In silico mouse study identifies tumor growth kinetics as biomarkers for the outcome of anti-angiogenic treatment

Qianhui Wu\textsuperscript{a}, Alyssa D. Arnheim, and Stacey D. Finley\textsuperscript{a,c,*}

\textsuperscript{a}Department of Biomedical Engineering, University of Southern California, Los Angeles, California, United States of America

\textsuperscript{b}Department of Biomedical Engineering, Boston University, Boston, Massachusetts, United States of America

\textsuperscript{c}Department of Chemical Engineering and Materials Science, University of Southern California, Los Angeles, California, United States of America

* Corresponding author

Stacey D. Finley

1042 Downey Way, DRB 140

Los Angeles, CA 90089

Phone: 213-740-8788

Email: sfinley@usc.edu
Abstract

Angiogenesis is a crucial step in tumor progression, as this process allows tumors to recruit new blood vessels and obtain oxygen and nutrients to sustain growth. Therefore, inhibiting angiogenesis remains a viable strategy for cancer therapy. However, anti-angiogenic therapy has not proved to be effective in reducing tumor growth across a wide range of tumors, and no reliable predictive biomarkers have been found to determine the efficacy of anti-angiogenic treatment. Using our previously established computational model of tumor-bearing mice, we sought to determine whether tumor growth kinetic parameters could be used to predict the outcome of anti-angiogenic treatment. A model trained with datasets from six *in vivo* mice studies was used to generate a randomized *in silico* tumor-bearing mouse population. We analyzed tumor growth in untreated mice (control) and mice treated with an anti-angiogenic agent and determined the Kaplan-Meier survival estimates based on simulated tumor volume data. We found that the ratio between two kinetic parameters, $k_0$ and $k_1$, which characterize the tumor’s exponential and linear growth rates, as well as $k_1$ alone, can be used as prognostic biomarkers of population survival outcome. Our work demonstrates a robust, quantitative approach for identifying tumor growth kinetic parameters as prognostic biomarkers and serves as a template that can be used to identify other biomarkers for anti-angiogenic treatment.

Background:

Tumor angiogenesis results in the vascularization of a tumor. This process facilitates tumor growth by allowing tumor cells to obtain oxygen and nutrients through the newly formed blood vessels. As excessive vascularization is often seen in many types of cancer, inhibiting angiogenesis is thought to decrease tumor growth. Therefore, anti-angiogenic treatment is pursued as an attractive therapeutic strategy in oncology [1, 2].
Bevacizumab is a humanized monoclonal antibody against vascular endothelial growth factor A (VEGF), a key angiogenic promoter in tumors [1]. This drug has been approved as a monotherapy or in combination with chemotherapy for many cancers, including renal cell carcinoma, metastatic colorectal cancer, non-small cell lung cancer, and metastatic cervical cancer [3]. It also gained accelerated approval for treatment of metastatic breast cancer through the US Food and Drug Administration (FDA) in 2008. However, subsequent results showed that bevacizumab failed to improve overall survival and that the drug elicited significant adverse side effects. Consequently, the FDA revoked its approval for use of bevacizumab for first-line metastatic breast cancer in late 2011 [4, 5]. Several Phase II and III clinical stage studies have also revealed contradicting results regarding the benefit of add-on bevacizumab in the neoadjuvant treatment setting for breast cancer patients [6–11]. Altogether, these studies illustrate that angiogenic therapy may not be effective across a wide range of patients. Indeed, breast cancer is a genetically and clinically heterogeneous cancer type, which makes identifying optimal therapies a challenge [12].

More broadly, there is a need for biomarkers to predict the response to treatment and identify the tumors for which anti-angiogenic treatment will be effective. A number of mechanistic biomarkers have been investigated for their ability to predict response to anti-angiogenic treatment and to determine an optimal treatment strategy. Promising biomarker candidates include the concentration ranges of circulating angiogenic molecules (such as plasma levels of VEGF) [13, 14], tissue markers (tumor microvessel density) [15–18], and imaging parameters (MRI-measured \( K^{\text{trans}} \)) [15, 19, 20]. However, currently no validated and robust biomarkers are available that can guide selection of patients for whom anti-angiogenic therapy is most beneficial [5, 15].
As an alternative, tumor growth kinetics may be used as biomarkers. There is a body of work that investigates how tumor growth kinetics can serve as prognostic biomarkers of the response to anti-angiogenic treatment [21–25]. Recently, a study showed that volume-based tumor growth kinetics may be a reliable indicator of treatment efficacy, and are in good agreement with standardized approaches for assessing response treatment [21]. Moreover, we developed a computational systems biology model to further investigate the relationship between tumor growth kinetics and the response to anti-angiogenic therapy [26]. The model predicts VEGF distribution and kinetics in tumor-bearing mice, where the dynamic tumor volume is a function of the pro-angiogenic complexes involving VEGF-bound receptors (the “angiogenic signal”). By fitting the model to in vivo experimental data, we estimated the kinetic parameters that characterize tumor growth. We then used the trained model to predict the effect of anti-VEGF treatment on tumor volume, using only the estimated parameter values. The model predictions of tumor growth in response to anti-VEGF treatment closely matched experimental data. In this study, we concluded that there is a strong correlation between particular intrinsic kinetic parameters and the response to anti-VEGF treatment in terms of the end relative tumor volume (RTV).

Taking advantage of our established model framework and its strong predictive power, in the present study, we use this model to further investigate the utility of tumor growth kinetics to serve as a biomarker for anti-angiogenic treatment outcome. We performed an in silico randomized mouse study and estimated the survival of tumor-bearing mice in response to anti-VEGF treatment. Here, we introduced variability in the mouse population by allowing the tumor growth kinetic parameter values to vary within defined ranges. By generating these large, heterogeneous in silico population of tumor-bearing mice, we can eliminate the likely bias caused by animals dropping out in experimental xenograft studies due to limitation of the tumor burden. In general, the average tumor size, particularly in the control group, can be
underestimated in an experimental study. This can result in an underestimation of the treatment effect, because large tumors are excluded from the analysis [27]. In contrast, computational modeling avoids these limitations and enables performance metrics (e.g. survival estimates) to be calculated [28]. Furthermore, computational systems biology is a powerful tool for studying how individual components contribute to the function and behavior of a large system, and has been applied to study cancer at multiple scales [29–31]. Such computational models have been used to identify predictive biomarkers and to enhance the efficacy of anti-angiogenic therapies [13, 32, 33].

In our previous work, we focused on the end RTV to evaluate the contribution of parameters to the treatment outcome. In this study, we use more reliable and appropriate readouts. We implemented our tumor growth kinetic model with time-to-event analysis [34]. Specifically, we simulated the Kaplan-Meier survival curves of the in silico mice based on the population tumor growth data. We then examined tumor growth kinetic parameters as potential prognostic biomarkers to distinguish the tumor response to anti-angiogenic treatment amongst the stratified groups, by comparing the predicted absolute tumor volume time courses and survival estimates.

Results:

In silico mouse population tumor growth in the whole-body model

We performed an in silico randomized mouse study using our whole-body mouse model (Figure 1). The model was previously fitted to each of six independent experimental datasets of control tumor volume in mice bearing MDA-MB-231 xenograft tumors and validated with a separate dataset [26]. The values of $k_0$ and $k_1$ (the rates of exponential and linear growth, respectively), and $\text{Ang}_0$ (the basal angiogenic signal at time, $t=0$) were estimated. Here, we simulated the tumor growth of the six in silico populations of mice (henceforth referred to as “Roland”, “Zibara”, “Tan”, “Volk2008”, “Volk2011a”, and “Volk2011b”), with and without anti-VEGF
treatment. For each population, the values of parameters $k_0$ and $k_1$ are randomly varied simultaneously with a uniform distribution within the ranges of their estimated values from our previous model fitting. Previously, a sensitivity analysis showed that the \( \text{Ang}_0 \) parameter was an influential parameter to the model output when the model was fitted; however, further analysis using partial least squares regression (PLSR) indicated that \( \text{Ang}_0 \) was not a strong predictor of response to treatment [26]. Therefore, in each case, \( \text{Ang}_0 \) is set as the median of the range of its estimated values. We generated 400 \textit{in silico} mice for each of the six cases.

Our simulations show that among the six cases, the anti-VEGF treatment has differential effects in reducing the tumor growth, as compared to the control group (\textbf{Figure 2}). For all cases, we used a single treatment protocol different from protocols used in each of the six experimental studies, in order to compare the predicted results without bias (termed \textit{“protocol A”}). For Roland, Tan, Volk2008, and Volk2011b (\textbf{Figure 2A,C,D,F}), the treated tumor volumes are less than the untreated tumors. Meanwhile, for Zibara and Volk2011a (\textbf{Figure 2B,E}), there is no apparent difference in the tumor volumes for the treated and control groups. Thus, the model simulations reveal distinct differences in the effect of anti-VEGF treatment.

We further studied the effect of anti-VEGF treatment on tumor growth using RTV, the ratio between the mean tumor volumes of the treated and control groups. We calculated the RTV at each time point for all simulated tumors (\textbf{Figure S1}). We also determined the RTV at the end of treatment (\textbf{Figure S2}). The RTV values in all cases are smaller than one, indicating that the anti-VEGF treatment limits tumor growth, similar to what has been observed experimentally [35–39]. For Zibara and Volk2011a, the endpoint RTV values are just slightly less than one (\textbf{Figure S2B,E}), which is an expected result based on the similar tumor growth curves between the control and treated groups (\textbf{Figure 2B,E}). Comparing the endpoint RTV among all six cases, the effect of anti-VEGF treatment in limiting tumor growth is the strongest for Volk2011b (RTV =
0.459 ± 0.054), followed by Roland (0.454 ± 0.096), Volk2008 (0.615 ± 0.066), and Tan (0.638 ± 0.049). This treatment effect is the least significant in Zibara (0.979 ± 0.009) and Volk2011a (0.987 ± 0.013).

**Kinetic parameters as potential predictor for stratified population response**

We investigated the relationship between the parameters that characterize tumor growth kinetics and the effect of the anti-VEGF treatment. Previously, our PLSR analysis indicated that for nearly all pairwise comparisons, if the RTV values for two datasets were significantly different, their $k_d/k_1$ ratios were also significantly different. This implies that the $k_d/k_1$ is a large contributor in predicting the endpoint RTV [26]. Additionally, plotting the RTV versus $k_0$, $k_1$, and $k_d/k_1$ shows some relationship between the endpoint RTV and the tumor growth parameters (Figure S2). Therefore, we investigated whether these tumor growth parameters could stratify the simulated mouse populations, and distinguish their tumor growth and survival estimates. To address this question, we used our simulated tumor growth data for each case, noting the number of *in silico* mice at each time point. We record the time at which a mouse is “sacrificed”, which happens when the tumor volume reaches 2 cm$^3$, as typically done in experimental studies [40]. This approach for modeling population survival allows us to closely mimic the practice in preclinical animal studies, and provides easily interpretable insights for researchers and clinicians.

We used the simulated population survival data to determine if $k_0$, $k_1$, or $k_d/k_1$ can be used to discriminate between tumors for which anti-VEGF treatment is effective or not. We found that in each case, a range of $k_d/k_1$ ratios, as well as $k_1$, can be used to distinguish the population response to the anti-VEGF treatment (Figure 3B,C). We term these “$ratio_{thresh}$” and “$k_{1,thresh}$”, the values of the growth kinetic parameters that separate the simulated mouse population into groups with significantly different survival estimates. In contrast, we did not find any values of $k_0$
alone that could be used to separate the simulated mouse population into groups whose survival estimates are statistically different for Roland, Zibara, and Volk2011b cases. For Tan and Volk2008, we only found one such $k_0$ value in each case (Figure 3A).

Interestingly, although the ranges of generated $k_0/k_1$ ratios and $k_1$ were different for each of the six sets of tumor growth data, we found that there is an overlap among the potential $\text{ratio}_{\text{thresh}}$ or $k_{1,\text{thresh}}$ values found in each of the six cases. The common range of $\text{ratio}_{\text{thresh}}$ is 9.757 to 17.982, and that of $k_{1,\text{thresh}}$ is $1.391 \times 10^{-6}$ to $1.931 \times 10^{-6}$. This means that separating the treatment group by any $k_{1,\text{thresh}}$ or $\text{ratio}_{\text{thresh}}$ value within its respective range will produce two groups of treated mice that have statistically different survival estimates. Specifically, the treated group with $k_0/k_1$ ratios larger than the $\text{ratio}_{\text{thresh}}$ value has a better survival estimate than the treated group with smaller ratios. The treated group with $k_1$ smaller than the $k_{1,\text{thresh}}$ value has a better survival estimate than the treated group with larger $k_1$.

We used the median $\text{ratio}_{\text{thresh}}$ value to illustrate this distinction. We compare the survival estimates for a total of six groups: 1) all mice in the control group; 2) all mice in the treatment group; 3) control group with $k_0/k_1 < \text{ratio}_{\text{thresh}}$; 4) control group with $k_0/k_1 > \text{ratio}_{\text{thresh}}$; 5) treatment group with $k_0/k_1 < \text{ratio}_{\text{thresh}}$; and 6) treatment group with $k_0/k_1 > \text{ratio}_{\text{thresh}}$. We generated the Kaplan-Meier survival curves for these groups for each of the six cases investigated (Figure 4). We also estimated the median survival of the six groups in each case (Table 1), the Mantel-Haenszel hazard ratio (HR), with 95% confidence interval (CI), and the $p$-values from the Mantel-Cox log rank test for the survival curve comparison (Table 2). When comparing two groups, if the HR is less than one, the first group has a lower death rate (see Methods). Together these analyses emphasize that mice with larger $k_0/k_1$ ratios survive for longer, with $p$-value < 0.05. Interestingly, for Zibara and Volk2011a, although the anti-VEGF treatment does not significantly reduce tumor growth and therefore does not yield a better survival estimate for
the treated groups compared to their control groups (Figures 2B,E and 4B,E), the stratified groups yield significantly different survival estimates. That is, the control and treated groups with $k_0/k_1$ ratios larger than $\text{ratio}_{\text{thresh}}$ have better survival estimates than those with smaller $k_0/k_1$ ratios.

We performed a similar analysis using the median $k_{1,\text{thresh}}$ value to show the distinction between the survival estimates (Figure S3). The control and treated groups with $k_1$ smaller than $k_{1,\text{thresh}}$ have better survival estimates than those with larger $k_1$ values. We also estimated the median survival of the six groups separated using the median $k_{1,\text{thresh}}$ (Table 3), the Mantel-Haenszel HR, and the $p$-values from the Mantel-Cox log rank test for the survival curve comparison (Table 4). From these analyses, mice with smaller $k_1$ survive longer than those with larger $k_1$, and the HR is smaller than one ($p<0.05$).

**Alternative treatment strategies to improve survival estimates**

We next sought to understand whether alternative treatment protocols can effectively reduce tumor volume for the Zibara and Volk2011a cases, since the baseline protocol we used did not significantly affect tumor volume. For the Zibara case, we simulated the original treatment protocol used in the experimental study (termed “protocol Z”). This protocol starts the 10 mg/kg biweekly treatment upon tumor engraftment (assuming the initial tumor volume to be 0.004 cm$^3$) [36]. The predicted tumor volumes are smaller in the treated group (Figure S4A), recapitulating the findings from the published experimental study. The predictions may suggest that in this case, starting the treatment earlier is more effective in limiting the tumor growth. For mice with $k_0/k_1$ ratios larger than the median $\text{ratio}_{\text{thresh}}$ value (13.689), or with $k_1$ smaller than the median $k_{1,\text{thresh}}$ value ($1.661\times10^{-6}$), the HR between the treated and control groups is smaller than one, and the survival curves are significantly different ($p<0.0001$) (Tables 2 and 4).
For Volk2011a, we simulated treatment termed “protocol V11a”, which starts the 10 mg/kg biweekly treatment when the tumor volume reaches 0.5 cm$^3$, a start time extracted from the published preclinical study [39]. After 12 weeks, the simulated mean tumor volumes in the treated group are significantly smaller than the control tumors (Figure S4B). However, the survival estimates were not significantly different ($p$>0.05). Again, the treated group with $k_0/k_1$ ratios larger than the median $ratio_{thresh}$ or with $k_1$ smaller than the median $k_{1,thresh}$ has a significantly better survival estimate than the opposite group ($p$$<$0.0001) (Tables 2 and 4). This phenomenon is similar to that observed in the Volk2011a case using protocol A, where the two groups separated according to the $k_0/k_1$ ratio or $k_1$ have distinct survival estimates, but there is no significant difference between the treated and control groups.

Finally, we explored whether another treatment protocol could significantly improve the survival estimates for the treated group compared to the control. We simulated protocol V11a-D, where biweekly treatment starts when the tumor volume reaches 0.5 cm$^3$, and the drug dosage is doubled to 20 mg/kg. This treatment protocol significantly limits the tumor growth (Figure S4C), and the survival curves are significantly better for the treated group compared to the control ($p$$<$0.0001). Overall, the treated and control groups have an HR of 0.2016 (95% CI: 0.1343-0.3027) (Table 2).

Validation of thresholds using an independent dataset

To validate the use of the range of $ratio_{thresh}$ and $k_{1,thresh}$ values that we found, we used a recently published independent set of data that measures tumor growth in mice with MDA-MB-231 xenografts, with or without bevacizumab treatment [41]. First, we fit the model to the measured tumor volumes without treatment. We obtained 12 sets of estimated parameter values for $k_0$, $k_1$, and $Ang_0$ that allow the model to best fit to the control data. We then validated the fitted model
by simulating anti-VEGF treatment and comparing to the experimental measurements. The predicted tumor growth with treatment matches closely to the experimental data (Figure 5A).

Using the same approach as described above, we generated 400 sets of tumor volumes for an in silico mouse population with and without treatment (referred to as “Mollard”). To do so, we randomly varied $k_0$ and $k_1$ from the ranges of the 12 sets of estimated parameter values from model fitting to the Mollard dataset, with $Ang_0$ held constant at the median of its estimated values. The simulated tumor volumes for the control and treated groups are shown in Figure 5B.

We generated the population survival data based on the simulated tumor growth profiles. We tested whether the common range of $ratio_{thresh}$ and $k_{1,thresh}$ values identified using the six datasets described above are able to separate the population survival data for this validation case (Mollard). For all $ratio_{thresh}$ values within the range, the survival estimate of the treated mice with $k_0/k_1$ ratios larger than the threshold is better than those with smaller $k_0/k_1$ ratios. Examples using the median $ratio_{thresh}$ and the median $k_{1,thresh}$ are shown in Figure 5C-D. We calculated the HR values, as well as the $p$-value from the Mantel-Cox log rank test among the treated and control groups, separated using the median of the common $ratio_{thresh}$ range (Table 2) or the common $k_{1,thresh}$ range (Table 4). Thus, we were able to validate the threshold values.

### Tumor growth dynamics among stratified populations

We explored the dynamics of the tumor growth for the groups separated by the threshold values to better understand why the anti-VEGF treatment has differential effects in the simulated mouse populations. As researchers have pointed out, log-transformation of tumor growth data provides information on the tumor growth rates (given by the slope of the curve) and is more suitable for detecting a transient biological or therapeutic effect [40, 42, 43]. Therefore, we
compared the mean RTV time courses (Figure S1) and the log-transformed mean tumor volume data (Figure S5) of the groups stratified by the median ratio\(_{\text{thresh}}\) (13.869) in each case.

For Roland, Tan, and Volk2008, the mean RTV of the group with larger \(k_0/k_1\) ratios (Figure S1A, C, D) is initially larger, and then becomes smaller relative to the opposite group. This switch occurs because in the group with larger \(k_0/k_1\) ratios, the difference between the treated and control tumor volumes is smaller at early times, and then becomes larger (Figure S5). Meanwhile, the actual tumor volumes for this group are both relatively low. As a result, this group survives longer (Figure 4). For the Mollard case used for validation, the differences between the treated and control tumor volumes in the group with larger \(k_0/k_1\) ratios are larger (Figure 6B, dotted curves), giving rise to the larger mean RTV (Figure 6A). However, the group with larger \(k_0/k_1\) ratios still survives longer because the actual tumor volumes are relatively low (Figure 5C).

The log-transformed tumor volume data also reveal that the tumor growth rates of control and treated groups diverge at different time points. For Roland, the gap between the linear tumor volume data of control and treated groups continually increases (Figure 7A). However, the log-transformed tumor volume data show that their growth rates mostly differentiate during days 14-40. The growth rates become similar during the later stage (after 40 days), as evidenced by the parallel curves of the log-transformed data (Figure 7B). Therefore, the increasingly large gap between the tumor volumes is a result of early differences in the tumor growth rates. A similar phenomenon is observed for Volk2011b, where the tumor growth rate of the treated group is suppressed transiently at early times but not in the later stage (Figure S5). In Zibara, Tan and Volk2008, the growth rates start to differentiate between day 30 and day 45, and only gradually become similar towards the end of the simulated time. Overall, analysis of the log-transformed growth curves reveal that the anti-VEGF treatment has differential effects in limiting tumor
growth, and the effects occur at different stages for the simulated cases. The treatment effect appears to be stronger for the group with $k_0/k_1$ ratios larger than the median $\text{ratio}_{\text{thresh}}$.

**Discussion**

In this study, we focus on identifying potential tumor growth kinetic parameters as biomarkers for the outcome of anti-VEGF treatment. We developed a computational approach to perform biomarker identification that incorporates model training, simulation of tumor growth within a heterogeneous population, and estimation and analysis of population response.

We applied the model to simulate anti-VEGF treatment and compared the effect of treatment across tumor-bearing mice generated from our previous fitting to six independent preclinical studies. For most simulated tumors, the anti-VEGF agent significantly reduces tumor volume compared to control. However, our simulations for Zibara and Volk2011a show that these populations do not respond to the treatment (Figure 2B,E), which is different than the effect seen experimentally. This difference occurs for two reasons. First, our simulated treatment protocol A is universal across the six cases, and is different from what was used in each of the original six experimental studies. Second, in our simulations, $k_0$ and $k_1$ are varied simultaneously and independently of each other, possibly resulting in more variability than what occurs in the experimental tumor growth.

Our study demonstrates that the $k_0/k_1$ ratio or $k_1$ alone can be utilized to stratify the population response with or without anti-VEGF treatment. This finding agrees with our previous finding through PLSR analysis that the ratio is a key predictor of the tumor response to anti-VEGF treatment [26]. Building on that framework, we found that the survival estimate of mice with larger $k_0/k_1$ ratios or smaller $k_1$ is better compared to those with smaller ratios or higher $k_1$. Interestingly, the result for the ratio is the opposite of the conclusion we drew previously (that a
larger ratio correlates with a poorer response to treatment). However, in that work, we focused only on whether the final RTV value was low. This highlights the fact that only evaluating the endpoint RTV of the treated and control group and neglecting the actual tumor volume data over time can lead to misinterpretation of the treatment effect. Indeed, researchers have recognized that while most preclinical studies focus on the end points of tumor growth, monitoring tumor growth kinetically may oftentimes be more insightful [42, 43].

We found that in two cases (Volk2011a simulated with protocol A and protocol V11a), no significant difference is observed in the survival estimates between the pairs of treated and control groups. However, even for these cases, two populations with significantly different survival estimates can be identified based on their $k_0/k_1$ ratios (Figure 4B,E) or $k_1$ value (Figure S3B,E). This indicates that even when the treatment is not effective in reducing tumor volume, there is still a difference in tumor growth dynamics between the two populations stratified based on the tumor’s growth kinetic parameters. Thus, we believe that the $k_0/k_1$ ratio or $k_1$ may be prognostic biomarkers to stratify populations for their survival estimate without the anti-angiogenic treatment. Interestingly, the parameters provide mechanistic insight into tumor growth. In particular, they highlight that slower linear growth (larger ratio or smaller $k_1$) results in less aggressive overall tumor growth (Figure S5) and therefore, better survival outcome.

Another interesting aspect is the utility of $k_1$ to serve as a prognostic biomarker. Although $k_1$ was not revealed as a strong predictor of the final RTV previously in the PLSR analysis, it is inversely correlated with the $k_0/k_1$ ratio, and therefore in our study, it also can be used to stratify the population survival outcome. The use of survival analysis in this study addresses one of the limitations in our previous work. That is, with the PLSR analysis, we were able to identify which parameters were related to treatment efficacy, but could not identify the specific relationship between the kinetic parameter values and effectiveness of the treatment.
Compared to the mean RTV data, the tumor volume data provide more useful insight into the tumor growth characteristics of the stratified population. In particular, the log-transformed tumor volume more clearly illustrates the source of the differences in the population survival estimates. Specifically, we found that larger $k_0/k_1$ ratios often yield slower tumor growth in a population, and therefore, lead to a better survival estimate of the population. This conclusion could not be made if we were to only analyze the RTV data. In addition, the log-transformed tumor volume data reveal that the effect of anti-VEGF treatment in tumor growth can be relatively transient (as observed in Roland and Volk 2011b) or gradual (as seen in Zibara, Tan, and Volk 2008).

Our study makes use of a predictive and useful computational model of tumor growth with and without anti-VEGF treatment. This is a pharmacokinetics-pharmacodynamics model with mechanistic detail that goes beyond what is found in other models. However, in the future, this model can be expanded to address limitations that are not currently accounted for. For example, we do not account for changes in tumor vascularity relative to tumor volume. In addition, vascular normalization is an important process that has been shown to affect tumor growth and can be regulated by anti-VEGF agents [32]; however, this process is not included in our model. These aspects can be implemented into the model as more quantitative data become available and enable us to characterize the dynamics of vessel normalization. The model can then be further extended to account for other characteristics of tumor progression, including tumor perfusion and metastatic potential. Furthermore, the range of threshold values for tumor stratification is constrained to be within the overlap of estimated parameter values from model training to each experimental dataset. It is possible that artifacts coming from experimental data quantification led to bias in the range of the fitted parameter values. This can also be improved when more quantitative data from experimental studies become available to allow for additional model training, which enables improved model predictions. We note that the biomarker
candidates identified in this study are best used to stratify populations for their survival outcome, whether the mice receive treatment or not, rather than to predict treatment efficacy. This utility of the growth parameters for stratification is primarily because the datasets used to train the model were the control time courses of tumor growth over several weeks. Our results would be of broader applicability if only pre-treatment data were adequate to train the model. We have attempted such approach in our previous study [26]; however, the fitting results were not satisfactory (predicted volumes varied widely, preventing us from making conclusive predictions). Despite this perceived limitation, our modeling approach generates hypotheses about potential biomarkers, and further experimental validation is needed to ensure the utility of the biomarkers identified here.

Our study demonstrates a time- and cost-effective way to generate large in silico mouse populations, predict anti-VEGF treatment outcome, and stratify the populations. This approach provides useful information that could facilitate efficient experimental design, such as predicting the effect of different treatment protocols (varying the dosage and the timing of the injections). Additionally, our modeling approach can be adapted for analysis of patient treatment outcome in clinical studies. With data from a small patient population, we can develop a model that is trained to patient-specific data and generate a larger in silico patient population. Analysis of the simulated tumor growth and survival data can be used to identify biomarkers that predict responders versus non-responders to anti-VEGF treatment, stratify the predicted population survival, and test the response to various treatment schedules.

Conclusion

We examined tumor growth kinetic parameters as potential biomarkers of anti-angiogenic treatment outcome. Using a computational model that simulates VEGF-dependent tumor growth in tumor-bearing mice, we generated an in silico mouse population and related the kinetic
parameters that characterize tumor growth to the response to anti-VEGF treatment. We found that the ratio between two tumor growth kinetic parameters, $k_0$ and $k_1$, as well as $k_1$ alone, can be prognostic biomarkers and that the simulated treatment protocol may have a better outcome for mice whose tumors have smaller linear growth rates. In fact, we found ranges of threshold values for the $k_0/k_1$ ratio and $k_1$ that distinguish tumors’ response to the anti-VEGF treatment. This study demonstrates an approach for identifying tumor growth kinetic parameters as potential biomarkers, and this model framework can be adapted to predict the efficacy of other anti-angiogenic strategies.

**Methods**

**Computational model**

This work directly uses our previously calibrated and validated three-compartment model of a tumor-bearing mouse [26]. In this previous work, we fit the model to six independent control datasets, and the model-predicted tumor growth curves match closely to the experimental data (fitting error range: 0.0405 – 0.1833). We provide a full description of this model in Supplemental File S1.

**Numerical implementation**

All model equations were implemented in MATLAB using the SimBiology toolbox. The model is provided as the SimBiology project file, as SBML, and as a MATLAB m-file (Supplemental File S2). Parameter fitting was performed using the `lsqnonlin` function in MATLAB. Kaplan-Meier survival estimation was performed using the `kmplot` function in MATLAB, and GraphPad Prism was used for statistical survival analyses.

**Simulation of in silico mouse population**
In our previous work, we fit the three-compartment model of a tumor-bearing mouse to six datasets. Here, we generated 400 sets of values for growth parameters $k_0$ and $k_1$, randomly selected from a uniform distribution within the range of the best-fit parameter sets from our previous study (Table S1). The $Ang_0$ value is set to be the median of the best fits in each case (Table S1). These sets were used to calculate tumor growth with or without anti-VEGF treatment, simulating a population of mice for each of the six datasets. In order to keep tumor growth profiles realistic, tumors that do not reach 0.1 cm$^3$ within 10 days upon tumor engraftment (assuming tumor volume to be 0.004 cm$^3$) were excluded from the analyses.

For each dataset, we simulated anti-VEGF treatment as intravenous injections lasting for one minute. We implement this injection by adding a net rate of secretion of the drug (bevacizumab, which binds to human VEGF isoforms) directly into the blood compartment (Figure 1). We simulated different anti-VEGF treatment protocols. Treatment protocol A is simulated universally across the six cases. In this protocol, weekly treatment starts when the tumor volume reached 0.1 cm$^3$, as the switch where angiogenesis is more strongly promoted occurs when the tumor reaches 1-2 mm in diameter. The treatment dosage is 10 mg/kg. The model was simulated for 12 weeks after treatment started. We also simulated alternate treatment protocols: Z, denotes biweekly treatment at dosage of 10 mg/kg starting when tumor volume is 0.004 cm$^3$; V11a, denotes biweekly treatment (twice a week) at a dosage of 10 mg/kg, starting when the tumor volume is 0.5 cm$^3$; and V11a-D, denotes biweekly treatment at a dosage of 20 mg/kg, starting when tumor volume is 0.5 cm$^3$. Information for all treatment protocols is shown in Table S3.

Relative tumor volume (RTV)

Based on the model-generated tumor growth data, the relative tumor volume (RTV), the ratio between the treated and control tumor volumes, is calculated at any simulated time point:

$$RTV = \frac{V_{\text{treatment}}}{V_{\text{control}}}$$
An RTV value less than one indicates that the treated tumor volume is smaller than the control.

**Kaplan-Meier survival estimation**

We applied time-to-event analysis to determine the survival of each mouse population [34]. An *in silico* mouse is recorded as “sacrificed” when its tumor reaches 2 cm$^3$ within the simulated time. Alternatively, a mouse is recorded as “censored” at a particular time point, $t$, if its tumor volume simulation remains below 2 cm$^3$ but ended before that time $t$. All other mice are retained in the study and recorded as “alive”. Survival curves were estimated by the Kaplan-Meier method using the *kmplot* function in MATLAB [44], and compared using the Mantel-Cox log rank test and Mantel-Haenszel hazard ratio in GraphPad Prism.

The hazard ratio (HR) compares the rate of death in two groups, with the assumption that the population hazard ratio is consistent over time. It is calculated using the Mantel-Haneszel approach, which is more accurate than the log rank approach [45]. As an example, an HR of 0.5 between two groups means that the death rate of the first group is half of that of the second group.

**Determinations of threshold values**

In order to determine threshold values for the $k_0/k_1$ ratio, we ordered the simulated mouse tumor volume data for each of the six populations according to the $k_0/k_1$ ratio. Then, we systematically tested each $k_0/k_1$ ratio (called “ratio$_{thresh}$”) value to see if there is a significant difference between the survival estimates for the mice with $k_0/k_1$ ratio above and below “ratio$_{thresh}$” in the log rank test ($p<0.05$). We performed a similar analysis for $k_0$ and $k_1$ individually to determine any $k_{0,thresh}$ and $k_{1,thresh}$ values.

**Validation of the predicted biomarker**
Upon identifying a potential predictive biomarker for the efficacy of anti-VEGF treatment, we validated our findings using an independent set of data that was not used to determine the range of the threshold value. To do so, we fit the control tumor growth for the independent data set and generated an in silico mouse population based on the fitted parameters.

Data extraction. For threshold validation, data from the published in vivo experimental study of MDA-MB-231 xenograft tumor growth in mice by Mollard et al. were used for parameter estimation [41]. The dataset for tumor growth in untreated tumors (control) was used for fitting the model, and the dataset for tumors treated with bevacizumab twice weekly was used for validation. Experimental data was extracted using the WebPlotDigitizer program [46]. The numerical values are provided in Table S1.

Parameter estimation. We trained the model to fit the control tumor growth dataset from [41] using the same approach as described in our previous work [26]. The values of tumor growth parameters $k_0$, $k_1$, and $Ang_0$ were estimated. In their study, Mollard and coworkers only reported the tumor volumes relative to day eight. However, the absolute tumor volumes are needed to determine how the tumor interstitial volume varies as a function of the total tumor volume. Therefore, we compared the relative tumor volume at each time point in the work by Mollard and coworkers to that of all the available control datasets (Figure S6). We then chose to use the interstitial volume equation from the Zibara data, given that the relative tumor volume closely matches that of the data in Mollard. Finally, we fit our tumor growth model to the Mollard control dataset using the lsqnonlin function in MATLAB to minimize the sum of squared residuals:

$$\min SSR(\Theta) = \min \sum_{i=1}^{n} (V_{exp,i} - V_{sim,i}(\Theta))^2$$
where \( V_{\text{exp},i} \) is the \( i \)th experimental data point of tumor volume, \( V_{\text{sim},i} \) is the \( i \)th simulated volume at the corresponding time point, and \( n \) is the total number of experimental data points. The minimization is subject to \( \Theta \), the set of upper and lower bounds on each of the free parameters.

The bounds for each parameter spanned at least one order of magnitude: \( 10^{-8} \) to \( 10^{-2} \) for \( k_0 \) and \( k_1 \), and \( 10^{-16} \) to \( 10^{-14} \) for \( \text{Ang}_0 \). After fitting to the control data, we validated the estimated parameters with measured tumor volumes not used in the fitting for model validation: the experimentally measured volumes of tumors treated with the anti-angiogenic drug bevacizumab. We simulated the dosing regimen used in the experiment performed by Mollard et al. using the estimated parameters obtained from fitting to the control data. This protocol involved three cycles of weekly intravenous injections lasting for one minute starting from day five. We used the combined SSR for the relative tumor volume between model prediction and the experimental data (both control and treatment) to identify the optimal parameters. Twelve parameter sets with the smallest errors were taken to be the “best” sets (Table S1) and the ranges of the estimated parameter values were used for subsequent model simulations (Table S2).

We extracted the absolute tumor volume at day 8 from previously reported data from Mollard and coworkers [47] to determine the survival estimates for a mouse population simulated based on the fitted growth kinetics parameter values.

**Declarations:**

**Ethics approval and consent to participate.** Not applicable

**Consent for publication.** Not applicable

**Data accessibility.** All data analyzed during this study, including the model used for data generation, are included in this published article and its supplementary information files.
Author’s contributions. SDF designed the research. ADA wrote MATLAB scripts and performed initial model simulations. QW performed model simulations and data analysis. QW and SDF wrote the manuscript. All authors have read and approved the final manuscript, and agree to be accountable for all aspects of the work.

Competing interests. The authors declare that they have no competing interests

Funding. The authors acknowledge the support of the US National Science Foundation (CAREER Award 1552065). The funders had no role in study design, data analysis, decision to publish, or preparation of the manuscript.

Acknowledgements. The authors thank members of the Finley research group for critical comments and suggestions.

References:


tumor kinetics as a prognostic biomarker for patients with EGFR mutant lung adenocarcinoma
undergoing EGFR tyrosine kinase inhibitor therapy: a case control study. *Cancer Imaging* 16

al. 2015 Performance of tumor growth kinetics as an imaging biomarker for response
assessment in colorectal liver metastases: correlation with FDG PET. *Abdom Imaging* 40:3043–
51.


of renal carcinomas: a novel algorithm suggests early discontinuation of bevacizumab resulted

growth rate of localized pancreatic adenocarcinoma in response to gemcitabine, bevacizumab,
and radiation therapy on MDCT. *Int J Radiat Oncol Biol Phys* 81:452–9
(doi:10.1016/j.ijrobp.2010.05.060).

26. Gaddy TD, Wu Q, Arnheim AD, Finley SD. 2017 Mechanistic modeling quantifies the

27. Martin EC, Aarons L, Yates JWT. 2016 Accounting for dropout in xenografted tumour
efficacy studies: integrated endpoint analysis, reduced bias and better use of animals. *Cancer


44. KMPlot - File Exchange - MATLAB Central.
45. GraphPad Statistics Guide.
46. WebPlotDigitizer - Extract data from plots, images, and maps.
### Tables:

**Table 1. Summary of median survival of population separated by median $\text{ratio}_{\text{thresh}}$**

<table>
<thead>
<tr>
<th>Median survival (days)</th>
<th>Roland†</th>
<th>Zibara‡</th>
<th>Tan†</th>
<th>Volk 2008‡</th>
<th>Volk 2011a‡</th>
<th>Volk 2011b‡</th>
<th>Zibara§</th>
<th>Volk 2011a¶</th>
<th>Volk 2011a¶¶</th>
<th>Mollard†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control (All)</td>
<td>53</td>
<td>55</td>
<td>58</td>
<td>64</td>
<td>69</td>
<td>68</td>
<td>55</td>
<td>69</td>
<td>69</td>
<td>85</td>
</tr>
<tr>
<td>Control ($k_0/k_1 &lt; \text{ratio}_{\text{thresh}}$)</td>
<td>44</td>
<td>44.5</td>
<td>45</td>
<td>40</td>
<td>44</td>
<td>42.5</td>
<td>44.5</td>
<td>42.5</td>
<td>44</td>
<td>77</td>
</tr>
<tr>
<td>Control ($k_0/k_1 &gt; \text{ratio}_{\text{thresh}}$)</td>
<td>86.5</td>
<td>84</td>
<td>Un*</td>
<td>89.5</td>
<td>87</td>
<td>88</td>
<td>84</td>
<td>80</td>
<td>87</td>
<td>96</td>
</tr>
<tr>
<td>Treatment (All)</td>
<td>Un*</td>
<td>63</td>
<td>77.5</td>
<td>Un*</td>
<td>69</td>
<td>Un*</td>
<td>63</td>
<td>71</td>
<td>Un*</td>
<td>Un*</td>
</tr>
<tr>
<td>Treatment ($k_0/k_1 &lt; \text{ratio}_{\text{thresh}}$)</td>
<td>74</td>
<td>46</td>
<td>50</td>
<td>50</td>
<td>44</td>
<td>79</td>
<td>46</td>
<td>42.5</td>
<td>78.5</td>
<td>Un*</td>
</tr>
<tr>
<td>Treatment ($k_0/k_1 &gt; \text{ratio}_{\text{thresh}}$)</td>
<td>Un*</td>
<td>Un*</td>
<td>Un*</td>
<td>Un*</td>
<td>87</td>
<td>Un*</td>
<td>Un*</td>
<td>87</td>
<td>Un*</td>
<td>Un*</td>
</tr>
</tbody>
</table>

*Un: Undefined. The median survival cannot be estimated if the survival estimation does not reach below 50%;

†protocol A: biweekly treatment at a dosage of 10 mg/kg, starting when tumor volume reaches 0.1 cm³;

‡protocol Z: biweekly treatment at a dosage of 10 mg/kg, starting when tumor volume is 0.004 cm³ (upon engraftment of tumor);

§protocol V11a: biweekly treatment at a dosage of 10 mg/kg, starting when tumor volume reaches 0.5 cm³;

¶protocol V11a-D: biweekly treatment at a dosage of 20 mg/kg, starting when tumor volume reaches 0.5 cm³.
Table 2. Statistics comparing the Kaplan-Meier survival curves of population separated by median ratio\_{\text{thresh}}: hazard ratio (95%CI) and log rank test p-values

<table>
<thead>
<tr>
<th>HR (95% CI) p-value</th>
<th>Roland\textsuperscript{†}</th>
<th>Zibara\textsuperscript{†}</th>
<th>Tan\textsuperscript{†}</th>
<th>Volk2008\textsuperscript{†}</th>
<th>Volk2011a\textsuperscript{†}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment (k/\nu &gt; \text{ratio}<em>{\text{thresh}}) vs. Treatment (k/\nu &lt; \text{ratio}</em>{\text{thresh}})</td>
<td>0.2073 (0.1059-0.4057) p=0.0001</td>
<td>0.2005 (0.0991-0.4055) p=0.0001</td>
<td>0.1623 (0.0872-0.3021) p=0.0001</td>
<td>0.0576 (0.0237-0.1399) p=0.0001</td>
<td>0.0216 (0.0098-0.0476) p=0.0001</td>
</tr>
<tr>
<td>Control (k/\nu &gt; \text{ratio}<em>{\text{thresh}}) vs. Control (k/\nu &lt; \text{ratio}</em>{\text{thresh}})</td>
<td>0.1214 (0.06974-0.2144) p=0.0001</td>
<td>0.1627 (0.0866-0.3056) p=0.0001</td>
<td>0.1445 (0.0808-0.2582) p=0.0001</td>
<td>0.0422 (0.0173-0.1035) p=0.0001</td>
<td>0.0296 (0.0098-0.0476) p=0.0001</td>
</tr>
<tr>
<td>Treatment (k/\nu &gt; \text{ratio}_{\text{thresh}}) vs. Treatment (All)</td>
<td>0.0675 (0.0234-0.1949) p=0.0001</td>
<td>0.1191 (0.0119-1.188) p=0.0697</td>
<td>0.101 (0.0249-0.4103) p=0.0013</td>
<td>0.5683 (0.33490.9643) p=0.0362</td>
<td>0.9921 (0.6117-1.609) p=0.9742</td>
</tr>
<tr>
<td>Treatment (k/\nu &gt; \text{ratio}_{\text{thresh}}) vs. Treatment (All)</td>
<td>0.2562 (0.1239-0.5299) p=0.0002</td>
<td>0.2538 (0.1191-0.5408) p=0.0004</td>
<td>0.239 (0.1236-0.4621) p=0.0001</td>
<td>0.6138 (0.3732-1.01) p=0.0546</td>
<td>0.576 (0.3863-0.8588) p=0.0068</td>
</tr>
<tr>
<td>Treatment (All) vs. Control (All)</td>
<td>0.2307 (0.1548-0.3438) p=0.0001</td>
<td>0.6742 (0.4384-1.037) p=0.0726</td>
<td>0.5845 (0.3934-0.8684) p=0.0079</td>
<td>0.6481 (0.4331-0.9699) p=0.0350</td>
<td>0.9959 (0.7038-1.409) p=0.9815</td>
</tr>
<tr>
<td>Treatment (k/\nu &gt; \text{ratio}_{\text{thresh}}) vs. Treatment (All)</td>
<td>1.569 (0.9721-2.41) p=0.0549</td>
<td>1.558 (0.9714-2.5) p=0.0658</td>
<td>1.794 (1.158-2.778) p=0.0089</td>
<td>6.405 (2.971-13.81) p=0.0001</td>
<td>7.657 (4.065-14.42) p&lt;0.0001</td>
</tr>
<tr>
<td>Treatment (k/\nu &gt; \text{ratio}_{\text{thresh}}) vs. Treatment (All)</td>
<td>0.0673 (0.0228-0.157) p=0.0001</td>
<td>0.2332 (0.1191-0.4566) p=0.0001</td>
<td>0.0110 (0.0039-0.315) p=0.0001</td>
<td>0.358 (0.0157-0.0815) p=0.0001</td>
<td>0.1862 (0.1043-0.3327) p=0.0001</td>
</tr>
<tr>
<td>Control (k/\nu &gt; \text{ratio}<em>{\text{thresh}}) vs. Control (k/\nu &lt; \text{ratio}</em>{\text{thresh}})</td>
<td>0.0368 (0.01750-0.3077) p=0.00001</td>
<td>0.1655 (0.0925-0.2962) p=0.0001</td>
<td>0.0110 (0.0039-0.315) p=0.0001</td>
<td>0.0418 (0.0204-0.0857) p=0.0001</td>
<td>0.6573 (0.5219-0.8029) p=0.0001</td>
</tr>
<tr>
<td>Treatment (k/\nu &gt; \text{ratio}<em>{\text{thresh}}) vs. Control (k/\nu &lt; \text{ratio}</em>{\text{thresh}})</td>
<td>0.1592 (0.0817-0.3102) p=0.00001</td>
<td>0.1707 (0.0557-0.5226) p=0.0001</td>
<td>0.7027 (0.4718-1.047) p=0.0825</td>
<td>0.0842 (0.0452-0.1568) p=0.00001</td>
<td>0.0782 (0.0561-0.1089) p=0.00001</td>
</tr>
<tr>
<td>Treatment (k/\nu &gt; \text{ratio}_{\text{thresh}}) vs. Treatment (All)</td>
<td>0.3364 (0.1662-0.6806) p=0.00001</td>
<td>0.2513 (0.1294-0.4881) p=0.00001</td>
<td>0.7306 (0.5008-1.066) p=0.1033</td>
<td>0.2524 (0.126-0.5058) p=0.00001</td>
<td>0.317 (0.176-0.571) p=0.00001</td>
</tr>
<tr>
<td>Treatment (All) vs. Control (All)</td>
<td>0.266 (0.1733-0.4083) p=0.00001</td>
<td>0.6742 (0.4384-1.037) p=0.0726</td>
<td>0.7766 (0.5543-1.088) p=0.1416</td>
<td>0.2016 (0.1343-0.3027) p=0.00001</td>
<td>0.0940 (0.0762-0.1159) p=0.00001</td>
</tr>
<tr>
<td>Treatment (k/\nu &lt; \text{ratio}_{\text{thresh}}) vs. Treatment (All)</td>
<td>3.938 (1.985-7.811) p=0.00001</td>
<td>2.119 (1.294-3.47) p=0.0005</td>
<td>15.64 (6.536-37.41) p=0.00001</td>
<td>5.074 (2.653-9.705) p=0.00001</td>
<td>0.5152 (0.3481-0.8458) p=0.00001</td>
</tr>
</tbody>
</table>

\textsuperscript{†}protocol A, \textsuperscript{‡}protocol Z, \textsuperscript{§}protocol V11a, \textsuperscript{¶}protocol V11a-D
### Table 3. Summary of median survival of population separated by median $k_{1,\text{thresh}}$

<table>
<thead>
<tr>
<th>Median survival (days)</th>
<th>Roland†</th>
<th>Zibara‡</th>
<th>Tan‡</th>
<th>Volk 2008§</th>
<th>Volk 2011a†</th>
<th>Volk 2011b‡</th>
<th>Zibara§</th>
<th>Volk 2011a¶</th>
<th>Volk 2011a¶¶</th>
<th>Mollard‡</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control (All)</td>
<td>53</td>
<td>55</td>
<td>58</td>
<td>64</td>
<td>69</td>
<td>68</td>
<td>55</td>
<td>69</td>
<td>69</td>
<td>87</td>
</tr>
<tr>
<td>Control ($k_1 &lt; k_{1,\text{thresh}}$)</td>
<td>59.5</td>
<td>84</td>
<td>86</td>
<td>Un*</td>
<td>87</td>
<td>Un*</td>
<td>84</td>
<td>87</td>
<td>87</td>
<td>92</td>
</tr>
<tr>
<td>Control ($k_1 &gt; k_{1,\text{thresh}}$)</td>
<td>35</td>
<td>45</td>
<td>43</td>
<td>41</td>
<td>44</td>
<td>45</td>
<td>45</td>
<td>44</td>
<td>44</td>
<td>72</td>
</tr>
<tr>
<td>Treatment (All)</td>
<td>Un*</td>
<td>63</td>
<td>77.5</td>
<td>Un*</td>
<td>69</td>
<td>Un*</td>
<td>63</td>
<td>71</td>
<td>Un*</td>
<td>Un*</td>
</tr>
<tr>
<td>Treatment ($k_1 &lt; k_{1,\text{thresh}}$)</td>
<td>Un*</td>
<td>Un*</td>
<td>Un*</td>
<td>Un*</td>
<td>88</td>
<td>Un*</td>
<td>108</td>
<td>108</td>
<td>Un*</td>
<td>Un*</td>
</tr>
<tr>
<td>Treatment ($k_1 &gt; k_{1,\text{thresh}}$)</td>
<td>55</td>
<td>46</td>
<td>46</td>
<td>47.5</td>
<td>44</td>
<td>82</td>
<td>44</td>
<td>44</td>
<td>78.5</td>
<td>Un*</td>
</tr>
</tbody>
</table>

† protocol A, ‡ protocol Z, § protocol V11a, ¶ protocol V11a-D
Table 4. Statistics comparing the Kaplan-Meier survival curves of population separated by median $k_{t,\text{thresh}}$: hazard ratio (95%CI) and log rank test $p$-values

<table>
<thead>
<tr>
<th>HR (95% CI)</th>
<th>Roland(^1)</th>
<th>Zibara(^2)</th>
<th>Tan(^3)</th>
<th>Volk2008(^4)</th>
<th>Volk2011a(^5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment ($k_t &lt; k_{t,\text{thresh}}$) vs. Treatment ($k_t &gt; k_{t,\text{thresh}}$)</td>
<td>0.0012 (0.0004-0.0041)</td>
<td>0.0904 (0.0456-0.1793)</td>
<td>0.0794 (0.0422-0.1491)</td>
<td>0.0241 (0.0118-0.0493)</td>
<td>0.0138 (0.0063-0.0301)</td>
</tr>
<tr>
<td>Control ($k_t &lt; k_{t,\text{thresh}}$) vs. Control ($k_t &gt; k_{t,\text{thresh}}$)</td>
<td>0.5882 (0.3549-0.9751)</td>
<td>0.0832 (0.0421-0.1643)</td>
<td>0.0809 (0.0430-0.152)</td>
<td>0.0241 (0.0118-0.0492)</td>
<td>0.01376 (0.0063-0.0301)</td>
</tr>
<tr>
<td>Treatment ($k_t &lt; k_{t,\text{thresh}}$) vs. Treatment (All)</td>
<td>0.134 (0.0805-0.2231)</td>
<td>0.0927 (0.0258-0.3324)</td>
<td>0.0843 (0.0287-0.2471)</td>
<td>0.1775 (0.0682-0.462)</td>
<td>0.9909 (0.5909-1.662)</td>
</tr>
<tr>
<td>Treatment (All) vs. Treatment (All)</td>
<td>0.5033 (0.2978-0.8505)</td>
<td>0.2096 (0.1054-0.4169)</td>
<td>0.2076 (0.1112-0.3874)</td>
<td>0.2094 (0.1143-0.3839)</td>
<td>0.515 (0.3422-0.7749)</td>
</tr>
<tr>
<td>Treatment (All) vs. Control (All)</td>
<td>0.2307 (0.1548-0.3438)</td>
<td>0.6742 (0.4384-1.037)</td>
<td>0.5845 (0.3934-0.8684)</td>
<td>0.6481 (0.4331-0.9699)</td>
<td>0.9959 (0.7038-1.409)</td>
</tr>
<tr>
<td>Treatment (All) vs. Control (All)</td>
<td>30.33 (12.08-76.16)</td>
<td>2.553 (1.538-4.239)</td>
<td>2.759 (1.728-4.405)</td>
<td>5.824 (3.356-10.1)</td>
<td>8.024 (4.379-14.7)</td>
</tr>
<tr>
<td>Treatment ($k_t &gt; k_{t,\text{thresh}}$) vs. Treatment (All)</td>
<td>0.0662 (0.0301-0.1456)</td>
<td>0.0904 (0.0456-0.1793)</td>
<td>0.0138 (0.0063-0.0301)</td>
<td>0.0265 (0.0118-0.0595)</td>
<td>0.1265 (0.0619-0.2588)</td>
</tr>
<tr>
<td>Treatment ($k_t &gt; k_{t,\text{thresh}}$) vs. Treatment (All)</td>
<td>0.0254 (0.0126-0.0513)</td>
<td>0.0832 (0.0421-0.1643)</td>
<td>0.0138 (0.0063-0.0301)</td>
<td>0.0138 (0.0063-0.0301)</td>
<td>0.6814 (0.5175-0.8971)</td>
</tr>
<tr>
<td>Treatment ($k_t &gt; k_{t,\text{thresh}}$) vs. Control (All)</td>
<td>0.0960 (0.0410-0.2244)</td>
<td>0.0927 (0.0258-0.3324)</td>
<td>0.5917 (0.3641-0.9618)</td>
<td>0.0636 (0.0329-0.1233)</td>
<td>0.0781 (0.0611-0.0999)</td>
</tr>
<tr>
<td>Treatment ($k_t &gt; k_{t,\text{thresh}}$) vs. Control (All)</td>
<td>0.1959 (0.0899-0.4268)</td>
<td>0.2096 (0.1054-0.4169)</td>
<td>0.5129 (0.3404-0.7728)</td>
<td>0.1783 (0.0858-0.3702)</td>
<td>0.5779 (0.3567-0.9361)</td>
</tr>
<tr>
<td>Treatment (All) vs. Control (All)</td>
<td>0.266 (0.1733-0.4083)</td>
<td>0.6742 (0.4384-1.037)</td>
<td>0.7766 (0.5543-1.088)</td>
<td>0.2016 (0.1343-0.3027)</td>
<td>0.0940 (0.0762-0.1159)</td>
</tr>
<tr>
<td>Treatment (All) vs. Control (All)</td>
<td>3.248 (1.77-5.959)</td>
<td>2.168 (1.336-3.516)</td>
<td>8.024 (4.379-14.7)</td>
<td>3.518 (1.913-6.472)</td>
<td>3.444 (1.863-6.365)</td>
</tr>
</tbody>
</table>

\(^1\)protocol A, \(^2\)protocol Z, \(^3\)protocol V11a, \(^4\)protocol V11a-D
Figures:

Figure 1. Schematic and overview of computational model of tumor-bearing mice. The three-compartment mouse model was previously trained and validated. The pro-angiogenic signal \( (\text{Ang}(t)) \) is calculated as the summation of the concentrations of VEGF-bound receptor complexes in the tumor endothelium. The dynamic tumor volume is a function of the angiogenic signal. In this study, we randomly varied tumor growth parameters within specified ranges to simulated tumor growth of several heterogeneous mouse populations.

Figure 2. Model-simulated tumor growth data of \textit{in silico} mouse populations. The whole-body mouse model previously fit to each of the six datasets individually was used to simulate tumor volume over time. To generate the simulated tumors, the tumor growth kinetic parameters \( k_0 \) and \( k_1 \) were randomly varied within the range of the estimated values. A total of 400 simulations were run for each case. The mean and 95% confidence interval at each time point are shown. A, Roland. B, Zibara. C, Tan. D, Volk2008. E, Volk2011a. F, Volk2011b. Asterisks indicate that the difference between the control and treatment group tumor volumes is statistically significant \((p<0.05)\).

Figure 3. Range of parameter and threshold values. In each of the six cases, values of \( k_{1,\text{thresh}} \) and \( \text{ratio}_{\text{thresh}} \) were found among all of the randomly generated values of \( k_1 \) or \( k_0/k_1 \) ratio used in the simulations. A, \( k_0 \). B, \( k_1 \). C, \( k_0/k_1 \) ratio. Bars: the ranges of all generated parameter values in each case. Boxes: the ranges of possible threshold values in each case. Shading: the common range of threshold values among the six cases.

Figure 4. Kaplan-Meier curves for the six simulated groups of tumor-bearing mice. Here, the \( \text{ratio}_{\text{thresh}} \) value is taken as the median from the common range found among the six cases (13.8693). A, Roland. B, Zibara. C, Tan. D, Volk2008. E, Volk2011a. F, Volk2011b. The
estimated survival curves of *in silico* mice subgroups within each group are shown in each plot: all mice, mice with ratio above or below the median $\text{ratio}_{\text{thresh}}$ in the control setting or with treatment.

**Figure 5.** Validation of $\text{ratio}_{\text{thresh}}$ and $k_{1,\text{thresh}}$ values with an independent set of data from [41]. **A,** Model fit to control data and validation with treatment data from [41]. The model was fit to normalized tumor volume, and the tumor growth kinetic parameters were estimated. The model is able to reproduce experimental data in the control group and predict the treatment data. **Line:** mean of best fits. **Shading:** range of standard deviation. **Squares:** experimental data. **Error values:** SSR for mean of the best fits. **B,** Model-simulated tumor growth of an *in silico* mouse population, with tumor growth kinetic parameters $k_0$ and $k_1$ for each simulation randomly varied within the range of their estimated values. The mean and 95% confidence interval at each time point are shown. Asterisks indicate that the difference between the control and treatment group tumor volumes is statistically significant ($p<0.05$). **C and D,** Estimated Kaplan-Meier survival curves of the simulated mouse population obtained using the model that was fitted to Mollard data. The population is separated using the median of the range of **C,** $\text{ratio}_{\text{thresh}}$ values (13.8693), or **D,** $k_{1,\text{thresh}}$ values ($1.661 \times 10^{-6}$).

**Figure 6.** Dynamics of tumor volume. **A,** Time course of relative tumor volume (RTV) for the Mollard case. The mean RTV for all *in silico* mice and mice with tumors whose $k_0/k_1$ ratio is smaller and larger than the median $\text{ratio}_{\text{thresh}}$ (13.8693) are shown. **B,** Log-transformed tumor volume data for all *in silico* mice and mice separated according to the tumor’s $k_0/k_1$ ratio.

**Figure 7.** Mean tumor growth. We plot the mean tumor volume for all *in silico* mice in control and treatment groups using the model fitted to Roland data. **A,** Linear scale and **B,** Log scale.
**Supplemental Figures:**

**Figure S1.** Time course of relative tumor volume (RTV). The mean RTV levels of all *in silico* mice and the groups with $k_0/k_1$ smaller or larger than the median $\text{ratio}_{\text{thresh}}$ (13.8693) are shown. A, Roland. B, Zibara. C, Tan. D, Volk2008. E, Volk2011a. F, Volk2011b.

**Figure S2.** Scatter plot of RTV at the end of simulations versus tumor growth kinetic parameters. Left to right columns: $k_0$, $k_1$, and $k_0/k_1$ ratio. A, Roland. B, Zibara. C, Tan. D, Volk2008. E, Volk2011a. F, Volk2011b. Color gradient represents the range of RTV values (from 0 to 1).

**Figure S3.** Kaplan-Meier curves for the six simulated groups of tumor-bearing mice. Here, the $k_{1,\text{thresh}}$ value is taken as the median from the common range found among the six cases (1.661×10⁻⁶). A, Roland. B, Zibara. C, Tan. D, Volk2008. E, Volk2011a. F, Volk2011b. The estimated survival curves of *in silico* mice subgroups within each group are shown in each plot: all mice, mice with $k_1$ smaller or larger than the median $k_{1,\text{thresh}}$, in the control setting or with treatment.

**Figure S4.** Model-simulated tumor growth data with alternative treatment protocols. The mean and 95% confidence interval at each time point are shown. A, Zibara case with treatment protocol Z. B, Volk2011a case with treatment protocol V11a. C, Volk2011a case with treatment protocol V11a-D (see Methods). Asterisks indicate that the difference between the control group and the treatment group tumor volumes is statistically significant (p<0.05).

Figure S6. Comparison of normalized experimental data. Control tumor volume on day eight is extrapolated from an exponential fit to the experimental data [35–39], and used to calculate the relative tumor volumes for the models fit to the control tumor volume from each of the six datasets. The resulting normalized control tumor volume datasets are compared to that from the Mollard study.

Supplementary Material

Supplemental Tables:
Table S1: Experimental measurements and estimated parameter values from fitting to Mollard dataset
Table S2: Parameter bounds and values used in simulations
Table S3: Description of treatment protocols

Supplemental Figures:
Figure S1. Time course for the relative tumor volume (RTV) vs. time
Figure S2. Relationship between RTV and tumor growth kinetic parameters
Figure S3. Kaplan-Meier curves for tumor-bearing mouse population separated using $k_{1,\text{thresh}}$
Figure S4. Tumor volume time course for Z and V11a, V11a-D protocols
Figure S5. Time course for log-transformed tumor volume
Figure S6. Experimental data for normalized control tumor volume extracted from published datasets

Supplemental Files:
File S1. Description of three-compartment model
<table>
<thead>
<tr>
<th></th>
<th><strong>File</strong></th>
<th><strong>Description</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>116</td>
<td>S2.</td>
<td>Zipped file containing the computational model, as a MATLAB SimBiology file, SBML,</td>
</tr>
<tr>
<td>117</td>
<td></td>
<td>and a MATLAB .m file</td>
</tr>
<tr>
<td>118</td>
<td>S3.</td>
<td>Supplemental Tables</td>
</tr>
<tr>
<td>119</td>
<td>S4.</td>
<td>Supplemental Figures</td>
</tr>
</tbody>
</table>
The figure shows the results of a study on tumor growth and survival in control and treatment groups over time.

**A. Normalized Tumor Volume**
- **Error** (Control): $1.5 \times 10^4$
- **Error** (Treatment): $6.71 \times 10^3$
- **Error** (Total): $2.17 \times 10^4$

**B. Tumor Volume (cm$^3$)**
- Comparison of tumor volume over time between control and treatment groups.

**C. Survival (%)**
- Survival rates are compared across different conditions.

**D. Survival (%)**
- Survival trends are shown for both control and treatment groups, with specific conditions indicated.
A) RTV vs. Time (days)

B) Log Tumor Volume (cm³) vs. Time (days)

- Control (All)
- Control (ratio < ratio_{thresh})
- Control (ratio > ratio_{thresh})
- Treatment (All)
- Treatment (ratio < ratio_{thresh})
- Treatment (ratio > ratio_{thresh})

The data shows the effect of different treatment ratios on tumor volume over time.