tumor cell kill with untargeted cytotoxic and ligand-directed agents (20-24). These examples highlight the previous success of mathematical models to qualitatively or quantitatively discern the underlying biologically and physically relevant, often non-linear processes present in cancer, which may otherwise be missed, and help optimize treatment delivery approaches.

The purpose of the present study is to demonstrate how standard clinical measures, including radiological imaging assessments and immunohistochemical (IHC) analysis of tissue samples, may inform a mechanistic mathematical model to elucidate patient and tumor-specific attributes that would likely benefit from the application of checkpoint inhibitor-based therapy. The mathematical model presented in this report is based on an extensive series of prior methodological reports (20-34), but with rigorous emphasis on parameters related to the cancer immune response to the treatment with checkpoint inhibitors in the setting of patients with solid tumors. In particular, we have developed a model that contains key mechanistic biological and physical processes involved in checkpoint inhibitor therapy, which has been objectively validated against imaging data of patient tumor burden and measured response, defined by both change in total tumor burden and patient survival (35). To date, this modeling work has only examined tumor response on a bulk scale, without confirming the mechanistic links between key model parameters and the underlying biology they describe. In this work, we demonstrate how model parameters may be informed by using cell-scale IHC analysis of biopsied tissues, which are available as early as at the start of treatment, for reliable prediction of therapeutic response. Thus, this represents a key step towards clinical translation, (i) by validating prior results in additional patient cohorts, and (ii) by demonstrating how IHC may be used with the model to predict patient response at times closer to the start of treatment. The predictions obtained through this model were compared against available retrospective clinical data from published trials with monoclonal antibodies blocking the PD-1/PD-L1 pathway for validation, thereby focusing on a quantitative relationship between tumor response and its underlying mechanisms. Taken together, these retrospective results indicate that our predictive mathematical model with the class of PD-1/PD-L1 checkpoint inhibitors has translational merit and that it should therefore be evaluated as an integral predictive biomarker in the setting of carefully designed prospective cancer investigational trials.

METHODS

Mathematical Model

The final mathematical model we used for comparison with clinical data is a single ODE, which determines changes in relative tumor mass over time after initiation of anti-PD1 therapy. It is a simplified and user-friendly version originated from a complex set of partial differential equations, which takes into account spatial relationships within the tumor microenvironment. The full mathematical derivation has been demonstrated elsewhere (35), but its underlying cancer biology is also schematically depicted in Figure 1. It is ultimately represented by:

$$\rho' = \frac{\rho_{\infty}}{1 - (1 - \rho_{\infty})e^{-(\alpha - \mu + \mu \cdot \Lambda) \cdot t}}, \quad (1)$$

where ρ' is the normalized tumor mass, t is time, α is a proliferation constant of tumor cells, Λ is the ratio of cancer cells (σ) to effector immune cells multiplied by the ability of immune cells (ψ) to kill cancer cells (namely, the anti-tumor immune state), μ is the effect of immunotherapy on the ability of immune cells to kill cancer cells, and ρ_{∞} represents the final long-term tumor burden, which may be calculated as:

$$\rho_{\infty} = 1 + \frac{\alpha - \mu}{\mu \cdot \Lambda}. \quad (2)$$

In words, Equation (1) describes the time-course tumor burden under immunotherapy intervention, as a function of three key parameters α , Λ , and μ , which represent the ability of malignant tumor cells to grow, the ability of immune cells to kill cancer cells within the tumor environment, and the potential efficacy of checkpoint inhibitor-based immunotherapy treatment. Equation (2) states that long-term tumor burden is a function of the relationship between tumor growth and kill rates [the greater of which sets the sign of response, i.e., progressive disease ($\alpha > \mu$) or favorable clinical response ($\alpha < \mu$)], as scaled by the parameter $\mu \cdot \Lambda$. These three parameters are intrinsically abstract terms, which have

multiple components in themselves. Λ is dependent on the density of malignant tumor versus immune cells at the initiation of immunotherapy ($t = 0$) and on the ability of each immune cell to kill tumor cells. Equation (1) assumes the following five assumptions: (i) Tumor mass grows exponentially with rate α in the absence of immune cell killing, (ii) cancer cells are killed by one type of net effector immune cell (termed ψ), (iii) the influx of immune cells and the concentration of checkpoint inhibitor within the tumor environment reaches a drug steady-state on timescales far faster (i.e., from hours to days) than clinical response (i.e., from weeks to months), (iv) all cancer cells are within an average intratumoral movement path length of effector immune cells, and the effect of cytokines in stimulating immune cell function/movement is small, thus spatial diversity effects are negligible in this particular case, and (v) the site of action of the checkpoint inhibitor takes place at the interface between the tumor cell and immune cell, with blocking of one of the immune checkpoint receptors, either on the cancer or on the immune cell, which are deemed necessary and sufficient for blockade of the immune inhibitory pathway.

While one must recognize that although these assumptions certainly represent a gross biological generalization of the true tumor-immune microenvironment complexity, they enable direct comparison with clinical response data, in which measurements of all the parameters required for a full description of immune system reaction to tumors are often not readily available in conventional clinical settings. In this retrospective mechanistic analysis, we start with a simple biophysical model to validate clinical findings and the mechanistic links between high-level model parameters and the underlying biology they describe, with the possible addition of further refining variables as an aspirational goal for future prospective translational or clinical cancer studies. The mathematical simulations and procedures conducted are fully described in the Supplementary Information (Figures S1 and S2); finally, we also address the implications of these arbitrary choices in the Discussion.

Clinical Evaluation and Fitting of the Mathematical Model

To evaluate Equation (1) with clinical data, we initially focused on clinical trials with the class of checkpoint inhibitors that specifically target the PD-1/PD-L1 pathway, although the model may be adaptable to other types of checkpoint inhibitors affecting the CTLA-4 pathway or perhaps even to other cell-based immunotherapies, such as CAR T-cells, as appropriate. In this analysis, tumor volumes over time measurements were obtained from six published clinical trial reports (1, 36-40). As an initial retrospective proof-of-concept, five clinical trials of anti-PD-1 therapy plus another clinical trial of anti-PD-L1 treatment (Table 1) were included, and the combined dataset representing several common solid tumor pathology types ($n = 189$ cancer patients); this dataset represents our model calibration cohort. An online tool [Web Plot Digitizer (41)] served to extract the individual patient tumor response data over time from the published clinical trials, and subsequently Equation (1) was fit to these individual and collective datasets to derive the parameters α , λ , and μ .

The fitting procedure was performed in two sequential steps. First, the fastest progressing patient in each trial was fit to a simple exponential function to estimate α for each cancer histopathology (i.e., approximation for uninhibited tumor doubling-time). Although this approximation represents a limiting factor in the study, it was necessary due to lack of pre-treatment tumor measurement in the calibration cohort, which would allow for more accurate estimation of the growth kinetic without clinical intervention. As presented, α may represent the fastest tumor growth potential in the cohort thus and overestimate the true average growth kinetic; however, because it likely represents the patient with the least response to treatment in each cohort, we take it to be a reasonable approximation of tumor growth kinetic unaffected by treatment. Second, this cancer-specific α was used to determine the

remaining parameters, Λ and μ , for each individual patient by fitting Equation (1) to time-course tumor burden data (Figure 2 and Figure S1). The curve fitting was performed by using non-linear least squares and the Mathematica function NonLinearModelFit (42). The workflow for deriving the model parameters from the clinical data is depicted in Figure S1.

The results from this model calibration was checked against an additional validation patient cohort from the University of Texas M.D. Anderson Cancer Center (MDACC). Briefly, data from patients ($n = 64$) with non-small cell lung cancer (NSCLC) treated with pembrolizumab (MK-3475) were obtained from the investigational study NCT02444741. Serial lung tumor measurements were taken on post-contrast chest CT images acquired with 2.5 mm slice thickness. All lesions that were previously indexed by the radiologist were included in order to best reproduce actual clinical practice (patients had a range of 1 to 12 indexed lesions; median, 3 indexed lesions). Lesion volumes were estimated as 3D spheres calculated from the geometric mean of long and short axes of each lesion at each follow-up, and then volumes were summed to generate a total volume for all indexed lesions at each follow-up time point. Because imaging data before the start of pembrolizumab treatment was available for each patient within this cohort, we were able to obtain a patient-specific pre-treatment growth rate α for each patient between the imaging immediately preceding start of treatment and at the time of first immunotherapy treatment by using the exponential growth-rate estimation $\rho(t) \approx e^{\alpha \cdot t}$ between tumor burdens measured via imaging collected before treatment and at start of treatment. These per-patient α values were then used in Equation (1) to obtain values for Λ and μ , as described above.

Sensitivity Analysis and Model Simulations

A sensitivity analysis was next performed on Equation (1) in two independent ways (Figure S2). First, each of the parameters λ and μ were individually altered by $\pm 10\%$ while holding the other parameters constant to simulate the change in relative tumor mass (ρ'). Second, individual values of ρ' in the calibration cohort were varied by a uniform random variable by ± 0.1 (i.e., 10%) and Equation (1) was re-fitted to these modified data to determine any changes in the derived values for λ and μ . To determine stability of derived λ and μ values over time, we also truncated the ρ' data at differing time points (namely, $t < 30, 60, 120, 200$ days), and subsequently re-fitted Equation (1) to this truncated data to determine λ and μ . The correlation between the derived values of λ and μ from the truncated datasets and the full composite dataset was then determined by using Spearman correlation analysis. A simulated long-term RECIST v1.1 score was then determined for each patient by projecting the expected tumor burden at $t = 700$ days (chosen as a time point beyond the last reported follow-up for most of the patients) by using these values of λ and μ derived from the truncated fits in order to compare the predicted RECIST scores to the known retrospective scores. Finally, full model simulations were performed at different levels of α , λ and μ to determine values of ρ' based on input parameter variation.

Validation of λ and μ by Tumor-infiltrating Immune Cells and Immunostaining

The parameters λ and μ derived from fitting the clinical response data were compared to pathologic biomarkers from additional anti-PD-1 or PD-L1 clinical trials (1, 3, 8, 13, 38-40, 43-48), as shown in Table S1. Specifically, we investigated the ability of λ and μ to correlate with CD8+ tumor-infiltrating lymphocytes (TILs) and tumor PD-L1 expression, respectively. To compare the values of λ and μ derived from fitting the clinical response data over time with the trials assessing TIL and PD-L1 expression, the extracted cancer patients were categorized into objective responders versus non-responders by applying the Response Evaluation Criteria in Solid Tumors (RECIST) v1.1 (49); note that resist categories

determined by lesion diameter (D) or our volume estimation are mathematically comparable by

$$\text{Volume} = \frac{4}{3} \pi \left(\frac{D}{2} \right)^3.$$

Specifically, patients from all RECIST categories (49) were condensed into two groups for analysis, with patients who had $\geq 30\%$ reduction in tumor burden (partial or complete response) grouped into the “favorable” response group, and patients with either $< 30\%$ disease reduction (stable disease or disease progression) into the “unfavorable” response group. The last recorded time point from the response curve from each individual patient served to designate the RECIST v1.1 category, as this provides the most accurate long-term treatment outcome.

After stratifying each patient into a RECIST v1.1 category, Λ and μ were compared between responding and non-responding patients by using a Wilcoxon rank sum test and compared to the literature either as a continuous variable (Λ) or by using thresholds (μ). For comparison with biomarkers reported in the literature, Λ was converted to an estimated intratumoral CD8+ T-cell count (for details, the interested reader is referred to (35)) by assuming each CD8+ T-cell would kill one tumor cell on average (the assumed mean “fitness” of the immune cell population (50)), and that there were 5,558 cells/mm² in the tumor microenvironment, as has been quantitatively measured in melanoma (51). For comparison with PD-L1 staining, μ was converted from its raw numerical value to a percentage; no other scaling of the variable was performed as the number of cancer cells bound by anti-PD-1/PD-L1 therapy action is the dominant term in the integral specified in Equation S4 in ref. (35); this approximation and its implications are further examined in Discussion. Assessment of objective response rates (ORR) at standard thresholds of 1% and 5% PD-L1 staining (corresponding to μ) were used in this study. Other alternative PD-L1 cutoffs present in some studies (Table 1 and Table S1) were not included in this analysis due to the relatively small numbers of patients. The biomarker trials used for validation of these parameters are also presented in Table S1.

RESULTS

Quantification of Model Parameters α , Λ and μ

Measured tumor burden data over time from the total collective calibration patient pool ($n = 189$) after initiating checkpoint inhibitor-based therapy were extracted from the six reported clinical trials (Table 1). Of this population, 55 patients (29%) had objective responses by RECIST v1.1 criteria while 134 patients (71%) demonstrated stable/progressive disease. The derived tumor proliferation constant, α , was determined to range from 0.0034 (tumor doubling-time of ~ 200 days) in renal cell carcinoma to 0.0622 (tumor doubling-time of ~ 11 days) in non-mismatch repair colon cancer with an average α of 0.018 and average tumor doubling-time of ~ 85 days (Table 1). In the validation NSCLC cohort, 25 patients (39%) had objective response, and 39 patients (61%) demonstrated stable/progressive disease, with individual patient proliferation constants (α) ranging from -0.0129 to 0.0602 (note that $\alpha < 0$ indicates a subset of patients that had a shrinking tumor burden before start of therapy).

Having determined α for each cancer histopathology, the clinical response curves to checkpoint inhibitors for each tumor type were fit to Equation (1) to determine Λ and μ . The average root-mean-square error of fitting Equation (1) to the clinical data was only 0.4%. A sample exponential fit to extracted melanoma time course data are shown (Figure S1, second panel, red curve). A sample of these curve fits for patients with melanoma from the trial by Topalian et al. (40) is depicted in Figure 2. The derived fits from all 189 patient response curves in the calibration cohort yielded a mean $\pm$ standard error of the mean (S.E.M.) value of Λ of 0.714 ± 0.257 in patients with partial/complete response and 0.0995 ± 0.0264 in patients with stable/progressive disease ($p = 0.119$; Wilcoxon 2-tail), while the mean $\pm$ S.E.M. value of μ was 0.054 ± 0.014 versus 0.013 ± 0.0012 in partial/complete response versus stable/progressive disease, respectively ($p < 0.001$). In the validation NSCLC cohort, the mean $\pm$ S.E.M.

value of Λ was 0.876 ± 0.102 in patients with partial/complete response and 0.0297 ± 0.469 in patients with stable/progressive disease ($p < 0.001$), while the mean $\pm$ S.E.M. value of μ was 0.0529 ± 0.00982 versus -0.0064 ± 0.0032 in partial/complete response versus stable/progressive disease ($p < 0.001$). These values are depicted in Figure 3. These results suggest that, while Λ and μ give significantly separated classification ranges for partial/complete response versus stable/progressive disease, the specific value of the binary classification threshold is likely a function of the unique disease-drug combination used, as observed in our prior studies (52). This point is explored further in the Discussion.

Confirmation of Model Stability

The parameters α , Λ and μ were held constant during the sensitivity analysis at three different values for each parameter. These values were the minimum, average, and maximum values derived from fitting Equation (1) to the data as described above. Changing Λ and μ from these values $\pm 10\%$ yielded a maximum change in ρ' of 9.2% at $t = 200$ days. Varying ρ' data points randomly up to ± 0.1 resulted in Spearman correlation coefficients of 0.766 and 0.919 for Λ and μ , respectively, derived from fitting these randomized data versus the actual clinical data. We note that, due to high stability observed to these perturbations in the large calibration cohort ($n = 189$ patients), we did not repeat this sensitivity analysis in the validation cohort. Λ values derived from fitting truncated ρ' data for the calibration cohort at various time points displayed Spearman correlation coefficients with Λ derived from fitting the full data ranging from 0.071 at $t < 30$ days to 0.730 at $t < 200$ days, while Spearman correlation coefficients for μ ranged from 0.910 ($t < 30$ days) to 0.921 ($t < 200$ days) for all truncated datasets (Table 2). In the validation cohort, Spearman correlation coefficients for Λ ranged 0.800 at $t < 30$ days to 0.771 at $t < 200$ days, and 0.800 at $t < 30$ days to 0.989 at $t < 200$ days for μ . Simulated RECIST v1.1 categorization (complete/partial response vs. stable/progressive disease) by using the values of Λ and μ derived from

fitting the truncated calibration cohort data resulted in 19% misclassification by using data from $t < 60$ days ($n = 177$) and 13% misclassification at $t < 200$ days ($n = 189$) when compared to the full dataset. In the validation cohort, we observed 18.6% misclassification at $t < 60$ days ($n = 43$) and 16.1% misclassification at $t < 60$ days ($n = 62$). Predicted normalized tumor mass after initiation of immunotherapy is depicted at different combinations of α , Λ and μ (Figure 4).

Analysis of Model Predictions

Using the binary classification of tumor response of partial/complete response as positive and stable/progressive disease as negative, receiver operator characteristic (ROC) analysis was performed to identify optimal predictive response thresholds for μ and Λ (Figure 3, insets), as higher values of μ and Λ are expected to correspond to a more favorable patient response. Identification of optimal predictive thresholds in the calibration cohort by maximizing the Youden's J statistic revealed cutoff thresholds where sensitivity (the proportion of correctly identified patients with partial/complete response) was higher for both Λ and μ (0.945 and 0.891, respectively), while specificity was reduced (0.381 for Λ and 0.567 μ). Response prediction thresholds were identified as 0.722 for Λ and 0.00905 for μ . Testing these response threshold values in the validation cohort revealed sensitivity of 0.6 for Λ and for 0.960 μ , and specificity of 0.743 for Λ and 0.769 for μ in the validation cohort. We found that, in the calibration cohort, positive predictive values (PPV; the probability a patient will be a responder when the model predicts they will be a responder: 0.381 for Λ and 0.458 for μ) were lower than the negative predictive values (NPV; 0.889 for Λ and 0.927 for μ), indicating fewer false-negatives (stable/progressive disease) than false-positives (partial/complete response). This trend was confirmed in the validation cohort (a PPV of 0.600 for Λ and 0.723 for μ ; an NPV of 0.744 for Λ and 0.968 for μ). Lastly, we note that overall accuracy was lower by both Λ and μ in the calibration cohort than were observed in the validation

cohort; this is presumably due to the misbalance of patients with ORR (29% = 55 patients) versus patients with stable/progressive disease (71%, = 134 patients) while sensitivity (correctly identified patients with ORR) was higher than specificity.

Comparison of λ and μ with Clinical and Pathologic Data

Conversion of λ to an estimated intratumoral CD8+ T-cell count yielded a predicted mean $\pm$ S.E.M. of $3,970 \pm 1,429$ cells/mm² in patients with partial/complete response to checkpoint inhibition and 553 ± 147 cells/mm² in patients with stable/progressive disease (Figure 5A) in the calibration cohort, and $4,871 \pm 567$ cells/mm² vs. $165 \pm 2,604$ cells/mm² in patients with partial/complete response vs. stable/progressive disease in the validation cohort. Intratumoral CD8+ T-cell counts referenced from Tumeh et al. (13) are derived from two scenarios, either encompassing data in the pre-treatment setting only ($n = 46$) or data including pathologic CD8+ T-cell counts encompassing both pre-treatment and on treatment ($n = 23$). In the pre-treatment setting, the average intratumoral CD8+ T-cell count was $2,632 \pm 518$ cells/mm² in patients with partial/complete response to checkpoint inhibitors and 322 ± 133 cells/mm² in patients demonstrating stable/progressive disease, while including on-treatment plus pre-treatment data yielded an average of $3,770 \pm 675$ cells/mm² in patients with partial/complete response and 501 ± 113 cells/mm² in patients with stable/progressive disease.

For comparison of μ with PD-L1 staining, data were extracted from a total of twelve published clinical trials ($n = 975$ cancer patients) with data on clinical response by using a PD-L1 staining cutoff of $\geq 1\%$ versus $< 1\%$ and 1,492 patients with a cutoff of $\geq 5\%$ versus $< 5\%$. The ORR of patients with these thresholds was 15% for PD-L1 staining $< 1\%$ and 28% for PD-L1 staining $\geq 1\%$, and 21% and 38% for $< 5\%$ and $\geq 5\%$, respectively. The ORR of patients in the calibration cohort by using the same thresholds in the

derived values of μ was therefore 10% and 45% for patients with $\mu < 1\%$ and $\geq 1\%$ respectively, and was 25% and 67% for patients with $\mu < 5\%$ and $\geq 5\%$, while in the validation cohort, the model-predicted ORR was 9% and 71% for patients with $\mu < 1\%$ and $\geq 1\%$ respectively, and was 29% and 100% for patients with $\mu < 5\%$ and $\geq 5\%$ (Figure 5B).

DISCUSSION

The translational mathematical model introduced here retrospectively investigates the molecular, cellular, and biophysical mechanism(s) behind patient response to immune checkpoint inhibitor therapy, and it explores the potential clinical value of the early prediction of treatment effectiveness. The reductionist approach introduced here simplifies the rather complex biological cross-talk between the immune system with cancer cells to a set of differential equations ultimately dependent on three intrinsic parameters that represent tumor proliferation, intratumoral immune cell presence/killing efficiency (considered as baseline immune cell infiltration), and checkpoint inhibition efficacy; this latter parameter represents the active role of the checkpoint inhibitors in initiating and amplifying the anti-tumor response.

From a mechanistic standpoint, the mathematical model is fully compatible with the current understanding of the biological effect of checkpoint blockade and clinical subtypes of immunotherapy-responsive tumors (12). First, the parameter values derived from our model for intratumoral immune cell infiltration (λ) and immunotherapy antibody efficacy (μ) correspond with the quantitative pathologic measures of intratumoral CD8+ T-cell infiltration and PD-L1 positive staining (Figure 5). Although parameters λ and μ are not direct representations of CD8+ T-cell counts or PD-L1 expression, they correlated with measured pathological biomarkers, thus providing strong experimental evidence that these mathematical parameters may be quantified from these biomarkers. This offers a unique method for clinical translation using clinically-available measured quantities to provide personalized prediction of patient response, a potential that we have often found particularly lacking in prior mechanistic modeling research. By demonstrating links between model parameters and known biology, our model may also be used to quantify the associated mechanistic underpinnings of treatment failure,

with the goal of suggesting clinical interventions that may overcome these deficiencies to improve patient outcome. Unfortunately, the absence of per-patient measurements of CD8+ T-cell counts or PD-L1 expression corresponding to individual patient response to treatment in the cohorts examined prevented analysis of how these parameters perform on a per-patient basis. We are currently collecting in-house a new data set containing both measures for single patients, and the results of this study will be published once complete. We are also investigating how these biomarkers may inform the model to prospectively identify pseudoprogression (a clinical phenomenon wherein lesions are observed to experience rapid volume increase followed by disease response) on a per-patient basis, based on our previous result that retrospectively known long-term parameter values are unique in pseudoprogression (35).

Moreover, the model demonstrates that the rate of immunotherapy response arises from a non-linear interactions between both adequate effector immune cell infiltration in tumors and immunotherapy efficacy within the tumor microenvironment (Figure 4). Mechanistic mathematical modeling across several cancer types (Table 1) demonstrates that the overlying master equation [i.e., Equation (1)] may well be universally applied, while the model parameter values presented here are likely tumor-specific. Indeed, this conclusion may be observed in Figure 3, where Λ and μ are distinctly different for ORR versus stable or progressive disease, but the values are different for the different cohorts; thereby capturing mechanistic differences across the different cancer types. The non-trivial interaction of these inputs predicted by this model is different in comparison to many established and current candidate biomarker studies (which focus on linear immunotherapy responses to just one of these markers, e.g., PD-L1 expression (14)), perhaps providing a plausible explanation why single biomarkers sometimes provide inconsistent predictions across multiple cancer types (14). Similarly, the unique insights presented herein allow for a more direct comparison of the model parameters to common pathologic

markers either reported in the literature or as yet to be reported, and provide a different perspective compared to other immune checkpoint inhibitor modeling studies (19, 53, 54).

The high sensitivity at early time points ($t < 60$ days) demonstrated here indicates this model may provide valuable early identification of cancer patients most likely to benefit from checkpoint inhibition therapy, with a high NPV indicating minimal false-negatives in the group predicted to receive the least benefit from therapy. Therefore, this attribute could serve as an early indicator of the efficacy of checkpoint inhibition by simply using either clinical or imaging measurements of index tumor lesions and by applying the mathematical model to project maximal effect for an individual patient. Further, the model demonstrates robustness and stability in the derived values of Λ and μ to random variations in p' , and by using values of Λ and μ derived from fitting limited data over time (Table 2). This result is in stark contrast to previous mathematical descriptions of immune-mediated killing of cancer cells, where a relatively small change in a single input parameter of merely $\pm 1\%$ could dramatically vary tumor response by up to $\pm 50\%$ (18).

Towards clinical translation of cancer models, a mathematical model with two simple terms (an exponential growth term and a regression term) has previously been used to successfully correlate and predict patient survival (55). The fundamental difference between this prior model (55) and the model reported here is that the quantification of cell death is more mechanistic, i.e., depending on the specific mechanisms (Figure 1) that underlie cancer cell and immunotherapy drug interactions, hence allowing one to gain more mechanistic insight into the treatment system. As demonstrated, model parameters Λ and μ can be used to stratify responding and non-responding patients (Figure 3), which has immediate potential utility as a biomarker for patient selection in prospective clinical trials. Moreover, the model

contains biological values that, at least in principle, may either be measured (e.g., tumor burden, growth rate) or changed clinically (μ : e.g., drug dosing or dosing schedule; λ : e.g., radiotherapy-induced increase in PD-L1 expression, radiotherapy abscopal effect). Thus, we expect that the model (i) may be informed by using standard-of-care measurements (e.g., Figure 5), and (ii) can provide quantitative information on which parameters must be changed to maximize therapeutic benefit, providing clinicians with potentially valuable decision-making information.

Of course, as with any biomathematical model, this model is based on several educated assumptions about the interactions of immune effectors (CD8+ T-cells) with cancer cells, which naturally does not reflect the sheer magnitude and complexity of the tumor immune microenvironment. These include the presence of immunosuppressive cells (such as neutrophils, myeloid-derived suppressor cells, Treg, and M2-polarized tumor-associated macrophages), which extensively infiltrate into tumor mass and hinder the efficacy of immune checkpoint inhibitors (56). Moreover, effects of other immune cells, such as CD4+ T-cells, are not explicitly included in the current model form; while it is likely that some of the effects from these other cells is captured within the model representation of CD8+ T-cells, this assumption warrants future model development where each immune component is explicitly represented. Some of these assumptions (e.g., lack of immune cell/antibody infiltration terms) are supported with *in vitro* or *in vivo* data such as high rates of immune cell binding by anti-PD1 antibodies in the peripheral circulation (40) and local intratumoral (as opposed to systemic myelogenous) expansion of CD8+ T-cells during checkpoint inhibition (13, 57). Other potential model parameters would logistically be quite challenging to measure in an outpatient clinic-ready setting routinely at this time, and their inclusion remains open for follow-up confirmatory retrospective studies and ultimately for future prospective clinical trials. Moreover, the two response parameters derived from the model, λ and μ , represent abstract terms, which may not precisely model intratumoral CD8+ lymphocyte

infiltration and tumor cell PD-L1 staining and potentially can be modified by relevant determinants, such as tumor mutational burden (58) or the human microbiome of the gut (59). The integration of additional prognostic biomarkers and master regulators that intrinsically (e.g., genetic stability) and extrinsically control the efficacy of immunotherapy represent open research areas for future translational investigation.

Finally, a few operational aspects of this work deserve further comment. Response was defined using RECIST v1.1 criteria, *in lieu* of the more recent adaptation of RECIST criteria to immunotherapy treatment (iRECIST), because the studies used to derive the model parameters in this mechanistic analysis did not incorporate these new criteria, and our use of published data for the calibration cohort limited or precluded us from obtaining *a priori* knowledge about non-target lesions. These inherent study limitations notwithstanding, we will continue to improve our model to more accurately predict outcomes in the immunotherapy specific setting of individualized cancer patients, potentially through the inclusion of additional biomarkers, as bioactive molecules such as IFN γ , CD206, CX3CR1, CD1D, and iNOS, along with cell-mediated mechanisms known to have an effect on immune checkpoint inhibitor therapy (60, 61). However, we do not anticipate that such improvements and refinements would change the basic findings and the potential value of the mechanistic mathematical model reported in this work. In going forward, one hopes that knowledge broadly applicable across multiple cancer types and/or immune microenvironment control will be serially incorporated to the dynamic model upon their future discovery and validation.

In conclusion, we present a mechanistic model of immune checkpoint inhibition that is able to describe various immunotherapy response profiles, with the inputs to the model correlating with common

pathologic biomarkers in current clinical use. An early and robust *a priori* predictor for checkpoint inhibition response and outcome might provide a glimpse of the immense potential for timely adjustments and therapeutic personalization. Future prospective investigations of this computational science-assisted approach will focus on readily available clinical data as inputs to the model and further refining the complex interplay between the immune system and the cancer environment to extract other important variables for immunotherapy efficacy. Ultimately, the merit of this approach will rely on its future ability to reliably predict early individual patient response with the goal of improving personalized cancer care.

Acknowledgements. We gratefully acknowledge support from the Andrew Sabin Family Fellowship, Center for Radiation Oncology Research, Sheikh Ahmed Center for Pancreatic Cancer Research, GE Healthcare, Philips Healthcare, and institutional funds from the University of Texas M.D. Anderson Cancer Center. We have also received Cancer Center Support Grants from the National Cancer Institute (P30CA016672 to the University of Texas M.D. Anderson Cancer Center and P30CA072720 the Rutgers Cancer Institute of New Jersey). This research has also been supported in part by grants from the National Science Foundation Grant DMS-1930583 (Z.W., V.C.), the National Institutes of Health (NIH) 1U01CA196403 (Z.W., V.C.), 1U01CA213759 (Z.W., V.C.), 1R01CA226537 (Z.W., R.P., W.A., V.C.), 1R01CA222007 (Z.W., V.C.), U54CA210181 (Z.W., V.C.), and the University of Texas System STARS Award (V.C.). B.C. acknowledges support through the SER Cymru II Programme, funded by the European Commission through the Horizon 2020 Marie Skłodowska-Curie Actions (MSCA) COFUND scheme and the Welsh European Funding Office (WEFO) under the European Regional Development Fund (ERDF). E.K. has also received support from the Project Purple, NIH (U54CA210181, U01CA200468, and U01CA196403), and the Pancreatic Cancer Action Network (16-65-SING). M.F. was supported through NIH/NCI center grant U54CA210181, R01CA222959, DoD Breast Cancer Research Breakthrough Level IV Award W81XWH-17-1-0389, and the Ernest Cockrell Jr. Presidential Distinguished Chair at Houston Methodist Research Institute. R.P. and W.A. received serial research awards from AngelWorks, the Gillson-Longenbaugh Foundation, and the Marcus Foundation. This work was also supported in part by grants from the National Cancer Institute to S.H.C. (R01CA109322, R01CA127483, R01CA208703, and U54CA210181 CITO pilot grant) and to P.Y.P. (R01CA140243, R01CA188610, and U54 CA210181 CITO pilot grant). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author Contributions. V.C. and E.J.K. conceived and designed the project. J.D.B., G.V.M., Z.W., E.J.K., and V.C. contributed to development and performed analysis of the model. J.D.B., G.V.M., Z.W., B.C., M.F., N.E., C.C., D.S.H., J.W., N.H., E.A.M., S.A.C., S.H.C., P.Y.P., S.K.L., S.G., R.L.S., R.P., W.A., E.J.K., and V.C. analyzed and interpreted the data. J.D.B., G.V.M., Z.W., B.C., M.F., N.E., C.C., D.S.H., J.W., N.H., E.A.M., S.A.C., S.H.C., P.Y.P., H.M.C., S.K.L., S.G., R.L.S., R.P., W.A., E.J.K., and V.C. wrote and/or edited the manuscript.

Competing interests. The authors declare no conflict of interest.

REFERENCES

1. Borghaei H, *et al.* (2015) Nivolumab versus docetaxel in advanced nonsquamous non-small-cell lung cancer. *N Engl J Med* 373(17):1627-1639.
2. Robert C, *et al.* (2015) Nivolumab in previously untreated melanoma without BRAF mutation. *N Engl J Med* 372(4):320-330.
3. Robert C, *et al.* (2015) Pembrolizumab versus ipilimumab in advanced melanoma. *N Engl J Med* 372(26):2521-2532.
4. Pardoll DM (2012) The blockade of immune checkpoints in cancer immunotherapy. *Nat Rev Cancer* 12(4):252-264.
5. Postow MA, Callahan MK, & Wolchok JD (2015) Immune checkpoint blockade in cancer therapy. *J Clin Oncol* 33(17):1974-1982.
6. Le DT, *et al.* (2017) Mismatch repair deficiency predicts response of solid tumors to PD-1 blockade. *Science* 357(6349):409-413.
7. Sharma P, Hu-Lieskovan S, Wargo JA, & Ribas A (2017) Primary, adaptive, and acquired resistance to cancer immunotherapy. *Cell* 168(4):707-723.
8. Brahmer JR, *et al.* (2012) Safety and activity of anti-PD-L1 antibody in patients with advanced cancer. *N Engl J Med* 366(26):2455-2465.
9. Jerby-Arnon L, *et al.* (2018) A Cancer Cell Program Promotes T Cell Exclusion and Resistance to Checkpoint Blockade. *Cell* 175(4):984-997 e924.
10. Moreira A, *et al.* (2019) Senescence markers: Predictive for response to checkpoint inhibitors. *Int J Cancer* 144(5):1147-1150.
11. Riaz N, *et al.* (2017) Tumor and microenvironment evolution during immunotherapy with nivolumab. *Cell* 171(4):934-949 e916.
12. Teng MW, Ngiow SF, Ribas A, & Smyth MJ (2015) Classifying cancers based on T-cell infiltration and PD-L1. *Cancer Res* 75(11):2139-2145.
13. Tumeh PC, *et al.* (2014) PD-1 blockade induces responses by inhibiting adaptive immune resistance. *Nature* 515(7528):568-571.
14. Carbone L, *et al.* (2015) Differential activity of nivolumab, pembrolizumab and MPDL3280A according to the tumor expression of programmed death-ligand-1 (PD-L1): Sensitivity analysis of trials in melanoma, lung and genitourinary cancers. *PLoS One* 10(6):e0130142.
15. Bunimovich-Mendrazitsky S, Halachmi S, & Kronik N (2016) Improving bacillus Calmette-Guérin (BCG) immunotherapy for bladder cancer by adding interleukin 2 (IL-2): A mathematical model. *Math Med Biol* 33(2):159-188.
16. Bunimovich-Mendrazitsky S, Shochat E, & Stone L (2007) Mathematical model of BCG immunotherapy in superficial bladder cancer. *Bull Math Biol* 69(6):1847-1870.
17. Kronik N, *et al.* (2010) Predicting outcomes of prostate cancer immunotherapy by personalized mathematical models. *PLoS One* 5(12):e15482.
18. de Pillis LG, Radunskaya AE, & Wiseman CL (2005) A validated mathematical model of cell-mediated immune response to tumor growth. *Cancer Res* 65(17):7950-7958.
19. Luksza M, *et al.* (2017) A neoantigen fitness model predicts tumour response to checkpoint blockade immunotherapy. *Nature* 551(7681):517-520.
20. Pascal J, *et al.* (2013) Mechanistic modeling identifies drug-uptake history as predictor of tumor drug resistance and nano-carrier-mediated response. *ACS Nano* 7(12):11174-11182.
21. Hosoya H, *et al.* (2016) Integrated nanotechnology platform for tumor-targeted multimodal imaging and therapeutic cargo release. *Proc Natl Acad Sci USA* 113(7):1877-1882.
22. Dogra P, *et al.* (2018) Establishing the effects of mesoporous silica nanoparticle properties on in vivo disposition using imaging-based pharmacokinetics. *Nat Commun* 9(1):4551.

- 592 23. Goel S, *et al.* (2019) Size-Optimized Ultrasmall Porous Silica Nanoparticles Depict Vasculature-
593 Based Differential Targeting in Triple Negative Breast Cancer. *Small* 15(46):e1903747.
- 594 24. Dogra P, *et al.* (2020) A mathematical model to predict nanomedicine pharmacokinetics and
595 tumor delivery. *Computational and structural biotechnology journal* 18:518-531.
- 596 25. Das H, *et al.* (2013) Impact of diffusion barriers to small cytotoxic molecules on the efficacy of
597 immunotherapy in breast cancer. *PLoS One* 8(4):e61398.
- 598 26. Pascal J, *et al.* (2013) Mechanistic patient-specific predictive correlation of tumor drug response
599 with microenvironment and perfusion measurements. *Proc Natl Acad Sci USA* 110(35):14266-
600 14271.
- 601 27. Wang Z, *et al.* (2016) Theory and experimental validation of a spatio-temporal model of
602 chemotherapy transport to enhance tumor cell kill. *PLoS Comput Biol* 12(6):e1004969.
- 603 28. Koay EJ, *et al.* (2014) Transport properties of pancreatic cancer describe gemcitabine delivery
604 and response. *J Clin Invest* 124(4):1525-1536.
- 605 29. Frieboes HB, *et al.* (2015) Predictive modeling of drug response in non-Hodgkin's lymphoma.
606 *PLOS ONE* 10(6):e0129433.
- 607 30. Wang Z, Butner JD, Kerketta R, Cristini V, & Deisboeck TS (2015) Simulating cancer growth with
608 multiscale agent-based modeling. *Seminars in cancer biology* 30:70-78.
- 609 31. Brocato TA, *et al.* (2018) Understanding the connection between nanoparticle uptake and
610 cancer treatment efficacy using mathematical modeling. *Sci Rep* 8(1):7538.
- 611 32. Cristini V, Koay E, & Wang Z (2017) *An introduction to physical oncology: How mechanistic
612 mathematical modeling can improve cancer therapy outcomes* (CRC Press, Taylor & Francis
613 Group, Oxfordshire, UK).
- 614 33. Dogra P, *et al.* (2019) Mathematical modeling in cancer nanomedicine: a review. *Biomedical
615 microdevices* 21(2):40.
- 616 34. Brocato TA, *et al.* (2019) Predicting breast cancer response to neoadjuvant chemotherapy based
617 on tumor vascular features in needle biopsies. *JCI insight* 4(8):e126518.
- 618 35. Butner JD, *et al.* (2020) Mathematical prediction of clinical outcomes in advanced cancer
619 patients treated with checkpoint inhibitor immunotherapy. *Sci Adv* 6(18):eaay6298.
- 620 36. Antonia SJ, *et al.* (2015) Phase I/II study of nivolumab with or without ipilimumab for treatment
621 of recurrent small cell lung cancer (SCLC): CA209-032. *J Clin Oncol* 33(suppl; abstr 7503).
- 622 37. Le DT, *et al.* (2015) PD-1 Blockade in tumors with mismatch-repair deficiency. *N Engl J Med*
623 372(26):2509-2520.
- 624 38. Motzer RJ, *et al.* (2015) Nivolumab for metastatic renal cell carcinoma: results of a randomized
625 phase II trial. *J Clin Oncol* 33(13):1430-1437.
- 626 39. Powles T, *et al.* (2014) MPDL3280A (anti-PD-L1) treatment leads to clinical activity in metastatic
627 bladder cancer. *Nature* 515(7528):558-562.
- 628 40. Topalian SL, *et al.* (2012) Safety, activity, and immune correlates of anti-PD-1 antibody in cancer.
629 *N Engl J Med* 366(26):2443-2454.
- 630 41. Rohatgi A (2010-2019) WebPlotDigitalizer: HTML5 based online tool to extract numerical data
631 from plot images. Available at: <https://automeris.io/WebPlotDigitizer> [Accessed April 11, 2019].
- 632 42. Wolfram Research, Inc. (2017) Mathematica, Version 11.2, Champaign, Illinois.
- 633 43. Garon EB, *et al.* (2015) Pembrolizumab for the treatment of non-small-cell lung cancer. *N Engl J
634 Med* 372(21):2018-2028.
- 635 44. Herbst RS, *et al.* (2014) Predictive correlates of response to the anti-PD-L1 antibody MPDL3280A
636 in cancer patients. *Nature* 515(7528):563-567.
- 637 45. Kefford R, *et al.* (2014) Clinical efficacy and correlation with tumor PD-L1 expression in patients
638 (pts) with melanoma (MEL) treated with the anti-PD-1 monoclonal antibody MK-3475. *J Clin
639 Oncol* 32(15 Suppl):3005-3005.

46. Spira AI, *et al.* (2015) Efficacy, safety and predictive biomarker results from a randomized phase II study comparing MPDL3280A vs docetaxel in 2L/3L NSCLC (POPLAR). *J Clin Oncol* 33(15 Suppl):8010-8010.
47. Taube JM, *et al.* (2014) Association of PD-1, PD-1 ligands, and other features of the tumor immune microenvironment with response to anti-PD-1 therapy. *Clin Cancer Res* 20(19):5064-5074.
48. Weber JS, *et al.* (2015) Nivolumab versus chemotherapy in patients with advanced melanoma who progressed after anti-CTLA-4 treatment (CheckMate 037): A randomised, controlled, open-label, phase 3 trial. *Lancet Oncol* 16(4):375-384.
49. Eisenhauer EA, *et al.* (2009) New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1). *Eur J Cancer* 45(2):228-247.
50. Mempel TR & Bauer CA (2009) Intravital imaging of CD8+ T cell function in cancer. *Clin Exp Metastasis* 26(4):311-327.
51. Erdag G, *et al.* (2012) Immunotype and immunohistologic characteristics of tumor-infiltrating immune cells are associated with clinical outcome in metastatic melanoma. *Cancer Res* 72(5):1070-1080.
52. Butner JD, *et al.* (2021) A mechanistic immunotherapy model provides patient-specific quantification of immune response and associated long-term tumor burden. *Nat Biomed Eng* 5(4):297-308.
53. Wilkie KP & Hahnfeldt P (2013) Tumor-immune dynamics regulated in the microenvironment inform the transient nature of immune-induced tumor dormancy. *Cancer Res* 73(12):3534-3544.
54. Serre R, *et al.* (2016) Mathematical modeling of cancer immunotherapy and its synergy with radiotherapy. *Cancer Res* 76(17):4931-4940.
55. Stein WD, *et al.* (2008) Tumor growth rates derived from data for patients in a clinical trial correlate strongly with patient survival: a novel strategy for evaluation of clinical trial data. *Oncologist* 13(10):1046-1054.
56. Jenkins RW, Barbie DA, & Flaherty KT (2018) Mechanisms of resistance to immune checkpoint inhibitors. *Br J Cancer* 118(1):9-16.
57. Ribas A, *et al.* (2016) PD-1 blockade expands intratumoral memory T cells. *Cancer Immunol Res* 4(3):194-203.
58. Goodman AM, *et al.* (2017) Tumor mutational burden as an independent predictor of response to immunotherapy in diverse cancers. *Mol Cancer Ther* 16(11):2598-2608.
59. Gopalakrishnan V, *et al.* (2018) Gut microbiome modulates response to anti-PD-1 immunotherapy in melanoma patients. *Science* 359(6371):97-103.
60. Park YJ, Kuen DS, & Chung Y (2018) Future prospects of immune checkpoint blockade in cancer: from response prediction to overcoming resistance. *Exp Mol Med* 50(8):109.
61. Gubin MM, *et al.* (2018) High-dimensional analysis delineates myeloid and lymphoid compartment remodeling during successful immune-checkpoint cancer therapy. *Cell* 175(5):1443.

Figure Legends.

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 (ψ_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 (35).

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 ref. (40), while solid lines represent best curve fits of Equation (1) to those data (with $\alpha^{-1} = 144$ 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 non-responders (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 non-response ($n = 134$) 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 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 (2 tails): Λ , $p = 0.119$ and $p < 0.001$ for literature (calibration) and NSCLC (validation) cohorts, respectively; μ , $p < 0.001$ for both literature (calibration) and NSCLC (validation) cohorts. Insets: ROC curves for patient response vs model parameters for both cohorts; Λ , literature cohort: sensitivity = 0.381, specificity = 0.945, accuracy = 0.545; 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.

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 which 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 colormap 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, 2 tails); squares: validation cohort, $p < 0.001$) was derived from Λ and literature CD8 intratumoral count was taken from IHC staining in ref. (13) in melanoma [diamonds; average CD8 counts including on-treatment values ($n = 23$)]. CD8+ T-cell counts from pre-treatment biopsies only ($n = 46$) demonstrated mean values ($\pm$ S.E.M.) of $2,632 \pm 518$ cells/mm² in patients with response to immunotherapy and 322 ± 133 cells/mm² in non-responding 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 (1, 3, 8, 13, 38-40, 43-48) for the literature data (green; $n = 975$ for 1% cutoff, $n = 1,492$ for 5% cutoff; see Table S1). Response to immune checkpoint inhibition was determined by RECIST v1.1 criteria. PR, partial response; CR, complete response.

Figures

Figure 1.

Figure 2.

Figure 3.

Figure 4.

Figure 5.

Table 1. Clinical trials with checkpoint inhibitors used to fit the mathematical model and derived values of tumor proliferation constant, α .

Manuscript Reference	Tumor Type Histopathology	Checkpoint Inhibitor Monoclonal Antibody	Constant α (days ⁻¹)	Calculated Tumor Doubling Time (α^{-1} , days)
(36)	CRC	Pembrolizumab (anti-PD1)	0.0622	11
(39)	UCC	Atezolizumab (anti-PDL1)	0.016	43
(36)	SCLC	Nivolumab (anti-PD1)	0.014	50
(40)	MM	Nivolumab (anti-PD1)	0.0069	100
(1)	NSCLC	Nivolumab (anti-PD1)	0.0069	100
(38)	RCC	Nivolumab (anti-PD1)	0.0034	204

CRC, colorectal carcinoma; UCC, urothelial cell carcinoma; SCLC, small cell lung cancer; MM, malignant melanoma; NSCLC, non-small lung cancer; RCC, renal cell carcinoma

Table 2. Spearman correlation coefficients between Λ and μ derived from fitting truncated datasets versus full dataset. t : days. Note that values of 1.000 are due to only a small number of patients ($n = 4$) that were imaged before $t = 30$ days in the validation cohort; these either did not have lesion volumes reassessed before the next reported time threshold ($t = 60$ days) or did not observe a change in monotonic relationships within this timeframe ($t = 30$ –120 days).

Calibration cohort					Validation Cohort				
Λ	$t < 60$	$t < 120$	$t < 200$	$t = \text{all days}$	Λ	$t < 60$	$t < 120$	$t < 200$	$t = \text{all days}$
$t < 30$	0.476	0.162	0.080	0.071	$t < 30$	1.000	1.000	0.800	0.800
$t < 60$		0.416	0.309	0.306	$t < 60$		0.812	0.658	0.823
$t < 120$			0.668	0.599	$t < 120$			0.676	0.750
$t < 200$				0.730	$t < 200$				0.771
μ	$t < 60$	$t < 120$	$t < 200$	$t = \text{all days}$	μ	$t < 60$	$t < 120$	$t < 200$	$t = \text{all days}$
$t < 30$	0.942	0.928	0.922	0.910	$t < 30$	1.000	0.800	0.800	0.800
$t < 60$		0.968	0.941	0.946	$t < 60$		0.974	0.960	0.963
$t < 120$			0.946	0.922	$t < 120$			0.971	0.961
$t < 200$				0.921	$t < 200$				0.989