ARTICLE

Growth-rate model predicts in vivo tumor response from in vitro data

Rocky Diegmiller[1] | Laurent Salphati[2] | Bruno Alicke[3] | Timothy R. Wilson[4] | Thomas J. Stout[5] | Marc Hafner[6]

[1]Department of Chemical and Biological Engineering and Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, New Jersey, USA

[2]Department of Drug Metabolism and Pharmacokinetics, Genentech Inc., South San Francisco, California, USA

[3]Department of Translational Oncology, Genentech Inc., South San Francisco, California, USA

[4]Department of Oncology Biomarker Development, Genentech Inc., South San Francisco, California, USA

[5]Department of Product Development Oncology, Genentech Inc., South San Francisco, California, USA

[6]Department of Oncology Bioinformatics, Genentech Inc., South San Francisco, California, USA

Correspondence
Marc Hafner, Department of Oncology Bioinformatics, Genentech Inc., 1 DNA Way, South San Francisco, CA 94080, USA.
Email: hafner.marc@gene.com

Abstract

A major challenge in oncology drug development is to elucidate why drugs that show promising results in cancer cell lines in vitro fail in mouse studies or human trials. One of the fundamental steps toward solving this problem is to better predict how in vitro potency translates into in vivo efficacy. A common approach to infer whether a model will respond in vivo is based on in vitro half-maximal inhibitory concentration values (IC_{50}), but yields limited quantitative comparison between cell lines and drugs, potentially because cell division and death rates differ between cell lines and in vivo models. Other methods based either on mechanistic modeling or machine learning require molecular insights or extensive training data, limiting their use for early drug development. To address these challenges, we propose a mathematical model integrating in vitro growth rate inhibition values with pharmacokinetic parameters to estimate in vivo drug response. Upon calibration with a drug-specific factor, our model yields precise estimates of tumor growth rate inhibition for in vivo studies based on in vitro data. We then demonstrate how our model can be used to study dosing schedules and perform sensitivity analyses. In addition, it provides meaningful metrics to assess association with genotypes and guide clinical trial design. By relying on commonly collected data, our approach shows great promise for optimizing drug development, better characterizing the efficacy of novel molecules targeting proliferation, and identifying more robust biomarkers of sensitivity while limiting the number of in vivo experiments.

Study Highlights

WHAT IS THE CURRENT KNOWLEDGE ON THE TOPIC?
Current methods to predict how a cancer model responds in vivo either yield limited quantitative comparisons or require extensive training data, limiting their use for early drug development.
WHAT QUESTION DID THIS STUDY ADDRESS?
How well can we predict in vivo efficacy from in vitro potency data and which pharmacokinetic parameters affect tumor growth inhibition?

WHAT DOES THIS STUDY ADD TO OUR KNOWLEDGE?
Because of its prediction power and reliance on in vitro data, our model can be used for estimating efficacy in early drug development, studying dosing schedules, performing sensitivity analyses, and providing meaningful metrics for biomarker association studies.
HOW MIGHT THIS CHANGE DRUG DISCOVERY, DEVELOPMENT, AND/OR THERAPEUTICS?
The simplicity of our method and availability of the input data may improve drug development and support the design of clinical trials.

INTRODUCTION

In the first phases of drug development in oncology, candidate molecules are tested in vitro on workhorse cell lines as a primary method of evaluating efficacy and potency. Whereas in vitro studies can give baseline estimates for the activity of a drug, animal studies are critical to provide insights into the efficacy of a drug in a living system, accounting for pharmacokinetic (PK) properties, intrinsic tumor growth, tumor vascularization, microenvironment, and toxicity.[1–6] However, commonly used approaches rely on heuristic processes to link in vitro potency to the efficacy of the drugs on tumor models. Many animal studies of drug response typically compare expected free concentration to in vitro potency metrics like half-maximal inhibitory concentration (IC_{50}), the concentration at which viability is half of the control condition.[7,8] However, IC_{50} values capture only a single point of the dose–response curve, ignoring differences in maximal in vitro efficacy[9] and, along with other metrics based on relative viability, is not robust to the number of cell divisions that take place throughout an experiment.[10]

Methods to improve on the prediction of in vivo and clinical responses based on in vitro data can be broadly categorized into mechanistic or statistical models. First, drug effects are modeled at the molecular level, accounting for known underlying biochemical processes in vivo[2,11–17] as reviewed in Danhof et al.[2] Other mathematical models have focused on population dynamics to predict how scheduling of dosing regimens and combinations of therapies can improve efficacy and address the emergence of resistance.[18,19] Second, statistical and machine learning methods aim to leverage large datasets collected on various cancer cell lines, animal studies, and clinical data.[7,20–24] Although both types of approaches provide useful insights and assist pharmaceutical development, they are highly dependent on the depth, amount, and quality of data collected on drugs, as well as how much in vivo and clinical data are available, limiting their use for early drug development. In this paper, we introduce a model based on growth rates that addresses these limitations and can be used broadly in the drug development process and translational research.

Because in vitro viability data are often readily available in early drug development, we built a mathematical model of tumor growth based on such data. We hypothesized that biochemical and molecular responses, as well as birth-death processes, can be abstracted, such that concentration-dependent in vitro growth inhibition can directly predict in vivo efficacy when combined with in vivo PK models. Instead of relative viability, we relied on growth rate (GR) inhibition values for normalizing the endpoint in vitro data to account for the variability in growth rates observed between cell lines and experimental settings.[9,10] We presumed that our model required limited training data and could be applied without knowledge of the underlying mechanism(s) of action of the drug being considered, as long as the primary effect of the drug was on cancer cells and not the immune system or host cells. Using data from drugs individually targeting the PI3K/AKT or MAPK pathways, we demonstrated accurate predictions of in vivo dose-escalation studies across various cell lines from different genotypes. The method outlined in this work is independent of the host organism and can be used with human PK data to support the design of clinical trials. An R package with our model is available at github.com/Genentech/TGRmodel.

RESULTS

Modeling the intrinsic tumor growth

Instead of relying on absolute tumor volume or cell population, our approach focuses on the growth rate under treatment relative to the untreated growth rate given that cell doubling times differ when grown in cultures or as xenografts. To model the acute and intrinsic response of tumor cells to treatment, which is

commonly measured in early drug development, we hypothesized that a given drug at a given concentration would proportionally reduce the growth rate of cells to the same extent both in vivo and in vitro. In addition, we assumed that the growth rates, whether positive or negative, remain constant throughout the experiments; that is, growth is exponential. Because this assumption is valid for the early intrinsic response, delayed resistance or cell extrinsic effects are beyond the scope of our model. To quantify the in vivo response, we used tumor growth rate inhibition values (TGR, see Methods S1). A TGR value of 1 corresponds to the growth of the vehicle-treated reference cohort, 0 to a stable disease (no net growth), and negative values to tumor shrinkage. TGR is related to the more common tumor growth inhibition (TGI) metric, but has the advantage of being independent of the untreated growth rate and experiment duration (see Methods S1). Furthermore, TGR can be generalized to nonexponential growth by replacing the exponential growth rate with any growth rate metric, such as the end point Gain Integrated in Time.[25]

For normalizing the in vitro data, we used the GR method[10] that quantifies the efficacy of the drug in terms of growth instead of end point relative viability. Metrics related to relative viability, such as IC_{50} values, are confounded by the number of cell divisions which can lead to artifactual results,[9] especially regarding efficacy.[10] In addition to the in vitro GR metrics, we used the PK profile – either simulated or measured – of the drug within the host to model the drug concentration profile throughout the experiment. Most studies consisted of one dose per day (q.d.), but our model can account for any dosing schedule, including irregular ones. As the in vitro GR function depends on concentration and the PK profile provides a function of the concentration over time, we calculated the instantaneous TGR (iTGR) values. Averaging the iTGR values over the duration of the experiment yielded a predicted TGR value ($\widehat{TGR}$). Modeling multiple dosages yielded a dose response profile that could be compared to in vivo dose-escalation efficacy studies to calibrate the model and account for factors affecting the effective in vivo concentration not captured by our PK model, such as tumor/plasma ratio different than 1 or interactions between the drug and the organism[1–3] (Figure 1a; see Methods S1 for details). This resulted in the equation:

$$\widehat{TGR}(d) = TGR_{\infty} + \left(1 - TGR_{\infty}\right) / \left(1 + \left(\alpha \cdot d / TGEC_{50}\right)^{h_{TGR}}\right),$$

With d, the dose and TGR_{∞}, $TGEC_{50}$, and h_{TGR}, parameters predicted by the in vitro and PK data and α trained on limited in vivo data (see Methods S1).

Fitting and prediction

We gathered data from in vivo efficacy studies comprising 23 cell lines from multiple diseases treated with one of seven drugs targeting either the PI3K/AKT or the MAPK pathways, including compounds currently in clinical trials or approved, and their corresponding in vitro data[26–33] (Data Files S1 and S2). After excluding studies with unexpectedly poor growth of the untreated tumors (mean TGR values above 1.2) or experiments with large variability between animals (SEM above 0.25), we split our in vivo data into a training set with studies containing at least three dosages and a test set with other studies. If no individual studies for a given drug had enough dosages, data from multiple studies using the same cell line were aggregated. Using in vitro and PK data from the training set (Data Files S2 and S3), we calculated a single calibration factor, α, for each drug as the geometric mean of α values measured across all studies of the same drug in the training set, independent of the cell lines used (Data File S4). The fitted data from the training set were highly correlated with the experimental data (Spearman's $\rho = 0.85$, $p < 2.2e-16$; Figure 2a). Then, we predicted the responses from cell lines that were not part of the training set, obtaining a highly significant prediction with a Spearman's ρ of 0.64 ($p = 4.8e-9$; Figure S1). When including data from cell lines found in the training set but from different studies than the training data, we obtained an overall Spearman's ρ of 0.74 ($p < 2.2e-16$; Figure 2b, Data File S5), with even higher correlation values for some drugs (Figure S2).

When comparing the quality of our prediction to the biological variability measured between studies with the same cell line and drug combination, we observed that 60.4% of the conditions in the training set and 56.4% of the test set were within one standard deviation of biological replicates (value of 0.20 shown as a gray band in Figure 2; Figure S3). It is worth noting that four outlier values (red arrow on the left) were from GDC-0032/taselisib studies that had surprisingly weaker responses compared to biological replicates. For the other discrepancies, our model tended to underpredict growth inhibition. In the case of MAPK pathway inhibitors, this error could potentially be due to the increased sensitivity of 3D cultures.[33,34]

Our model predictions also compared well against a common approach for estimating drug efficacy in vivo: the fraction of time that the free drug concentration is above the in vitro IC_{50}, ($FT_{IC_{50}}$), or other arbitrary cutoff based on relative viability.[35,36] Although we found that $FT_{IC_{50}}$ values correlated with tumor growth (Spearman's $\rho = -0.62$, $p = 2.5e-16$; Figure S4), this metric has limited information content (Normalized Shannon's entropy of 0.39); the large majority of $FT_{IC_{50}}$ values are either 0 or 1, even if other cutoff values (e.g., IC_{90}) or a scaling factor
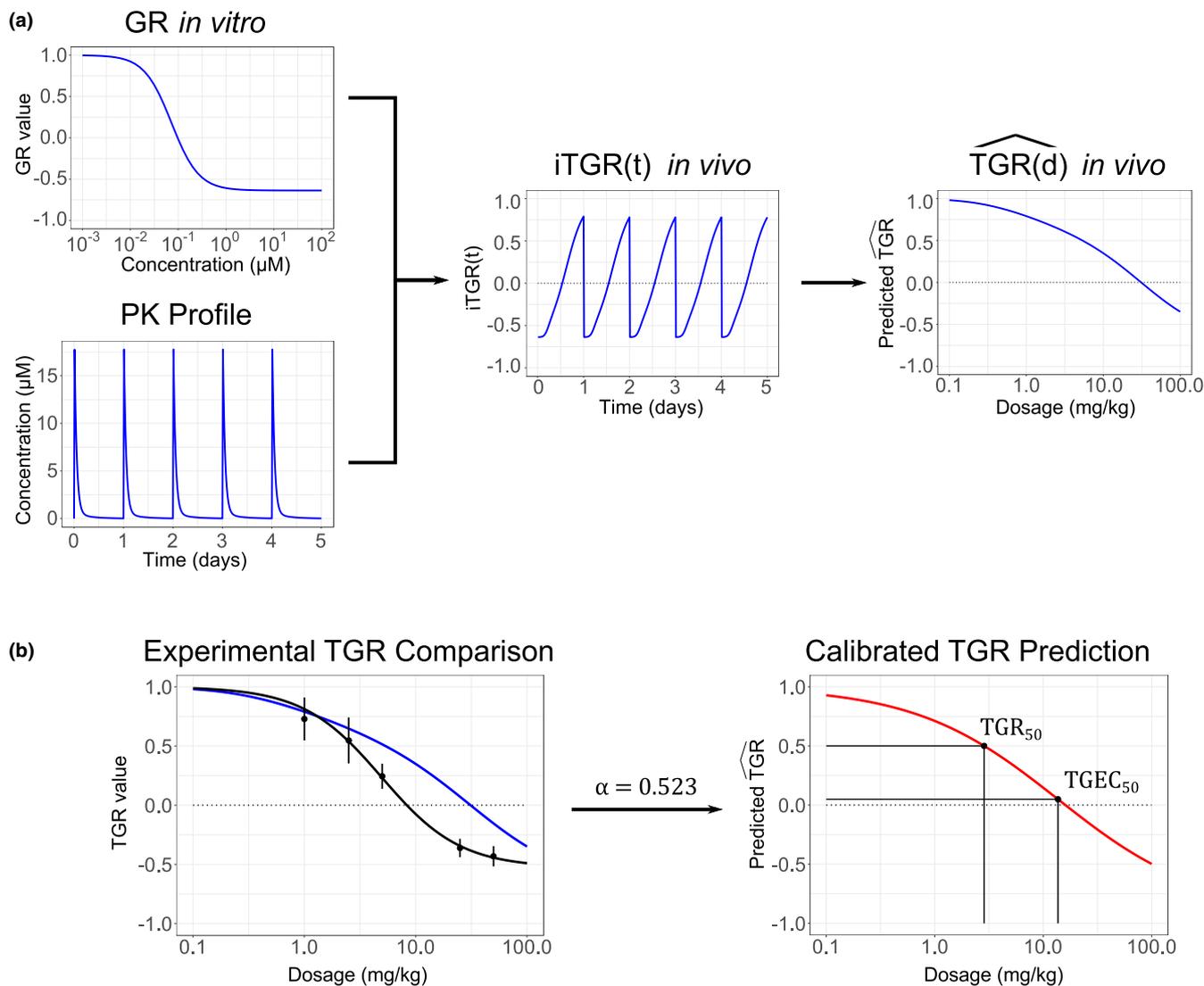
FIGURE 1 Schematic for TGR predictions based on in vitro GR and PK data. (a) Using in vitro GR values and the PK profile, we predict the drug effects on tumor growth rate over time at a given dosage. Then, we average the effect on growth across the experiment to obtain the average tumor growth rate inhibition at each dosage. (b) When comparing the experimental data (gray dots; fit denoted with black line; black dots at each dosage indicate average value with standard deviation) to the prediction (in blue), we infer a drug-specific calibration factor (α) that is applied for predictions (right). GR, growth rate; iTGR, instantaneous tumor growth rate; PK, pharmacokinetic; TGR, tumor growth rate.

were used. In contrast, TGR prediction values were much better distributed (entropy of 0.87), closer to the experimental distribution (entropy of 0.92) and allowed for more quantitative comparisons between drugs and cell lines. Moreover, the normalized mutual information between the experimental data and the predicted TGR values was substantially higher than for FT_{IC50} values (0.52 vs. 0.33).

Sensitivity to dosing schedule and drug parameters

Next, we tested how dosing schedules affect efficacy in vivo. As an example, we attempted to predict the response

of MCF-7 breast cancer cells to GDC-0068/ipatasertib given twice a day (b.i.d.) using a model trained on q.d. data (triangles). Our prediction (Figure 3a, plain line) laid within the standard error across the experimental measurements at two separate b.i.d. dosages (circles). It should also be noted that doubling the frequency did not result in a strong increase in efficacy, a difference that other metrics like FT_{IC50} could not capture well (red lines). In addition, sensitivity analysis could predict which b.i.d. dosage for a given drug and cell line pair yielded the same efficacy in terms of tumor growth as a given q.d. dosage (see Supporting Information). For GDC-0068, a b.i.d. schedule with a dosage of 35% of the q.d. one yielded the same efficacy. We found that this ratio depended on the elimination rate (k_e): a three-fold increase in k_e allowed for
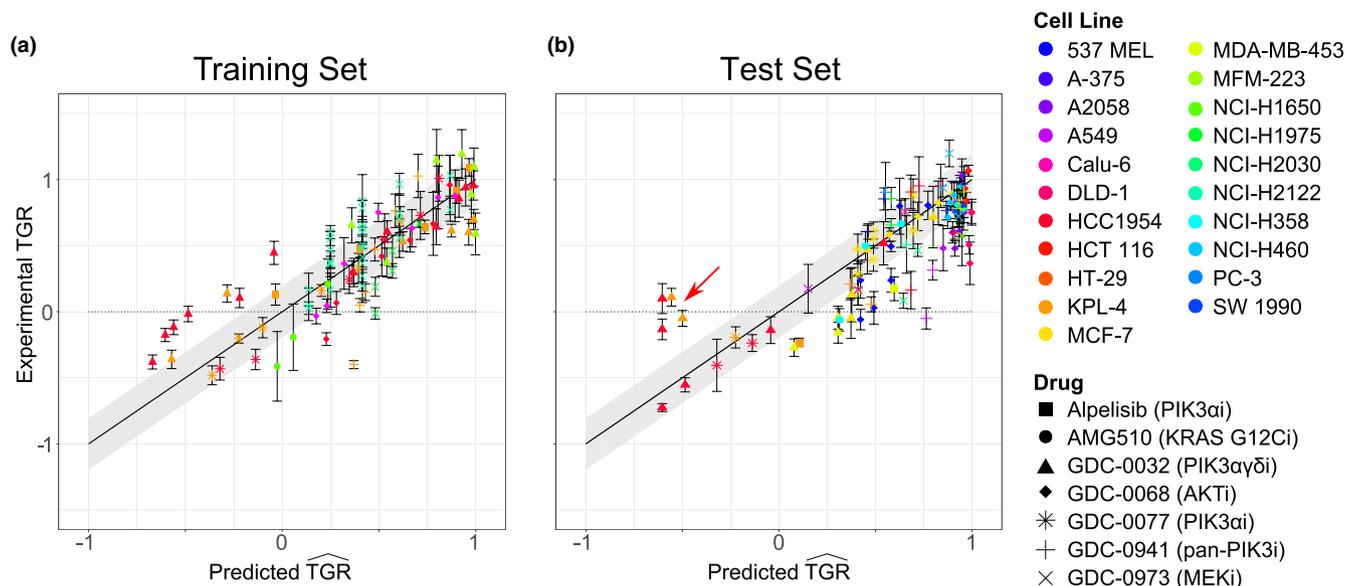
FIGURE 2 Training and prediction of TGR values. (a) Measured TGR values against fitted $\widehat{\text{TGR}}$ values for the training data. The shape of the marker indicates the drug, whereas the color indicates the cell line. All error bars are standard deviations of the experimental measurements. The gray band around the line $y = x$ indicates the standard deviation extracted from all studies with biological replicates. (b) Experimental TGR values against predicted $\widehat{\text{TGR}}$ values for the test data. Same legend as (a). TGR, tumor growth rate; $\widehat{\text{TGR}}$, predicted tumor growth rate value.
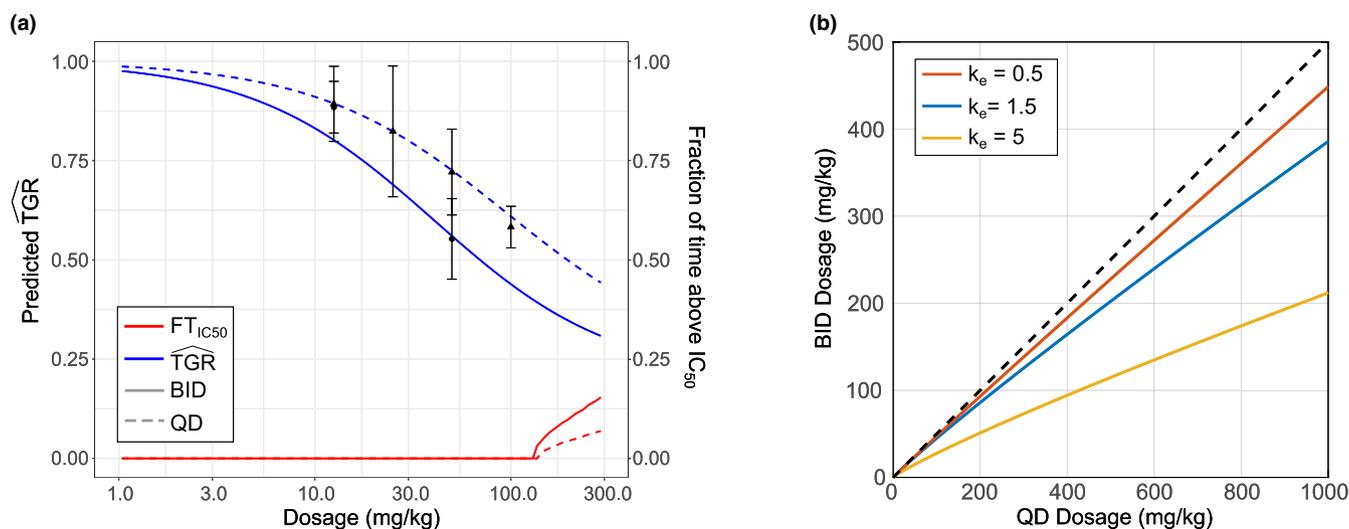
FIGURE 3 Alternative dosing schedules and comparison with time above IC_{50}. (a) Predicted TGR values for GDC-0068/ipatasertib on MCF-7 tumor cells administered as b.i.d. (blue) or q.d. (dotted blue), experimental results (circles and triangles, respectively), and fraction of time above IC_{50} (FT_{IC50}) for each dosing schedule (in red and dotted red, respectively). (b) Equivalent b.i.d. dosage as a function of q.d. dosage for GDC-0068 on MCF-7 cells based on their respective $\widehat{\text{TGR}}$ predicted fits for various values of elimination rate k_e. IC_{50}, half-maximal inhibitory concentration; TGR, tumor growth rate; $\widehat{\text{TGR}}$, predicted tumor growth rate value.

ratios as low as around 20%, whereas a three-fold decrease pushed ratios much closer to 50% (Figure 3b). Our model also allowed us to test treatment regimens with variable dosage and calculate the overall response (Figure S5). These examples illustrate how our model enables the optimization of dosing regimens and treatment strategies based on drug properties in a quantitative manner.

By performing sensitivity analyses on all drug parameters, we found that lower values for either elimination rate (k_e), in vitro potency (GEC_{50}), or in vitro maximal efficacy (GR_∞) yielded higher efficacy, as expected (see Figure S6). In addition, the Hill coefficient, which is assumed to capture how heterogeneous the response is across cells,[37] was a good example for elucidating the dependencies between
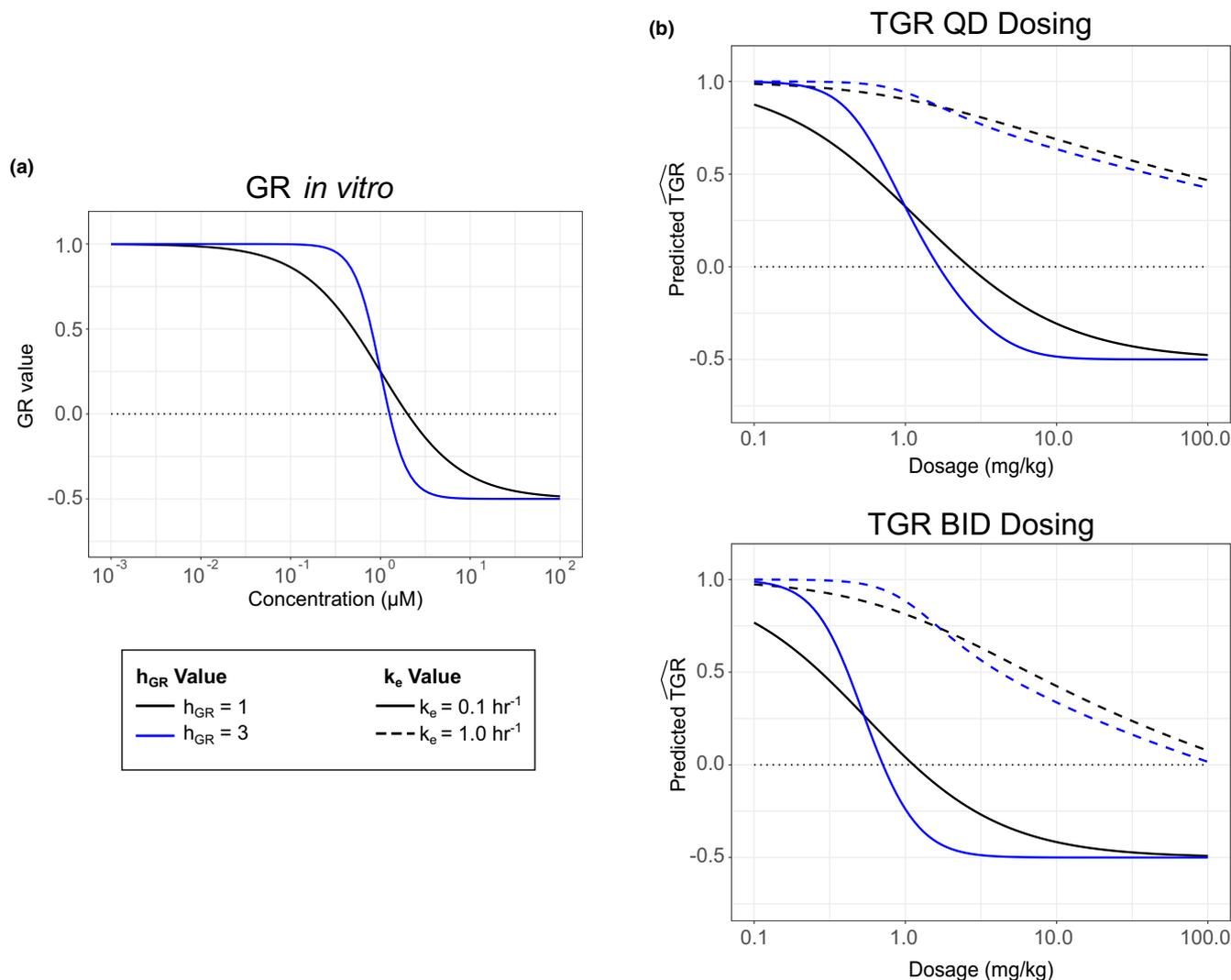
FIGURE 4 Heuristics in optimizing drug properties. (a) Example of in vitro dose–response curves for drugs with $GEC_{50} = 1\,\mu M$, $GR_{\infty} = -0.5$, and $h_{GR} = 1$ (black) or 3 (blue). (b) Predicted in vivo dose–response curves for drugs in (a) for differing values of k_e ($k_e = 0.1\,h^{-1}$, solid; $k_e = 1\,h^{-1}$, dotted), as well as dosing frequency (top, q.d.; bottom, b.i.d.). The example drugs follow a one-compartment PK model with $k_a = 2\,h^{-1}$, MW = 450 g/mol, and VF = 1 L/kg. GR, growth rate; TGR, tumor growth rate; $\widehat{\text{TGR}}$, predicted tumor growth rate value.

drug parameters and tumor response. Considering two drugs with different Hill coefficients but similar in vitro potency and maximal efficacy (Figure 4a), the shallow in vitro curve ($h_{GR} = 1$; black curve) was predicted to have higher efficacy at low doses ($TGEC_{50} < 1$ mg/kg). When increasing the dose above 1 mg/kg, the steep curve ($h_{GR} = 3$; in blue) quickly reached its maximal effect, whereas the drug with the shallow curve did not. For the same efficacy at 1 mg/kg, a five-fold increase in dosage almost reached maximal efficacy for the steep curve, whereas a 30-fold increase was necessary for the shallow curve (Figure 4b, top). Changing from q.d. to b.i.d. magnified this difference: a shallower curve favored b.i.d. dosing for dosages below $TGEC_{50}$ (Figure 4b, bottom). When increasing the elimination rate k_e, the sensitivity on h_{GR} value became less pronounced: if k_e increased from 0.1 h^{-1} (plain line) to 1.0 h^{-1} (dashed line), the difference in h_{GR} was irrelevant

across all doses (Figure 4b). In general, drugs with shallow response curves had a greater effect at low doses than drugs with steep response curves and benefit from increasing dosage, whereas drugs with high h_{GR} needed dosages above $TGEC_{50}$ to inhibit growth but reached maximal efficacy without substantial additional increases in dosage. Overall, such analysis can provide key heuristics to further guide drug development and study design.

Tumor response depends on *PIK3CA* mutational status for GDC-0077 more than for GDC-0032

One common use of in vitro drug response is to identify biomarkers of sensitivity,[38] but it remains unclear which metrics are the most informative or which in vitro

concentrations are meaningful, especially when comparing different drugs.[37] Our model addresses some of these limitations: based on the maximum tolerated dosage in vivo, we could compare the predicted growth inhibition of xenografts and assess if biomarkers identified from in vitro data showed significant differences.[15,16] To illustrate this, we looked at differences between hormone receptor positive (HR+) breast cancer lines bearing either wild-type (WT) or mutated *PIK3CA* in their responses to two PI3K inhibitors: GDC-0032/taselisib, a β-sparing inhibitor (i.e., inhibits the PI3Kα, γ, and δ isoforms),[29,39–43] and GDC-0077/inavolisib, which is selective for the PI3Kα isoform.[27,30]

The median IC_{50} value for GDC-0032 was 0.089 μM across four mutant *PIK3CA* cell lines, significantly lower than the median IC_{50} of 2.2 μM for seven WT *PIK3CA* cell lines ($p = 6.1e\text{-}3$, Wilcoxon's Rank-Sum test, Data S6). Similarly, GDC-0077 was found to be more potent on *PIK3CA* mutant cell lines compared to WT ones (median IC_{50} of 0.11 μM vs. 109 μM; $p = 9.5e\text{-}3$, Wilcoxon's Rank-Sum test; Data File S6). These comparisons suggested substantial differences between *PIK3CA* mutant and WT lines for both drugs; however, they provided limited insights on the magnitude of the differences in vivo.[42] Interestingly, we did not predict a significant difference in TGR values between *PIK3CA* mutant and WT lines at any dosage for GDC-0032 (Figure 5a, left). In contrast, we predicted significantly lower TGR values for the *PIK3CA* mutant lines for dosages above 0.1 mg/kg for GDC-0077 ($p < 0.05$, Wilcoxon's Rank-Sum test; Figure 5a right). In addition, we found that at 25 mg/kg, a dosage tolerated for both compounds,[27,29] our model predicted only a 20% difference in average TGI value for GDC-0032 between mutant and WT cell lines at 21 days (94.7% vs. 74.3%; $p = 0.16$, Wilcoxon's Rank-Sum test; Figure 5b, left), whereas a greater and significant difference was predicted for GDC-0077, with TGI values of 73.6% versus 24.4% ($p = 0.019$, Wilcoxon's Rank-Sum test; Figure 5b, right).

Using human PK values provides guidance for response in patients

Next, we input human PK values into our model to estimate the response of patients to GDC-0032 based on their tumor genotype. Here, we used PK values measured in clinical trials[39] with the in vitro cell line data (Data File S6) to predict single agent efficacy for patients with breast cancer. For the HER2-amplified (HER2+) subtype, our model suggested that *PIK3CA* mutant tumors would best respond to GDC-0032, with stable disease for doses around 5 mg (human predicted TGR ~ 0) and tumor volume reduction at higher doses (TGR < −0.25 above

9 mg; Figure 6, left). In HER2+/*PIK3CA* WT tumors, we predicted no tumor volume reduction at tolerated doses (TGR ~ 0.1 at 16 mg). For the HR+/HER2– subtype, we expected that patients with tumors bearing mutated *PIK3CA* would not show a reduction in tumor volume even at high doses (median TGR ~ 0.2 at 16 mg), whereas *PIK3CA* WT tumors were likely to progress (TGR ~ 0.5; Figure 6, right), showing more separation in humans than in mice (Figure 5a). In summary, our model suggested a clinical benefit could be expected in the tolerated dose range for patients with HER2+/*PIK3CA* mutated tumors, whereas a stable disease may be achieved in other patients, although less likely in patients with HR+/HER2– *PIK3CA* WT tumors. When comparing these predictions to the phase I clinical trial data for patients with breast cancer published in Juric et al.,[39] we observed that five patients had a partial response (decrease in tumor size of at least 30%); all of them had *PIK3CA* mutated tumors, among which two were HER2+. The remaining patients (7/12) had a best sum of the longest diameter (SLD) change between −30% and +20% (stable disease), with the four patients deriving the least benefit being HR+/HER2–. As predicted by our model, increasing doses above 12 mg did not show additional benefit. In a small HER2+ cohort treated with 6 mg (ClinicalTrials.gov identifier: NCT01296555; Figure S7), most patients (9/13), including all with a *PIK3CA* WT tumor, had a stable disease, whereas two patients with *PIK3CA* mutated tumors showed a partial response. In conclusion, although the response of patients with HR+/HER2– cancer to GDC-0032 was marginally better than predicted, our results for patients with HER2+ cancer were consistent with clinical outcome and correctly predicted a higher likelihood of tumor regression in patients with *PIK3CA* mutated tumors.

DISCUSSION

In this work, we developed a model based on growth rate inhibition to predict the effect of antiproliferative drugs on tumor growth in vivo. We hypothesized and showed that accurate predictions of in vivo efficacy can be obtained from combining growth rates measured in vitro and PK parameters without the need of understanding the underlying biochemical and molecular processes. If PK measures are not yet available, PK properties can be estimated through computational inference, or known metabolic differences among organisms.[3,4,6,44] Introducing a drug-specific calibration factor calculated with training data from only a few – potentially a single – in vivo efficacy studies yielded accurate predictions. Our model allowed us to predict the efficacy of an alternate dosing schedule or variable dosage, as well as investigate sensitivity of in vivo
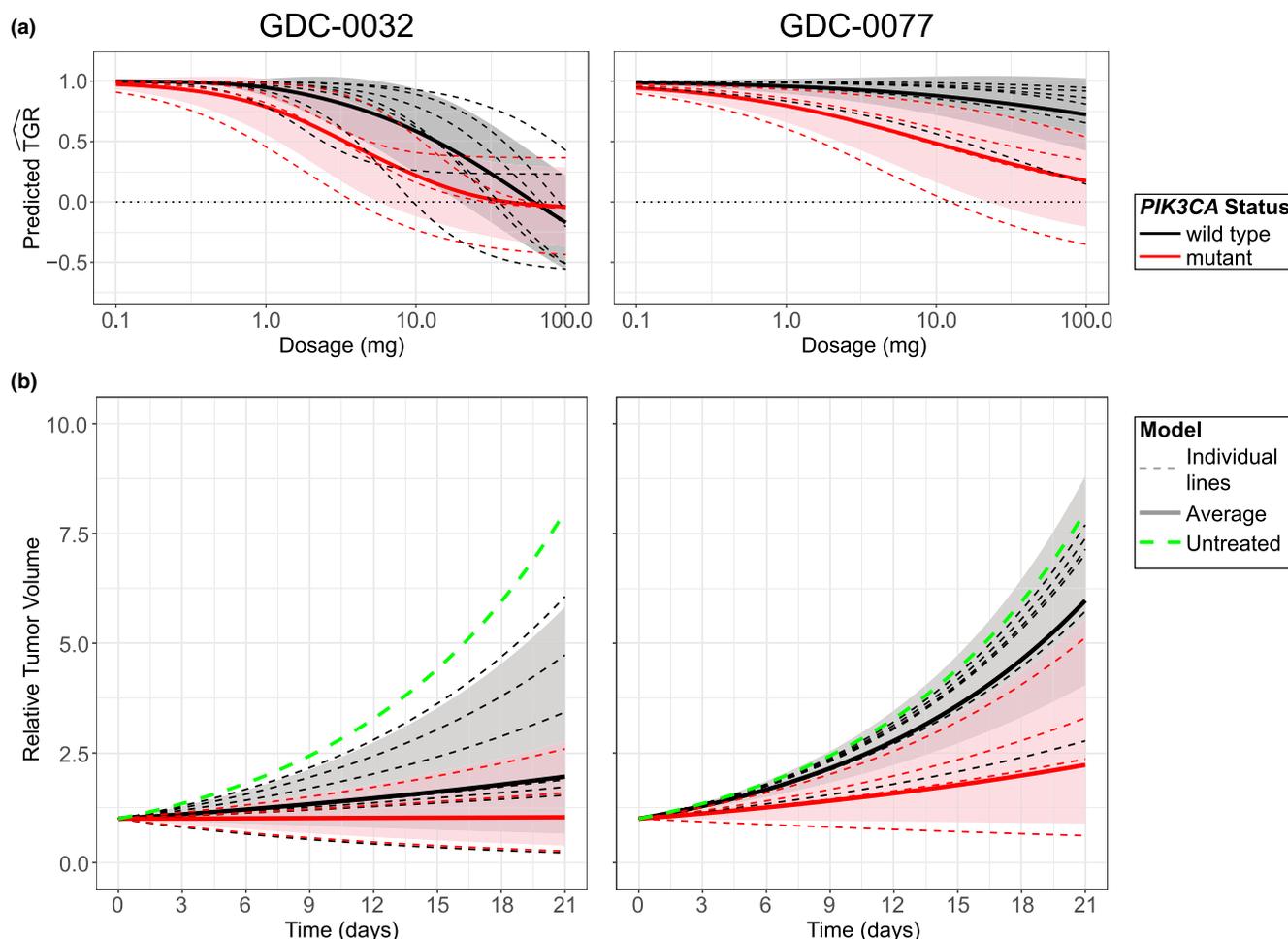
FIGURE 5 Response of HR+ breast cancer cells to two PI3K inhibitors. (a) Predicted in vivo dose–response curves of multiple breast cancer cell lines from HR+ subtypes with either wild-type (black) or mutant (red) *PIK3CA* gene to PI3K inhibitors GDC-0032 (left) and GDC-0077 (right). (b) Comparative growth plots at 25 mg/kg for each drug/cancer cell type as in (a). The untreated growth rate was taken such that the tumor size doubles every 7 days, denoted by the dotted green line. In all plots, the gray and pink areas represent one standard deviation from the black and red averages. TGR, tumor growth rate.

efficacy to the drug parameters. Finally, we showed how our model can inform the dosage at which difference in response between genotypes can be expected in xenograft models and clinical trials, providing meaningful metrics to perform biomarker analyses. An R package for predicting TGR values is available under open source license at github.com/Genentech/TGRmodel.

The correlation between our predicted drug response and the experimental data was high, including for cell lines not present in the training data, but we observed three types of discrepancies. First, our analysis of the prediction error revealed potency shifts specific to each drug, suggesting that in vivo properties from the drugs, such as tumor/plasma ratio or microenvironment effects, are not fully captured by the PK model (Figure S8). We found that these shifts could be corrected by a single drug-specific calibration factor trained on limited in vivo data, which contrasts with machine learning approaches

requiring extensive datasets. Second, we observed an overestimation of the effect of some drugs for negative TGR values. This can potentially be explained by necrotic and scar tissue replacing volume lost by cell death and limited clearance of dead cells – two phenomena not happening in vitro – resulting in an overestimation of the tumor volume when measured by calipers or Response Evaluation Criteria in Solid Tumors (RECIST) measurements. In addition, our model assumed a constant growth or shrinkage rate, which means that emergence of resistance cannot be predicted. Third, we observed an underestimation of the effect of some drugs in the high TGR value range. One way of interpreting differences between predicted and experimental values is to consider our model as the null hypothesis for intrinsic tumor growth: deviation from our model could uncover adaptive response or interactions between the drug and cell extrinsic factors such as hypoxia or the microenvironment. One such case could be the observed
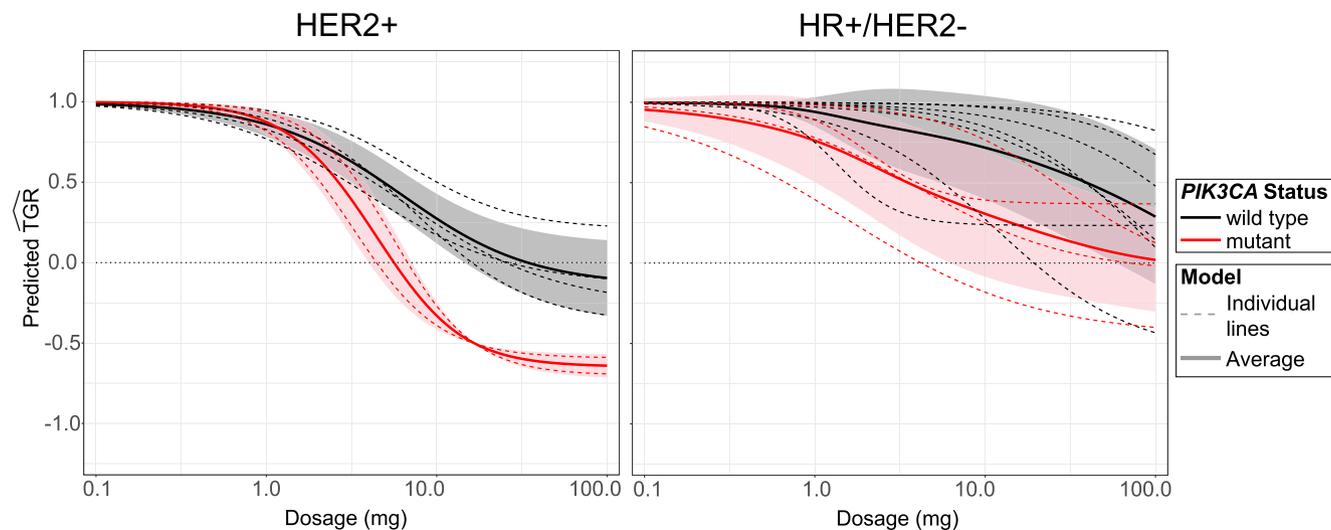
FIGURE 6 Predicted response of breast cancer cells in humans. Predicted dose–response curves of multiple breast cancer cell lines from HER2+ (left) or HR+/HER2− (right) subtypes with either wild type (black) or mutant (red) *PIK3CA* gene to PI3K inhibitor GDC-0032 in humans. TGR, tumor growth rate.

difference between 2D and 3D cultures in the response to some pan-PI3K[45] and MAPK pathway inhibitors.[33,34] Ultimately, new in vitro experiments with low serum, low-adhesion, or in 3D culture may be required to understand these discrepancies and improve predictions. Although our modeling approach was not tested for drugs that affect the cell environment and ignores effects on the immune components from the host, experiments with primary cultures and immune cell co-culture may address this limitation.[46]

A common method to estimate efficacy in vivo is the time that the free concentration is above a given in vitro viability cutoff (e.g., IC_{50}). However, we found that, beyond the intrinsic biases of in vitro relative viability metrics,[10] this method suffered from limited dynamical range, did not capture changes in efficacy between drug schedules, and was not amenable to sensitivity analyses. Other published approaches for predicting in vivo drug efficacy can be classified in two groups: complex models of drug-tumor interactions that rely on target engagement and molecular understanding of cellular growth or machine learning methods that require extensive training data. Overall, these approaches yielded predictions with correlation values up to 0.5,[2,7,12,20,21,24] which are lower than our method. In addition, our approach strikes a pragmatic balance in terms of data requirements: our growth-rate model relies on end point in vitro phenotypic data instead of direct measure of drug-target interaction or time-course experiments,[12] and does not require extensive training data to be as predictive as other approaches.[24] Therefore, our approach is more widely applicable and can potentially be extended to drug combinations[14,18,47,48] or new drugs with limited knowledge of their cell-intrinsic effects.

By predicting tumor response using mathematical equations instead of a statistical model, our approach supports drug development and translational research in four manners. First, our model can be used to optimize in vivo studies: predicted response can suggest meaningful dosages for mouse studies and inform the outcome of different dosing schedules. Second, sensitivity analysis can guide drug design by identifying parameters that will have the most impact on in vivo efficacy: situations in which improving PK properties is most important or when tumor response is limited by in vitro asymptotic efficacy. Third, in vitro data can be converted into in vivo TGR values, addressing the question of which in vitro metrics are the most informative for clinical development.[37] To illustrate this, we showed how our model predicts better differential response for GDC-0077 based on *PIK3CA* mutational status compared to GDC-0032 – something that in vitro IC_{50} values do not.[42] Finally, by directly inputting human PK parameters (either measured or predicted), our method can infer the expected response in patients and guide the choice of dosages for clinical trials. We illustrated this by assessing differences in response to GDC-0032 based on aggregated predictions from multiple cell lines of the same genotype. We found that GDC-0032 may be able to shrink HER2+/*PIK3CA* mutant tumors but should only be able to stabilize disease for HER2+/*PIK3CA* WT tumors and HR+/HER2−/*PIK3CA* mutant tumors. This result was consistent with clinical trial results and previous work.[39,41,42] Although there is not yet a direct comparison for GDC-0077 efficacy in patients with *PIK3CA* mutant tumors versus those with WT tumors (currently enrolling exclusively *PIK3CA*-mutated cancers), the improved

toxicity profile of GDC-0077/inavolisib in the clinic suggests more specificity for mutant PI3Kα than WT PI3Kα,[49] consistent with our prediction.

In conclusion, the model developed in this work can be readily used to predict baseline efficacy of candidate drugs during drug development and the early phases of clinical trials because of its reliance on common in vitro data and absence of drug-specific assumptions. Application of our method will allow for better leveraging of in vitro efficacy data, reducing the number of animals necessary for adequate quantification of drug efficacy, and strengthening drug efficacy predictions for the clinic – all of which can save substantial time and resources in translational research.

AUTHOR CONTRIBUTIONS

R.D., L.S., and M.H. wrote the manuscript. R.D., L.S., and M.H. designed the research. R.D. and M.H. performed the research. R.D., L.S., B.A., T.R.W., T.J.S., and M.H. analyzed the data. B.A. contributed analytical tools.

ACKNOWLEDGEMENTS

CONFLICT OF INTEREST

L.S., B.A., T.R.W., T.J.S., and M.H. are employees of Genentech and Roche shareholders. R.D. declared no competing interests for this work.

ORCID

Rocky Diegmiller https://orcid.org/0000-0002-6115-7891
Marc Hafner https://orcid.org/0000-0003-1337-7598

REFERENCES

1. Au JL-S, Jang SH, Zheng J, et al. Determinants of drug delivery and transport to solid tumors. *J Control Release*. 2001;74:31-46.
2. Danhof M, Delange E, Dellapasqua O, Ploeger B, Voskuyl R. Mechanism-based pharmacokinetic-pharmacodynamic (PK-PD) modeling in translational drug research. *Trends Pharmacol Sci*. 2008;29:186-191.
3. Nigade PB, Gundu J, Pai KS, Nemmani KVS. Prediction of tumor-to-plasma ratios of basic compounds in subcutaneous xenograft mouse models. *Eur J Drug Metab Pharmacokinet*. 2018;43:331-346.
4. Saleem A, Price PM. Early tumor drug pharmacokinetics is influenced by tumor perfusion but not plasma drug exposure. *Clin Cancer Res*. 2008;14:8184-8190.
5. Schuhmacher J. Determination of the free fraction and relative free fraction of drugs strongly bound to plasma proteins. *J Pharm Sci*. 2000;89:14.
6. Mager DE, Woo S, Jusko WJ. Scaling pharmacodynamics from in vitro and preclinical animal studies to humans. *Drug Metab Pharmacokinet*. 2009;24:16-24.
7. Wang D, Hensman J, Kutkaite G, et al. A statistical framework for assessing pharmacological responses and biomarkers using uncertainty estimates. *eLife*. 2020;9:e60352.
8. Gubler H, Schopfer U, Jacoby E. Theoretical and experimental relationships between percent inhibition and IC_{50} data observed in high-throughput screening. *J Biomol Screen*. 2013;18:1-13.
9. Hafner M, Niepel M, Sorger PK. Alternative drug sensitivity metrics improve preclinical cancer pharmacogenomics. *Nat Biotechnol*. 2017;35:500-502.
10. Hafner M, Niepel M, Chung M, Sorger PK. Growth rate inhibition metrics correct for confounders in measuring sensitivity to cancer drugs. *Nat Methods*. 2016;13:521-527.
11. Benzekry S, Lamont C, Beheshti A, et al. Classical mathematical models for description and prediction of experimental tumor growth. *PLoS Comput Biol*. 2014;10:e1003800.
12. Bouhaddou M, Yu LJ, Lunardi S, et al. Predicting in vivo efficacy from *in vitro* data: quantitative systems pharmacology modeling for an epigenetic modifier drug in cancer. *Clin Transl Sci*. 2020;13:419-429.
13. Yin A, Moes DJAR, Hasselt JGC, Swen JJ, Guchelaar H. A review of mathematical models for tumor dynamics and treatment resistance evolution of solid tumors. *CPT Pharmacometrics Syst Pharmacol*. 2019;8:720-737.
14. Kirouac DC, Schaefer G, Chan J, et al. Clinical responses to ERK inhibition in BRAF V600E-mutant colorectal cancer predicted using a computational model. *NPJ Syst Biol Appl*. 2017;3:14.
15. Mould D, Walz A, Lave T, Gibbs J, Frame B. Developing exposure/response models for anticancer drug treatment: special considerations. *CPT Pharmacometrics Syst Pharmacol*. 2015;4:12-27.
16. Zhu AZ. Quantitative translational modeling to facilitate preclinical to clinical efficacy & toxicity translation in oncology. *Future Sci OA*. 2018;4:FSO306.
17. Mager DE, Wyska E, Jusko WJ. Diversity of mechanism-based pharmacodynamic models. *Drug Metab Dispos*. 2003;31:510-518.
18. Irurzun-Arana I, McDonald TO, Trocóniz IF, Michor F. Pharmacokinetic profiles determine optimal combination treatment schedules in computational models of drug resistance. *Cancer Res*. 2020;80:3372-3382.
19. Chakrabarti S, Michor F. Pharmacokinetics and drug interactions determine optimum combination strategies in computational models of cancer evolution. *Cancer Res*. 2017;77:3908-3921.
20. Adam G, Rampášek L, Safikhani Z, Smirnov P, Haibe-Kains B, Goldenberg A. Machine learning approaches to drug response prediction: challenges and recent progress. *NPJ Precis Oncol*. 2020;4:19.
21. Azuaje F. Computational models for predicting drug responses in cancer research. *Brief Bioinform*. 2016;18:820-829. doi:10.1093/bib/bbw065
22. Voskoglou-Nomikos T, Pater JL, Seymour L. Clinical predictive value of the in vitro cell line, human xenograft, and mouse allograft preclinical cancer models. *Clin Cancer Res*. 2003;9:4227-4239.

23. Johnson JI, Decker S, Zaharevitz D, et al. Relationships between drug activity in NCI preclinical in vitro and in vivo models and early clinical trials. *Br J Cancer*. 2001;84:1424-1431.

24. Ma J, Fong SH, Luo Y, et al. Few-shot learning creates predictive models of drug response that translate from high-throughput screens to individual patients. *Nat Can*. 2021;2:233-244.

25. Forrest WF, Alicke B, Mayba O, et al. Generalized additive mixed modeling of longitudinal tumor growth reduces bias and improves decision making in translational oncology. *Cancer Res*. 2020;80:5089-5097.

26. Lin J, Sampath D, Nannini MA, et al. Targeting activated Akt with GDC-0068, a novel selective Akt inhibitor that is efficacious in multiple tumor models. *Clin Cancer Res*. 2013;19:1760-1772.

27. Song KW, Edgar KA, Hanan EJ, et al. RTK-dependent inducible degradation of mutant PI3Kα drives GDC-0077 (Inavolisib) efficacy. *Cancer Discov*. 2022;12:204-219.

28. Wong H, Choo EF, Alicke B, et al. Antitumor activity of targeted and cytotoxic agents in murine subcutaneous tumor models correlates with clinical response. *Clin Cancer Res*. 2012;18:3846-3855.

29. Ndubaku CO, Heffron TP, Staben ST, et al. Discovery of 2-{3-[2-(1-Isopropyl-3-methyl-1H-1,2–4-triazol-5-yl)-5,6-dihydrobenzo[f]imidazo[1,2-d][1,4]oxazepin-9-yl]-1H-pyrazol-1-yl}-2-methylpropanamide (GDC-0032): a β-sparing phosphoinositide 3-kinase inhibitor with high unbound exposure and robust in vivo antitumor activity. *J Med Chem*. 2013;56:4597-4610.

30. Han C, Kelly SM, Cravillion T, Savage SJ, Nguyen T, Gosselin F. Synthesis of PI3K inhibitor GDC-0077 via a stereocontrolled N-arylation of α-amino acids. *Tetrahedron*. 2019;75:4351-4357.

31. Salphati L, Wong H, Belvin M, et al. Pharmacokinetic-pharmacodynamic modeling of tumor growth inhibition and biomarker modulation by the novel phosphatidylinositol 3-kinase inhibitor GDC-0941. *Drug Metab Dispos*. 2010;38:1436-1442.

32. Fritsch C, Huang A, Chatenay-Rivauday C, et al. Characterization of the novel and specific PI3Kα inhibitor NVP-BYL719 and development of the patient stratification strategy for clinical trials. *Mol Cancer Ther*. 2014;13:1117-1129.

33. Canon J, Rex K, Saiki AY, et al. The clinical KRAS(G12C) inhibitor AMG 510 drives anti-tumour immunity. *Nature*. 2019;575:217-223.

34. Ekert JE, Johnson K, Strake B, et al. Three-dimensional lung tumor microenvironment modulates therapeutic compound responsiveness in vitro – implication for drug development. *PLoS One*. 2014;9:e92248.

35. Laborde L, Oz F, Ristov M, Guthy D, Sterker D, McSheehy P. Continuous low plasma concentrations of everolimus provides equivalent efficacy to oral daily dosing in mouse xenograft models of human cancer. *Cancer Chemother Pharmacol*. 2017;80:869-878.

36. McInnes IB, Byers NL, Higgs RE, et al. Comparison of baricitinib, upadacitinib, and tofacitinib mediated regulation of cytokine signaling in human leukocyte subpopulations. *Arthritis Res Ther*. 2019;21:183.

37. Fallahi-Sichani M, Honarnejad S, Heiser LM, Gray JW, Sorger PK. Metrics other than potency reveal systematic variation in responses to cancer drugs. *Nat Chem Biol*. 2013;9:708-714.

38. Seashore-Ludlow B, Rees MG, Cheah JH, et al. Harnessing connectivity in a large-scale small-molecule sensitivity dataset. *Cancer Discov*. 2015;5:1210-1223.

39. Juric D, Krop I, Ramanathan RK, et al. Phase I dose-escalation study of Taselisib, an oral PI3K inhibitor, in patients with advanced solid tumors. *Cancer Discov*. 2017;7:704-715.

40. Zumsteg ZS, Morse N, Krigsfeld G, et al. Taselisib (GDC-0032), a potent β-sparing small molecule inhibitor of PI3K, Radiosensitizes head and neck squamous carcinomas containing activating *PIK3CA* alterations. *Clin Cancer Res*. 2016;22:2009-2019.

41. Dickler MN, Saura C, Richards DA, et al. Phase II study of Taselisib (GDC-0032) in combination with Fulvestrant in patients with HER2-negative, hormone receptor–positive advanced breast cancer. *Clin Cancer Res*. 2018;24:4380-4387.

42. Moore HM, Savage HM, O'Brien C, et al. Predictive and pharmacodynamic biomarkers of response to the phosphatidylinositol 3-kinase inhibitor Taselisib in breast cancer preclinical models. *Mol Cancer Ther*. 2020;19:292-303.

43. Dent S, Cortés J, Im YH, et al. Phase III randomized study of taselisib or placebo with fulvestrant in estrogen receptor-positive, PIK3CA-mutant, HER2-negative, advanced breast cancer: the SANDPIPER trial☆. *Ann Oncol*. 2021;32:197-207.

44. Hao K, Qi Q, Wan P, et al. Prediction of human pharmacokinetics from preclinical information of Rhein, an antidiabetic nephropathy drug, using a physiologically based pharmacokinetic model. *Basic Clin Pharmacol Toxicol*. 2014;114:160-167.

45. Muranen T, Selfors LM, Worster DT, et al. Inhibition of PI3K/mTOR leads to adaptive resistance in matrix-attached cancer cells. *Cancer Cell*. 2012;21:227-239.

46. Poggi A, Villa F, Fernadez JLC, Costa D, Zocchi MR, Benelli R. Three-dimensional culture models to study innate anti-tumor immune response: advantages and disadvantages. *Cancer*. 2021;13:3417.

47. Shen J, Li L, Yang T, Cohen PS, Sun G. Biphasic mathematical model of cell–drug interaction that separates target-specific and off-target inhibition and suggests potent targeted drug combinations for multi-driver colorectal cancer cells. *Cancer*. 2020;12:436.

48. Zimmer A, Katzir I, Dekel E, Mayo AE, Alon U. Prediction of multidimensional drug dose responses based on measurements of drug pairs. *Proc Natl Acad Sci*. 2016;113:10442-10447.

49. Juric D, Oliveira M, Kalinsky K, et al. Abstract OT1-08-04: a first-in-human phase Ia dose escalation study of GDC-0077, a p110a-selective and mutant-degrading PI3K inhibitor, in patients with PIK3CA-mutant solid tumors. *Cancer Res*. 2020;80(suppl 4):OT1-08-04.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.