Anti-tumor activity of targeted and cytotoxic agents in murine subcutaneous
tumor models correlates with clinical response

Harvey Wong¹, Edna F. Choo¹, Bruno Alicke², Xiao Ding¹, Hank La¹, Erin McNamara²,
Frank-Peter Theil³*, Jay Tibbitts³, Lori S. Friedman², Cornélis E.C.A. Hop¹, and Stephen
E. Gould².

Running Title: Clinical Translation of Murine Tumor Model Efficacy by PK-PD Analysis

Keywords: PK-PD Modeling, Xenografts, Allografts, Translational Biology

Author Affiliations: Departments of Drug Metabolism and Pharmacokinetics¹,
Translational Oncology², and Pharmacokinetic and Pharmacodynamic Sciences³,
Genentech, Inc., South San Francisco, CA, USA

Potential Conflict of Interest: All authors are employees of Genentech, Inc.

Address for correspondence:

Harvey Wong, PhD,
Drug Metabolism and Pharmacokinetics, Genentech,
1 DNA Way, MS 412a, South San Francisco, CA 94080
Phone: 650-225-5739, Fax: 650-467-3487
E-mail: wong.harvey@gene.com

AND

Stephen E. Gould, PhD.
Translational Oncology, Genentech,
1 DNA Way, MS 50, South San Francisco, CA 94080
Phone: 650-467-8025
E-mail: gould.stephen@gene.com

*Current affiliation, UCB Pharma
STATEMENT OF TRANSLATIONAL RELEVANCE

The current work investigates whether or not preclinical murine subcutaneous tumor models are predictive of clinical response. We have performed a retrospective pharmacokinetic-pharmacodynamic (PK-PD) analysis of xenograft/allograft efficacy data for 8 agents (both molecularly targeted and cytotoxic) with known clinical response data. Simulations of tumor growth inhibition (TGI) in xenografts/allografts using human exposures at clinically relevant doses and schedules were performed in order to account for species differences in pharmacokinetics. Strikingly, a significant correlation was observed between simulated TGI driven by human pharmacokinetics and clinical response but not when using TGI observed at maximum tolerated doses in mice. Furthermore, agents that lead to greater than 60% TGI in preclinical models, at clinically relevant exposures, appeared more likely to lead to clinical response. These data provide a framework for the successful use of murine subcutaneous tumor models for selecting active agents to advance into the clinic.
ABSTRACT

Purpose: Immunodeficient mice transplanted with subcutaneous tumors (xenograft or allograft) are widely used as a model of preclinical activity for the discovery and development of anti-cancer drug candidates. Despite their widespread use, there is a widely held view that these models provide minimal predictive value for discerning clinically active versus inactive agents. In order to improve the predictive nature of these models, we have performed a retrospective population pharmacokinetic-pharmacodynamic (PK-PD) analysis of relevant xenograft/allograft efficacy data for 8 agents (molecularly targeted and cytotoxic) with known clinical outcome. Experimental Design: PK-PD modeling was performed to first characterize the relationship between drug concentration and anti-tumor activity for each agent in dose ranging xenograft or allograft experiments. Next, simulations of tumor growth inhibition (TGI) in xenografts/allografts at clinically relevant doses and schedules were performed by replacing the murine PK which was used to build the PK-PD model with human PK obtained from the literature in order to account for species differences in pharmacokinetics. Results: A significant correlation ($r=0.91$, $p=0.0008$) was observed between simulated xenograft/allograft TGI driven by human pharmacokinetics and clinical response but not when TGI observed at maximum tolerated doses in mice was correlated with clinical response ($r=0.36$, $p=0.34$). Conclusions: Based on these analyses, agents that led to greater than 60% tumor growth inhibition in preclinical models, at clinically relevant exposures, are more likely to lead to responses in the clinic. A proposed strategy for the use of murine subcutaneous models for compound selection in anti-cancer drug discovery is discussed.
INTRODUCTION

The discovery and development of new drugs is a challenging process and finding new anti-cancer agents is particularly difficult (1). The most prominent reason for failure is lack of efficacy highlighting the need for improved “preclinical tools” to aid in the identification of active anti-cancer drug candidates. Immune-deficient mice transplanted with subcutaneous tumors (xenografts or allografts) have served as a major or “workhorse” model for the selection of anti-cancer drug candidates for more than 30 years. The primary advantages of this model system are the ease at which anti-tumor activity can be monitored, the relative diversity of models available, and that the tumors are of human origin in the case of xenografts. Despite their widespread use, there has been much debate on their predictive value (2-6). The literature debating the clinical relevance of these xenograft models is extensive and has often concluded that these models inadequately predict clinical activity. Much of the literature has focused on the predictive value of xenograft studies for cytotoxic chemotherapeutic agents and has applied somewhat arbitrary cut-offs to assess preclinical response (i.e., T/C = 40%). One of the more extensive examinations of preclinical xenograft models predicting clinical response was performed by the National Cancer Institute (NCI) (6) and suggested that xenograft studies are moderately predictive of clinical response for agents that showed activity in at least one third of the xenograft models tested. In their analysis of 39 cytotoxic agents, anti-tumor activity in non-small cell lung cancer (NSCLC) xenograft mice was the most predictive of clinical response in the same disease, and overall, breast cancer xenografts were the most predictive of a clinical response in any disease. Importantly, doses of anti-cancer agents used in this analysis were at, or just below, the
maximum tolerated dose (MTD) in mice (6). Similar findings were reported by the National Cancer Institute of Canada Clinical Trials Group where xenograft models were not predictive of clinical response for 31 cytotoxic agents in colon and breast cancer but were predictive in NSCLC and ovarian cancer (7). As with the NCI study no attempt was made to correct for drug exposure between the preclinical and clinical studies. Importantly, available information on newer molecularly targeted agents is limited.

Preclinical pharmacokinetic-pharmacodynamic (PK-PD) modeling can play an important role in the drug discovery and development process by providing an integrated assessment of the relationships between drug concentrations in the body and efficacy (8, 9). This is accomplished through the use of differential equations to quantitatively describe drug behavior in the biological system of interest. PK-PD modeling is particularly helpful in understanding biological systems where there are apparent disconnects in drug concentration and effect related to temporal delays in drug action. In addition, mechanism-based PK-PD models allow for known species differences in factors such as pharmacokinetics or in vivo drug potency to be quantitatively assessed as well as enabling a more robust characterization of the concentration/anti-tumor response relationship from inherently variable preclinical efficacy data. Finally, establishment of PK-PD models enables the ability to make prospective simulations of untested dose and schedules. The primary objective of the current study was to evaluate whether xenograft/allograft data was predictive of clinical response following correction of differences between preclinical and clinical drug exposures through PK-PD modeling (depicted in Figure 1). A second objective was to assess the predictive value of
xenografts/allografts for molecular targeted anti-cancer agents. Our observations are consistent with the notion that xenograft efficacy data are predictive of clinical response when appropriately corrected for therapeutic exposure (2, 4, 5).
MATERIALS AND METHODS

Xenograft/allograft efficacy studies

Three million (KPL-4), five million (Colo205, K-562, Cal51x1.1s, Calu-6) or ten million (Fibroblast NR6-Del(746-752), A549, H460, H520) cells were resuspended in Hank’s Balanced Salt Solution or in Hank’s Balanced Salt Solution mixed 1:1, v/v with Matrigel (BD Biosciences, San Jose, CA), and implanted subcutaneously into the right flank or right thoracic region of female naïve SCID mice (Taconic, Oxnard, CA, USA) or NCR (Taconic, Oxnard, CA, USA), athymic (Harlan Laboratories Inc., Indianapolis, IN) or CD-1 nude mice (Charles River Laboratories, Hollister, CA, USA). For MX-1 xenografts, D5123 xenografts and medulloblastoma allografts, tumor fragments (1 to 3 mm³) were implanted subcutaneously into the right flank of naïve female SCID or nude mice. Tumor fragments (1 to 3 mm³) were implanted subcutaneously into the right lateral thorax of naïve beige nude XID mice (Harlan Laboratories Inc., Indianapolis, IN) for 786-O xenografts. Tumor sizes and body weights were recorded twice weekly, and the mice were regularly observed over the course of the study. Tumors were measured in two dimensions (length and width) and tumor volume (TV) was calculated using the following formula:

\[ TV (\text{mm}^3) = (\text{length} \times \text{width}^2) \times 0.5 \]

Mice were randomly assigned to groups of five to ten individuals bearing similarly-sized tumors. The mean tumor volume at the start of dosing for all studies ranged from 127 to 340 mm³. Mice were promptly euthanized if their tumor volume exceeded 2000 mm³ or
if their body weight dropped by more than 20% of the starting weight. Mice were housed and maintained according to the animal use guidelines of Genentech, Inc, conforming to California State legal and ethical practices.

MX-1 (triple negative) breast cancer cells were purchased from the NCI-Frederick Cancer DCT Tumor Repository and maintained as a passaged tumor fragment (Frederick, MD, USA). KPL-4 breast cancer cells were obtained from J. Kurebayashi (10) and implanted into FcγRI−/−FcγRIII−/−RAG2−/−Tg (human FcγRIIIa) mice (11). The NR6-del (746-752) fibroblast cell line (12), was originally obtained from V. Chazin (13) and engineered to express the EGFR del(746-752) mutant at Genentech (South San Francisco, CA, USA) (12). All other cell lines were obtained from American Type Culture Collection (ATCC; Manassas, VA, USA). Cell lines were maintained in high-glucose DMEM:Ham’s F-12 (50:50) supplemented with 10% FBS and 2 mM L-glutamine. Cal51x1.1s is an in vivo-passaged subclone of Cal51 (ATTC, Manassas, VA, USA). D5123 colon adenocarcinoma tissue was acquired from National Disease Interchange (Philadelphia, PA, USA). Medulloblastoma tumor fragments used to establish medulloblastoma allografts were generated at Genentech as described previously (14).

Vismodegib (GDC-0449), erlotinib (Tarceva®) and trastuzumab (Herceptin®) were provided by Genentech (South San Francisco, CA, USA). Clinical grade sunitinib (Sutent®; Pfizer, New York, NY, USA), dasatinib (Sprycel®; Bristol-Myers Squibb, Princeton, NJ, USA), 5-Fluorouracil (5-FU) (Adrucil®; Sicor Pharmaceuticals Inc., Irvine, CA, USA), carboplatin (Paraplatin®; Bristol-Myers Squibb, Princeton, NJ, USA) and docetaxel (Taxotere®; Sanofi-Aventis, Bridgewater, NJ, USA) were used in the
studies. All other reagents or material used in this study were purchased from Sigma-Aldrich (St Louis, MO) unless otherwise stated. The dosing regimens used in efficacy studies are listed in Table S1. The highest dose used in the listed regimens above were at or near the maximum tolerated dose (MTD) in mice for each anti-cancer agent with the exception of trastuzumab for which MTD was not defined.

**Pharmacokinetic (PK) studies in mice**

Pharmacokinetic studies were performed in female nude mice. The mouse strain, route of administration, and formulation used was selected to match the efficacy studies. Doses of each anti-cancer agent were chosen in order to adequately characterize the range over which the xenograft/ allograft efficacy studies were performed. Blood samples (~1 mL) were collected as described (14) and analyzed by LC-MS-MS. Serum rather than plasma was collected for trastuzumab and analyzed by ELISA assay.

Pharmacokinetic parameters for each anti-cancer agent in mouse were estimated using SAAM II (Saam Institute, University of Washington, Seattle, WA). Estimated parameters were used to simulate systemic concentrations of anti-cancer agents for the various regimens in the xenograft studies described above during the PK-PD modeling process.

**Pharmacokinetic-pharmacodynamic (PK-PD) modeling**

Tumor volumes from xenograft efficacy studies were fitted to an indirect response model as described by the following differential equation and Figure 1A.
\[
\frac{d(TV)}{dt} = k_{ng} \, (TV) - K(TV)
\]

(Equation 1)

\(TV\) (mm\(^3\)) is defined as the tumor volume, \(t\) (day) is time, \(k_{ng}\) (day\(^{-1}\)) is the net growth rate constant and \(K\) (day\(^{-1}\)) is the rate constant describing the anti-tumor effect of the anti-cancer agent used in the xenograft study.

For studies where the anti-tumor effect of the anti-cancer agent of interest was saturable, \(K\) was defined by Equation 2.

\[
K = \frac{K_{max} \times C^n}{KC_{50} \, n \times C^n}
\]

(Equation 2)

\(K_{max}\) (day\(^{-1}\)) is defined as the maximum value of \(K\), \(C\) is the plasma concentration of the anti-cancer agent, \(KC_{50}\) is the plasma concentration of the anti-cancer agent where \(K\) is 50% of \(K_{max}\) and \(n\) is the Hill coefficient. In all studies in this manuscript, \(n\) was fixed to 1.

For studies where the anti-tumor effect was linear, \(K\) was defined by Equation 3:

\[
K = E \times C^n
\]

(Equation 3)

\(E\) (µM\(^{-1}\) day\(^{-1}\)) is defined as a constant that relates the anti-cancer agent plasma concentration with \(K\).
For some anti-cancer agents, a delay in anti-tumor effect was required to characterize the time course of tumor volume changes from xenograft efficacy studies. In these cases, a modified version of a PK-PD model employing a series of transit compartments representing irreversibly damaged cells similar to that described by Simeoni and coworkers was used (15) (See Figure 1B). The following series of differential equations (Equations 4-6) along with Equation 1 was used to describe the described tumor transit PK-PD model.

\[
\frac{d(TV_1)}{dt} = K(TV) - k_{tran}(TV_1) \quad \text{(Equation 4)}
\]

\[
\frac{d(TV_2)}{dt} = k_{tran}(TV_1) - k_{tran}(TV_2) \quad \text{(Equation 5)}
\]

\[
\frac{d(TV_3)}{dt} = k_{tran}(TV_2) - k_{tran}(TV_3) \quad \text{(Equation 6)}
\]

$TV_1$, $TV_2$ and $TV_3$ are transit compartments representing tumor cells irreversibly damaged by treatment with an anti-cancer agent and $k_{tran}$ is the rate constant describing movement of cells through the transit compartments. In this model, tumor volume is represented as the sum of $TV$, $TV_1$, $TV_2$ and $TV_3$. As with the previous model, $K$ for a saturable anti-tumor effect was described using Equation 2 and $K$ for a linear anti-tumor effect was described using Equation 3. For all xenograft studies, model selection was based upon both the growth characteristics of the tumor in treatment groups and the Akaike's information criterion.
The S-ADAPT II (v.1.56) program, an augmented version of ADAPT II with population analysis capabilities (16, 17) was used to fit individual tumor volumes from all dose levels of a particular anti-cancer agent simultaneously to the described PK-PD models. Concentrations of anti-cancer agents were simulated based on their estimated pharmacokinetic parameters in mouse and the described dosing regimens. Inter-subject variability was assumed to be log-normally distributed and fitted using an exponential variance model. Residual variability error was modeled using a proportional error model. In cases where inter-subject variance was small and could not be estimated reliably, the inter-subject variance was fixed to 0.00001. Population parameter estimates are presented as the estimate followed by the %SE in parentheses. All estimates of population PD parameters were estimated with good precision with CV’s <25%. All estimates of intra and inter-subject variance had CV’s <45%.

Simulations of clinical regimens were performed as described in Figure 1C using the population PK-PD model and pharmacodynamic parameter estimates from xenograft studies in combination with human pharmacokinetic parameters obtained from the literature (See Table S2). Due to the atypical kinetics of vismodegib (GDC-0449) in humans, the human pharmacokinetics of vismodegib was simulated by driving mean steady-state concentrations to 16.1 µM using the reported median time to steady-state of 14 days (18).

Data analysis
The percent tumor growth inhibition (%TGI) at a specific time point was calculated as described in the following equation.

\[
\%TGI = \frac{TV_{\text{vehicle}} - TV_{\text{treatment}}}{TV_{\text{vehicle}} - TV_{\text{initial}}} \times 100
\]  
(Equation 7)

\(TV_{\text{vehicle}}\) is the tumor volume for the vehicle treated animals at a specified endpoint time, \(TV_{\text{initial}}\) is the initial tumor volume at the start of the treatment and \(TV_{\text{treatment}}\) is the tumor volume of the treatment groups at a specified endpoint time. Since %TGI is a time dependent efficacy endpoint, it is calculated at a fixed time point. For the 50 mg/kg oral erlotinib studies in non-small cell lung cancer xenografts, %TGI was calculated at the end of the dosing period for each study.

Pearson correlation coefficients (r) and their associated p-values (two-tailed), were determined using GraphPad Prism version 4.02 for Windows (GraphPad Software, San Diego California USA).
RESULTS

Characterization of preclinical dose – efficacy relationships

In order to characterize the relationship between drug and anti-tumor activity we performed dose-ranging efficacy studies using either xenografts derived from cultured cell lines, passaged patient-derived xenografts, or murine allografts for 8 agents. Where possible models were selected based on prior knowledge of sensitivity with the aim of selecting those exhibiting a response that was representative versus being hyper-sensitive or inherently resistant. Dose ranges were chosen to provide the full range of efficacy with the upper end being the maximum tolerated dose, where achievable. In almost all cases anti-tumor activity was dose-dependent with the highest dose tested leading to near tumor stasis or regression during the dosing period (Figure S1 or Figure 2A). Administration of vismodegib to patient derived D5123 colorectal xenografts, however, led to tumor growth delay as a maximal response. Administration of erlotinib to NR6-Del(746-752) (containing erlotinib sensitive EGFR Del(746-752) mutation) allografts resulted in stasis at oral daily doses of 40 and 50 mg/kg (Figure 2A). Little to no anti-tumor activity was observed in studies where oral daily doses of 50 mg/kg of erlotinib were dosed in mice bearing the wild-type EGFR lines A549, Calu-6, H460 and H520, (Figure S2).

Effects of anti-cancer agents are adequately described by PK-PD models

The impact of drug on tumor growth was determined through a population PK-PD modeling approach. Two separate mechanistic PK-PD models (Figure 1A and B)
characterizing drug effect (see Materials and Methods for details) were used to fit tumor volume data from preclinical studies. The indirect response PK-PD model (8) described in Figure 1A assumes that drug acts to decrease tumor volume by increasing the elimination rate constant, $K$. The second model depicted in Figure 1B is similar to the first, but is used for drugs with a more prolonged delay in drug effect. The delay in effect is captured through the addition of a series of transit compartments representing irreversibly damaged tumor cells (8, 15). The choice of model used for each anti-cancer agent was based on model fitting statistics (see Materials and Methods) and the growth characteristics of individual xenograft tumors from treated animals. In general, tumor growth was variable based upon the estimates of inter-individual variance of $k_{ng}$ and was the main source of variability in the xenograft studies. Among the xenograft/allograft models tested there was an approximate 4-fold range in tumor growth with KPL-4 ($k_{ng} = 0.0463 \text{ day}^{-1}$) being the slowest and NR6-Del(746-752) ($k_{ng} = 0.200 \text{ day}^{-1}$) being the fastest growing model (Table S3). In general, there was good concordance between the individual predicted tumor volumes obtained from the population PK-PD fits and observed tumor volumes indicating that the PK-PD models used adequately characterize the effect of drug on tumor volume (Figure 2C).

Direct comparisons of pharmacodynamic parameters reflective of anti-tumor activity could not be made across all anti-cancer agents tested since different PK-PD models were used to characterize each agent. However, comparisons of the relative anti-tumor activity for particular agents in different xenografts/allografts could be made for sunitinib, vismodegib, and erlotinib. The maximum anti-tumor effect ($K_{max}$) of sunitinib in the
786-O renal cell carcinoma and the Colo205 colorectal cancer xenografts were similar however sunitinib was approximately 2-fold more potent in the 786-O xenografts (>2-fold lower $K_{C50}$ [concentration of the anti-cancer agent where $K$ is 50% of $K_{max}$]). Visomodegib’s $K_{max}$ was approximately 18-fold higher and $K_{C50}$ was 2-fold lower in medulloblastoma allografts that exhibit constitutive activation of the Hedgehog (Hh) pathway, compared to D5123 xenografts characterized by Hh ligand activation (14). In the case of erlotinib, we characterized the anti-tumor effect in EGFR mutant (Del(746-752)) murine allografts. As mentioned above, little or no anti-tumor activity was observed in wild-type EGFR xenografts (Figure S2).

**Anti-tumor activity in xenografts correlates with clinical response**

In order to relate preclinical response to clinical response data, we replaced mouse PK parameters in the PK-PD models connecting preclinical efficacy and drug effect with human PK parameters derived from clinical trials and performed simulations matching both dose and regimen as depicted in Figure 1C (see Table S2 for references for human PK). A representative simulation with human PK is shown for erlotinib in Figure 2B. The simulation using the exposure of erlotinib observed in patients results in 102% TGI which was similar to the 99% TGI observed at MTD in mice (Table 1). The resulting simulated percent tumor growth inhibition (%TGI) using this hybrid PK-PD model was determined for each agent assuming a 21-day study (Table 1). The study duration was chosen since it is a common duration for xenograft studies and also corresponds to a common duration for 1 cycle of chemotherapy. Among the anti-cancer agents, simulations with vismodegib (in medulloblastoma allografts), erlotinib and dasatinib
resulted in the highest anti-tumor activity (Table 1). Of note, simulations with erlotinib and docetaxel were the only 2 out of 10 simulations where the %TGI was similar when the activity was driven by human PK versus that obtained with mouse exposures at, or near, the MTD. For all other simulations the %TGI was lower when driven by human PK.

With the exception of clinical activity for vismodegib in tumors exhibiting ligand dependent Hedgehog pathway activation, overall response (complete + partial response) was obtained from single agent trials in the literature (see Table S2 for references). We found that when efficacy observed at the maximum tolerated dose (MTD) in our preclinical studies was compared to overall response, there was no evidence of a correlation between anti-tumor activity in mice and man (r=0.36, p=0.34; Figure 3A). However, when simulated human %TGI derived from PK-PD modeling of our preclinical studies was compared to overall response rates from clinical trials there was a significant correlation (Figure 3B) (r=0.91, p=0.0008). Agents that led to greater simulated human %TGI in preclinical tumors led to greater response rates in the clinic. These data illustrate the importance of correcting for differences in drug pharmacokinetics and tolerability between mice and man.

Simulations using alternate clinical doses and schedules are in-line with clinical response
In order to explore whether our PK-PD models derived from preclinical efficacy could correctly rank-order the clinical outcome of agents given on alternative dose and
schedules, we performed additional simulations using human pharmacokinetics for docetaxel and 5-FU. Docetaxel and 5-FU were chosen for these simulations as both agents have been used clinically on several alternative dose and schedules. Two regimens were simulated for docetaxel, a once every three week regimen (100 mg/m$^2$ every 3 weeks; Docetaxel Regimen A) and a weekly regimen (40 mg/m$^2$ every week; Docetaxel Regimen B) (Figure 4A). For 5-FU, three regimens were simulated; a weekly regimen (600 mg/m$^2$ every week; 5-FU Regimen A), an intense 5 day regimen (500 mg/m$^2$ on days 1 to 5 of a 35 day regimen; 5-FU Regimen B) and a continuous constant infusion regimen (300 mg/m$^2$/day; 5-FU Regimen C) (Figure 4B). Simulated human %TGI was compared at Day 21 for docetaxel and Day 35 for 5-FU as these represented the longest duration for one cycle of treatment for the simulated regimens and would allow for equitable comparison. For docetaxel, simulated human %TGI for both regimens was comparable (Figure 4C) and aligned well with similar overall response in the clinic observed for both regimens in metastatic breast cancer (33-34%) (19). Similarly, simulated %TGI for 5-FU ranked well with the overall response observed in colorectal cancer trials with the continuous constant infusion regimen performing better than the weekly and the intense 5-day regimens (20-22) (Figure 4D).
DISCUSSION

Evaluation of anti-cancer agents in immune-deficient mice transplanted with subcutaneous tumors largely of human origin has been an important part of oncology drug discovery and development for the past 30 years. Differences in growth rates, immune competence, stromal content, and orthotopic location are often cited as factors contributing to the poor predictive value of murine xenografts. Indeed, mean tumor-volume-doubling times for untreated xenograft tumors in this manuscript were approximately 10 days or less. Growth is much slower in cancer patients with tumor-volume-doubling times in the order of months for breast and colon cancer and years for prostate cancer (23). Other important differences between preclinical tumor models in mice and humans include species differences in both pharmacokinetics and drug tolerability (2, 4, 24). Given these differences we wished to calibrate preclinical efficacy, normalized for species differences in drug exposure, to clinical response. Establishing methods that increase the correlation between preclinical and clinical efficacy will help increase the success rate of drugs brought forward ultimately leading to improved patient outcomes while simultaneously reducing drug development costs.

Population pharmacokinetic-pharmacodynamic modeling offers a powerful approach to improve the translatability of xenograft efficacy. Given the high variability observed in tumor growth, population PK-PD analysis provides a means to quantify the inherent variability in the xenograft model and provide a more robust characterization of the concentration/anti-tumor response relationship. Simulations of xenograft response using the established preclinical PK-PD models along with human pharmacokinetics enable a
more exact correction of species differences in drug exposure. The PK-PD analysis presented in this manuscript revealed a strong correlation between simulated human %TGI and overall response from clinical trials suggesting xenograft/allograft efficacy is predictive of clinical response (Figure 3B). Importantly, this correlation was not apparent when efficacy at maximally tolerated doses in preclinical models was compared to clinical response data demonstrating the importance of normalizing preclinical and clinical drug exposure (Figure 3A). Of interest, the top-performing agents in preclinical models and clinical response were all agents that target cancers with either activated oncogenes, erlotinib (EGFRmut, NSCLC) and dasatinib (BCR/ABL, chronic myeloid leukemia), or inactivated tumor suppressors, vismodegib (PTCH, medulloblastoma) (Figure 3B).

Of the anti-cancer agent/xenograft combinations tested, two cases were representative of clinical failures. Sunitinib failed to demonstrate meaningful clinical response in a Phase II colorectal cancer trial (25) and accordingly it had one of the lowest simulated %TGI (64% simulated TGI; Figure 3B). Simulated %TGI for vismodegib/D5123 was the lowest of all simulations performed; 40% TGI (not plotted in Figure 3B; Table 1). To date, vismodegib has shown great promise in cancers associated with ligand-independent activation of the Hh-pathway (18, 26-28). However, the drug has yet to show signs of clinical activity in tumors exhibiting ligand-dependent Hh-pathway activation (29, 30). Drugs with positive clinical trial results but the lowest clinical overall responses were 5-FU and trastuzumab (Table 1, Figure 3B) and were characterized by 75 and 73% simulated %TGI respectively. Combined, these data suggest that agents that can achieve
60% TGI or higher have an increased likelihood of eliciting a clinical response assuming that drug exposures used to drive preclinical efficacy are achievable in man. Of note, this 60% TGI bar is similar to the commonly cited T/C (treated/control) tumor size ratio of 40% that defines the threshold below which anti-cancer agents are considered active in conventional preclinical models (6). This cutoff is, however, only a guide and purposefully set conservatively to decrease the chances of discarding otherwise active drugs.

Observations from our study are consistent with the notion that xenograft efficacy is predictive of clinical response when drug concentrations in mice are appropriately corrected for therapeutic exposure (2, 4, 5). Overall, the correlation between %TGI derived from preclinical models and clinical response were markedly improved following PK-PD analysis (Figure 3). Two notable examples where correcting for exposure had a dramatic impact on the simulated %TGI are 5-FU and sunitinib in Colo205. Anti-tumor activity of 5-FU is anticipated to be lower due to higher circulating levels of thymidine in xenograft mice compared to human (31). However, 5-FU exhibited strong activity in Colo205 xenografts when used at its MTD (116%) (Figure 3A). Despite not being able to correct for differences in circulating thymidine, substantially less activity was predicted in simulations using drug exposures and schedules that are tolerated in the clinic and placed the preclinical activity (75% TGI) more in line with observed clinical response (Figure 3B). Similarly, sunitinib was very effective in treating Colo205 xenografts leading to 104% TGI at MTD but when corrected for exposures achieved in man led to 64% TGI. These differences in activity illustrate the power of this PK-PD.
approach for helping in the identification of novel anti-cancer agents. Finally, our finding that simulations of alternate dose and schedules of docetaxel and 5-FU in xenografts were predictive of the clinical response further illustrates the utility of this approach in preclinical drug development (Figure 4).

An important aspect of this investigation is the examination of preclinical efficacy from molecularly targeted anti-cancer agents. To our knowledge, there have been no comprehensive studies evaluating the predictive value of xenograft/allograft models for molecularly targeted anti-cancer agents. For such agents that target cancers driven by either oncogenic activation or tumor suppressor inactivation, the genetic features of the tumor would be expected to be more predictive of clinical response than tumor histology. Erlotinib is an epidermal growth factor receptor (EGFR) kinase inhibitor that is an approved treatment for non-small cell lung cancer (32). Mutations in the EGFR kinase domain serve to hyperactivate EGFR resulting in a dependence on this kinase for survival. Robust activity was observed for erlotinib in allografts implanted with a fibroblast line engineered to express a sensitizing mutation in EGFR (exon 19 del; Figure 2A) (33, 34). In contrast, little or no activity was observed in EGFR wild type or null non-small cell lung cancer xenografts (Figure S2). These differences in preclinical activity are consistent with what is known about genetic features of non-small lung cancers and their associated clinical response to EGFR inhibitors (32, 34). Vismodegib (also known as GDC-0449) was also tested in mice implanted with two molecularly distinct tumor types, medulloblastoma allografts and D5123 colorectal xenografts. Medulloblastoma allografts were derived from tumors that arose in the cerebellum of
Ptch+/− mice whose proliferation is driven by constitutive activation of the Hedgehog pathway (35, 36). D5123 tumors are colorectal tumors of patient derived origin that exhibit ligand-dependent Hedgehog pathway activation (36, 37). Vismodegib’s anti-tumor activity is clearly different in these two types of tumors with maximal anti-tumor activity being more robust in medulloblastoma allografts versus D5123 xenografts (115% TGI, 49% TGI, respectively, Figure S1 and Table 1). Preclinical efficacy from xenograft mice for vismodegib appears predictive of what has been observed thus far in the clinic (Table 1, Figure 3B). Similar to medulloblastoma allografts, robust activity has been reported in medulloblastoma and basal cell carcinoma both of which exhibit constitutive activation of the Hh pathway (18, 28). No activity has been reported thus far in colorectal and ovarian cancer trials both of which exhibit ligand dependent Hedgehog pathway activation (29, 30). As a whole, preclinical data from erlotinib and vismodegib demonstrate the importance of the genetic/molecular features of preclinical models when correlating preclinical efficacy to clinical response for molecularly targeted anti-cancer agents.

An important goal of this analysis was to provide a strategy to improve the identification of anti-cancer compounds using xenograft/allograft models in a drug discovery setting. From this analysis several key concepts emerge. First, these data highlight the importance of selecting preclinical models that best reflect the disease in which the drug is being evaluated in the clinic. This is perhaps best illustrated by the example of vismodegib which shows striking activity only in the setting of constitutive activation of the Hh-pathway and not ligand-dependent activation. These differences related to an
inherent difference in sensitivity to inhibition of the Hh-pathway (14). In the absence of specific predictive diagnostic hypotheses to guide model selection, models representative of the intended clinical population that exhibit a range of responses should be used (38). The use of both cell-line derived xenograft models and more specialized low-passage patient-derived models may enable such broad phenotypic exploration. Second, robust exposure gates should be set for advancement of agents from Phase I to Phase II clinical development. These gates should be consistent with the exposures required to drive greater than 60% TGI in relevant preclinical models.

The described strategy attempts to leverage an understanding of tumor biology coupled with PK-PD analysis in order to improve the identification and selection of new anti-cancer agents. It is hoped that the exposure-adjusted calibration of xenograft/allograft models to clinical response will improve the predictive value of these models and increase the speed at which novel drugs can bring benefit to cancer patients. Finally, the described strategy should not be viewed as exclusive to xenograft/allograft models but can be broadly applied to other preclinical models used in oncology.
REFERENCES


22. Lokich JJ, Ahlgren JD, Gullo JJ, Philips JA, Fryer JG. A prospective randomized comparison of continuous infusion fluorouracil with a conventional bolus schedule in


ACKNOWLEDGEMENTS

All authors are employees of Genentech. The authors thank members of the DMPK and translational oncology at Genentech that played a part in the generation of the data used in this manuscript. Vismodegib (GDC-0449) was discovered by Genentech under a collaboration agreement with Curis, Inc.
LEGENDS FOR FIGURES

**Figure 1**  The PK-PD modeling and simulation process used to translate preclinical data is depicted in Figure 1. Tumor growth kinetics in xenograft/allograft studies were described using either a standard indirect response (1A) or a tumor transit (1B) PK-PD model. In these models, $TV$ is defined as the tumor volume, $TV_1$, $TV_2$ and $TV_3$ are transit compartments representing tumor cells irreversibly damaged by treatment with an anti-cancer agent, $k_{ng}$ (day$^{-1}$) is the net growth rate constant, $K$ (day$^{-1}$) is the rate constant describing the anti-tumor effect of the anti-cancer agent and $k_{tran}$ is the rate constant describing movement of cells through the transit compartments. The simulation process integrating human pharmacokinetics is described in Figure 1C. PD parameters relating drug concentration to anti-tumor effect data are estimated for each anti-cancer drug following fitting of either model 1A or 1B to xenograft/allograft tumor volume data using mouse pharmacokinetics. Simulations of clinical regimens were performed using the preclinical PK-PD model, the estimated PD parameters from preclinical xenograft/allograft studies, and human pharmacokinetics from the literature (see Table S2).

**Figure 2**  An example of a PK-PD analysis of a xenograft/allograft efficacy study is shown. 2A shows a representative plot of mean tumor volume versus time for dosing ranging study with erlotinib in NR6-Del(746-752) allografts. 2B shows a representative simulation of clinical regimen of erlotinib using human
pharmacokinetics as described in Figure 1C. Figure 2C shows plots of individual PK-PD model predicted tumor volume vs. observed tumor volume for the following xenograft/allograft dose ranging studies: Erlotinib in NR6-Del(746-752) (A- ERLOTINIB); Vismodegib in medulloblastoma (B- VISMODEGIB); Vismodegib in D5123 (C- VISMODEGIB); Sunitinib in Colo205 (D- SUNITINIB); Sunitinib in 786-O (E- SUNITINIB); Dasatinib in K-562 (F- DASATINIB); Trastuzumab in KPL-4 (G- TRASTUZUMAB); Docetaxel in Cal51x1.1s (H- DOCETAXEL); Carboplatin in MX-1 (I- CARBOPLATIN); 5-FU in Colo205 (J- 5-FU).

**Figure 3** Plots of overall response versus observed %TGI at MTD in xenograft/allograft studies (3A) and overall response versus simulated human %TGI (3B) are shown for both molecular targeted agents and chemotherapy. The correlation coefficient (r) and p-values are shown for each plot.

**Figure 4** Simulations of clinical regimens using human pharmacokinetics and PK-PD models for docetaxel in Cal51x1.1 (4A), and 5-FU in Colo205 (4B). For docetaxel, regimen A is 100 mg/m² IV every three weeks and regimen B is 40 mg/m² IV every week. For 5-FU, regimen A is 600 mg/m² IV every week, regimen B is 500 mg/m² on Days 1 to 5, and regimen C is 300 mg/m²/day as a constant IV infusion. Simulated human %TGI is compared to overall response from single agent clinical trials using the same doses and schedules for docetaxel (4C) and 5-FU (4D). The overall response presented for 5-FU
regimen B in Figure 4D is the average single agent response for this 5-FU regimen based on two studies (21, 22).
<table>
<thead>
<tr>
<th>Anti-Cancer Agent</th>
<th>Xenograft/Allograft</th>
<th>Mouse %TGI at MTD</th>
<th>Human Simulated %TGI on Day 21</th>
<th>Clinical Disease</th>
<th>Clinical Response (Overall Response)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Erlotinib</td>
<td>Fibroblast NR6-del (746-752)</td>
<td>99%</td>
<td>102%</td>
<td>NSCLC*, exon 19 del</td>
<td>71%</td>
</tr>
<tr>
<td>Sunitinib</td>
<td>786-O (Renal Cell Adenocarcinoma)</td>
<td>106%</td>
<td>81%</td>
<td>Metastatic Renal Cell Carcinoma</td>
<td>47%</td>
</tr>
<tr>
<td>Sunitinib</td>
<td>Colo205 (Colorectal)</td>
<td>104%</td>
<td>63%</td>
<td>Colorectal Cancer</td>
<td>1.2%</td>
</tr>
<tr>
<td>Dasatinib</td>
<td>K-562 (Chronic Myeloid Leukemia)</td>
<td>117%</td>
<td>104%</td>
<td>Chronic Myeloid Leukemia</td>
<td>90%</td>
</tr>
<tr>
<td>Trastuzumab</td>
<td>KPL-4 (HER2+ Breast)</td>
<td>80%</td>
<td>73%</td>
<td>HER2+ Breast</td>
<td>15%</td>
</tr>
<tr>
<td>Vismodegib (GDC-0449)</td>
<td>Medulloblastoma, mutational driven tumor</td>
<td>115%</td>
<td>104%</td>
<td>Advanced Basal Cell Carcinoma</td>
<td>55%</td>
</tr>
<tr>
<td>Vismodegib (GDC-0449)</td>
<td>D5123 (Colorectal), ligand-driven tumor</td>
<td>49%</td>
<td>40%</td>
<td>Ovarian Cancer/ Colorectal with FOLFOX or FOLFIRI</td>
<td>No Signal</td>
</tr>
<tr>
<td>Docetaxel</td>
<td>Cal51x1.1s (Triple negative breast)</td>
<td>90%</td>
<td>88%</td>
<td>Metastatic Breast Cancer</td>
<td>33%</td>
</tr>
<tr>
<td>Carboplatin</td>
<td>MX-1 (Triple negative breast)</td>
<td>97%</td>
<td>88%</td>
<td>Advanced Breast Cancer</td>
<td>31%</td>
</tr>
<tr>
<td>5-FU</td>
<td>Colo205 (Colorectal)</td>
<td>116%</td>
<td>75%</td>
<td>Colorectal Cancer</td>
<td>10.3%</td>
</tr>
</tbody>
</table>

* NSCLC – Non small cell lung cancer
A. Indirect Response PK-PD Model

B. Tumor Transit PK-PD Model

\[ k_{ng} \quad TV \quad K \quad TV_1 \quad TV_2 \quad TV_3 \]

Tumor Volume = \( TV + TV_1 + TV_2 + TV_3 \)

C. Clinical Regimen Simulation

MOUSE PK

XENOGRAFT EFFICACY

PHARMACODYNAMIC PARAMETERS

FITTING

Effect

CONCENTRATION

SIMULATION AT CLINICAL DOSES

%TGI at Clinical Exposures

HUMAN PK

Tumor Vol

CONCENTRATION

Time

Time

FIGURE 1
**FIGURE 2**

A. Erlotinib - NR6-Del

B. Erlotinib - Simulation

C. Tumor volume comparison for various treatments:
   - A. ERLOTINIB
   - B. VISMODEGIB
   - C. VISMODEGIB
   - D. SUNITINIB
   - E. SUNITINIB
   - F. DASATINIB
   - G. TRASTUZUMAB
   - H. DOCETAXEL
   - I. CARBOPLATIN
   - J. 5-FU

Correlation coefficients:
- A: r = 0.97
- B: r = 0.98
- C: r = 0.98
- D: r = 0.94
- E: r = 0.98
- F: r = 0.96
- G: r = 0.96
- H: r = 0.97
- I: r = 0.97
- J: r = 0.97
FIGURE 3

A

Overall Response (%)

%TGI at MTD

r = 0.36
p = 0.34

\[ \Delta \text{ erlotinib} \]
\[ \times \text{ vismodegib - medullo} \]
\[ \ast \text{ sunitinib - Colo205} \]
\[ \bullet \text{ sunitinib - 786-O} \]
\[ + \text{ dasatinib} \]
\[ \diamond \text{ trastuzumab} \]
\[ - \text{ docetaxel} \]
\[ \circ \text{ carboplatin} \]
\[ \square \text{ 5-FU} \]

B

Overall Response (%)

Simulated %TGI

r = 0.91
p = 0.0008

\[ \Delta \text{ erlotinib} \]
\[ \times \text{ vismodegib - medullo} \]
\[ \ast \text{ sunitinib - Colo205} \]
\[ \bullet \text{ sunitinib - 786-O} \]
\[ + \text{ dasatinib} \]
\[ \diamond \text{ trastuzumab} \]
\[ - \text{ docetaxel} \]
\[ \circ \text{ carboplatin} \]
\[ \square \text{ 5-FU} \]
FIGURE 4

A. Docetaxel Simulation in Cal51x1.1s

B. 5-FU Simulation in Colo205

C. Docetaxel

D. 5-FU
Anti-tumor activity of targeted and cytotoxic agents in murine subcutaneous tumor models correlates with clinical response

Harvey Wong, Edna F. Choo, Bruno Alicke, et al.

*Clin Cancer Res* Published OnlineFirst May 30, 2012.

Updated version
Access the most recent version of this article at:
doi:10.1158/1078-0432.CCR-12-0738

Supplementary Material
Access the most recent supplemental material at:
http://clincancerres.aacrjournals.org/content/suppl/2012/05/30/1078-0432.CCR-12-0738.DC1

Author Manuscript
Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, use this link http://clincancerres.aacrjournals.org/content/early/2012/05/29/1078-0432.CCR-12-0738. Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.