Published OnlineFirst May 30, 2012; DOI: 10.1158/1078-0432.CCR-12-0738

Cancer Therapy: Preclinical

Clinical Cancer Research

Antitumor Activity of Targeted and Cytotoxic Agents in Murine Subcutaneous Tumor Models Correlates with Clinical Response

Harvey Wong[1], Edna F. Choo[1], Bruno Alicke[2], Xiao Ding[1], Hank La[1], Erin McNamara[2], Frank-Peter Theil[3], Jay Tibbitts[3], Lori S. Friedman[2], Cornelis E.C.A. Hop[1], and Stephen E. Gould[2]

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 anticancer 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. To improve the predictive nature of these models, we have carried out a retrospective population pharmacokinetic–pharmacodynamic (PK–PD) analysis of relevant xenograft/allograft efficacy data for eight agents (molecularly targeted and cytotoxic) with known clinical outcome.

Experimental Design: PK–PD modeling was carried out to first characterize the relationship between drug concentration and antitumor 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 carried out by replacing the murine pharmacokinetics, which were used to build the PK–PD model with human pharmacokinetics obtained from literature 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: On the basis of these analyses, agents that led to greater than 60% TGI 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 anticancer drug discovery is discussed. *Clin Cancer Res; 18(14); 3846–55. ©2012 AACR.*

Introduction

The discovery and development of new drugs is a challenging process and finding new anticancer 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 anticancer drug candidates. Immune-deficient mice transplanted with subcutaneous tumors (xenografts or allografts) have served as a major or workhorse model for the selection of anticancer drug candidates for more than 30 years. The primary advantages of this model system are the ease at which antitumor 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 cutoffs to assess preclinical response (i.e., T/C = 40%). One of the more extensive examinations of preclinical xenograft models predicting clinical response was carried out by the National Cancer Institute (NCI; ref. 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, antitumor activity in non–small cell lung cancer

Authors' Affiliations: Departments of [1]Drug Metabolism and Pharmacokinetics, [2]Translational Oncology, and [3]Pharmacokinetic and Pharmacodynamic Sciences, Genentech, Inc., South San Francisco, California

Note: Supplementary data for this article are available at Clinical Cancer Research Online (http://clincancerres.aacrjournals.org/).

Current address for F.-P. Theil: UCB Pharma, Brussels, Belgium.

Corresponding Authors: Harvey Wong, Department of 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, Department of Translational Oncology, Genentech, 1 DNA Way, MS 50, South San Francisco, CA 94080. Phone: 650-467-8025; E-mail: gould.stephen@gene.com

doi: 10.1158/1078-0432.CCR-12-0738

Published OnlineFirst May 30, 2012; DOI: 10.1158/1078-0432.CCR-12-0738

Translational Relevance

This work investigates whether or not preclinical murine subcutaneous tumor models are predictive of clinical response. We have carried out 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 carried out 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, seemed 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.

(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 anticancer 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 in which 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 biologic system of interest. PK–PD modeling is particularly helpful in understanding biologic systems in which 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/antitumor response relationship from inherently variable preclinical efficacy data. Finally, establishment of PK–PD models enables the ability to make prospective simulations of untested doses and schedules. The primary objective of this study was to evaluate whether xenograft/allograft data were predictive of clinical response following correction of differences between preclinical and clinical drug exposures through PK–PD modeling (depicted in Fig. 1). A second objective was to assess the predictive value of xenografts/allografts for molecular targeted anticancer 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), 5 million (Colo205, K-562, Cal51x1.1s, Calu-6), or 10 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) and implanted subcutaneously into the right flank or right thoracic region of female naive severe combined immunodeficient (SCID) mice (Taconic) or NCR (Taconic), athymic (Harlan Laboratories Inc.) or CD-1 nude mice (Charles River Laboratories). For MX-1 xenografts, D5123 xenografts and medulloblastoma allografts, tumor fragments (1–3 mm^3) were implanted subcutaneously into the right flank of naive female SCID or nude mice. Tumor fragments (1–3 mm^3) were implanted subcutaneously into the right lateral thorax of naive beige nude XID mice (Harlan Laboratories Inc.) 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 2 dimensions (length and width) and tumor volume (TV) was calculated using the following formula:

$$TV(mm^3) = (length \times width^2) \times 0.5$$

Mice were randomly assigned to groups of 5 to 10 individuals bearing similarly sized tumors. The mean tumor volume at the start of dosing for all studies ranged from 127 to 340 mm^3. Mice were promptly euthanized if their tumor volume exceeded 2,000 mm^3 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. KPL-4 breast cancer cells were obtained from J. Kurebayashi, Kawasaki Medical School, Kurashiki, Okayama, Japan (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, Department of Medicine, University of California, Los Angeles, CA (13) and engineered to express the EGF receptor (EGFR) Del (746-752) mutant at Genentech (12). All other cell lines were obtained from American Type Culture Collection. Cell lines were maintained in high-glucose Dulbecco's Modified Eagle's Medium:Ham's F-12 (50:50) supplemented with

Published OnlineFirst May 30, 2012; DOI: 10.1158/1078-0432.CCR-12-0738

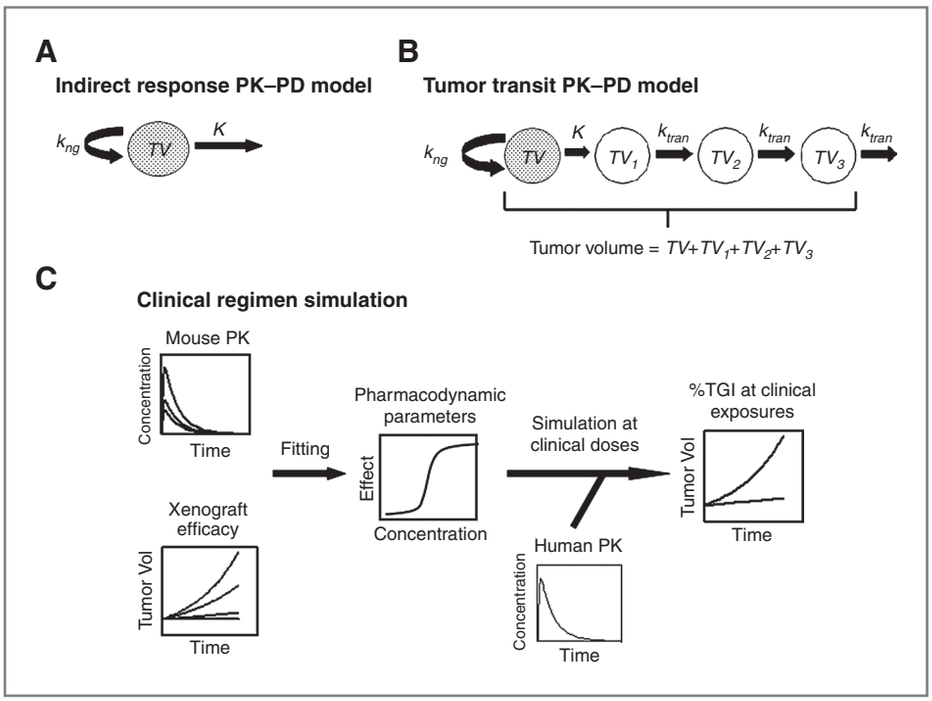

Figure 1. The PK–PD modeling and simulation process used to translate preclinical data is depicted. Tumor growth kinetics in xenograft/allograft studies were described using either a standard indirect response (A) or a tumor transit (B) PK–PD model. In these models, *TV* is defined as the tumor volume,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

10% FBS and 2 mmol/L L-glutamine. Cal51x1.1s is an *in vivo* passaged subclone of Cal51 (ATTC). D5123 colon adenocarcinoma tissue was acquired from National Disease Interchange. 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. Clinical grade sunitinib (Sutent; Pfizer), dasatinib (Sprycel; Bristol-Myers Squibb), 5-Fluorouracil (5-FU; Adrucil; Sicor Pharmaceuticals Inc.), carboplatin (Paraplatin; Bristol-Myers Squibb), and docetaxel (Taxotere; Sanofi-Aventis) were used in the studies. All other reagents or material used in this study were purchased from Sigma-Aldrich unless otherwise stated. The dosing regimens used in efficacy studies are listed in Supplementary Table S1. The highest dose used in the listed regimens above were at or near the MTD in mice for each anticancer agent, with the exception of trastuzumab for which MTD was not defined.

Pharmacokinetic studies in mice

Pharmacokinetic studies were carried out in female nude mice. The mouse strain, route of administration, and formulation used was selected to match the efficacy studies. Doses of each anticancer agent were chosen to adequately characterize the range over which the xenograft/allograft efficacy studies were done. Blood samples (~1 mL) were collected as described (14) and analyzed by liquid chromatography/tandem mass spectrometry. Serum rather than plasma was collected for trastuzumab and analyzed by ELISA assay. Pharmacokinetic parameters for each anticancer agent in mouse were estimated using SAAM II (Saam Institute, University of Washington, Seattle, WA). Estimated parameters were used to simulate systemic concentrations of anticancer agents for the various regimens in the xenograft studies described above during the PK–PD modeling process.

PK–PD modeling

Tumor volumes from xenograft efficacy studies were fitted to an indirect response model as described by the following differential equation and Fig. 1A.

$$\frac{d(TV)}{dt} = k_{ng}(TV) - K(TV) \qquad \text{(A)}$$

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 antitumor effect of the anticancer agent used in the xenograft study.

Published OnlineFirst May 30, 2012; DOI: 10.1158/1078-0432.CCR-12-0738

For studies in which the antitumor effect of the anticancer agent of interest was saturable, K was defined by Equation B.

$$K = \frac{K_{\max} \times C^n}{KC_{50}^n + C^n} \quad \text{(B)}$$

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

For studies in which the antitumor effect was linear, K was defined by Equation C:

$$K = E \times C^n \quad \text{(C)}$$

E ($L^*\mu mol^{-1*}day^{-1}$) is defined as a constant that relates the anticancer agent plasma concentration with K.

For some anticancer agents, a delay in antitumor 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 using a series of transit compartments representing irreversibly damaged cells similar to that described by Simeoni and coworkers was used (ref. 15; See Fig. 1B). The following series of differential equations (Equations D–F) along with Equation A was used to describe the described tumor transit PK–PD model.

$$\frac{d(TV_1)}{dt} = K(TV) - k_{tran}(TV_1) \quad \text{(D)}$$

$$\frac{d(TV_2)}{dt} = k_{tran}(TV_1) - k_{tran}(TV_2) \quad \text{(E)}$$

$$\frac{d(TV_3)}{dt} = k_{tran}(TV_2) - k_{tran}(TV_3) \quad \text{(F)}$$

TV_1, TV_2, and TV_3 are transit compartments representing tumor cells irreversibly damaged by treatment with an anticancer agent and k_{tran} (day^{-1}) 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 antitumor effect was described using Equation B and K for a linear antitumor effect was described using Equation C. 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 anticancer agent simultaneously to the described PK–PD models. Concentrations of anticancer agents were simulated based on their estimated pharmacokinetic parameters in mouse and the described dosing regimens. Intersubject 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 in which intersubject variance was small and could not be estimated reliably, the intersubject variance was fixed to 0.00001. Population parameter estimates are presented as the estimate followed by the %SE in parentheses. All estimates of population pharmacodynamic parameters were estimated with good precision with CVs less than 25%. All estimates of intra- and intersubject variance had CVs less than 45%.

Simulations of clinical regimens were carried out as described in Fig. 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 Supplementary 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 μmol/L 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.

$$\%TGI = \frac{TV_{vehicle} - TV_{treatment}}{TV_{vehicle} - TV_{initial}} \times 100 \quad \text{(G)}$$

$TV_{vehicle}$ is the tumor volume for the vehicle-treated animals at a specified endpoint time, $TV_{initial}$ is the initial tumor volume at the start of the treatment, and $TV_{treatment}$ is the tumor volume of the treatment groups at a specified endpoint time. Because %TGI is a time-dependent efficacy endpoint, it is calculated at a fixed time point. For the 50 mg/kg oral erlotinib studies in NSCLC xenografts, %TGI was calculated at the end of the dosing period for each study.

Pearson correlation coefficients (r) and their associated P values (2-tailed) were determined using GraphPad Prism version 4.02 for Windows (GraphPad Software).

Results

Characterization of preclinical dose–efficacy relationships

To characterize the relationship between drug and antitumor activity, we carried out 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 hypersensitive or inherently resistant. Dose ranges were chosen to provide the full range of efficacy with the upper end being the MTD, where achievable. In almost all cases, antitumor activity was dose dependent with the highest dose tested leading to near tumor stasis or regression during the dosing period (Supplementary Fig. S1 or Fig. 2A). Administration of vismodegib to patient-derived D5123

Published OnlineFirst May 30, 2012; DOI: 10.1158/1078-0432.CCR-12-0738

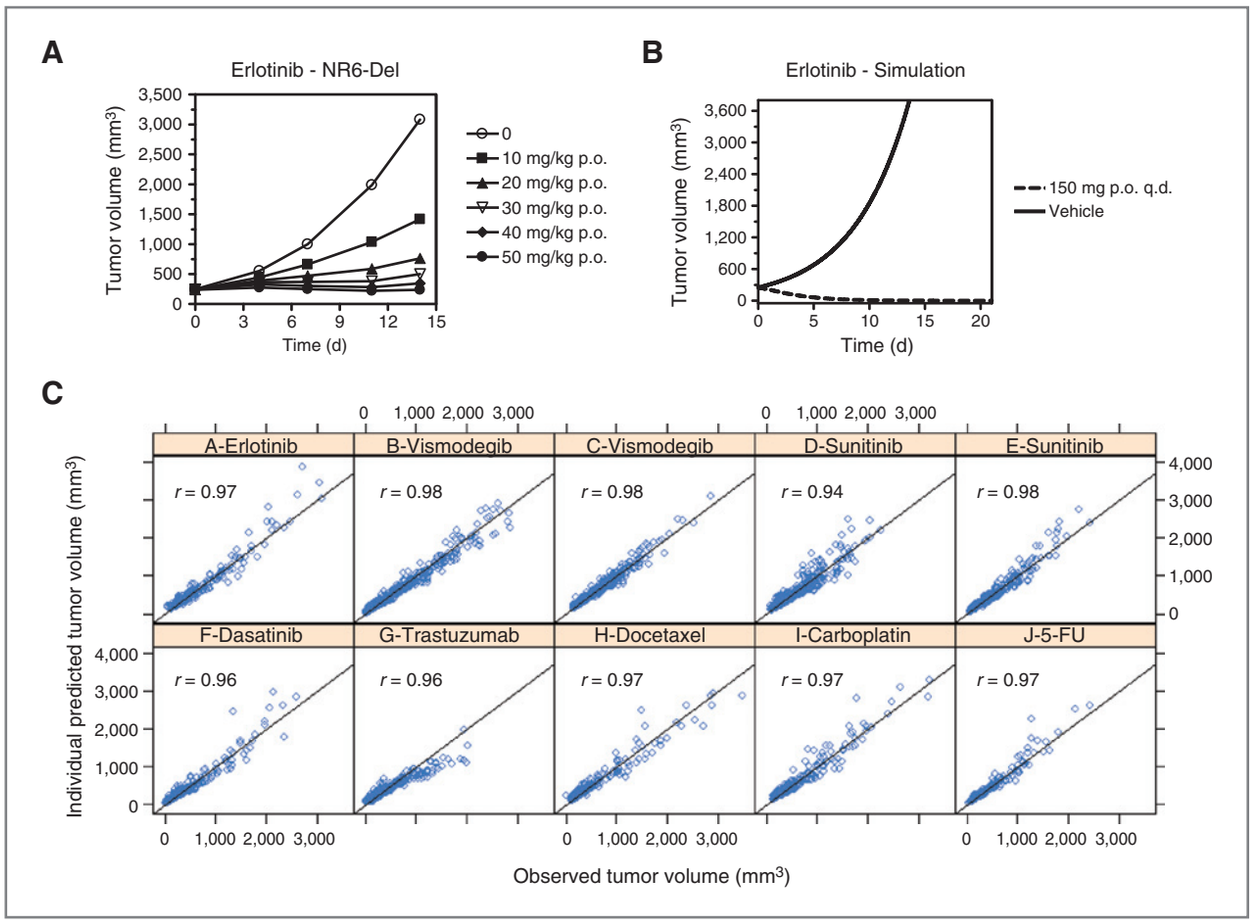

Figure 2. An example of a PK–PD analysis of a xenograft/allograft efficacy study is shown. A, a representative plot of mean tumor volume versus time for dosing-ranging study with erlotinib in NR6-Del(746-752) allografts. B, a representative simulation of clinical regimen of erlotinib using human pharmacokinetics as described in Fig. 1C. C, plots of individual PK–PD model predicted tumor volume versus 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). p.o., per oral; q.d, every day.

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 (Fig. 2A). Little to no antitumor activity was observed in studies in which 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 (Supplementary Fig. S2).

Effects of anticancer 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 (Fig. 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 Fig. 1A assumes that drug acts to decrease tumor volume by increasing the elimination rate constant, K. The second model depicted in Fig. 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 anticancer 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 interindividual 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 day^{-1}) being the slowest and NR6-Del(746-752) [k_{ng} –0.200 day^{-1}] being the fastest growing model (Supplementary 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

Published OnlineFirst May 30, 2012; DOI: 10.1158/1078-0432.CCR-12-0738

Table 1. Tumor growth inhibition at maximum tolerated dose in mouse, human simulated tumor growth inhibition using xenograft/allograft PK–PD model, and clinical response for 8 anticancer agents

Anticancer agent	Xenograft/ Allograft	Mouse %TGI at MTD	Human simulated %TGI on day 21	Clinical disease	Clinical response (overall response) %
Erlotinib	Fibroblast NR6-Del (746–752)	99	102	NSCLC, exon 19 del	71
Sunitinib	786-O (renal cell adenocarcinoma)	106	81	Metastatic renal cell carcinoma	47
Sunitinib	Colo205 (colorectal)	104	63	Colorectal cancer	1.2
Dasatinib	K-562 (chronic myeloid leukemia)	117	104	Chronic myeloid leukemia	90
Trastuzumab	KPL-4 (HER2+ Breast)	80	73	HER2+ Breast	15
Vismodegib (GDC-0449)	Medulloblastoma, mutational driven tumor	115	104	Advanced basal cell carcinoma	55
Visomodegib (GDC-0449)	D5123 (colorectal), ligand-driven tumor	49	40	Ovarian cancer/colorectal with FOLFOX or FOLFIRI	No signal
Docetaxel	Cal51x1.1s (triple negative breast)	90	88	Metastatic breast cancer	33
Carboplatin	MX-1 (triple negative breast)	97	88	Advanced breast cancer	31
5-FU	Colo205 (Colorectal)	116	75	Colorectal cancer	10.3

PK–PD models used adequately characterize the effect of drug on tumor volume (Fig. 2C).

Direct comparisons of pharmacodynamic parameters reflective of antitumor activity could not be made across all anticancer agents tested because different PK–PD models were used to characterize each agent. However, comparisons of the relative antitumor activity for particular agents in different xenografts/allografts could be made for sunitinib, vismodegib, and erlotinib. The maximum antitumor 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 KC_{50} (concentration of the anticancer agent in which K is 50% of K_{max})]. Visomodegib's K_{max} was approximately 18-fold higher and KC_{50} was 2-fold lower in medulloblastoma allografts that exhibit constitutive activation of the Hedgehog (Hh) pathway, compared with D5123 xenografts characterized by Hh ligand activation (14). In the case of erlotinib, we characterized the antitumor effect in EGFR-mutant [Del(746-752)] murine allografts. As mentioned above, little or no antitumor activity was observed in wild-type EGFR xenografts (Supplementary Fig. S2).

Antitumor activity in xenografts correlates with clinical response

To relate preclinical response to clinical response data, we replaced mouse pharmacokinetic parameters in the PK–PD models connecting preclinical efficacy and drug effect with human pharmacokinetic parameters derived from clinical trials and conducted simulations matching both dose and regimen as depicted in Fig. 1C (see Supplementary Table S2 for references for human PK). A representative simulation with human pharmacokinetics is shown for erlotinib in Fig. 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 %TGI using this hybrid PK–PD model was determined for each agent assuming a 21-day study (Table 1). The study duration was chosen because it is a common duration for xenograft studies and also corresponds to a common duration for 1 cycle of chemotherapy. Among the anticancer agents, simulations with vismodegib (in medulloblastoma allografts), erlotinib, and dasatinib resulted in the highest antitumor activity (Table 1). Of note, simulations with erlotinib and docetaxel were the only 2 of 10 simulations in which the %TGI was similar when the activity was driven by human pharmacokinetics versus that obtained with mouse exposures at, or near, the MTD. For all other simulations, the %TGI was lower when driven by human pharmacokinetics.

With the exception of clinical activity for vismodegib in tumors exhibiting ligand-dependent Hh pathway activation, overall response (complete + partial response) was obtained from single-agent trials in the literature (see Supplementary Table S2 for references). We found that when efficacy observed at the MTD in our preclinical studies was compared with overall response, there was no evidence of a correlation between antitumor activity in mice and man ($r = 0.36$, $P = 0.34$; Fig. 3A). However, when simulated

Published OnlineFirst May 30, 2012; DOI: 10.1158/1078-0432.CCR-12-0738

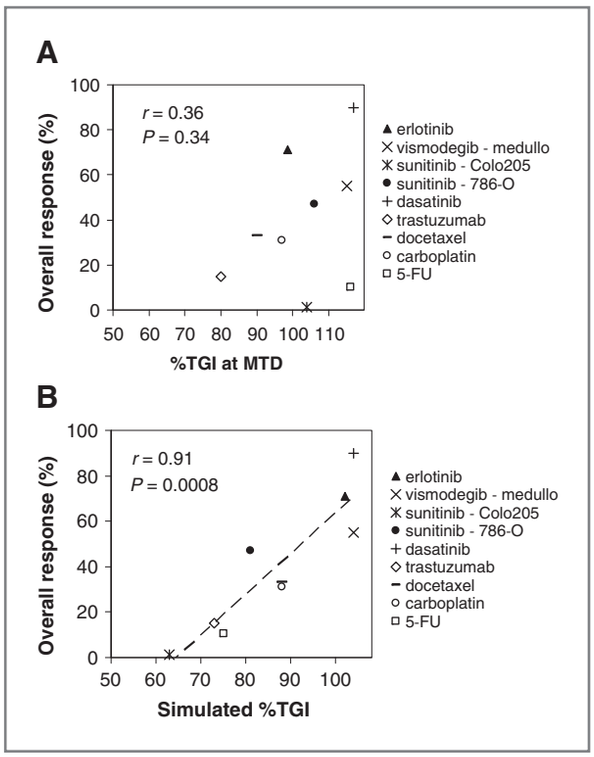

Figure 3. Plots of overall response versus observed %TGI at MTD in xenograft/allograft studies (A) and overall response versus simulated human %TGI (B) are shown for both molecular targeted agents and chemotherapy. The correlation coefficient (*r*) and *P* values are shown for each plot.

human %TGI derived from PK–PD modeling of our preclinical studies was compared with overall response rates from clinical trials, there was a significant correlation (Fig. 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

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 carried out 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 3 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; Fig. 4A). For 5-FU, 3 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/d; 5-FU Regimen C; Fig. 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 (Fig. 4C) and aligned well with similar overall response in the clinic observed for both regimens in metastatic breast cancer (33%–34%; ref. 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 (refs. 20–22; Fig. 4D).

Discussion

Evaluation of anticancer 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 article 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 PK–PD 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/antitumor 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 article revealed a strong correlation between simulated human %TGI and overall response from clinical trials, suggesting that xenograft/allograft efficacy is predictive of clinical response (Fig. 3B). Importantly, this correlation was not apparent when efficacy at maximally tolerated doses in preclinical models was compared with clinical response data showing the importance of normalizing preclinical and clinical drug exposure (Fig. 3A). Of interest, the top-performing agents in preclinical models and clinical response were all agents that target cancers with

Published OnlineFirst May 30, 2012; DOI: 10.1158/1078-0432.CCR-12-0738

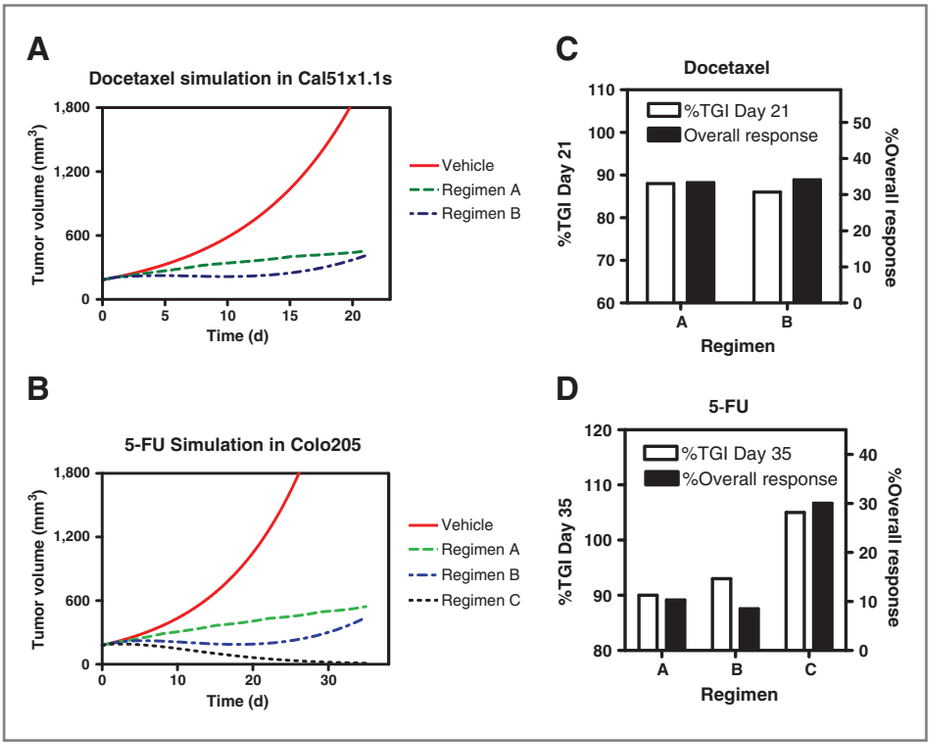

Figure 4. Simulations of clinical regimens using human pharmacokinetics and PK–PD models for docetaxel in Cal51x1.1 (A) and 5-FU in Colo205 (B). For docetaxel, regimen A is 100 mg/m^2 i.v. every 3 weeks and regimen B is 40 mg/m^2 i.v. every week. For 5-FU, regimen A is 600 mg/m^2 i.v. every week, regimen B is 500 mg/m^2 on days 1 to 5, and regimen C is 300 mg/m^2/d as a constant intravenous infusion. Simulated human %TGI is compared with overall response from single-agent clinical trials using the same doses and schedules for docetaxel (C) and 5-FU (D). The overall response presented for 5-FU regimen B in panel D is the average single-agent response for this 5-FU regimen based on 2 studies (21, 22).

either activated oncogenes, erlotinib (EGFR^{mut}, NSCLC), and dasatinib (BCR/ABL, chronic myeloid leukemia), or inactivated tumor suppressors, vismodegib (PTCH, medulloblastoma; Fig. 3B).

Of the anticancer agent/xenograft combinations tested, 2 cases were representative of clinical failures. Sunitinib failed to show meaningful clinical response in a phase II colorectal cancer trial (25), and, accordingly, it had one of the lowest simulated %TGI (64% simulated TGI; Fig. 3B). Simulated %TGI for vismodegib/D5123 was the lowest of all simulations carried out; 40% TGI (not plotted in Fig. 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, Fig. 3B) and were characterized by 75% and 73% simulated %TGI, respectively. Combined, these data suggested 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 anticancer 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 (Fig. 3). Two notable examples in which correcting for exposure had a dramatic impact on the simulated %TGI are 5-FU and sunitinb in Colo205. Antitumor activity of 5-FU is anticipated to be lower because of higher circulating levels of thymidine in xenograft mice compared with human (31). However, 5-FU exhibited strong activity in Colo205 xenografts when used at its MTD (116%; Fig. 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 (Fig. 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 anticancer 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 usefulness of this approach in preclinical drug development (Fig. 4).

An important aspect of this investigation is the examination of preclinical efficacy from molecularly targeted anticancer agents. To our knowledge, there have been no comprehensive studies evaluating the predictive value of xenograft/allograft models for molecularly targeted

Published OnlineFirst May 30, 2012; DOI: 10.1158/1078-0432.CCR-12-0738

anticancer 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 EGF receptor (EGFR) kinase inhibitor that is an approved treatment for NSCLC (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; Fig. 2A; refs. 33, 34). In contrast, little or no activity was observed in EGFR wild type or null NSCLC xenografts (Supplementary Fig. 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 2 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 Hh pathway (35, 36). D5123 tumors are colorectal tumors of patient-derived origin that exhibit ligand-dependent Hh pathway activation (36, 37). Vismodegib's antitumor activity is clearly different in these 2 types of tumors with maximal antitumor activity being more robust in medulloblastoma allografts versus D5123 xenografts (115% TGI, 49% TGI, respectively, Supplementary Fig. S1 and Table 1). Preclinical efficacy from xenograft/allograft mice for vismodegib seems predictive of what has been observed thus far in the clinic (Table 1, Fig. 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 Hh pathway activation (29, 30). As a whole, preclinical data from erlotinib and vismodegib show the importance of the genetic/molecular features of preclinical models when correlating preclinical efficacy to clinical response for molecularly targeted anticancer agents.

An important goal of this analysis was to provide a strategy to improve the identification of anticancer 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 are 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 to improve the identification and selection of new anticancer 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.

Disclosure of Potential Conflicts of Interest

All authors are employees of Genentech, Inc. F.-P. Theil is currently the head of Nonclinical Development at UCB Pharma, J. Tibbitts has ownership interest (including patents) in Roche, and L.S. Friedman has ownership interest (including patents) in Genentech, a member of the Roche group.

Authors' Contributions

Conception and design: H. Wong, E.F. Choo, B. Alicke, F.-P. Theil, C.E.C.A. Hop, S.E. Gould
Development of methodology: H. Wong, F.-P. Theil, S.E. Gould
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): H. Wong, B. Alicke, X. Ding, E. McNamara, J. Tibbitts, S.E. Gould
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): H. Wong, B. Alicke, X. Ding, S.E. Gould
Writing, review, and/or revision of the manuscript: H. Wong, E.F. Choo, B. Alicke, F.-P. Theil, J. Tibbitts, L.S. Friedman, S.E. Gould
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): H. La
Study supervision: E. McNamara, F.-P. Theil, C.E.C.A. Hop, S.E. Gould

Acknowledgments

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 article. Vismodegib (GDC-0449) was discovered by Genentech under a collaboration agreement with Curis, Inc.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received March 2, 2012; revised April 23, 2012; accepted May 14, 2012; published OnlineFirst May 30, 2012.

References

1. Kola I, Landis J. Can the pharmaceutical industry reduce attrition rates? Nat Rev Drug Discov 2004;3:711–5.
2. Peterson JK, Houghton PJ. Integrating pharmacology and *in vivo* cancer models in preclinical and clinical drug development. Eur J Cancer 2004;40:837–44.
3. Kelland LR. "Of mice and men": values and liabilities of the athymic nude mouse model in anticancer drug development. Eur J Cancer 2004;40:827–36.
4. Troiani T, Schettino C, Martinelli E, Morgillo F, Tortora G, Ciardiello F. The use of xenograft models for the selection of cancer

Published OnlineFirst May 30, 2012; DOI: 10.1158/1078-0432.CCR-12-0738

treatments with the EGFR as an example. Crit Rev Oncol Hematol 2008;65:200–11.
5. Kerbel RS. Human tumor xenografts as predictive preclinical models for anticancer drug activity in humans: better than commonly perceived-but they can be improved. Cancer Biol Ther 2003;2:S134–9.
6. Johnson JI, Decker S, Zaharevitz D, Rubinstein LV, Venditti JM, Schepartz S, et al. Relationships between drug activity in NCI preclinical *in vitro* and *in vivo* models and early clinical trials. Br J Cancer 2001;84:1424–31.
7. 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;9:4227–39.
8. Mager DE, Wyska E, Jusko WJ. Diversity of mechanism-based pharmacodynamic models. Drug Metab Dispos 2003;31:510–8.
9. Danhof M, de Jongh J, De Lange EC, Della Pasqua O, Ploeger BA, Voskuyl RA. Mechanism-based pharmacokinetic-pharmacodynamic modeling: biophase distribution, receptor theory, and dynamical systems analysis. Annu Rev Pharmacol Toxicol 2007;47:357–400.
10. Kurebayashi J, Otsuki T, Tang CK, Kurosumi M, Yamamoto S, Tanaka K, et al. Isolation and characterization of a new human breast cancer cell line, KPL-4, expressing the Erb B family receptors and interleukin-6. Br J Cancer 1999;79:707–17.
11. Junttila TT, Parsons K, Olsson C, Lu Y, Xin Y, Theriault J, et al. Superior *in vivo* efficacy of afucosylated trastuzumab in the treatment of HER2-amplified breast cancer. Cancer Res 2010;70:4481–9.
12. Carey KD, Garton AJ, Romero MS, Kahler J, Thomson S, Ross S, et al. Kinetic analysis of epidermal growth factor receptor somatic mutant proteins shows increased sensitivity to the epidermal growth factor receptor tyrosine kinase inhibitor, erlotinib. Cancer Res 2006;66:8163–71.
13. Chazin VR, Kaleko M, Miller AD, Slamon DJ. Transformation mediated by the human HER-2 gene independent of the epidermal growth-factor receptor. Oncogene 1992;7:1859–66.
14. Wong H, Alicke B, West KA, Pacheco P, La H, Januario T, et al. Pharmacokinetic-pharmacodynamic analysis of vismodegib in preclinical models of mutational and ligand-dependent hedgehog pathway activation. Clin Cancer Res 2011;17:4682–92.
15. Simeoni M, Magni P, Cammia C, De Nicolao G, Croci V, Pesenti E, et al. Predictive pharmacokinetic-pharmacodynamic modeling of tumor growth kinetics in xenograft models after administration of anticancer agents. Cancer Res 2004;64:1094–101.
16. D'Argenio DZ, Schumitzky A. ADAPT II user's guide: pharmacokinetic/pharmacodynamic systems analysis software 1997. Los Angeles: Biomedical Simulations Resource.
17. Bauer RJ, Guzy S. Monte Carlo parametric expectation maximization (MC-PEM) method for analyzing population pharmacokinetic/pharmacodynamic (PK/PD) data. In:D'Argenio DZ, editors. Advanced methods of pharmacokinetic and pharmacodynamic systems analysis. Boston: Kluwer; 2004. p. 135–63.
18. Von Hoff D, Lorusso P, Rudin C, Reddy J, Yauch R, Tibes R, et al. Inhibition of the Hedgehog pathway in advanced basal-cell carcinoma. N Engl J Med 2009;361:1164–72.
19. Tabernero J, Climent MA, Lluch A, Albanell J, Vermorken JB, Barnadas A, et al. A multicentre, randomised phase II study of weekly or 3-weekly docetaxel in patients with metastatic breast cancer. Ann Onc 2004;15:1358–65.
20. Kalofonos HP, Nicolaides C, Samantas E, Mylonakis N, Aravantinos G, Dimopoulos MA, et al. A phase III study of 5-fluorouracil versus 5-fluorouracil plus interferon alpha 2b versus 5-fluorouracil plus leucovorin in patients with advanced colorectal cancer: a Hellenic Cooperative Oncology Group (HeCOG) study. Am J Clin Oncol 2002;25:23–30.
21. Poon MA, O'Connell MJ, Moertel CG, Wieand HS, Cullinan SA, Everson LK, et al. Biochemical modulation of fluorouracil: evidence of significant improvement of survival and quality of life in patients with advanced colorectal carcinoma. J Clin Oncol 1989;7:1407–18.
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 metastatic colorectal carcinoma: a Mid-Atlantic Oncology Program Study. J Clin Oncol 1989;7:425–32.
23. Klein CA. Parallel progression of primary tumours and metastases. Nat Rev Cancer 2009;9:302–12.
24. Houghton PJ, Adamson PC, Blaney S, Fine HA, Gorlick R, Haber M, et al. Testing of new agents in childhood cancer preclinical models: meeting summary. Clin Cancer Res 2002;8:3646–57.
25. Saltz LB, Rosen LS, Marshall JL, Belt RJ, Hurwitz HI, Eckhardt SG, et al. Phase II trial of sunitinib in patients with metastatic colorectal cancer after failure of standard therapy. J Clin Oncol 2007;25:4793–9.
26. LoRusso PM, Rudin CM, Reddy JC, Tibes R, Weiss GJ, Borad MJ, et al. Phase I trial of hedgehog pathway inhibitor vismodegib (GDC-0449) in patients with refractory, locally advanced or metastatic solid tumors. Clin Cancer Res 2011;17:2502–11.
27. Amin SH, Tibes R, Kim JE, Hybarger CP. Hedgehog antagonist GDC-0449 is effective in the treatment of advanced basal cell carcinoma. Laryngoscope 2010;120:2456–9.
28. Rudin C, Hann C, Laterra J, Yauch R, Callahan C, Fu L, et al. Treatment of medulloblastoma with Hedgehog pathway inhibitor GDC-0449. N Engl J Med 2009;361:1173–8.
29. Kaye SB, Fehrenbacher L, Holloway R, Horowitz N, Karlan B, Amit A, et al. A phase 2, randomized, placebo-controlled study of hedgehog (HH) pathway inhibitor GDC-0449 as maintenance therapy in patients with ovarian cancer in 2(nd) or 3(rd) complete remission (CR). Ann Oncol 2010;21:(Supplement 8) viii11.
30. Berlin J, Bendell J, Hart LL, Firdaus I, Gore I, Hermann RC, et al. A phase 2, randomized, double-blind, placebo-controlled study of hedgehog pathway inhibitor (HPI) GDC-0449 in patients with previously untreated metastatic colorectal cancer (mCRC). Ann Oncol 2010;21:(Supplement 8) viii10.
31. Houghton JA, Williams LG, Loftin SK, Cheshire PJ, Morton CL, Houghton PJ, et al. Factors that influence the therapeutic activity of 5-fluorouracil [6RS]leucovorin combinations in colon adenocarcinoma xenografts. Cancer Chemother Pharmacol 1992;30:423–32.
32. Sharma SV, Bell DW, Settleman J, Haber DA. Epidermal growth factor receptor mutations in lung cancer. Nat Rev Cancer 2007;7:169–81.
33. Mulloy R, Ferrand A, Kim Y, Sordella R, Bell DW, Haber DA, et al. Epidermal growth factor receptor mutants from human lung cancers exhibit enhanced catalytic activity and increased sensitivity to gefitinib. Cancer Res 2007;67:2325–30.
34. Lynch TJ, Bell DW, Sordella R, Gurubhagavatula S, Okimoto RA, Brannigan BW, et al. Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. N Engl J Med 2004;350:2129–39.
35. Wong SY, Reiter JF. Wounding mobilizes hair follicle stem cells to form tumors. Proc Natl Acad Sci U S A 2011;108:4093–8.
36. Scales S, de Sauvage F. Mechanisms of Hedgehog pathway activation in cancer and implications for therapy. Trends Pharmacol Sci 2009;30:303–12.
37. Yauch R, Gould S, Scales S, Tang T, Tian H, Ahn C, et al. A paracrine requirement for hedgehog signalling in cancer. Nature 2008;455:406–10.
38. Houghton PJ, Morton CL, Tucker C, Payne D, Favours E, Cole C, et al. The pediatric preclinical testing program: description of models and early testing results. Pediatr Blood Cancer 2007;49:928–40.

Published OnlineFirst May 30, 2012; DOI: 10.1158/1078-0432.CCR-12-0738

Clinical Cancer Research

Antitumor 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 2012;18:3846-3855. Published OnlineFirst May 30, 2012.