Nonclinical Evaluations of Small-Molecule Oncology Drugs: Integration into Clinical Dose Optimization and Toxicity Management

Donna M. Dambach[1,2], Natalie E. Simpson[3], Thomas W. Jones[2,4], Richard J. Brennan[5], Richard Pazdur[3], and Todd R. Palmby[3]

Abstract

Multidisciplinary approaches that incorporate nonclinical pharmacologic and toxicologic characterization of small-molecule oncology drugs into clinical development programs may facilitate improved benefit–risk profiles and clinical toxicity management in patients. The performance of the current nonclinical safety-testing scheme was discussed, highlighting current strengths and areas for improvement. While current nonclinical testing appears to predict the clinical outcome where the prevalence of specific adverse effects are high, nonclinical testing becomes less reliable for predicting clinical adverse effects that occur infrequently, as with some kinase inhibitors. Although adverse effects associated with kinase inhibitors can often be predicted on the basis of target biology, drugs can be promiscuous and inhibit targets with poorly defined function and associated risks. Improvements in adverse effect databases and better characterization of the biologic activities of drug targets may enable better use of computational modeling approaches in predicting adverse effects with kinase inhibitors. Assessing safety of a lead candidate in parallel with other drug properties enables incorporation of a molecule's best features during chemical design, eliminates the worst molecules early, and permits timely investigation/characterization of toxicity mechanisms for identified liabilities. A safety lead optimization and candidate identification strategy that reduces intrinsic toxicity and metabolic risk and enhances selectivity can deliver selective kinase inhibitors that demonstrate on-target adverse effects identified nonclinically. Integrating clinical and nonclinical data during drug development can facilitate better identification and management of oncology drugs. Follow-up nonclinical studies may be used to better understand the risks in a given patient population and minimize or manage these risks more appropriately. *Clin Cancer Res; 22(11); 2618–22. ©2016 AACR.*

See all articles in this *CCR Focus* section, "New Approaches for Optimizing Dosing of Anticancer Agents."

Introduction

Multidisciplinary approaches to improve the efficiency of oncology drug development include integration of pharmacologic and toxicologic characterization with clinical trial experience. Dose optimization for targeted oncology drugs, such as kinase inhibitors (KI), has proven challenging when using the traditional clinical paradigm for cytotoxic drugs (1). Patients treated with the recommended therapeutic dose of some approved targeted drugs frequently undergo dose reductions due to toxicity. Further dose optimization is often needed in the postmarketing setting. Identifying key best practices in the nonclinical evaluation of pharmacology, including selectivity and secondary pharmacology, and toxicology and integration of this information into the design and conduct of clinical trials may facilitate future development of drugs with improved benefit-risk profiles and better management of toxicities in patients.

The development of two Abelson (ABL) KIs, ponatinib and nilotinib, highlights the challenges associated with adverse effects (AE) that would have benefited from more robust characterization in clinical trials to support initial FDA approval. Both drugs have a favorable risk–benefit profile but also a history of labeling changes. Clinical trials and animal toxicology studies did not adequately characterize the clinically relevant AEs of thromboembolism and vascular occlusion, which have been reported at increasing frequency since FDA approval. These two drugs appear to have an increased risk for these cardiovascular effects compared with other ABL KIs. An exploratory analysis by FDA of grade 3–4 cardiovascular effects obtained using human safety data showed that vascular effects observed with ponatinib and nilotinib cluster away from other ABL KIs and traditional VEGFR inhibitors. Another clustering analysis of published data, in which the activity of various KIs against 300 human kinases was compared, demonstrated a broad range of selectivity among KIs (2). Ponatinib and nilotinib target multiple kinases, including those with known effects on endothelial cell survival and function and vascular permeability and maintenance, and differentially target some kinases compared with other ABL inhibitors. Conduct of further nonclinical characterization of the underlying mechanism(s) leading to these cardiovascular AEs in patients may aid in identification of risk factors, safety biomarkers, or appropriate

[1]Department of Safety Assessment, Genentech, San Francisco, California. [2]International Consortium for Innovation and Quality in Pharmaceutical Development DruSafe Leadership Group, Washington, District of Columbia. [3]Office of Hematology and Oncology Products, Center for Drug Evaluation and Research, U.S. Food and Drug Administration, Silver Spring, Maryland. [4]Toxicology and Pathology, Eli Lilly and Company, Indianapolis, Indiana. [5]Preclinical Safety, Sanofi, Waltham, Massachusetts.

Corresponding Author: Todd R. Palmby, Office of Hematology and Oncology Products, Center for Drug Evaluation and Research, U.S. Food and Drug Administration, 10903 New Hampshire Avenue, Silver Spring, MD 20993. Phone: 301-796-4254; Fax: 301-796-9909; E-mail: todd.palmby@fda.hhs.gov

doi: 10.1158/1078-0432.CCR-15-2645

management of these effects by treating physicians. While the analyses were exploratory, this example highlights the potential opportunities for integrating nonclinical and clinical information at all stages of a drug product's life cycle.

This article presents experiences and considerations for the utility of nonclinical safety assessments to inform clinical AEs and trial design. Issues relating to approaches to oncology dose-finding trials and dose–exposure exploration and challenges in early-phase oncology trials are discussed elsewhere in this *CCR Focus* section (3–5) and in ref. 6.

Is It Safe: Understanding the Performance of Nonclinical Safety Assessment Models in Predicting Human Outcomes?

Although clinical trials are generally regarded as safe for patients, the current nonclinical safety-testing scheme has come under increasing criticism in recent years for failing to predict serious AEs in humans. There is concern that the use of healthy, young animals in the current testing paradigm may underestimate risk to clinical trial subjects or the broader population of patients exposed postregistration for certain drug classes. Despite the fact that nonclinical safety data are used to support nearly all clinical development activities, there remains a paucity of the human outcome data necessary to quantitatively evaluate nonclinical safety assessment performance. By piecing together key performance attributes (i.e., sensitivity and specificity) drawn from the literature, a quantitative framework model was developed to illustrate how the current nonclinical testing scheme operates. The model indicates that the positive predictive value (PPV) of nonclinical safety testing, that is how well the testing scheme predicts serious adverse human outcomes, is strongly influenced by the nature (i.e., pharmacologic class) of the compounds being tested. Overall, the prevalence or incidence of compounds associated with serious adverse human outcomes entering clinical development is low. As a result and despite commonly held expectations, the PPV associated with nonclinical safety testing is, on average, less than 50%. Instead, the model emphasizes the importance of negative predictive value (NPV) in increasing confidence that novel drug candidates will be safe in human testing. When compounds are tested in the clinic under conditions demonstrated to be safe using nonclinical models (i.e., safe exposure levels), there is a high probability that they will be safe in humans. This would explain how nonclinical testing works to protect clinical trial subjects despite the fact that the PPV (the ability to predict specific adverse outcomes) is relatively low. However, the framework model also anticipates a different performance profile when testing specific compound classes associated with a higher inherent risk of human toxicity (such as oncology agents including KIs). In testing these agents, where the prevalence of specific adverse human outcomes (gastrointestinal and hematologic toxicity) is well above 50%, the PPV of nonclinical testing would be expected to be high as well. On the flipside, the lower NPV under these circumstances means that nonclinical testing becomes a less reliable predictor of safe clinical testing conditions. As an example, the framework model suggests that cardiovascular toxicity represents a unique "blind-spot" in the nonclinical testing of KIs. Despite a high prevalence of human cardiovascular toxicity associated with this compound class, using examples derived from the literature, the framework model demonstrates overlapping pretest and posttest probabilities (also known as, prevalence and PPV, respectively) suggesting a failure of nonclinical testing to resolve any uncertainty regarding cardiovascular toxicity potential at the individual molecule level. This would suggest that work is needed to identify new nonclinical safety approaches capable of better anticipating the human cardiovascular toxicity risk associated with this class of drugs.

Nonclinical to Clinical Correlation in AEs of KIs

As of February 8, 2015, 30 small-molecule KIs had been approved by the FDA (7), the majority of which (27) are for oncology indications. Although KIs are generally safer than cytotoxic chemotherapeutic agents, clinical practice revealed a number of AEs associated with kinase inhibition, some of them serious and dose limiting (8).

In many cases, clinical AEs are expected because of the biologic functions of the kinases targeted. Toxicities arising in rapid-turnover tissues due to perturbation of signaling through highly conserved cell proliferation and survival pathways are well predicted in nonclinical studies, such as dermatotoxicity with EGFR inhibitors and gastrointestinal toxicity with KIs targeting a number of kinases (8, 9).

Cardiovascular AEs have been serious and dose limiting for many approved KIs and nonclinical prediction of these toxicities is often difficult. Effects on QT interval and arrhythmia caused by inhibition of ion channels such as hERG (nilotinib), may be predictable by standard *in vitro* screening assays. However, serious functional cardiovascular AEs such as hypotension, bleeding, left ventricular function loss, and cardiac failure with VEGF inhibitors, cardiomyopathy and cardiac failure with mitogen-activated protein kinase kinase (MEK) inhibitors, and left ventricular dysfunction, thromboembolism, vascular occlusion, and congestive heart failure with ABL inhibitors are reflected only by more subtle nonclinical findings (10).

Some ocular toxicities, such as retinal vein occlusion (MEK) or conjunctivitis (EGFR), may have species-specific correlates or manifest differently in the different species (8). Visual disturbances that are subjectively reported by the patient, such as is seen with crizotinib, cannot be monitored in standard nonclinical toxicology studies, but may be detectable by specialized electroretinography measurements. Similarly, hypothyroidism observed clinically with some KIs manifests as a regression of thyroid vasculature and increased levels of thyroid-stimulating hormone in mice (8).

Hepatic and renal toxicities of KIs are not well predicted in animal studies; approximately 40% to 50% of human hepatotoxicants, and 50% of human nephrotoxicants were detected in rat or dog. Causes of hepatotoxicity are multifactorial and the interplay of the various factors is not well understood (8). There is a high rate of false positives for renal toxicity in animal studies (8). Several other rare but serious AEs, including interstitial lung disease, idiopathic pulmonary fibrosis, and reversible posterior leukoencephalopathy have been seen with KIs, but not predicted nonclinically. Human-specific risk factors such as smoking, comorbidities, or human-specific functions of the targeted kinase may contribute to these differences in AE detection.

The conserved nature of the ATP-binding pocket targeted by most KIs makes achieving target specificity challenging. Many KIs are active against more than one and sometimes several unrelated

kinases (11). The pattern and severity of AEs is dependent on the particular spectrum of target activities for a given molecule. The pathways/networks in which the affected kinases participate and the range of biologic functions controlled via these networks determine the pattern of AE. In some cases, patterns of target activity may predict specific AEs (12). As more kinases are characterized, and the roles and interplay of the signaling networks in which they participate more completely understood, it should become possible to design KI drugs with improved efficacy and AE profiles.

Safety Lead Optimization of KIs: Learning from Nonclinical to Clinical Translation of AEs

Safety lead optimization (LO) focuses on identification of the most promising drugs through developing and implementing assay strategies and hypothesis-driven issue investigation that enable informed decision making. Front-loading the assessment of safety in parallel with other drug properties (e.g., efficacy, pharmacokinetics, pharmaceutics) enables chemists to incorporate the best overall features of a molecule during chemical design, eliminates the worst molecules early, and permits timely investigation/characterization of toxicity mechanisms for identified liabilities. The desired outcomes are to progress the drug with the best safety profile for a therapeutic area; remove the most toxic drugs from the portfolio prior to entry into humans to reduce clinical attrition due to toxicity; and establish a well-characterized hazard and translational risk profile to inform clinical trial designs. This is accomplished through a framework that balances the multiple considerations to identify a drug with the overall best drug characteristics, and a cogent understanding of mechanisms of toxicity. The framework components include establishing a target candidate profile for each program that defines the qualities of a successful drug based on the intended therapeutic area, including the risk tolerance for liabilities, evaluating potential liabilities that may be pharmacology-mediated (on-target) and that are chemically mediated (off-target), and characterizing identified liabilities. LO and investigation relies upon the integrated use of a variety of technologies and models (*in silico*, *in vitro*, *in vivo*) that have achieved a sufficient level of qualification or validation to provide confidence in their use. This framework is employed at Genentech to identify small-molecule drugs that are highly selective, have minimal intrinsic toxicity, have low metabolic risk (low rates of metabolism and reactive metabolite formation), and have well-characterized safety profiles (13). To assess whether this approach was achieving the desired outcomes, investigators at Genentech evaluated the impact of their safety LO strategy on their portfolio of 37 target-distinct KIs with regard to advancing drugs clinically and the translatability of AEs.

Eleven drugs (~30%) advanced into clinical trials for oncology indications. Of those, two (18%) have been removed from the clinical portfolio due to on-target, dose-limiting toxicities that precluded achieving meaningful efficacy. For these two drugs, the dose-limiting toxicities were a recognized class effect; the toxicities were identified nonclinically *in vitro*, but were not observed in animal studies. While the liability was known, it was not possible to use nonclinical platforms to reliably model a clinical therapeutic index or dosing regimen.

Comparison of the common (>10%) or significant AEs observed clinically to those identified in nonclinical models revealed that 87% of clinical AEs were defined as on-target and in all cases these on-target toxicities were identified in nonclinical models. There were several AEs identified nonclinically that did not translate to humans. These included effects on the lymphoid and hematopoietic systems (two drugs), potential reproductive effects (which cannot be similarly assessed clinically; one drug), cardiac effects (two drugs), and a rodent-specific finding (one drug). Finally, several clinical AEs did not have nonclinical correlates. These included fatigue or asthenia (three drugs), gastrointestinal effects (nausea, emesis, diarrhea; one drug), ocular effects (one drug; known class effect), and elevations in creatine kinase (one drug).

On the basis of the available data from the KI portfolio at Genentech for several oncology targets, the safety LO and candidate identification strategy that reduces intrinsic toxicity and metabolic risk and enhances selectivity has delivered highly selective KIs that tend to show on-target AEs that are identified nonclinically. The experience gained from advancing KIs for oncology can be applied toward developing KIs for non-oncology therapeutic targets.

Enhancing the Safety of KI Oncology Drugs

The goal of integrating clinical and nonclinical data during drug development is to facilitate better identification and management of the drugs that are being introduced into clinical use. Our knowledge of the kinome, specifically which kinases may make good targets for efficacy in specific settings and which should be avoided because of toxicity, is still incomplete. The methodology used to evaluate potencies does not always translate well. In addition, it is critical to understand what these kinases are doing in normal tissues as a means to assess potential toxicities because mechanisms of how kinases act in the tumor are often quite different than the way they act in normal tissues.

Given the current clinical approach in oncology to maximize dose levels expected to provide benefit, it is not surprising that patients experience toxicity, which must be managed. Whether dose escalation or dose de-escalation is needed should not be viewed as a failure, but rather as a mechanism to obtain the most information from the clinical trials as efficiently as possible. Nonclinical studies actually have fairly good predictability for common clinical effects of KIs; however, AEs are infrequently observed in clinical trials that are not readily detected in traditional nonclinical studies, for example, thromboembolism and edema. In these cases, directed nonclinical studies may be needed to observe these effects, which may be used to initially characterize rather than predict clinical toxicity. Furthermore, cardiomyopathy induced by KIs has been documented to include comorbidities such as a patient's age, previous anthracycline treatment, and hypertension. In normal healthy animals used for toxicity testing these risk factors would not be detectable. Nonetheless, animal models can be effective in evaluating a specific issue in a proactive fashion, or retroactively when a problematic finding is encountered during clinical development. Nonclinical studies may aid in determining a relationship with schedule, dose, or pharmacokinetics.

For example, rats were used to address questions related to clinical management of a MEK inhibitor–induced cardiomyopathy. Decreased left ventricular ejection fraction in normal rats is

minimal, making it a difficult model to use in evaluating changes due to administration of the MEK inhibitor. Spontaneously hypertensive rats have reduced left ventricular ejection fraction, making it easier to observe an effect of the MEK inhibitor and better represent what occurs clinically. This model demonstrated that the reduced ventricular ejection fraction induced by the MEK inhibitor was reversible when treatment was stopped, was no worse when the animals were retreated, and was preventable by control of hypertension with an ACE inhibitor, but not with a β-blocker. As such, follow-up nonclinical studies to enhance our understanding of the risks in a given patient population and to minimize or manage the risks more appropriately will be of value.

Discussion and Recommendations

Small-molecule anticancer drug safety lead optimization and follow-up nonclinical studies intended to improve clinical toxicity management can improve the efficiency and success of development programs for these therapies (Fig. 1). The nonclinical characterization of activity and safety during development of targeted small molecules intended to treat patients with cancer has evolved on the basis of past experiences that highlight successes and areas for improvement. Advancements in drug selection and optimization using rigorous methods to screen for specific characteristics and activities have dramatically reduced drug attrition rates in oncology due to toxicity in early-phase clinical trials. Examples include in-depth evaluation of the target biology, including related family members, and use of a variety of newer tools such as systems biology approaches and genetic models. Also included is the evaluation of secondary targets identified in screening assays, which may require more in-depth analysis of potential AEs. Finally, there is more proactive use of clinically relevant and translatable toxicity monitoring in nonclinical toxicology studies. Application of newer technologies in nonclinical development programs, such as whole-genome sequencing or proteomic approaches, are commonly used in a retrospective fashion to investigate mechanisms and identify correlations when AEs are observed in nonclinical studies that were not expected.

Choosing the best available nonclinical model to characterize a specific AE in humans can be difficult. Experience with some animal models is limited, and these models may not predict rare clinical AEs well. This does not mean that the model does not have utility, but that information obtained from the model should be used appropriately in guiding clinical development. If the sensitivity and specificity of a particular model are not well characterized, a high false-positive rate could result, which may impede development of potentially beneficial therapies. Therefore, it is advisable to obtain clinical information to guide hypothesis development around what needs to be evaluated to develop meaningful nonclinical models as screens for specific clinical AEs.

The relevance of nonclinical models for determining optimal sequencing of drug combinations (e.g., one drug to be administered before or after another) is questionable. Pharmacology data can be useful to predict synergistic effects, but combination toxicology studies are rarely performed prospectively for

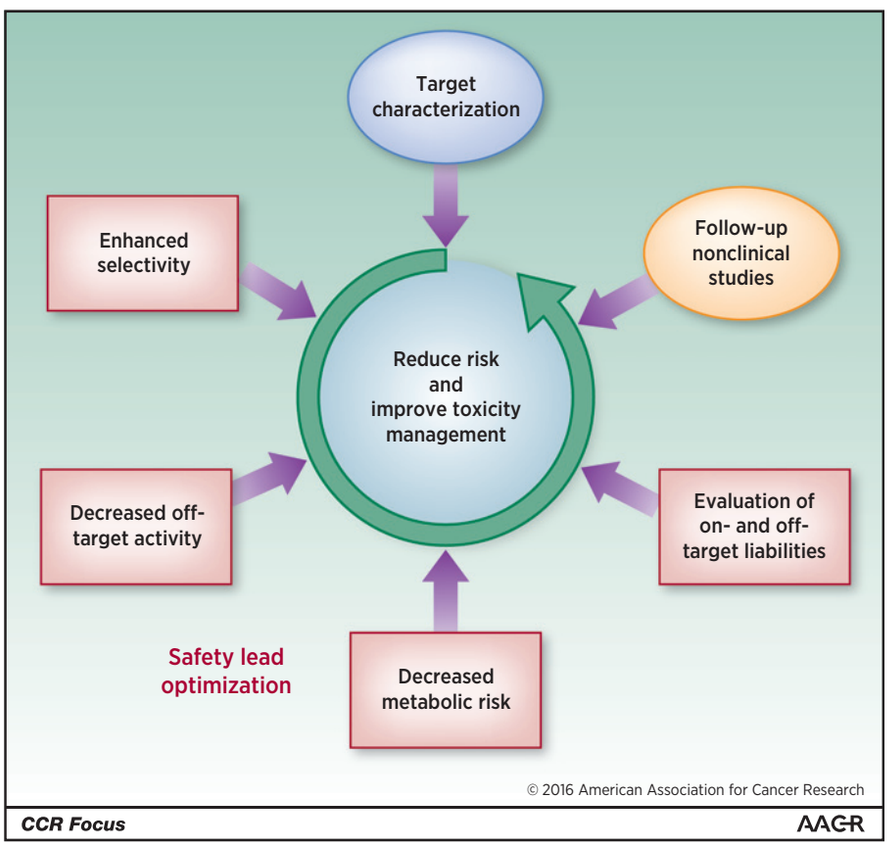

Figure 1.
Small-molecule anticancer drug lead optimization and risk evaluation.

combinations used to treat patients with cancer. Development of combination treatments often relies largely on clinical information obtained on each drug in the combination. If there is a concern about toxicity with a combination and uncertainty that it can be safely managed in the clinical trial, then a combination toxicology program may be informative and should be considered, especially if it could affect the clinical development of the combination.

A great deal of interest has been expressed in using nonclinical models to assess optimal dosing schedules for both efficacy and toxicity prior to clinical development. Some approaches involve attempts at relating fractional target inhibition (e.g., IC_{50} or IC_{90}) to specific pharmacodynamics or toxicity markers in different models. However, there is some reluctance to investigate alternative dosing schedules because it is often assumed that continuous maximal target inhibition will result in maximal antitumor activity, so the impact on efficacy of minimizing toxicity by altering the schedule is uncertain. In addition, the translatability of tumor-bearing mice is questionable for efficacy considerations and the nonclinical toxicity as it relates to the schedule may not translate clinically. Because translational pharmacodynamic tumor biomarkers to assess efficacy are rare, information on clinical efficacy must typically be derived from patients.

Examples of multidisciplinary approaches to successfully integrate information from nonclinical studies to inform clinical development encourage a focus on manageability rather than predictability. Knowledge about the reversibility of specific clinically relevant AEs acquired from focused nonclinical studies designed to address a specific concern may translate into patient protection against toxicity during dose escalation. Bayesian models are being adopted to connect prior information from nonclinical studies to clinical trial design. This approach requires a relatively high degree of confidence in the prior beliefs, which is an area in need of future improvement to better apply these models and further increase the efficiency of oncology clinical development.

Disclosure of Potential Conflicts of Interest

D.M. Dambach has ownership interest in Roche/Genentech. T.W. Jones has ownership interest in Eli Lilly and Company. R.J. Brennan has ownership interest in Sanofi. No potential conflicts of interest were disclosed by the other authors.

Authors' Contributions

Conception and design: D.M. Dambach, R. Pazdur, T.R. Palmby
Development of methodology: R. Pazdur
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): T.W. Jones, R.J. Brennan
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): T.W. Jones, R.J. Brennan, R. Pazdur
Writing, review, and/or revision of the manuscript: D.M. Dambach, N.E. Simpson, T.W. Jones, R.J. Brennan, R. Pazdur, T.R. Palmby
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): D.M. Dambach, T.W. Jones, R.J. Brennan

Acknowledgments

The authors thank William Kluwe of Novartis, Haw-Jyh Chiu and John Leighton of the FDA, and Douglas Keller of Sanofi for their contributions to this manuscript.

Received February 15, 2016; revised March 31, 2016; accepted April 11, 2016; published online June 1, 2016.

References

1. Minasian L, Rosen O, Auclair D, Rahman A, Pazdur R, Schilsky RL. Optimizing dosing of oncology drugs. Clin Pharmacol Ther 2014;96:572–9.
2. Uitdehaag JCM, de Roos JADM, van Doornmalen AM, Prinsen MBW, de Man J, Tanizawa Y, et al. Comparison of the cancer gene targeting and biochemical selectivities of all targeted kinase inhibitors approved for clinical use. PLoS One 2014;9:1–13.
3. Janne PA, Kim G, Shaw AT, Sridhara R, Pazdur R, McKee AE. Dose finding of small-molecule oncology drugs: optimization throughout the development life cycle. Clin Cancer Res 2016;22:2613–7.
4. Nie L, Rubin EH, Mehrotra N, Pinheiro J, Fernandes LL, Roy A, et al. Rendering the 3 + 3 design to rest: more efficient approaches to oncology dose-finding trials in the era of targeted therapy. Clin Cancer Res 2016;22:2623–9.
5. Bullock JM, Rahman A, Liu Q. Lessons learned: dose selection of small molecule–targeted oncology drugs. Clin Cancer Res 2016;22:2630–8.
6. Massett HA, Mishkin G, Rubinstein L, Ivy SP, Denicoff A, Godwin E, et al. Challenges facing early phase trials sponsored by the National Cancer Institute: an analysis of corrective action plans to improve accrual. Clin Cancer Res. In press.
7. USFDA approved protein kinase inhibitors compiled by Robert Roskoski Jr. [about 18 screens] [cited 2016 Feb 8]. Available from: http://www.brimr.org/PKI/PKIs.htm.
8. Keller DA, Brennan RJ, Leach KL. Clinical and nonclinical adverse effects of kinase inhibitors. In: Urban L, Patel VF, Vaz RJ, editors. Antitargets and drug safety. Weinheim (Germany): Wiley-VCH Verlag GmbH & Co. KGaA; 2015. p. 365–400.
9. Brown AP, Dunstan RW, Courtney CL, Criswell KA, Graziano MJ. Cutaneous lesions in the rat following administration of an irreversible inhibitor of erbB receptors, including the epidermal growth factor receptor. Toxicol Pathol 2008;36:410–9.
10. Vaz RJ, Patel VF. Cardiac side effects associated with kinase proteins and their signaling pathways. In: Urban L, Patel VF, Vaz RJ, editors. Antitargets and drug safety. Weinheim (Germany): Wiley-VCH Verlag GmbH & Co. KGaA; 2015. p. 401–12.
11. Anastassiadis T, Deacon SW, Devarajan K, Ma H, Peterson JR. Comprehensive assay of kinase catalytic activity reveals features of kinase inhibitor selectivity. Nat Biotechnol 2011;29:1039–45.
12. Olaharski AJ, Gonzaludo N, Bitter H, Goldstein D, Kirchner S, Uppal H, et al. Identification of a kinase profile that predicts chromosome damage induced by small molecule kinase inhibitors. PLoS Comput Biol 2009;5:e1000446.
13. Dambach DM, Misner D, Brock M, Fullerton A, Proctor W, Maher J, et al. Safety lead optimization and candidate identification: integrating new technologies into decision-making. Chem Res Toxicol 2016;29:452–72.

Clinical Cancer Research

Nonclinical Evaluations of Small-Molecule Oncology Drugs: Integration into Clinical Dose Optimization and Toxicity Management

Donna M. Dambach, Natalie E. Simpson, Thomas W. Jones, et al.

Clin Cancer Res 2016;22:2618-2622.