

ISSUE BRIEF

Conference on Clinical Cancer Research

OCTOBER 2010

ATGATCCATGGTCA
CTATCGAGTCGTCAACTA
ATGGCCATTCAGTCCCAATCATT
TCCCTGTATCGATCCGTCAATTC

CGCTGCTTAGTTTTGGATGATCCATGGTCA
CACCATTCCGACTATCGAGTCGTCAACTA
GGGCTACCCATGGCCATTCAATCATT
TTACATTCCCTGATCGATCCGTCAATTC

CGCTGCTTAGTTTTGGATGATCCATGGTCA
CACCATTCCGACTATCGAGTCGTCAACTA
GGGCTACCCATGGCCATTCAATCATT
TTACATTCCCTGATCGATCCGTCAATTC



CONTENTS

COMMENTARY

PAGES 6632 - 6633

Achieving the Goals of Effective, Safe, and Individualized Cancer Care

Samantha A. Roberts, Friends of Cancer Research

Erin K. Karnes, Engelberg Center for Healthcare Reform at Brookings Institute

Jeffrey D. Allen, Friends of Cancer Research

Joshua S. Benner, Engelberg Center for Healthcare Reform at Brookings Institute

Ellen V. Sigal, Friends of Cancer Research

Mark McClellan, Engelberg Center for Healthcare Reform at Brookings Institute

REPORT 1

PAGES 6634 - 6640

Adaptive Clinical Trial Designs for Simultaneous Testing of Matched Diagnostics and Therapeutics

Howard I. Scher, Memorial Sloan-Kettering Cancer Center

Shelley Fuld Nasso, Susan G. Komen for the Cure Advocacy Alliance

Eric H. Rubin, Merck Research Laboratories

Richard Simon, Biometric Research Branch, National Cancer Institute

REPORT 2

PAGES 6641 - 6645

Identification and Elucidation of the Biology of Adverse Events: The Challenges of Safety Assessment and Translational Medicine

Kenneth W. Turteltaub, Battelle Memorial Institute and Lawrence Livermore National Laboratory

Myrtle A. Davis, Pharmacology Branch, National Cancer Institute

Leigh Ann Burns-Naas, Drug Safety Research and Development, Pfizer Inc.

Michael P. Lawton, Drug Safety Research and Development, Pfizer Inc.

Adam M. Clark, FasterCures, Scientific and Federal Affairs

Jack A. Reynolds, AnaBios Corporation

REPORT 3

PAGES 6646 - 6650

Integrating Pain Metrics into Oncology Clinical Trials

Charles S. Cleeland, MD Anderson Cancer Center

Ann O'Mara, Palliative Care Research and Community Oncology Prevention, National Cancer Institute

Martin Zagari, Health Economics and Outcome Research, Amgen Inc.

Carole Baas, Physical Sciences in Oncology, National Cancer Institute

REPORT 4

PAGES 6651 - 6657

Using Patient-Initiated Study Participation in the Development of Evidence for Personalized Cancer Therapy

Laurie Fenton Ambrose, Lung Cancer Alliance

Jamie Freedman, GlaxoSmithKline

Kenneth Buetow, Bioinformatics and Information Technology, National Cancer Institute

Stephen Friend, Sage Bionetworks

Richard L. Schilsky, University of Chicago Comprehensive Cancer Center

Achieving the Goals of Effective, Safe, and Individualized Cancer Care

Samantha A. Roberts¹, Erin K. Karnes², Jeffrey D. Allen¹, Joshua S. Benner², Ellen V. Sigal¹, and Mark McClellan²

At the 2010 Conference on Clinical Cancer Research, held on October 20, 2010, in Washington, DC, co-convened by Friends of Cancer Research and the Engelberg Center for Health Care Reform at the Brookings Institution, participants explored 4 pressing challenges in the field. Articles summarizing the panel's recommendations on each of these topics are featured in this issue of *Clinical Cancer Research* (1–4).

Recent years have been marked by numerous important discoveries in clinical cancer research, bringing new therapeutic options to patients in great need. As these discoveries have been translated from bench to bedside, another important trend has emerged: Increasingly, stakeholders across public and private sectors have identified common goals in clinical cancer research and joined together to drive real progress toward safer, more effective, and more individualized cancer prevention, diagnosis, and treatment strategies.

This is a reflection of some important underlying trends in discovery and development. First, collaborative efforts are increasingly required to tackle the most pressing challenges facing clinical cancer research today, including big questions such as how we ensure that drug development is efficient; that resulting products are safe, effective, and as personalized as possible; and that regulatory and reimbursement policies facilitate and reward innovation that is valuable to patients. Second, information technology and other types of technical progress have made such collaborations easier. However, many obstacles remain.

Overcoming these obstacles has increasingly been the focus of our collaborative work on innovation in cancer care. October 2010 marked the third year that the Engelberg Center for Health Care Reform at the Brookings Institution and Friends of Cancer Research have convened members of the cancer clinical research community to discuss some of the most significant opportunities and challenges related to their shared goals. With support from the American Association for Cancer Research, the American Society of Clinical Oncology, and Susan G. Komen for the Cure, this

conference facilitates substantive multisector collaboration among leading representatives from government, academia, the patient community, and industry. By bringing all of these varied perspectives to the table and organizing expert panels to focus on discrete topics, these conferences have built a track record of producing results. Concepts presented at the 2009 Conference on Clinical Cancer Research led to the publication of 4 articles and stimulated progress on important topics, including ways to streamline data collection for supplemental indications of cancer treatments, use of progression-free survival as an endpoint in phase III oncology trials, development of drug combinations, and an accelerated pathway for approval of targeted cancer treatments (5–8).

One of the most exciting and promising aspects of modern cancer drug development is the potential to personalize treatments by developing drugs that inhibit specific molecular targets. Success stories of personalized cancer treatments include anti-epidermal growth factor receptor (EGFR) therapies, such as erlotinib, which target EGFR-overexpressing tumors, and anti-human epidermal growth factor receptor 2 (HER2) therapies, such as trastuzumab, which target HER2-overexpressing breast cancers. The key to developing such targeted therapies lies in identifying responsive patient populations and tumor characteristics. Due to the molecular heterogeneity of most tumors, however, this has proven extremely challenging. It is often not possible to identify predictive biomarkers before the start of phase III trials of anticancer therapeutics. As a result, many drugs fail to show a statistically robust treatment effect in these trials even though they might be very effective if used in the correct patients. It is clear that new approaches are needed to develop matched diagnostics and therapeutics. The first of the 4 articles developed from presentations at the 2010 conference uses castrate-resistant prostate cancer as a case study to present a potential adaptive phase III trial design in which an appropriate patient population is identified early in the trial, allowing the efficacy of a test therapeutic to be evaluated within that population later in the same trial (1). Such an approach maintains the rigorous statistical standards needed to evaluate drugs, is more consistent with our current knowledge of tumor biology, and can speed progress in getting effective anticancer treatments to responsive patients.

In addition to efficacy, drug safety is a major factor in regulatory decision making. Anticancer drug toxicity can be severe, leading to drug discontinuation or even death, and is often responsible for the failure of a drug candidate to receive marketing approval. Despite the importance of

Authors' Affiliations: ¹Friends of Cancer Research; and ²Engelberg Center for Health Care Reform, Brookings Institution, Washington, District of Columbia

Corresponding Author: Samantha A. Roberts, Friends of Cancer Research, Health Policy, 1800 M Street NW, Suite 1050 South, Washington, DC 20036. Phone: 202-944-6717; Fax: 202-944-6704; E-mail: sroberts@focr.org

doi: 10.1158/1078-0432.CCR-11-1104

©2011 American Association for Cancer Research.

safety information in drug development, current methods of preclinical toxicity testing are outdated and rely heavily on animal models that are often not predictive of adverse events in humans. The incorporation of modern systems biology technologies, such as genomics and proteomics, into drug safety testing could greatly improve our evaluation of new drug candidates. Such an approach could enable an understanding of the biology of adverse events, identify biomarkers predictive of specific adverse events, and potentially identify the patients most at risk for an adverse event. The second article presents 2 case studies that demonstrate the potential of systems biology approaches in toxicity testing (2). Integration of these modern techniques into drug safety testing could greatly improve the efficiency and accuracy of drug development.

As treatments for cancer become more effective and patients are living longer, the issue of pain in cancer becomes an increasingly important topic. Cancer-related pain is a frequently reported symptom that can have a significant and long-lasting impact on quality of life. This pain can result from the cancer itself or from the cancer treatment. Although pain is widely recognized as a significant issue in cancer, integrating pain metrics into clinical and regulatory decision making is challenging due in part to the subjective nature of pain. Furthermore, there is a high level of uncertainty regarding what kind of pain-related data the U.S. Food and Drug Administration (FDA) would find sufficient to contribute to labeling or approval decisions, making many sponsors reluctant to incorporate pain measurements into drug development programs. As a result, few clinical trials include pain palliation or pain prevention as either a primary or secondary endpoint. The third article explores the feasibility of developing objective standards for pain measurement and identifies the need to develop new tools to measure pain (3). Several methodologic challenges need to be addressed in the form of an FDA guidance to facilitate the measurement of pain in oncology clinical trials. Including such measurements and incorporating the resulting information into drug labels would greatly benefit the cancer community, as patients could live not only longer, but happier and more productive lives.

Approval and labeling of new cancer drugs by the FDA relies upon safety and efficacy data from population-based

trials. However, data suggest that an average of only 1 in 4 patients receiving an approved cancer drug regimen significantly benefit, whereas the remainder of patients experience little to no benefit and may experience potentially toxic side effects. Although this clearly points to a need for a better understanding of factors associated with treatment response, generating timely and actionable evidence of this sort through prospective clinical trials can be difficult. The fourth and final article uses a case study in non-small cell lung cancer to examine the feasibility of directly engaging patients to participate in a proposed prospective study of molecular determinants of treatment response (4). With a focus on previously marketed drugs, the goal of such a study would be to inform labeling changes and clinical practice such that cancer patients receive treatments that are more personalized and therefore more likely to result in benefit rather than harm.

Each of these articles marks an early but significant step toward resolving real barriers to more effective, safe, and individualized cancer care. More importantly, as a collection, they illustrate the potential for innovation and collaboration within the cancer community and give reason for optimism that these goals can be achieved. To date, the ideas presented have been conceived of and refined by groups of collaborating stakeholders, with the benefit of input from other stakeholders as part of the annual conference. In order for these ideas to continue to come to fruition in a meaningful and practical way, more steps to resolve barriers will be required. In turn, these efforts will require collaboration reflecting active engagement of stakeholders, ranging from patients and consumers to scientists and regulatory officials. We hope to keep working together to make that happen.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

The authors gratefully acknowledge the drafting support provided by Ed Walters of the Engelberg Center for Health Care Reform at The Brookings Institution.

Received May 6, 2011; revised July 5, 2011; accepted August 24, 2011; published online November 1, 2011.

References

1. Scher HI, Fuld Nasso S, Rubin EH, Simon R. Adaptive clinical trial designs for simultaneous testing of matched diagnostics and therapeutics. *Clin Cancer Res* 2011;17:6634–40.
2. Turteltaub KW, Davis MA, Burns-Naas LA, Lawton MP, Clark AM, Reynolds JA. Identification and elucidation of the biology of adverse events: the challenges of safety assessment and translational medicine. *Clin Cancer Res* 2011;17:6641–5.
3. Cleeland CS, O'Mara A, Zagari M, Baas C. Integrating pain metrics into oncology clinical trials. *Clin Cancer Res* 2011;17:6646–50.
4. Ambrose LF, Freedman J, Buetow K, Friend S, Schilsky RL. Using patient-initiated study participation in the development of evidence for personalized cancer therapy. *Clin Cancer Res* 2011;17:6651–7.
5. Abrams J, Erwin R, Fyfe G, Schilsky RL. Data submission standards and evidence requirements. *Oncologist* 2010;15:488–91.
6. Amit O, Bushnell W, Dodd L, Roach N, Sargent D. Blinded independent central review of the progression-free survival endpoint. *Oncologist* 2010;15:492–5.
7. Clark A, Ellis M, Erlichman C, Lutzker S, Zwiebel J. Development of rational drug combinations with investigational targeted agents. *Oncologist* 2010;15:496–9.
8. McClellan M, Benner J, Schilsky R, Epstein D, Woosley R, Friend S, et al. An accelerated pathway for targeted cancer treatments. *Nat Rev Drug Discov* 2011;10:79–80.

Adaptive Clinical Trial Designs for Simultaneous Testing of Matched Diagnostics and Therapeutics

Howard I. Scher¹, Shelley Fuld Nasso², Eric H. Rubin³, and Richard Simon⁴

Abstract

A critical challenge in the development of new molecularly targeted anticancer drugs is the identification of predictive biomarkers and the concurrent development of diagnostics for these biomarkers. Developing matched diagnostics and therapeutics will require new clinical trial designs and methods of data analysis. The use of adaptive design in phase III trials may offer new opportunities for matched diagnosis and treatment because the size of the trial can allow for subpopulation analysis. We present an adaptive phase III trial design that can identify a suitable target population during the early course of the trial, enabling the efficacy of an experimental therapeutic to be evaluated within the target population as a later part of the same trial. The use of such an adaptive approach to clinical trial design has the potential to greatly improve the field of oncology and facilitate the development of personalized medicine. *Clin Cancer Res*; 17(21); 6634–40. ©2011 AACR.

Introductory Note

At the 2010 Conference on Clinical Cancer Research, convened by Friends of Cancer Research and the Engelberg Center for Health Care Reform at the Brookings Institution, participants explored 4 pressing challenges in the field. Articles summarizing the panel's recommendations on each of these topics are featured in this issue of *Clinical Cancer Research* (1–4).

Key Role of Companion Diagnostics in Oncology Drug Development

Nearly all cancer drugs being developed today are designed to inhibit molecular targets that have been identified as being dysregulated in human tumors. Genomics has established that the dysregulated pathways and mutated genes in tumors originating in a particular primary site are highly variable. To optimally evaluate and utilize a targeted approach requires the concurrent development of diagnostics that enable the identification of those tumors that are most likely to be sensitive to the anticancer effects of a particular drug or drug combination. The reality of code-

veloping a matched diagnostic and therapeutic has profound implications for the clinical trial designs used in drug development. Trials of cytotoxic drugs typically enroll unselected patients at a particular point in the continuum of a disease in the hope that the response of tumors that are sensitive to the treatment will be sufficient to show benefit for the population as a whole. Although this approach may lead to broad labeling indications, it also exposes patients with nonsensitive tumors to unnecessary toxicities and increases the possibility of discarding a drug that may dramatically benefit a subset of patients. Consequently, this strategy is not viable for molecularly targeted agents, in which the activity is likely to be restricted and determined more by the genomic alteration(s) within a tumor at the time treatment is being considered than by the primary site in which the tumor originated. The use of anatomically based (i.e., primary site of disease), "all comers" approaches to develop targeted approaches has typically led to failure in phase III studies, or demonstration of "success" based on statistically significant but clinically questionable benefits (5).

Although developing the right drug for a specific patient has great value to the individual and is critical for controlling the costs of health care, it dramatically increases the complexity of the drug development process. For many drugs, the complexities of identifying a predictive biomarker and the practical complexities of developing analytically valid diagnostic tests for the biomarker are grossly underestimated. Knowing when to start the development of the diagnostic is also an issue, particularly when the effectiveness of the drug in any population is uncertain. Developing the right drug for the right subset of patients requires new clinical trial designs and new paradigms of data analysis.

Efforts to codevelop a matched diagnostic and therapeutic face other challenges as well. Even with extensive

Authors' Affiliations: ¹Department of Medicine, Genitourinary Oncology Service, Memorial Sloan-Kettering Cancer Center, and Weill Cornell Medical College, New York, New York; ²Susan G. Komen for the Cure Advocacy Alliance, Washington, District of Columbia; ³Oncology Clinical Research, Merck Research Laboratories, Whitehouse Station, New Jersey; and ⁴Biometric Research Branch, National Cancer Institute, Bethesda, Maryland

Corresponding Author: Howard I. Scher, Department of Medicine, Sidney Kimmel Center for Prostate and Urologic Cancers, Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, New York, NY. Phone: 646-422-4323; Fax: 212-988-0851; E-mail: scherh@mskcc.org

doi: 10.1158/1078-0432.CCR-11-1105

©2011 American Association for Cancer Research.

preclinical investigations, it is often difficult to identify a predictive biomarker and evaluate it in the phase I trial of a given drug. As a result, it becomes necessary to have a test for, and develop preliminary data showing, the predictive power of the candidate marker in the context of phase II investigations, so that properly focused phase III trials can be designed, conducted, and completed. In the unusual case in which a single predictive biomarker has been identified and a validated assay has been developed prior to the start of the phase III trial, targeted enrichment designs and stratification designs can be used (6, 7). For example, in the targeted enrichment design used in the development of trastuzumab, only "marker-positive" patients were included. In the stratification design, patients were not excluded on the basis of marker status, but the size of the trial was adequately powered for the anticipated frequency of marker-positive patients and the overall 5% type I error allocated between the comparison of treatments overall and the comparison within marker-positive patients. Adaptive phase II designs, such as the design recently used in the Biomarker-Integrated Approaches of Targeted Therapy for Lung Cancer Elimination (BATTLE) clinical trial in non-small cell lung cancer (8) and the Investigation of Serial Studies to Predict Your Therapeutic Response with Imaging and Molecular Analysis 2 (I-SPY2) trial in breast cancer (9), are useful for identifying the most promising predictive biomarker in phase II development, but they require large sample sizes. The outcome-adaptive randomization weights used in the BATTLE study design complicate the interpretation of results, and the determination of whether they improve efficiency has not been established (10). The BATTLE study did, however, show the feasibility of a biopsy-based, hypothesis-driven biomarker trial, and the follow-up phase II trial, BATTLE 2, will test the predictive value of the gene signatures prospectively. In 2010, the U.S. Food and Drug Administration (FDA) issued a draft Guidance to Industry on Adaptive Design Clinical Trials for Drugs and Biologics (11).

Due to the complexity of cancer biology, it is often not possible to firmly establish the biomarker(s) most likely to predict sensitivity to a particular drug or class of drug by the time pivotal phase III trials are set to begin. Recently, however, several adaptive clinical trial designs have been published that show how to design the trial(s) so that the most suitable target population of patients is adaptively identified during the trial and the effectiveness of the drug is evaluated in that population in a rigorously defined and statistically valid manner (12–14). For example, when the biomarker assay has been validated and standardized, and performance characteristics are known, the adaptive signature design (12) and cross-validated adaptive signature design (14) are carefully crafted adaptive phase III non-Bayesian approaches that preserve the desired type I error rate while identifying an optimal target population. Neither design results in a change in randomization weights or in eligibility criteria (both of which could require statistical adjustments to avoid introduction of bias), which makes them better suited for licensing registration trials than the

Bayesian methods used in the phase II BATTLE trial. Interestingly, although the FDA Guidance on adaptive trial designs acknowledges that a Bayesian framework can be useful for planning purposes to evaluate model assumptions and decision criteria, they recommend that the study design be planned in a framework to control the overall study type I error rate (11).

These are, however, complex designs that have not been tested in practice. Challenges to the use of these designs are that the treatment comparisons can only be conducted after completion of the study, that the developed predictive signature may be based on a combination of factors with unclear biological meaning, and that it may be difficult to interpret the results if there are imbalances in other baseline prognostic factors between treatment arms in the marker-positive subgroup. Although these designs are in some ways conservative, they are nevertheless dramatically different from the kinds of designs used for the vast majority of clinical trials being conducted today.

Here, we describe how adaptive methods can be used for indication determination in a manner that provides the level of confidence in conclusions that we expect from phase III registration trials and in a manner consistent with the FDA Guidance on adaptive design. The current draft of the Guidance defines a clinical study using an adaptive design as one that "includes a prospectively planned opportunity for modification of one or more specified aspects of the study design and hypotheses based on analysis of data (usually interim data) from subjects in the study. Analyses of the accumulating study data are carried out at prospectively planned time points within the study, can be performed in a fully blinded manner or in an unblinded manner, and can occur with or without formal statistical hypothesis testing" (11). In some cases, adaptive designs require fewer patients but much more upfront planning.

To illustrate the careful planning necessary for proper use of adaptive methods in this context, a detailed illustration of the use of the adaptive signature design of Freidlin and Simon (12) is provided. The approach includes 3 components: (i) a statistically valid identification, based on the first stage of the trial, of the subset of patients who are most likely to benefit from the new agent; (ii) a properly powered test of overall treatment effect at the end of the trial with all randomized patients; and (iii) a test of treatment effect for the subset identified in the first stage but only with patients randomized in the remainder of the trial. The design is adaptive in the sense of the FDA Guidance because the primary plan for the final analysis is influenced by the results of the trial. The adaptive signature design (12) and the more recently published cross-validated adaptive signature design (14) were developed for use in gene expression profiling settings when there are enormous numbers of candidate measurements that can be combined to provide a classifier of which patients are likely (or unlikely) to benefit from a new treatment relative to a control regimen. These designs can be used much more broadly, however, regardless of the candidate predictors, and are discussed in greater generality by Simon (15).

Background

Application of the original adaptive signature paradigm to a real clinical development setting has many complexities, which will be illustrated here for castration-resistant prostate cancer (CRPC), the advanced, lethal form of the disease. Molecular profiling studies show that reactivation of androgen receptor (AR) function is a consistent feature of CRPC (16), in part, through AR overexpression and overexpression of androgen-synthetic enzymes leading to increased intratumoral androgens (17, 18). The clinical significance of these findings has been validated in trials of abiraterone acetate, an inhibitor of androgen synthesis in the testis, adrenal gland, and tumor (19, 20), and MDV3100, a novel AR antagonist selected for activity in prostate cancer model systems with overexpressed AR (21). Abiraterone was recently shown to confer a survival benefit after chemotherapy in patients with CRPC and is now approved by the FDA for this indication (22). MDV3100 has shown activity comparable with that of abiraterone in postchemotherapy CRPC (23), and a phase III registration trial for this population has been fully accrued. Noteworthy in the trials of both agents was the similarity of response in matched patient populations, which ranged from dramatic prostate-specific antigen declines with durable radiographic control in some to intrinsic resistance in others, suggesting the presence of predictive biomarkers in tumors. A number of other agents targeting different points in the AR signaling pathway are currently in development (24), and although predictive biomarkers of sensitivity have been postulated, none has warranted the development of a validated assay or begun the formal process of clinical qualification (11). As biotechnology continues to provide the tools to characterize tumors at the genomic scale and basepair resolution, it is likely that relevant predictive markers will be identified.

Further adding to the complexity of developing drugs for CRPC is the recent demonstration that 3 additional agents, with different mechanisms of action—Provenge (sipuleucel-T; Dendreon), Jevtana (cabazitaxel; Sanofi), and Alpharadin (radium-223; Bayer)—also confer a survival benefit (25–28). With the expanded armamentarium, it will become increasingly more difficult to show the survival benefit of an agent without enriching for patients most likely to respond. Identification and development of validated assays for predictive biomarkers of sensitivity are likely to play a significant role in the ultimate approval of these and future therapies.

The Adaptive Signature Approach

The adaptive signature approach provides for a final analysis consisting of 2 parts: first, outcomes for all patients randomized to receive the new drug will be compared with outcomes for all patients randomized to receive the control. If this comparison is significant at a more stringent than usual 2-sided significance level of α_0 , then the new drug is considered broadly effective. Otherwise, a single subset analysis is conducted. The patients in the trial are randomly

partitioned into a training set and a validation set. The training set is used to develop a "classifier" that identifies the subset of patients who seem to benefit from the new treatment compared with the control. This classifier can be based on a combination of all the clinical and biomarker candidate variables measured before treatment. When this single classifier is completely specified using only the training set, it is applied to classify patients in the validation set with respect to whether they are predicted to benefit from the new treatment. Outcomes for patients in this subset of the validation set who were randomized to receive the new treatment are compared with outcomes for patients in this subset who were randomized to receive the control regimen. Only the patients in the validation set are used for this comparison. Because the training set was used to develop the classifier, it cannot be used to evaluate it. If this difference is significant at the reduced 2-sided significance level of $0.05-\alpha_0$, then the new treatment is considered effective for the subset of patients defined by the classifier developed in the training set. The cross-validated adaptive signature design is a more statistically powerful version of this approach (14). However, even application of the original adaptive signature paradigm to a real clinical development setting has many complexities, an example of which will be illustrated here for CRPC.

A Phase III Adaptive Trial Design

The design we describe for the clinical trial is an application of the adaptive signature approach of Freidlin and Simon (12) and could be used with many more candidate predictive markers. This design is appropriate for settings in which (unlike the case of HER2 overexpression and trastuzumab development) there is not yet a single predictive biomarker candidate, in which there is high confidence by the time of initiation of the phase III clinical trial of the drug.

Eligible patients are individuals with progressive CRPC for whom a targeted therapeutic approach is being developed, and for whom tumor material is available. The requirement of sufficient tumor for analysis at entry ensures near-complete ascertainment of the biomarker or biomarker panel. Formalin-fixed, paraffin-embedded (FFPE) samples that were obtained either as part of the routine testing to establish diagnosis or during radical prostatectomy are typically the most readily available; therefore, for practical reasons, assays that can be conducted on FFPE specimens are preferred. For biomarkers present at a higher frequency in progressive metastatic CRPC (relative to primary tumors that are noncastrate), a repeat biopsy of the metastatic lesion will be required, whereas for those assays that can be conducted reliably only in frozen tumor, a repeat biopsy of either metastatic or primary tumor immediately before trial entry will be necessary. Tumor specimens are stored for future assay. After confirmation that sufficient tumor is available for analysis, a patient is randomized to treatment with compound X or control. A key aspect of this design is that neither the predictive biomarker nor the

analytically validated tests are needed until the time of the final analysis of the trial.

The primary endpoint for the study is overall survival, which is the primary regulatory endpoint for new drug approval for CRPC. A total of 935 patients will be accrued, and the final analysis will be conducted when there are 700 total deaths. This will provide approximately 90% statistical power for detecting a 25% reduction in hazard of death for compound *X* relative to control at a 2-sided statistical significance level of 1%. The remaining 4% of type I error will be used for evaluating the statistical significance of treatment effect on survival in the adaptively defined biomarker subset that is anticipated to derive greater benefit than the population as a whole. This will provide approximately 80% statistical power for detecting a 37% reduction in the hazard of death in the adaptively defined subset of the validation set, which consists of only 33% of the validation set, as described in more detail below. By splitting the traditional 5% significance threshold into a portion to be used for the overall comparison and a portion to be used for the comparison within the subset, the type I error rate of the trial is preserved at 5%.

The type I error rate of 5% can be partitioned into a part for the overall analysis and a part for the subset analysis in a variety of ways. One could attempt to optimize the split to minimize the total sample size subject to constraints on the statistical power for both the overall analysis and the subset analysis. We have not attempted such an optimization. We have allocated most of the 5% to the subset analysis because the power of the subset analysis drives the overall sample size, particularly when a minority of patients benefit from the new treatment. By taking into account the correlation between the 2 analyses, less stringent significance levels could be used (29).

The final analysis will be conducted in the following manner. A log-rank test will be used to compare survival times in the 2 treatment arms for all randomized patients. If the 2-sided significance level is less than 0.01 and favors compound *X*, then compound *X* will be considered effective for the randomized population as a whole. If not, then the following analysis will be conducted with the fallback design of the adaptive signature approach developed by Freidlin and Simon (12).

A predictive classifier $P(B_1, B_2, B_3, B_4)$ will be developed that identifies whether a patient with biomarker values B_1 , B_2 , B_3 , and B_4 (each representing the result of a specific validated assay) is likely to benefit from drug *X* compared with control *C*. For the purpose of illustration, we have arbitrarily specified 4 individual markers that can be used for building the classifier. The number is arbitrary as long as the markers and the algorithm for building the classifier with the candidate markers are specified before the data are examined and as long as an analytically validated assay is available for measuring each marker. The value of the adaptive signature design is greatest when the number of candidate markers is large. The P classifier will be developed using a randomly selected training set of patients consisting of 33% of the cases. The split proportion of 33% of the

patients for development and training of the classifier and 67% for evaluation of the classifier is somewhat arbitrary but influences the ability to develop a good classifier and to adequately compare the treatment in the subset of the validation set determined by the classifier. Dobbin and Simon (30) have studied the optimal splitting of data sets into a training set and a validation set for prognostic classifiers, but similar studies have not been reported for predictive classifiers as used in the adaptive signature design. We believe that a training set consisting of approximately 233 events should be adequate for developing a predictive classifier in which accuracy is close to that of the optimal classifier that could be developed with an infinite-sized training set, but a quantitative evaluation of this along the lines described by Dobbin and Simon should be pursued. Reducing the size of the validation set further constrains the statistical power of the subset analysis, as shown below in the paragraph describing how the power for the subset analysis in the validation set drives the total size of the study. The advantage of the more recently developed cross-validated adaptive signature design is that a fixed training-validation split is not required. (14)

The algorithm for developing the classifier is described in the Appendix, which follows the Discussion section. The value of the classifier function $C(B_1, B_2, B_3, B_4)$ equals 1 if the patient with those biomarker values is likely to benefit from *X*, and equals 0 otherwise. The set IND of combinations of biomarker values (B_1, B_2, B_3, B_4) for which the classifier equals 1 is the indication for treatment *X* should the subset analysis be statistically significant. As part of the final analysis, this indication will be described graphically, analytically, by decision tree, and as a classification function.

The training set data are extensively analyzed to develop a single completely specified classifier. Predictive classifier development is different from traditional subset analysis. Although the development algorithm may involve evaluation of subsets determined by single variables, a classifier must be developed that integrates all such information into a single function of all the baseline variables to predict whether a patient will benefit from receiving the new treatment relative to the control. Although a large body of literature exists on prognostic signatures, very little literature is available on predictive 2-treatment classifiers. A single completely specified classifier should be developed with the training data. If multiple classifiers were developed, they would have to be evaluated in the validation set and that would require additional portions of the type I error to be allocated to evaluate them.

The estimated improvement in survival for *X* versus *C* in the indicated population IND will be estimated by classifying each patient in the trial who was not included in the training set used to develop the classifier. Let S denote the set of patients in this "test set" classified as likely to benefit from *X* using $C(B_1, B_2, B_3, B_4)$. Kaplan-Meier survival curves will be computed for the patients in S who received *X* and for the patients in S who received *C*. The difference between these 2 survival curves will be summarized with a log-rank statistic (LR) and a log hazard ratio (LHR) and a 95%

confidence interval for LHR. If the log-rank statistic LR is significant at the 4% level of the χ^2 distribution with one degree of freedom and the HR of X versus C is less than 1, then the treatment X will be considered effective in improving survival of patients with an indication specified by the set IND defined based on the classifier C (B1, B2, B3, B4) as described above.

The statistical power of the biomarker-specified subset analysis depends on the proportion of patients who are included in the adaptively defined subset S. To have 80% power for detecting a 37% reduction in the hazard of death for X versus C (a reasonable target effect size given historical results with predictive biomarker-based treatments such as trastuzumab), approximately 157 deaths are required in the classifier-positive subset of the test set of patients (i.e., patients not used for developing the classifier). If one third of patients are classifier positive, then 471 total deaths are required in the test set. The test set will contain about two thirds of the patients and events. The total number of deaths at the time of final analysis will be 700, and hence this power target should be achievable.

Discussion

For the goal of developing the right drug for the right patient to become more than a cliché, sponsors, investigators, and regulators must recognize that some of the conventional wisdom used to guide clinical trial design and analysis in the era of broadly targeted cytotoxic agents is no longer appropriate. Indeed, the continued use of traditional clinical trial designs is likely to hamper the development of new drugs that are highly effective for molecularly well-defined subsets of patients.

The use of conventional, primary site-based approaches to develop targeted cancer therapeutics is in many cases not consistent with our knowledge of the underlying biology of a tumor, exposes patients to toxic drugs from which they are not expected to benefit, and may result in long delays for the approval and ultimately the availability of drugs that offer substantial benefit to molecularly characterized subsets of patients. Clearly, in this new era, issues previously considered to be standard, such as the role of subset analysis, the role of stratification, the need to have broad eligibility criteria, and the use of adaptive methods, must be critically reexamined. However, new methods for clinical trial design and analysis must be no less rigorous than conventional designs in their use of randomized controls, clinically meaningful endpoints, and protection against type I error.

Methods for adaptive characterization and validation of the patients most likely to benefit from a new treatment in phase III oncology trials have been developed in recent years (12–14). The specific designs are adaptive in distinct ways, but most have focused on intratrial modification of the number of patients to be included (sample size reestimation) or modification of the randomization weights (response-adaptive). Controversies with these designs include the question of whether adaptive sample size reestimation is more effective than traditional

sequential analysis methods, whether response-adaptive methods provide statistical analyses that are robust to time trends in unmeasured prognostic factors, and whether response-adaptive methods improve efficiency (10). As a result, response-adaptive designs are rarely used in phase III clinical trials.

A more promising area is the adaptive characterization of patients enrolled in phase III trials who are most likely (or least likely) to benefit from a new treatment. This adaptive determination of the treatment indication represents a paradigm shift in phase III clinical trial design with the potential for a major impact on oncology drug development and a major benefit to patients. However, there is a need for dialogue among academic investigators, government and industry sponsors, and regulators on how best to use this methodology. It was for this reason that this area of adaptive clinical trial design was chosen for focus by the members of the adaptive design panel of the 2010 Conference on Clinical Cancer Research.

Use of adaptive methods to identify the patients who are most likely or least likely to benefit from a new regimen requires substantial prospective planning. The methods cannot be used reliably in an exploratory *post hoc* manner. In fact, if done improperly they can introduce bias and risk "disqualifying" a trial as adaptive in the view of the FDA.

The following are some of the key features of the clinical trial we designed:

1. Use of an acceptable regulatory endpoint such as overall survival as the primary endpoint for final analysis.
2. Use of a randomized design with an appropriate control arm.
3. Obtaining tumor specimens prior to randomization for all patients registered on the trial. Tumor assays may be conducted at a later time, but prior to data analysis, if the analytically validated tests are not available when the clinical trial is initiated.
4. Use of an intermediate endpoint for interim futility analysis is considered necessary but not sufficient to ensure a treatment effect on the primary endpoint, even though it is not a validated surrogate of the primary endpoint.
5. Use of analytically validated tests for measuring all candidate predictive biomarkers.
6. Prespecification of the algorithm to be used for developing the classifier in the training set, and prespecification of how the validation analysis will be conducted.
7. Adequately powering the clinical trial for validation of a substantial treatment effect in the adaptively identified subset.
8. At the time of final analysis the patients are randomly partitioned into a portion (e.g., one third) for training a classifier that identifies which patients are most and least likely to benefit from the new treatment, and a portion (e.g., two thirds) for validating that classifier. The full set of patients is used for the overall

comparison of the new treatment with the control, but only the validation set is used for the comparison of new treatment versus control in the adaptively determined subset. The type I error is allocated among those 2 comparisons.

Oncology therapeutics development is in an era of fundamental change. The discoveries of both the molecular basis for human cancers and the heterogeneity of cancer provide a great opportunity to develop more effective treatments and to properly deliver them to the right patients. Many challenges remain to be addressed, and some familiar paradigms require reevaluation. What does not change is the need for clinical trials that are fundamentally science based, statistically sound, and responsive to the urgency for reducing mortality and morbidity from cancer.

Appendix

A wide variety of classifier development algorithms are possible. The algorithm should be described completely in the protocol. For the study being illustrated here, the classifier will be developed by the following algorithm:

A proportional hazards model will be fit to the data for the combined treatment X and control group. Denote this model by

$$\log[\lambda(t, B1, B2, B3, B4, v)/\lambda_0(t)] = \delta v + \beta_1 B1 + \beta_2 B2 + \beta_3 B3 + \beta_4 B4 + v(\gamma_1 B1 + \gamma_2 B2 + \gamma_3 B3 + \gamma_4 B4)$$

where v is a binary treatment indicator ($v = 1$ for X , $v = 0$ for C), δ is the regression coefficient that represents the main effect of treatment on survival, the β s reflect the prognostic effects of the biomarkers, and the γ s are the interaction effects that represent the predictive effects of the biomarkers. The left-hand side of the equation represents the log hazard relative to the baseline hazard. The markers will only be binary if a cut point is predefined based on preliminary data. Otherwise, no cut point will be imposed on the modeled values.

For a patient with biomarker values ($B1, B2, B3, B4$), the LHR if the patient receives treatment X minus the LHR if the patient receives the control C is

$$\Delta(B1, B2, B3, B4) = \delta + \gamma_1 B1 + \gamma_2 B2 + \gamma_3 B3 + \gamma_4 B4$$

By fitting the model to the data, we obtain estimates of the regression coefficients and a covariance matrix for these estimates. Hence, for any vector of biomarker values, we can compute $\hat{\Delta}(B1, B2, B3, B4)$ in which the regression coefficients are replaced by their estimates, and we can compute the variance $V[\hat{\Delta}(B1, B2, B3, B4)]$. A

binary classifier will be defined by

$$C(B1, B2, B3, B4) = 1 \text{ if } \hat{\Delta}(B1, B2, B3, B4) / \sqrt{V[\hat{\Delta}(B1, B2, B3, B4)]} \leq c$$

The patient is classified as likely to benefit from X if the standardized LHR of X relative to C is less than or equal to constant c . The constant will be determined by 10-fold cross-validation within the training set to maximize the log-rank statistic for treatment effect within the training set patients classified as likely to benefit from X . Application of this algorithm to the training data provides a completely specified classifier that can be used to classify each of the patients in the validation set. Each patient in the validation set has biomarker values $B1, B2, B3$, and $B4$. By plugging in these values to the classifier, the patient is classified as likely or unlikely to benefit from X relative to C . The patients in the validation set who are classified as likely to benefit from X are the subset to be analyzed. In that subset, outcomes for patients who received X are compared with outcomes for those who received the control C .

The classifier illustrated here is based on a proportional hazards regression analysis of 4 biomarker values. Alternatively, variable selection strategies could be used to include only variables that seem informative for distinguishing outcome on X from outcome on C . When the number of candidate variables is large, variable selection is essential. It should be recognized, however, that the objective is to accurately classify patients as to whether they will benefit from X , not to document with statistical significance the importance of individual variables.

Disclosure of Potential Conflicts of Interest

Dr. Scher has received research funding from Ortho Biotech Oncology Research and Development and Medivation. He has served as a consultant or in an advisory role for Ortho Biotech Oncology Research and Development and Medivation. No potential conflicts of interest were disclosed by the other authors.

Acknowledgments

Rajeshwari Sridhara contributed on behalf of the U.S. FDA. The authors thank Samantha Roberts for a critical review of the manuscript and Margaret McPartland for editorial assistance.

Grant Support

This work was supported in part by the MSKCC SPORE in Prostate Cancer (P50 CA92629), the Department of Defense Prostate Cancer Research Program (PC051382), The Research and Therapeutics Program for Prostate Cancer, and The Prostate Cancer Foundation.

Received May 6, 2011; revised September 22, 2011; accepted September 22, 2011; published online November 1, 2011.

References

1. Roberts SA, Karnes EK, Allen JD, Benner JS, Sigal EV, McClellan M. Achieving the goals of effective, safe, and individualized cancer care. *Clin Cancer Res* 2011;17:6632-3.
2. Turteltaub KW, Davis MA, Burns-Naas LA, Lawton MP, Clark AM, Reynolds JA. Identification and elucidation of the biology of adverse events: the challenges of safety assessment

- and translational medicine. *Clin Cancer Res* 2011;17:6641–5.
3. Cleeland CS, O'Mara A, Zagari M, Baas C. Integrating pain metrics into oncology clinical trials. *Clin Cancer Res* 2011;17:6646–50.
 4. Ambrose LF, Freedman J, Buetow K, Friend S, Schilsky RL. Using patient-initiated study participation in the development of evidence for personalized cancer therapy. *Clin Cancer Res* 2011;17:6651–7.
 5. Rubin EH, Anderson KM, Gause CK. The BATTLE trial: a bold step toward improving the efficiency of biomarker-based drug development. *Cancer Discov* 2011;1:17–20.
 6. Simon R. The use of genomics in clinical trial design. *Clin Cancer Res* 2008;14:5984–93.
 7. Mandrekar SJ, Sargent DJ. Clinical trial designs for predictive biomarker validation: theoretical considerations and practical challenges. *J Clin Oncol* 2009;27:4027–34.
 8. Printz C. BATTLE to personalize lung cancer treatment. Novel clinical trial design and tissue gathering procedures drive biomarker discovery. *Cancer* 2010;116:3307–8.
 9. Barker AD, Sigman CC, Kelloff GJ, Hylton NM, Berry DA, Esserman LJ. I-SPY 2: an adaptive breast cancer trial design in the setting of neoadjuvant chemotherapy. *Clin Pharmacol Ther* 2009;86:97–100.
 10. Korn EL, Freidlin B. Outcomeadaptive randomization: is it useful? *J Clin Oncol* 2011;29:771–6.
 11. United States Food and Drug Administration: 2010 draft guidance for industry on adaptive design clinical trials for drugs and biologics. Available from: <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm201790.pdf>.
 12. Freidlin B, Simon R. Adaptive signature design: an adaptive clinical trial design for generating and prospectively testing a gene expression signature for sensitive patients. *Clin Cancer Res* 2005;11:7872–8.
 13. Jiang W, Freidlin B, Simon R. Biomarker-adaptive threshold design: a procedure for evaluating treatment with possible biomarker-defined subset effect. *J Natl Cancer Inst* 2007;99:1036–43.
 14. Freidlin B, Jiang W, Simon R. The cross-validated adaptive signature design. *Clin Cancer Res* 2010;16:691–8.
 15. Simon R. Clinical trials for predictive medicine: new challenges and paradigms. *Clin Trials* 2010;7:516–24.
 16. Scher HI, Sawyers C. Biology of progressive castration resistant prostate cancer: directed therapies targeting the androgen receptor signaling axis. *J Clin Oncol* 2005;23:8253–61.
 17. Holzbeierlein J, Lal P, LaTulippe E, Smith A, Satagopan J, Zhang L, Ryan C, et al. Gene expression analysis of human prostate carcinoma during hormonal therapy identifies androgen-responsive genes and mechanisms of therapy resistance. *Am J Pathol* 2004;164:217–27.
 18. Stanbrough M, Bubley GJ, Ross K, Golub TR, Rubin MA, Penning TM, et al. Increased expression of genes converting adrenal androgens to testosterone in androgen-independent prostate cancer. *Cancer Res* 2006;66:2815–25.
 19. Danila DC, Morris MJ, de Bono JS, Ryan CJ, Denmeade SR, Smith MR, et al. Phase II multicenter study of abiraterone acetate plus prednisone therapy in patients with docetaxel-treated castration-resistant prostate cancer. *J Clin Oncol* 2010;28:1496–501.
 20. Reid AH, Attard G, Danila DC, Oommen NB, Olmos D, Fong PC, et al. Significant and sustained antitumor activity in post-docetaxel, castration-resistant prostate cancer with the cyp17 inhibitor abiraterone acetate. *J Clin Oncol* 2010;28:1489–95.
 21. Tran C, Ouk S, Clegg NJ, Chen Y, Watson PA, Arora V, et al. Development of a second-generation antiandrogen for treatment of advanced prostate cancer. *Science* 2009;324:787–90.
 22. de Bono JS, Logothetis CJ, Molina A, Fizazi K, North S, Chu L, et al. Abiraterone and increased survival in metastatic prostate cancer. *N Engl J Med* 2011;364:1995–2005.
 23. Scher HI, Beer TM, Higano CS, Anand A, Taplin ME, Efstathiou E, et al. Antitumor activity of MDV3100 in castration-resistant prostate cancer: a phase 1-2 study. *Lancet* 2010;375:1437–46.
 24. Knudsen KE, Scher HI. Starving the addiction: new opportunities for durable suppression of AR signaling in prostate cancer. *Clin Cancer Res* 2009;15:4792–8.
 25. Kantoff PW, Higano CS, Shore ND, Berger ER, Small EJ, Penson DF, et al. Sipuleucel-T immunotherapy for castration-resistant prostate cancer. *N Engl J Med* 2010;363:411–22.
 26. de Bono JS, Oudard S, Ozguroglu M, Hansen S, Machiels JP, Kocak I, et al. Prednisone plus cabazitaxel or mitoxantrone for metastatic castration-resistant prostate cancer progressing after docetaxel treatment: a randomised open-label trial. *Lancet* 2010;376:1147–54.
 27. Nilsson S, O'Bryan-Tear CG, Bolstad B, Lokna A, Parker CC. Alkaline phosphatase (ALP) normalization and overall survival in patients with bone metastases from castration-resistant prostate cancer (CRPC) treated with radium-223. [abstract]. In: Proceedings of the 2011 ASCO Annual Meeting; 2011 Jun 3–7; Chicago, IL: ASCO; 2011. Abstract nr 4620.
 28. Positive outcome of interim analysis of pivotal Alpharadin study: primary endpoint met in phase III ALSYMPCA study [press release]. Oslo, Norway: Algeta ASA; 2011.
 29. Wang SJ, O'Neill RT, Hung HM. Approaches to evaluation of treatment effect in randomized clinical trials with genomic subset. *Pharm Stat* 2007;6:227–44.
 30. Dobbin KK, Simon RM. Sample size planning for developing classifiers using high-dimensional DNA microarray data. *Biostatistics* 2007;8:101–17.

Identification and Elucidation of the Biology of Adverse Events: The Challenges of Safety Assessment and Translational Medicine

Kenneth W. Turteltaub¹, Myrtle A. Davis³, Leigh Ann Burns-Naas², Michael P. Lawton², Adam M. Clark⁴, and Jack A. Reynolds⁵

Abstract

There has been an explosion of technology-enabled scientific insight into the basic biology of the causes of adverse events. This has been driven, in part, by the development of the various "omics" tools (e.g., genomics, proteomics, and metabolomics) and associated bioinformatics platforms. Meanwhile, for decades, changes in preclinical testing protocols and guidelines have been limited. Preclinical safety testing currently relies heavily on the use of outdated animal models. Application of systems biology methods to evaluation of toxicities in oncology treatments can accelerate the introduction of safe, effective drugs. Systems biology adds insights regarding the causes and mechanisms of adverse effects, provides important and actionable information to help understand the risks and benefits to humans, focuses testing on methods that add value to the safety testing process, and leads to modifications of chemical entities to reduce liabilities during development. Leveraging emerging technologies, such as genomics and proteomics, may make preclinical safety testing more efficient and accurate and lead to better safety decisions. The development of a U.S. Food and Drug Administration guidance document on the use of systems biology in clinical testing would greatly benefit the development of drugs for oncology by communicating the potential application of specific methodologies, providing a framework for qualification and application of systems biology outcomes, and providing insight into the challenges and limitations of systems biology in the regulatory decision-making process. *Clin Cancer Res*; 17(21); 6641–5. ©2011 AACR.

Introductory Note

At the 2010 Conference on Clinical Cancer Research, co-convened by Friends of Cancer Research and the Engelberg Center for Health Care Reform at the Brookings Institution, participants explored 4 pressing new challenges in the field. Articles summarizing the panel's recommendations on each of these topics are featured in this issue of *Clinical Cancer Research* (1–4).

Gaps in Current Testing and Safety Assessment Paradigms

The toxicity of new oncology drugs is a leading cause of pharmaceutical attrition and a major impediment to

efficient and successful drug development. Safety, or lack thereof, is also a major factor in regulatory decisions involving drug approval, labeling, risk evaluation, and mitigation and even withdrawal from the marketplace. The current battery of preclinical safety studies required to support the clinical development of new drugs and marketing approval is mapped out in International Conference on Harmonisation (ICH) guidelines that include ICH M3, E14, and S1 to S9 (5). In addition, various other documents from regulatory agencies provide recommendations regarding specific toxicities or adverse events, such as hepatotoxicity (6). However, these testing methods and risk assessments have not kept pace with the rapid evolution of technology, biomedical research, and knowledge generation. For example, the studies required to meet international regulatory guidelines for drug development and approval rely almost exclusively on *in vivo* animal testing protocols and endpoint assessments that have changed little in decades. These current *in vivo* methods as they are being used do not fully predict complex, serious, and low-incidence effects in humans, and in many cases are not amenable to generating knowledge that leads to mechanistic insight into the causes or biology of adverse events (7). Importantly, without relevant knowledge about the pathophysiology of potential adverse events, we are unable to predict or understand the

Authors' Affiliations: ¹Battelle Memorial Institute and Lawrence Livermore National Laboratory, Livermore; and ²Drug Safety Research and Development, Pfizer Inc., San Diego, California; ³Pharmacology Branch, National Cancer Institute, Bethesda, Maryland; ⁴FasterCures, Scientific and Federal Affairs, Washington, District of Columbia; and ⁵AnaBios Corporation, San Diego, California

Corresponding Author: Battelle Memorial Institute; Lawrence Livermore National Laboratory, Biosciences and Biotechnology Division Center for Accelerator, 7000 East Avenue, L-452, Livermore, CA 94550. Phone: 925-423-8152; Fax: 925-422-2282; E-mail: turteltaub2@llnl.gov

doi: 10.1158/1078-0432.CCR-11-1106

©2011 American Association for Cancer Research.

occurrence of low-incidence or idiosyncratic events in humans, a task that is perhaps the most challenging one we face in drug development. Although the long history of *in vivo* animal studies has served the scientific and regulatory community well, there is a timely and compelling need to incorporate changes into the earlier components of drug discovery and development that can lead to more focused animal studies. It is clear that no single new method or testing paradigm will replace entirely the need for *in vivo* testing, but adopting new science and technology on a case-by-case or fit-for-purpose basis from an array of emerging methods in the safety scientist's toolbox has the potential to improve research and development productivity, enable the ongoing efforts to understand and mitigate adverse events, and most importantly, facilitate and expedite the access of new therapies for patients.

Driven by rapidly emerging technologies, a nascent transformation of the safety sciences has taken place from empirical, subjective, and observation-based disciplines to scientifically grounded, objective, and data-driven sciences. This evolution has spawned new methods and experimental tools that are capable of defining the biologic basis of adverse events at the cellular, molecular, and biochemical level. These tools include platforms such as genomics, proteomics, metabolomics, and bioinformatics. Together, they enable the practice of systems biology. Systems biology creates the capability to elucidate complex, highly networked, and pleiotropic pathways of toxicities and to identify specific biomarkers of impending undesirable events (8, 9). This provides the opportunity for the contemporary toxicologist to take a more active and visible role in safety-related decisions. Historically, due to the gap in our knowledge of most toxicities, many safety decisions were based solely on the perceived risk of a toxicity, and often disregarding the potential benefit of a drug. Elucidating the biology of an adverse event allows the supplanting of the perception of risk with specific data that form the underpinning of a robust decision on risk and benefit (10). With this new knowledge, toxicologists can contribute to a systematic and objective decision-making process that identifies patients at risk for serious adverse events and at the same time enables access to individuals that might receive maximum benefit.

Emerging technologies, particularly in the areas of systems biology, biomarkers, and imaging, have begun to be incorporated into clinical development programs. Modifications to the conduct of clinical trials include screening of Investigational New Drug (IND) applications, microdosing protocols, adaptive clinical trials, translational medicine, and risk management planning or risk mitigation strategies. Regulatory agencies have extended explicit overtures and shown a readiness to embrace change through the Critical Path Initiative in the United States and the Innovative Medicines Initiative in Europe, for example. On a smaller scale, a great deal has been learned about the utility and limitations of data derived from new science and enabling technologies from the U.S. Food and Drug Administration's

(FDA) voluntary exploratory data submission program (11) and active participation in a number of scientifically driven public-private consortia. The recent announcement of NIH and FDA grants directed at improving drug development and regulatory sciences is a timely testimony to the importance of these topics. The FDA has also sponsored a number of scientific meetings to solicit broad input into the challenges of creating and using drug safety knowledge in the mining of adverse events databases and prediction of adverse events. These laudable efforts create the opportunity to articulate a coordinated framework for policy change that can be understood and engaged by the pharmaceutical industry and broadly communicated to patients and the public. The treatment of cancers has undergone significant advances in the past few years, but as patients are living longer with their diseases, the onerous effects of drug treatment begin to emerge. Current topics in the sequelae of cancer therapy could provide the momentum and focus to urgently apply new technologies to preclinical toxicology.

Systems Biology

Systems biology is an attractive complementary approach to preclinical testing. It has been defined as the iterative and integrative study of biologic systems as they respond to perturbations (12). Systems toxicology comprises the integration of molecular endpoints and conventional toxicity endpoints into a systems biology approach. In a sense, contemporary systems biology is a renaissance of physiology, a traditional integrative discipline. Biologic research has enjoyed decades of success in dissecting the structures and functions of individual molecular and cellular components comprising an organism. However, the inherent complexity of biologic systems, due not only to the large number of their constituents but also to the intricate web of interactions among these constituents, has proved to be difficult to understand with reductionist approaches. Research must be conducted at a more global, systems level if we are to gain understanding of the overall behavior of the biologic networks that maintain normal physiology and the perturbations in these networks that lead to toxicity and disease. Environmental stressors, including physical and chemical agents, exert adverse effects by initially impinging on specific molecular or cellular targets. The ensuing responses triggered from the initial interactions and subsequently propagated along the normal molecular, cellular, or systemic networks, will ultimately affect the health of the intact organism. The application of computational systems biology in risk assessment focuses on developing quantitative simulation models of the dose-response relationships for network perturbations by chemical stressors and drugs (13, 14).

Driven by systems biology approaches, significant progress has been made in the elucidation and characterization of cellular response networks—the interconnected pathways composed of complex biochemical interactions of

genes, proteins, and small molecules that maintain normal cellular function, control communication between cells, and allow cells to adapt to perturbations in their environment (4). The myriad potential sites of interaction and impact that any given perturbation might have on a cell or organ function and the resulting complexity of gaining insight into how these can affect the entire system can be envisioned (2). This complexity can only be overcome and be of utility through the systematic and integrated approach to manipulation, modeling, and measuring the wide spectrum of activities. Mining of complex and disparate databases is essential to generate nonintuitive insights and testable hypotheses of the causes and sequelae of undesirable perturbations. Two case studies that exemplify the potential of systems biology are discussed below.

Case study 1: Drug-induced vascular injury

This case demonstrates how a systems biology approach can elucidate the pathophysiology of complex and dynamic biologic processes, create testable hypotheses related to these phenomena, and identify potential candidate biomarkers that can be assessed and validated as an indicator of the toxicity (15).

No sensitive and reliable biomarker currently exists for monitoring of the vascular lesions induced by chemicals in preclinical models (16). Moreover, the pathogenesis of these lesions in animals is still unclear. Using modern "omics" technologies, knowledge generation and intelligent networking tools, and targeted modeling methods, the pathophysiology of a well-known but enigmatic phenomenon of chemically induced vascular injury has been elucidated. Not only was the application of a systems biology approach essential to the characterization of the signals and pathways of these events, but long-sought-after candidate biomarkers were also identified. This research endeavor generated over one million data points that were shared with the FDA under their voluntary genomics submission program. After a rigorous analysis of the data, FDA scientists reached essentially the same conclusions about the pathophysiology of drug-induced ischemia and subsequent reperfusion.

Phosphodiesterase 4 (PDE4) inhibitors are a class of drugs that can provide novel therapies for asthma and chronic obstructive pulmonary disease. Their development is frequently hampered by the induction of vascular toxicity in rat mesenteric tissue during preclinical studies. Histopathologically, mesenteric vascular injury is characterized by perivascular edema and mixed inflammatory cell infiltration associated with medial necrosis and hemorrhage (17). Whereas these vascular lesions in rats have been well characterized histologically, little is known about their pathogenesis, and in turn, sensitive and specific biomarkers for preclinical and clinical monitoring do not exist. Development of potentially novel life-saving therapies has therefore been hindered due to the lack of biomarkers for drug-induced vascular injury to confirm that a candidate drug is safe for administration to humans (18). To investigate the early molecular mechanisms

underlying vascular injury, time-course studies were performed in which rats were treated for 2 to 24 hours with high doses of a candidate PDE4 inhibitor. Transcriptomics analyses in mesenteric tissue were performed using oligonucleotide microarray and real-time reverse transcriptase PCR technologies, and compared with histopathologic observations. In addition, protein measurements were performed in serum samples to identify soluble biomarkers of vascular injury. The results show that molecular alterations preceded the histologic observations of inflammatory and necrotic lesions in mesenteric arteries. Some gene expression changes suggest that the development of the lesions could follow a primary modulation of the vascular tone in response to the pharmacologic effect of the compound. Activation of genes coding for pro- and antioxidant enzymes, cytokines, adhesion molecules, and tissue inhibitor of metalloproteinase 1 (TIMP-1) indicates that biomechanical stimuli may contribute to vascular oxidant stress, inflammation, and tissue remodeling. This leads to the proposed time-dependent mechanism of toxicity of PDE inhibitors: (i) ischemia reperfusion-like injury initiates the toxic response followed by the induction of oxidative stress; (ii) release of cytokines such as interleukin-6 (IL-6), TNF, and IL-1 β activate the innate immune response; and (iii) the release of specific molecular mediators, such as leukotriene B₄, platelet activating factor, C5a, and oxidized low-density lipoprotein, induces an inflammatory response that leads to vascular necrosis. Indeed, TIMP-1 appeared to be an early and sensitive predictive biomarker of the inflammatory and tissue remodeling components of PDE4 inhibitor-induced vascular injury (19). Importantly, some of the candidate biomarkers identified by these studies are now being assessed and potentially validated in animal and human experiments and may lead to the renewed development of a very important class of potential therapeutics.

Case study 2: Oncology drug-induced cardiovascular toxicity

As multiple types of cancer transition from acute to chronic diseases, the cardiotoxicity of anticancer treatments has emerged as a serious clinical problem (20). Cardiotoxicity can manifest in a variety of ways depending on the type of anticancer treatment being used. For example, anthracyclines generate free radicals, causing permanent myocyte cellular destruction that is related to the cumulative lifetime dose, which limits the usefulness of anthracyclines in oncology (21). In contrast, the monoclonal antibody trastuzumab can mediate transient cardiotoxicity by disrupting cardiomyocyte cellular signaling pathways (22). An understanding of the mechanism of drug-induced cardiotoxicity is crucial in devising methods to treat or prevent this toxicity. Ideally, the use and application of systems biology approaches could provide an opportunity to facilitate or improve (i) selection of the most effective therapies, (ii) identification of specific patients at risk for chemotherapy-induced adverse events, (iii) dose selection and scheduling decisions, and (iv) identification of early

signals of emerging adverse events that may enable prompt clinical responses.

A systems biology approach was used to delineate the signaling pathways involved in imatinib-induced cardiotoxicity. Imatinib is a tyrosine kinase inhibitor that is active against the tyrosine kinase Bcr-Abl in chronic myelogenous leukemia (CML) as well as in the tyrosine kinase C-kit in gastrointestinal stromal tumors (GIST). Unfortunately, imatinib treatment is associated with a 4% incidence of heart failure (23). To identify the mechanism of this toxicity, researchers incubated rat cardiomyocytes with imatinib and determined that imatinib induced the endoplasmic reticulum stress response, leading to cell death (24). The researchers further demonstrated that imatinib mediates cardiomyocyte death through its interaction with Bcr-Abl by transducing into the cardiomyocytes the Abl T315I point mutation which renders the kinase resistant to imatinib (25). Subsequently, a redesigned imatinib was engineered that inhibited only the C-kit tyrosine kinase and no longer had activity against Bcr-Abl (26). Although this redesigned compound was ineffective in treating CML, it retained its activity against GIST and no longer exhibited cardiotoxicity in mouse models. These results demonstrate the potential of systems biology in combination with rational drug design to engineer drugs so that their adverse effects are minimized or eliminated while their desired anticancer effects are preserved.

A Pathway Forward

The current state of safety sciences and the related emerging technologies represent an unprecedented and timely opportunity to make a profound impact on drug development and regulatory decision making. By defining, characterizing, validating and integrating new methods and science into the regulatory decision-making framework, this enterprise will improve public health decision making and enhance the efficiency of bringing new drugs to patients. Overcoming current challenges of safety assessment through new technologies will help us (i) understand the translation (if any) of nonclinical safety signals to the patient population; (ii) aid in the development of safer drugs, beginning at the design phase; and (iii) enhance our understanding of the potential safety impact of a drug on a particular individual by understanding relationships to key personal "omic" signatures. These accomplishments would improve the efficiency of drug research and development and increase the probability of success, adding value to

patient communities by improving access to promising new therapies. Moreover, these changes could also have a profound impact on the business model of the pharmaceutical and chemical industries and help to stem the occurrence of unanticipated adverse effects in late-stage clinical trials or in the postmarketing phase, which can quickly halt the development or availability of novel therapeutics.

The use of systems biology to characterize the inherent risks of pharmaceuticals can markedly improve drug development and postmarketing processes in close collaboration with the FDA and other regulatory agencies. Although resource constraints, computational limitations, and the complexities inherent to human disease may limit the utility of systems biology, data resulting from science and methods centered in systems biology can be readily validated using *in vivo* models and rapidly assessed in humans. The evolving concept of systems biology is starting to be adopted by the pharmaceutical industry and integrated into the safety sciences. A recent example of a substantial move forward in this field is the European Union Framework 6 Project on Predictive Toxicology, which compared conventional toxicologic endpoints for several investigational compounds to transcriptomics, proteomics, and metabolomics profiles (27). These researchers found that the use of systems biology was instrumental in determining mechanisms of nephrotoxicity and hepatotoxicity as well as identifying potential predictive biomarkers. To realize the true potential of systems biology and improve drug safety, a more systematic integration into drug discovery and development is needed. As a direct result of this panel discussion, the FDA has spearheaded an oncology pilot study on the cardiotoxicity of tyrosine kinase inhibitors with the hope that a guidance document can be generated to encourage and accelerate the adoption of systems biology in the development of drugs for oncology.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

John Leighton contributed on behalf of the US Food and Drug Administration. James Sonnett of Battelle Memorial Institute, Jon Cook of Pfizer, and Jeff Allen and Samantha Roberts of Friends of Cancer Research assisted with drafting.

Received May 6, 2011; revised September 9, 2011; accepted September 12, 2011; published online November 1, 2011.

References

1. Roberts SA, Kames EK, Allen JD, Benner JS, Sigal EV, McClellan M. Achieving the goals of effective, safe, and individualized cancer care. *Clin Cancer Res* 2011;17:6632-3.
2. Scher HI, Fuld Nasso S, Rubin EH, Simon R. Adaptive clinical trial designs for simultaneous testing of matched diagnostics and therapeutics. *Clin Cancer Res* 2011;17:6634-40.
3. Cleeland CS, O'Mara A, Zagari M, Baas C. Integrating pain metrics into oncology clinical trials. *Clin Cancer Res* 2011;17:6646-50.
4. Ambrose LF, Freedman J, Buetow K, Friend S, Schilsky RL. Using patient-initiated study participation in the development of evidence for personalized cancer therapy. *Clin Cancer Res* 2011;17:6651-7.

5. International Conference on Harmonisation. Safety guidelines [database on the Internet]. Geneva (Switzerland): International Conference on Harmonisation. [cited 2011 Jul 5]. Available from: <http://www.ich.org/products/guidelines/safety/article/safety-guidelines.html>.
6. US Food and Drug Administration. Guidance for industry- drug-induced liver injury: premarketing clinical evaluation [monograph on the Internet]. Silver Spring (MD): FDA; July 2009 [cited 2011 Jul 6]. Available from: <http://www.fda.gov/downloads/Drugs/Guidance-ComplianceRegulatoryInformation/Guidances/UCM174090.pdf>.
7. Olson H, Betton G, Robinson D, Thomas K, Monro A, Kolaja G, et al. Concordance of the toxicity of pharmaceuticals in humans and in animals. *Regul Toxicol Pharmacol* 2000;32:56–67.
8. Raschi E, Vasina V, Ursino MG, Boriani G, Martoni A, De Ponti F. Anticancer drugs and cardiotoxicity: insights and perspectives in the era of targeted therapy. *Pharmacol Ther* 2010;125:196–218.
9. Pujol A, Mosca R, Farrés J, Aloy P. Unveiling the role of network and systems biology in drug discovery. *Trends Pharmacol Sci* 2010;31:115–23.
10. Gibb S. Toxicity testing in the 21st Century. A vision and a strategy. *Reprod Toxicol* 2008;25:136–8.
11. US Food and Drug Administration. Voluntary Exploratory Data Submission (VXDS) [database on the Internet]. Silver Springs (MD): FDA. [cited 2011 Jul 8]. Available from: <http://www.fda.gov/drugs/scienceresearch/researchareas/pharmacogenetics/ucm083673.htm>.
12. Auffray C, Imbeaud S, Roux-Rouquié M, Hood L. From functional genomics to systems biology: concepts and practices. *C R Biol* 2003;326:879–92.
13. Butcher EC, Berg EL, Kunkel EJ. Systems biology in drug discovery. *Nat Biotechnol* 2004;22:1253–9.
14. Plavec I, Sirenko O, Privat S, Wang Y, Dajee M, Melrose J, et al. Method for analyzing signaling networks in complex cellular systems. *Proc Natl Acad Sci U S A* 2004;101:1223–8.
15. Marrer E, Dieterle F. Impact of biomarker development on drug safety assessment. *Toxicol Appl Pharmacol* 2010;243:167–79.
16. Kerns W, et al. Drug-induced vascular injury—a quest for biomarkers. *Toxicol Appl Pharmacol* 2005;203:62–87.
17. Zhang J, Snyder RD, Herman EH, Knapton A, Honchel R, Miller T, et al. Histopathology of vascular injury in Sprague-Dawley rats treated with phosphodiesterase IV inhibitor SCH 351591 or SCH 534385. *Toxicol Pathol* 2008;36:827–39.
18. Zhang J, Defelice AF, Hanig JP, Colatsky T. Biomarkers of endothelial cell activation serve as potential surrogate markers for drug-induced vascular injury. *Toxicol Pathol* 2010;38:856–71.
19. Daguès N, Pawlowski V, Guigon G, Ledieu D, Sobry C, Hanton G, et al. Altered gene expression in rat mesenteric tissue following *in vivo* exposure to a phosphodiesterase 4 inhibitor. *Toxicol Appl Pharmacol* 2007;218:52–63.
20. Ewer MS, Ewer SM. Cardiotoxicity of anticancer treatments: what the cardiologist needs to know. *Nat Rev Cardiol* 2010;7:564–75.
21. Lefrak E, Pitha J, Rosenheim S, Gottlieb JA. A clinicopathologic analysis of adriamycin cardiotoxicity. *Cancer* 1973;32:302–14.
22. Negro A, Brar BK, Lee KF. Essential roles of Her2/erbB2 in cardiac development and function. *Recent Prog Horm Res* 2004;59:1–12.
23. Cohen MH, Williams G, Johnson JR, Duan J, Gobburu J, Rahman A, et al. Approval summary for imatinib mesylate capsules in the treatment of chronic myelogenous leukemia. *Blood* 2005;105:935–42.
24. Kerkelä R, Grazette L, Yacobi R, Iliescu C, Patten R. Cardiotoxicity of the cancer therapeutic agent imatinib mesylate. *Nat Med* 2006;12:908–16.
25. Gorre ME, Mohammed M, Ellwood K, Hsu N, Paquette R, Rao PN, et al. Clinical resistance to STI-571 cancer therapy caused by BCR-ABL gene mutation or amplification. *Science* 2001;293:876–80.
26. Fernández A, Sanguino A, Peng Z, Ozturk E, Chen J, Crespo A, et al. An anticancer C-Kit kinase inhibitor is reengineered to make it more active and less cytotoxic. *J Clin Invest* 2007;117:4044–54.
27. Suter L, Schroeder S, Meyer K, Gautier JC, Amberg A, Wendt M, et al. EU Framework 6 Project: Predictive Toxicology (PredTox)—overview and outcome. *Toxicol Appl Pharmacol* 2011;252:73–84.

Integrating Pain Metrics into Oncology Clinical Trials

Charles S. Cleeland¹, Ann O'Mara², Martin Zagari⁴, and Carole Baas³

Abstract

Cancer-related pain is highly prevalent and often severe, and as a result is often one of the defining experiences for patients with malignancy. Patients and patients' families almost always live with the ever-present reality that cancer treatment and progression may be accompanied by pain. For patients nearing the end of life, most fear that their final days will be spent living with the terrible effects of the disease, the most important of which is pain. Despite this, there is far less systematic research on the mechanisms of cancer-related pain or on the development of new agents to reduce or eliminate pain in cancer patients compared with research to combat the disease itself. Further, even when the focus of research is treatment of the tumor, the effects of anticancer treatments on pain are often under-reported in publications and other forums. To illustrate the relative drought in the cancer pain control area, there have been no new drugs approved for cancer-related pain in recent years. A number of methodologic and logistical challenges that hinder the ability to assess pain response in clinical trials are discussed in this article. Possible ways to address these challenges are also discussed. *Clin Cancer Res*; 17(21); 6646–50. ©2011 AACR.

Introductory Note

At the 2010 Conference on Clinical Cancer Research, co-convened by Friends of Cancer Research and the Engelberg Center for Health Care Reform at the Brookings Institution, participants explored 4 pressing challenges in the field. Articles summarizing the panel's recommendations on each of these topics are featured in this issue of Clinical Cancer Research (1–4).

The Need to Integrate Measures of Pain in Cancer Clinical Trials

As more effective drug products become available to treat cancer, survival rates for many types of cancers have improved. Patients are not only living longer with cancer, but they are also living longer with symptoms associated with both cancer and its treatment. Cancer-related pain and other symptoms, such as fatigue, disrupted sleep, and psychosocial distress, have a significant impact on functioning and health-related quality of life (5–8). With the

availability of more effective treatment options, oncology product development programs are targeting add-on, second-line, or advanced disease indications in addition to first-line therapy. As a result, in addition to including objective measures such as overall survival and tumor response, oncology clinical trials are also targeting improvements in patient-reported cancer-related symptoms.

Cancer-related pain is a frequently reported and distressing symptom associated with many malignancies. A recent systematic review indicates that approximately half of patients with solid tumors have pain, and that, of those with pain, one third report pain that is moderate to severe (9). Analgesics are the mainstay of therapy in treating cancer-related pain (10). However, chemotherapeutic agents that demonstrate evidence of pain reduction or of a delay in the onset of pain in addition to meeting standards of efficacy could provide a significant treatment benefit for the patient. As a result, some oncology clinical trials have included measures of pain in study designs (11–13). However, the number of patients enrolled in oncology trials that examine pain as an outcome is a fraction of those receiving care, even in the setting of a clinical trial, resulting in a paucity of quality of evidence of treatment effects on cancer pain. Carefully designed trials with cancer pain relief as a primary or secondary outcome are required in patients with well-defined disease and pain.

Adequately Measuring Subjective Pain in Clinical Trials

Adequately and reliably measuring and interpreting subjective endpoints such as pain can be challenging. Randomized clinical trials in oncology from 1996 through 2001

Authors' Affiliations: ¹Department of Symptom Research, The University of Texas MD Anderson Cancer Center, Houston, Texas; ²Palliative Care Research and Community Oncology Prevention, and ³Physical Sciences in Oncology, National Cancer Institute, National Institutes of Health, Bethesda Maryland; and ⁴Health Economics and Outcome Research, Amgen Inc., Thousand Oaks, California

Corresponding Author: Charles S. Cleeland, Department of Symptom Research, The University of Texas MD Anderson Cancer Center, 1515 Holcombe Boulevard, Unit 1450, Houston, TX 77030. Phone: 713-745-3470; Fax: 713-745-3475; E-mail: ccleeland@mdanderson.org

doi: 10.1158/1078-0432.CCR-11-1109

©2011 American Association for Cancer Research.

have included at least 125 different pain outcome measures (14). Different pain outcome measures may evaluate different aspects of pain, such as pain intensity, pain interference (the extent to which pain interferes with daily living), or pain relief. The most commonly used scale is a numeric rating scale, where patients typically rate their pain from 0 to 10, with 0 representing no pain and 10 corresponding to worst pain imaginable. Other scales (verbal descriptor scales) use graded adjectives (mild to severe) to represent increases in severity. Finally, the visual analog scale may be used in acute (postoperative) pain studies, where patients indicate on a line their pain intensity. Other domains that might be assessed include the degree to which pain interferes with physical and emotional function. The Brief Pain Inventory (BPI) is an example of a measure that samples both severity and interference, as well as the degree to which trial participants believe a therapy has helped their pain (15). Measures of pain relief are similar to those for pain intensity, except that patients are asked to rate how much relief they have experienced relative to their previous pain levels following some intervention. Some pain metrics provide a more qualitative assessment of cancer pain, such as the McGill Pain Questionnaire (MPQ) which scores sensory, affective, evaluative, and miscellaneous pain (16). The pain metrics listed above, and others, are described in detail in a systematic review of 164 studies of cancer pain comprising over 35,000 adult cancer patients (17). This report discusses the utility, validity, and reliability of these measures and provides recommendations for selecting the appropriate pain measure depending on the situation. More recently, consensus reviews both in the United States and Europe indicate that, at a minimum, a numeric rating scale of pain severity is acceptable as a unidimensional measure of pain, and is quite easy to implement in most clinical trials (18, 19). Thus, despite the subjective nature of pain, reliable methods exist to measure pain and the challenge now is to integrate these methods into clinical trials in a way that results in meaningful improvements for patients suffering from cancer pain.

The Challenge of Incorporating Pain Metrics into Oncology Clinical Trials

In 2009, the U.S. Food and Drug Administration (FDA) "Guidance for Industry: Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims" was published (20). The guidance describes how FDA reviews patient-reported outcome (PRO) measures that are used to support claims in approved product labeling. The guidance defines a PRO as any report of the status of a patient's health condition that comes directly from the patient without interpretation of the patient's response by a clinician or anyone else. The guidance notes that, like other endpoints in a clinical trial, a PRO measure must be well-defined and reliable and show evidence that it is an adequate measure of the specific concept it was designed to measure. Based on the FDA PRO Guidance principles, symptoms known only to the

patient, such as cancer-related pain, are best evaluated by a self-reported measure. The measure should be comprehensible, interpretable, and appropriate for the target population. In addition to selecting the actual pain measure, based on the FDA PRO Guidance, it is also important to consider how the pain measure will be incorporated in the clinical trial. The endpoint model describes the relationship of all endpoints, both PRO and non-PRO, in a clinical trial (e.g., primary, coprimary, or secondary). The endpoint model is critical in the implementation of the pain measure in the clinical trial.

Although the FDA PRO Guidance lays out general principles for developing PRO measures and endpoints, it is not specific to the issue of pain measurement in oncology. There are many uncertainties and methodologic challenges to consider in incorporating pain endpoints in oncology clinical trials. For example, if a sponsor plans to include pain as an efficacy endpoint in a pivotal trial, the sponsor will need to determine whether pain should be used as a primary or coprimary endpoint as opposed to a secondary endpoint. If pain is an endpoint, the endpoint model must be constructed to appropriately interpret study results with consideration for the impact of treatment on tumor burden in addition to pain. The sponsor must find some way to measure and differentiate between "cancer-related pain" and "treatment-related pain" (e.g., chemotherapy-induced peripheral neuropathy) in the context of the proposed trial. In addition, the sponsor must determine how to design the trial to include the appropriate frequency of pain assessment in order to answer the trial question but not burden patients. Specific enrollment criteria, including level of pain, also need to be considered for pain palliation trials, and there is no firm definition of what constitutes "significant pain."

In addition to these issues, sponsors must also address analgesic use by patients within the trial. Analgesic use must be monitored, and it may be difficult to differentiate pain relief provided by the analgesic from pain relief provided by the cancer treatment. Additionally, it may be hard to adequately measure and compare pain severity and pain relief when patients enrolled in the trial are taking different baseline and rescue analgesics. Because of these issues, the sponsor will need to use a pain endpoint that includes assessments of both pain and analgesic use.

There are statistical considerations for including pain as an efficacy endpoint in oncology clinical trials as well. For example, a pain palliation trial must include a placebo arm, and blinding will most likely not be possible. Therefore, the sponsor will need to find or develop strategies to minimize unblinding in palliation trials. The sponsor will also need to find strategies to minimize missing data, particularly when data are self-reported.

Potential Scenarios for Adding Pain Metrics to Chemotherapy Clinical Trials

In order to explore specific examples of how pain metrics could be introduced into a clinical trial and how those

metrics could serve as decision-making criteria for the regulatory approval process, we present here 3 hypothetical clinical trial scenarios in which changes in pain or use of analgesics might be an outcome. In 2 of the scenarios, pain progression or palliation are assessed as secondary outcomes in a clinical trial of a new chemical entity, whereas in the third scenario pain palliation is a primary outcome for a new agent designed to reduce pain when the agent is added to an approved, second-line therapeutic. For these scenarios, metastatic castration-resistant prostate cancer (CRPC) is used as an example because large numbers of CRPC patients have significant pain for long periods of time (21). In addition, pain has been used in the past as a trial outcome for CRPC patients, and some trials with this patient population have used pain relief as a primary endpoint (22, 23). Although CRPC is used as an example here, these scenarios can be generalized to other cancer types. The design and measurement challenges may be addressed differently depending on the scenario. These scenarios are not intended to reflect regulatory thinking.

Case 1: Pain progression

In this scenario, a chemotherapeutic agent that may prevent pain progression in addition to treating HRPC is being tested relative to placebo in chemotherapy-naïve patients. This blinded, randomized trial will enroll patients with the following pain medication use profile: patients who have received no more than 1 day of opioids in the previous 14 days, or patients who have received no more than 6 days of nonopioids in the prior 14 days. Presumably, these are patients whose pain required episodic rather than more continuous treatment.

This trial could use a patient-reported outcome assessment with a 0-to-10 numerical rating scale. Data collection would occur daily for 4 weeks and every 12 weeks thereafter until disease progression is confirmed, and subsequent analysis will attempt to determine if a single daily pain assessment is representative of the pain experienced by patients. The time to pain progression in this scenario is defined as an increase in the worst daily pain of more than 2 points as measured by the numerical rating scale, by the initiation of opioid analgesic use in those patients who had not taken opioids for cancer-related pain at the study's initiation, an increase in opioid use to more than 3 days over a 14-day period in patients who had used opioids as needed prior to the study, or the start of bone-directed radiotherapy for pain palliation. The time to pain progression will serve as a secondary outcome for the purposes of seeking regulatory approval for this new agent.

Case 2: Pain palliation

This scenario adds pain metrics to a clinical trial being run on a second-line chemotherapeutic agent in combination with prednisone in patients with hormone-refractory prostate cancer (HRPC) who have experienced failure of taxane-based therapy. While overall survival will

be the primary endpoint, pain progression and pain response will be important secondary endpoints. This trial is designed as a randomized, open-label multicenter study with one arm consisting of the new agent combined with prednisone and the other arm consisting of mitoxanthrone and prednisone. The patient population will have experienced documented disease progression during or within 6 months after prior hormone therapy and taxane therapy.

In this trial, pain and neuropathy might be assessed using the MPQ, which measures important neurosensory symptoms that patients might not describe as pain, such as numbness, as well as pain severity, and records analgesic use quantified as an analgesic score derived from a patient-kept analgesic diary. Pain will be assessed prior to every treatment cycle and at the end of the study with the goal of determining if pain and analgesic use assessment should be part of the inclusion criteria for the study.

Case 3: Pain palliation with product add-on

The third scenario is designed to use pain metrics to assess the efficacy of a medication designed to ameliorate pain in combination with an approved chemotherapeutic. In this trial, patients with stable baseline pain and analgesic use who have relapsed after first-line therapy will be randomized to receive second-line therapy in combination with either the new pain palliation drug candidate or placebo. In this trial, the primary endpoint will be the extent of permanent pain palliation as measured using a combination of the BPI short form (BPI-SF) and an analgesic log. Secondary endpoints will assess whether patients receiving drug, as opposed to placebo, have a longer time to pain progression or have less pain-induced interference with their ability to walk, work, and sleep.

This study will attempt to answer a number of questions relating to pain metrics, including how to define minimum, maximum, and stable pain, and how to define stable analgesic use in the context of which analgesics are used, such as nonsteroidal anti-inflammatory drugs or long-acting opioids. This scenario also calls for determining the optimal frequency for pain assessment and quantifying the degree and duration of pain reduction and analgesic use that is clinically meaningful.

Conclusions and Next Steps

Cancer-related pain is arguably the physical ailment most feared by cancer patients (24, 25). Yet while clinical trials designed to assess the efficacy of new therapies for cancer include a variety of measures to assess a patient's physical response to therapy, these trials often do not include pain as either a primary or secondary outcome. Furthermore, clinical trials in oncology often fail to assess other symptoms or aspects of quality of life. Especially with patients who have more advanced disease, simple PRO measures of additional symptoms, such as fatigue, sleep disturbance,

gastrointestinal function, and mood impairment, add significantly to what patients and providers can expect of treatments (26).

As discussed in the PRO Guidance, PRO measures of symptom reduction can be direct indicators of treatment benefit, but barriers still exist to including pain endpoints in trials. In addition to the methodologic challenges discussed above, sponsors face logistical challenges in measuring pain in oncology trials. There is a high degree of uncertainty regarding what pain measurement endpoints the FDA will accept and what changes they will find clinically meaningful. This level of uncertainty, coupled with the expense associated with the measurement of pain in clinical trials, can make sponsors reluctant to measure pain palliation or prevention in oncology. Increased dialogue between the FDA and sponsors is recommended early in product development to plan the most efficient path forward for PRO measurement. The development of an oncology-specific pain-measurement guidance that details the standards for trial design, the number of trials required to incorporate pain measurements into labels, the pain instruments that FDA will

accept, and standards for statistical analysis, as well as other methodologic issues, would greatly benefit the cancer community. Such a guidance would facilitate the incorporation of pain relief into oncology trials with the ultimate result of cancer patients not only living longer but experiencing a higher quality of life.

Disclosure of Potential Conflicts of Interest

Dr. Cleeland has received a commercial research grant from AstraZeneca Pharmaceuticals and has served as a consultant on the advisory board of Amgen Inc. and Abbott Laboratories. No potential conflicts of interest were disclosed by the other authors.

Acknowledgments

Laurie Burke and Ann Marie Trentacosti contributed on behalf of the U.S. Food and Drug Administration. Ethan Basch of Memorial Sloan-Kettering Cancer Center and Jeff Allen and Samantha Roberts of Friends of Cancer Research assisted with drafting.

Received May 6, 2011; revised August 24, 2011; accepted August 24, 2011; published online November 1, 2011.

References

1. Roberts SA, Kames E, Allen J, Benner JS, Sigal EV, McClellan M. Achieving the goals of effective, safe, and individualized cancer care. *Clin Cancer Res* 2011;17:6632-3.
2. Scher HI, Nasso SF, Rubin EH, Simon R. Adaptive clinical trial designs for simultaneous testing of matched diagnostics and therapeutics. *Clin Cancer Res* 2011;17:6634-40.
3. Turteltaub K, Davis M, Burns-Naas LA, Lawton M, Clark A, Reynolds JA. Identification and elucidation of the biology of adverse events: the challenges of safety assessment and translational medicine. *Clin Cancer Res* 2011;17:6641-5.
4. Ambrose LF, Freedman J, Buetow K, Friend S, Schilsky RL. Using patient-initiated study participation in the development of evidence for personalized cancer therapy. *Clin Cancer Res* 2011;17:6651-7.
5. Cleeland CS. Symptom burden: multiple symptoms and their impact as patient-reported outcomes. *J Natl Cancer Inst Monogr* 2007;16-21.
6. Harrington CB, Hansen JA, Moskowitz M, Todd BL, Feuerstein M. It's not over when it's over: long-term symptoms in cancer survivors—a systematic review. *Int J Psychiatry Med* 2010;40:163-81.
7. Patrick DL, Ferketich SL, Frame PS, Harris JJ, Hendricks CB, Levin B, et al. National Institutes of Health State-of-the-Science Panel. National Institutes of Health State-of-the-Science Conference Statement: Symptom management in cancer: pain, depression, and fatigue, July 15-17, 2002. *J Natl Cancer Inst* 2003;95:1110-7.
8. Esther Kim JE, Dodd MJ, Aouizerat BE, Jahan T, Miaskowski C. A review of the prevalence and impact of multiple symptoms in oncology patients. *J Pain Symptom Manage* 2009;37:715-36.
9. van den Beuken-van Everdingen MH, de Rijke JM, Kessels AG, Schouten HC, van Kleef M, Patijn J. Prevalence of pain in patients with cancer: a systematic review of the past 40 years. *Ann Oncol* 2007;18:1437-49.
10. Dworkin RH, Turk DC, Basch E, Berger A, Cleeland C, Farrar JT, et al. Considerations for extrapolating evidence of acute and chronic pain analgesic efficacy. *Pain* 2011;152:1705-8.
11. Berthold DR, Pond GR, Soban F, de Wit R, Eisenberger M, Tannock IF. Docetaxel plus prednisone or mitoxantrone plus prednisone for advanced prostate cancer: updated survival in the TAX 327 study. *J Clin Oncol* 2008;26:242-5.
12. Belani CP, Pereira JR, von Pawel J, Pluzanska A, Gorbounova V, Kaukel E, et al. TAX 326 study group. Effect of chemotherapy for advanced non-small cell lung cancer on patients' quality of life. A randomized controlled trial. *Lung Cancer* 2006;53:231-9.
13. Burris HA III, Moore MJ, Andersen J, Green MR, Rothenberg ML, Modiano MR, et al. Improvements in survival and clinical benefit with gemcitabine as first-line therapy for patients with advanced pancreas cancer: a randomized trial. *J Clin Oncol* 1997;15:2403-13.
14. Carr DB, Goudas LC, Balk EM, Bloch R, Ioannidis JPA, Lau J. Evidence report on the treatment of pain in cancer patients. *J Natl Cancer Inst Monogr* 2004;32:23-31.
15. Cleeland CS, Ryan KM. Pain assessment: global use of the Brief Pain Inventory. *Ann Acad Med* 1994;23:129-38.
16. Melzack R. The McGill Pain Questionnaire: major properties and scoring methods. *Pain* 1987;1:277-99.
17. Jensen MP. The validity and reliability of pain measures in adults with cancer. *J Pain* 2003;4:2-21.
18. Dworkin RH, Turk DC, Wyrwich KW, Beaton D, Cleeland CS, Farrar JT, et al. Interpreting the clinical importance of treatment outcomes in chronic pain clinical trials: IMMPACT recommendations. *J Pain* 2008;9:105-21.
19. Hjermstad MJ, Fayers PM, Haugen DF, Caraceni A, Hanks GW, Loge JH, et al. European Palliative Care Research Collaborative (EPCRC). Studies comparing numerical rating scales, verbal rating scales, and visual analogue scales for assessment of pain intensity in adults: a systematic literature review. *J Pain Symptom Manage* 2011;41:1073-93.
20. U.S. Department of Health and Human Services. Guidance for industry. Patient-reported outcome measures: use in medical product development to support labeling claims [monograph on the Internet]. Washington: U.S. Department of Health and Human Services; 2009. Available from: <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM193282.pdf>. [cited 15 Mar 2011].

21. Basch EM, Trentacosti A, Burke LB, Kane RC, Atkinson T, Kwitkowski VE, et al. Measurement of pain in metastatic prostate cancer: FDA perspective [abstract]. In: 2010 Genitourinary Cancers Symposium (ASCO, ASTRO, SUO); 2011 Mar 15; San Francisco, CA. Alexandria (VA): ASCO; 2010. Abstract nr 79.
22. Tannock IF, de Wit R, Berry WR, Horti J, Pluzanska A, Chi KN, et al. TAX 327 Investigators. Docetaxel plus prednisone or mitoxantrone plus prednisone for advanced prostate cancer. *N Engl J Med* 2004; 351:1502–12.
23. Tannock IF, Osoba D, Stockler MR, Ernst DS, Neville AJ, Moore MJ, et al. Chemotherapy with mitoxantrone plus prednisone or prednisone alone for symptomatic hormone-resistant prostate cancer: a Canadian randomized trial with palliative end points. *J Clin Oncol* 1996;14:1756–64.
24. Swarm R, Abernethy AP, Angheliescu DL, Benedetti C, Blinderman CD, Boston B, et al. Adult cancer pain. *J Natl Compr Canc Netw* 2010;8: 1046–86.
25. Levin DN, Cleeland CS, Dar R. Public attitudes toward cancer pain. *Cancer* 1985;56:2337–9.
26. Turk DC, Dworkin RH, Allen RR, Bellamy N, Brandenburg N, Carr DB, et al. Core outcome domains for chronic pain clinical trials: IMMPACT recommendations. *Pain* 2003;106:337–45.

Using Patient-Initiated Study Participation in the Development of Evidence for Personalized Cancer Therapy

Laurie Fenton Ambrose¹, Jamie Freedman², Kenneth Buetow³, Stephen Friend⁴, and Richard L. Schilsky⁵

Abstract

Personalized cancer therapy offers the promise of delivering the right treatments to the right patients to improve patient outcomes and quality of life, while reducing exposure to ineffective therapies and the cost of cancer care. Realizing this promise depends in large part on our ability to generate timely and sufficiently detailed information regarding factors that influence treatment response. Generating this evidence through the traditional physician investigator-initiated clinical trial system has proved to be challenging, given poor recruitment rates and low compliance with requests for biospecimen collection. As a result, our current understanding of treatment response is inadequate, particularly for cancer therapies that have been in use for many years. Patient-initiated study participation may offer a new model for evidence generation that capitalizes on strong patient interest in furthering research to inform better and more tailored cancer therapies. In this approach, patients are engaged and recruited directly by the sponsor of an Institutional Review Board–approved study, and patients subsequently drive the participation of their health care providers to facilitate collection of required data and tissue samples. The ultimate goal of these studies is to generate evidence of sufficient quality to inform regulatory decisions (i.e., labeling changes for marketed therapies to reflect populations most likely to respond) and treatment selection. Here, we describe a hypothetical prospective observational study in non–small cell lung cancer that could serve as a model for patient-initiated study participation applied to understand molecular determinants of treatment response. Key elements discussed include study design, patient engagement, and data/biospecimen collection and management principles. *Clin Cancer Res*; 17(21); 6651–7. ©2011 AACR.

Introductory Note

At the 2010 Conference on Clinical Cancer Research, co-convened by Friends of Cancer Research and the Engelberg Center for Health Care Reform at the Brookings Institution, participants explored 4 pressing challenges in the field. Articles summarizing the panel's recommendations on each of these topics are featured in this issue of *Clinical Cancer Research* (1–4).

Predicting Response or Non-Response to Approved Oncology Therapies

Approval of new cancer drugs by the U.S. Food and Drug Administration (FDA) relies upon safety and efficacy data from population-based trials. To date, such trials have

typically employed tumor classification systems that do not fully account for the growing body of genomic knowledge regarding tumor diversity (5). When drugs evaluated in these trials are approved and become standard of care, the implications of failing to account for tumor diversity become apparent. Standard-of-care cancer therapies may benefit only one in four patients, leaving upwards of 75% of patients without effective initial therapies and at risk of experiencing only toxic effects (6).

The goal of personalized cancer therapy can be achieved through the development of new therapies or the selective use of existing therapies in patients more likely to benefit. Designing new targeted therapies requires a clear understanding of the tumor biology and how it varies in the patient population. In cancers for which this understanding is still developing, an alternative approach is to study variations in response to available treatments in search of biomarkers that predict favorable outcomes. In cases where adequate evidence can be developed, a primary goal would be to modify the label of a marketed drug to specify the patient subgroups most likely to benefit or those unlikely to benefit. Such a post-approval labeling change happened recently in the case of cetuximab, a member of the class of cancer drugs known as epidermal growth factor receptor (EGFR) inhibitors (7).

For older drug products, existing data may not contain the needed genomic information to identify markers of response

Authors' Affiliations: ¹Lung Cancer Alliance, Washington, District of Columbia; ²GlaxoSmithKline, Collegeville, Pennsylvania; ³Bioinformatics and Information Technology, National Cancer Institute, Bethesda, Maryland; ⁴Sage Bionetworks, Seattle, Washington; and ⁵University of Chicago Comprehensive Cancer Center, Chicago, Illinois

Corresponding Author: Richard L. Schilsky, Biological Sciences Division, University of Chicago Medical Center, 5841 S. Maryland Avenue, MC 2115, Chicago, IL 60637. Phone: 773-834-3914; Fax: 773-834-3915; E-mail: rschilsk@medicine.bsd.uchicago.edu

doi: 10.1158/1078-0432.CCR-11-1110

©2011 American Association for Cancer Research.

or nonresponse. Moreover, in the traditional physician investigator-based patient recruitment model, the clinical trials necessary to expand our knowledge are plagued by low enrollment and poor compliance for biospecimen collection (8). When genomic data are to be collected in pivotal trials, the current practice relies on optional genomic patient consent (either prospective or retrospective), which results in convenience, and potentially biased, genomic sample data collection (9). The genomic consent rates vary from trial to trial, making study results very difficult, if not impossible, to interpret when the compliance rates are low (10).

One potentially promising avenue for developing such evidence rapidly lies in directly engaging patients to participate in studies that collect detailed information about their tumors, treatments, and clinical outcomes. We define patient-initiated study participation as a model in which patients are engaged and recruited directly by the sponsor of an Institutional Review Board (IRB)-approved study, and patients in turn drive the participation of their physicians and other health care providers to facilitate collection of required data and tissue samples. As part of such a study, patients who receive cancer care from their usual providers would volunteer to donate certain biospecimens and clinical information to the study sponsor prior to treatment initiation and over the course of treatment. The goal of these studies is to use patient biospecimens and other data to identify molecular markers of treatment response that can be used to select treatment for future patients.

Recognizing the promise of this approach, advocacy groups such as the Love/Avon Army of Women, the Lung Cancer Alliance and others have begun to mobilize their networks to generate data through patient-initiated participation (11–15). For example, the Love/Avon Army of Women Initiative is attempting to recruit one million healthy women (including breast cancer survivors and women at risk for breast cancer) to participate in breast-cancer related studies (11). As of May 2011, over 354,000 women and men had registered online and 50 studies have been launched after successful matches were made

between interested participants and researchers (16). These initiatives speak to the motivation and commitment of patients and their families to advancing cancer research in general and personalized cancer care in particular. By providing patients with tools to enroll themselves and their providers in studies (thereby flipping the traditional provider-initiated approach), these efforts demonstrate the promise of patient-initiated participation for rapid accrual of large amounts of detailed exposure and outcomes data to answer a range of important questions in cancer care, including how to better target therapies. To ensure that these efforts result in actionable information, what is needed now is a clearer understanding of how such data can be most effectively collected (e.g., through improved education for involved parties) and used to inform the decisions of doctors, patients, regulatory authorities, and payers.

Data Required to Identify Patient Subsets

Developing evidence to support targeting available treatments to a subgroup of patients requires collecting detailed and high-quality data. These data can be thought of in layers of comprehensive, longitudinally linked information so that treatments can be tracked over time and within subgroups of patients. In addition to basic information like demographics, clinical laboratory results, and medical history, needed layers will likely include normal tissue samples, tumor and other biological specimens, detailed information on treatment exposure, adverse events, and clinical outcomes (Fig. 1).

Few data sources currently have the breadth and depth of information necessary to support analyses with sufficient statistical power to identify biomarkers of response or nonresponse. Furthermore, changing FDA-approved labels and recommended standards of care requires robust evidence built on high-quality data and an acceptable study design. We envision collection of these data through post-approval studies in which genomic data from biospecimens

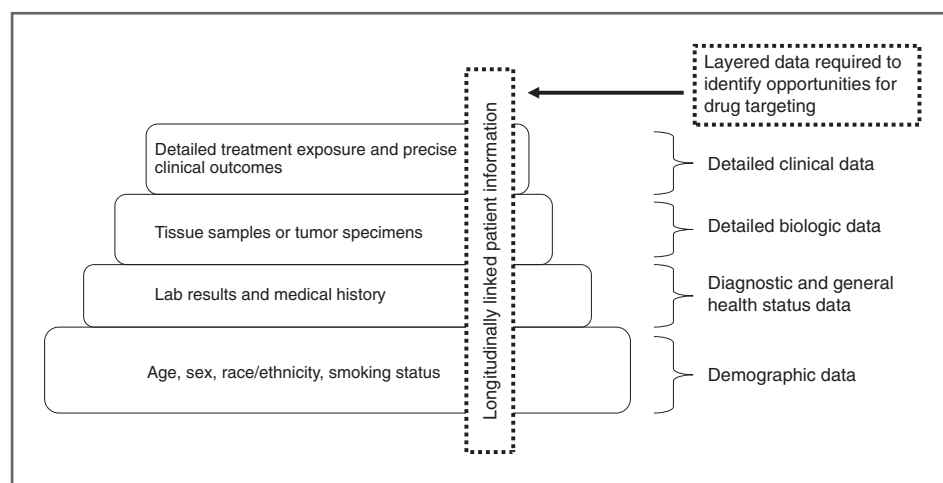


Figure 1. Layered data required to assess patient subsets.

are used to identify biomarkers predictive of clinical outcomes. Several factors can affect the ability to generate useful data from a genomic biomarker trial: (i) the prevalence of the target biomarker in the population; (ii) the prognostic impact of the biomarker to distinguish clinical outcomes in the population; (iii) concordance of biomarker expression between primary and metastatic tumor tissue; (iv) qualification and validation (analytical and clinical) of the biomarker assay; (v) availability of tissue specimens containing the biomarker; and (vi) quality and quantity of the tissue samples for biomarker analysis. Each of these factors will need to be addressed in any study of this nature.

Hypothetical Proposed Study

To help illustrate the issues, challenges, and potential solutions in using data generated through patient-initiated study participation for the purpose of informing labeling changes for existing cancer therapies, we examine study design considerations within the context of treatment for non-small cell lung cancer (NSCLC). First-line treatment for NSCLC typically consists of chemotherapy with a 2-drug regimen containing either cisplatin or carboplatin and another agent, which is typically vinorelbine, paclitaxel, docetaxel, gemcitabine, or pemetrexed. Experience with these regimens indicates that only approximately 30% of patients respond favorably (17).

To identify molecular signatures that explain variation in treatment response, several initiatives, including the Sage Bionetworks Non-Responder Project, are working to design studies that identify predictive markers of nonresponse. The Non-Responder Project has chosen several candidate tumors to study, including NSCLC, with an initial pilot study in acute myelogenous leukemia (5) based on 4 "first principles" for tumor selection (Fig. 2).

Informed by this and other related efforts, the objective of the proposed study is to identify one or more molecular

markers of nonresponse to first-line platinum-containing therapies for metastatic NSCLC, with the goal of supporting the revision of FDA-approved labels and recommended standard of care for these drugs.

The proposed study would begin when a patient with NSCLC is nearing a treatment decision and becomes aware of the opportunity to participate in the study by means of a website description or other form of outreach. This patient would approach his or her physician for support to enroll in the study. After enrollment, biospecimens would be collected from the patient at a designated research center and then the patient would return to the care of their oncologist. Meanwhile, tumor specimens would be analyzed in a Clinical Laboratory Improvement Amendments–certified laboratory for known and clinically actionable genetic variants. If clinically actionable results are identified from the research analyses, they would be returned to the patient and treating oncologist. Within the context of this hypothetical study, the study sponsor would determine what constitutes clinically actionable information on the basis of currently available evidence. Only clinically actionable results would routinely be shared with patients and providers; however, full results would be available upon request. Together, the patient and oncologist would select the most appropriate treatment approach (which may or may not rely on the results of tests performed in the study) from among the standard targeted or platinum-based therapies, and the provider would collect and report additional data on clinical outcomes over time. These longitudinal data would eventually be compiled, linked to other sources of electronic clinical data, and made available to qualified researchers.

Details regarding the proposed study design, including population characteristics, sample size, and key endpoints, are provided in Fig. 3. Analyses would be prespecified in the IRB-approved study protocol. Included in this study design would be the necessary and appropriate statistical analysis along with the network-biology modeling done to identify

Figure 2. Sage Bionetworks Nonresponder Project: First Principles for Tumor Selection.

Sage Bionetworks Nonresponder Project: First Principles for Tumor Selection
<ul style="list-style-type: none"> ▪ The treatment under investigation should have substantial response and nonresponse rates (>20% in either group). ▪ The disease must have clear, robust definitions of response and nonresponse that are clinically important. (A nonresponse biomarker should have the potential to change clinical practice.) ▪ Routine clinical management of the disease guarantees access to high-quality tissue specimens. (Use of archival tissue from diagnostic samples introduces risk when assessing treatments given at relapse.) ▪ The nonresponse group should ideally be defined as patients refractory to treatment rather than those who respond then relapse early. (If early relapse is caused by a resistant subpopulation at diagnosis, genomic analysis of tissue at diagnosis may or may not be informative, depending on the size of the resistant pool.)

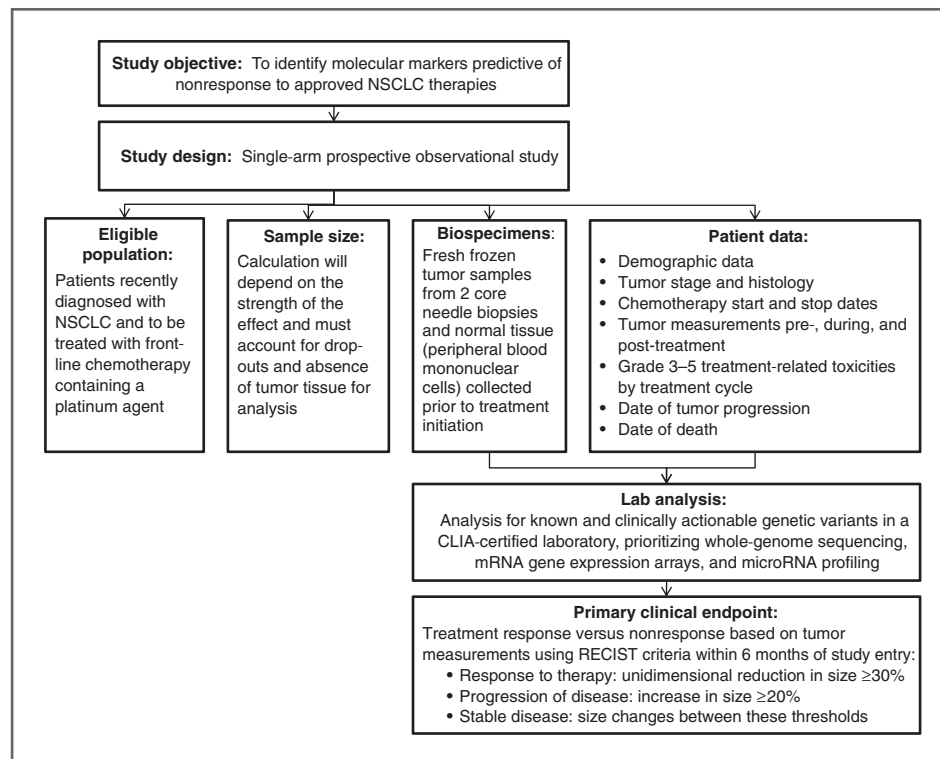


Figure 3. Design for hypothetical observational study. RECIST, Response Evaluation Criteria in Solid Tumors.

not just isolated markers for nonresponder populations but also sets of genes or "gene signatures" capable of identifying nonresponders. The goal for lung cancer might be to identify patients who have more than an 85% chance of not responding with a certainty of this outcome of 90%. This certainty around the likelihood that a patient may not respond would need to be set at a predetermined level of stringency to enable clinicians to use this information to determine whether to forego the original approved therapy and instead provide the patient an opportunity to receive an investigational regimen. The criteria for foregoing standards of care would be tumor and regimen specific and would need to be agreed to upfront with regulators and clinicians before the study is started.

Feasibility of Patient-Initiated Study Participation

While patient-initiated study participation offers promising opportunities for more efficient and dynamic clinical trial enrollment, a number of feasibility issues must be considered during data collection so that resulting data are relevant for regulatory and other forms of decision making.

Patient engagement

Patient-initiated study participation begins with raising awareness of participation opportunities, achieving patient/family engagement, and supporting patients/families through the process of enrolling in the study. Patient advocacy groups are positioned to play a key role in edu-

cating patients and their families regarding the importance of study participation and how the clinical research process works. In order to ensure optimal patient participation, we recommend that patient-initiated study participation efforts do not impose any sort of fees on patients. Instead, organizers of such efforts should absorb any associated costs or such costs should be incurred as part of routine cancer care.

As with any biomedical study, sponsors of studies employing patient-initiated participation must be careful to identify potential ethical concerns, address them as much as possible through study design, and ensure they are clearly communicated during the consent process. In some ways, a study such as the one proposed poses a narrower range of ethical concerns because it is not a treatment trial, but rather an observational study intended to enhance our understanding of response to established treatments. Even so, important issues to consider may include, but are certainly not limited to, the timing and nature of informed consent (e.g., if consent is obtained by study sponsors without the patient's provider present to offer guidance), clear communication of what results will and will not be returned to patients and providers, and anticipated risks of biospecimen donation that are above and beyond those associated with routine cancer care.

Role of health care providers

Health care providers can help to inform patients of the importance of study participation, streamline the consent and biospecimen collection process, and counsel patients

regarding potential health risks. In addition, health care providers assume certain formal responsibilities in the context of patient-initiated study participation. Once a patient decides to participate, they rely on their health care providers to provide the sponsor with medical records and possibly to perform relevant procedures, such as biopsies. As described above, the costs associated with these procedures should not be imposed on patients, and they also should not fall to providers. These costs should be covered as part of routine cancer care or should be absorbed by the study sponsor, as appropriate. To facilitate effective provider cooperation, organizations leading patient-initiated study participation efforts should consider proactively identifying interested providers and providing them with detailed information about the initiative and what level of provider involvement is expected.

Sample collection

When necessary, collection of biologic samples must address specific challenges. In general, normal tissue (e.g., blood, skin, or hair follicles) is easier to collect than tumor specimens. However, even these samples may require more complex sample collection schemes (e.g., peripheral blood mononuclear cells from whole blood) that call upon specialized collection methods and expertise at the clinical sites. Tumor samples are generally more difficult to collect because they require invasive procedures and because the quality of the specimens may be highly variable. Certain anatomic sites (e.g., skin or lymph nodes) are more amenable than others for collection of tumor specimens. Primary lung cancer specimens are very hard to collect because of location. If an assay for archival tissue is available from the original surgically obtained tumor specimen, such a sample might allow the highest yield if deemed scientifically appropriate to meet the study objectives.

During sample collection and all subsequent phases of storage and analysis, great care must be taken to ensure that biospecimens are of high quality. In an effort to improve the quality and standardization of biospecimens collected for cancer research, the Office of Biorepositories and Biospecimen Research within the National Cancer Institute (NCI) has launched several initiatives, including the development of best practices for biospecimen collection, processing, storage, retrieval, and dissemination (18). Although adherence to these best practices is voluntary, these standards and recommendations should be consulted in any patient-initiated study participation effort that involves biospecimen collection.

Sample compilation and storage

After samples and other data are obtained from patients, processes must be developed to efficiently compile and integrate them. Efforts that rely on patients to directly transfer data (e.g., computed tomography scans) to the study organizer will be more direct and simple to accomplish. Obtaining biospecimens from

patients may be more challenging because patients are typically not the "owners" of these samples and coordination must occur with health care providers. Such coordination may be more feasible if sample collection occurs at designated collection facilities and if relationships have been previously established with a core set of providers. As patient data/samples are collected, they should ideally be stored in a way that preserves the ability to link to other sources of electronic clinical data (e.g., from electronic health records) while protecting confidentiality. This measure is critical to creating the type of layered data necessary to identify markers of response and nonresponse.

Data access

To fully realize the goal of patient-initiated study participation, we recommend that data be compiled and made available free of charge in a standardized electronic format to all qualified researchers, rather than restricting access to a particular investigator or team. This availability will enable the widest possible access to patient data, and therefore the greatest possibility for important discoveries.

Patient privacy and data security

Ethical use of the data and samples requires review to ensure protection of human subjects, as well as assurance of patient privacy and data security. To this end, it is necessary to establish a "trust fabric" that grants access only where appropriate and only to data components that have been authorized [Health Insurance Portability and Accountability Act (HIPAA)] or consented to (Office for Human Research Protections) by the patient. Patients should be clearly advised that their donated data would be accessible to researchers and that the product of the research may be commercialized. The level of identification risk associated with donating their data must be transparently communicated to the participating patients and informed consent obtained. Because HIPAA assigns responsibility for protections to local groups that hold patient information, this trust fabric should recognize the need for local control of data release.

Patient privacy should be protected by removal of all HIPAA "identifiers" and by agreements that no parties may seek reidentifying information except for research covered by the informed consent. Double deidentification may provide further privacy protection with the use of 2 levels of coding between HIPAA "identifiers" and information relevant for research purposes (e.g., health outcomes or genetic/genomic test results). Use of this approach increases the stringency of privacy protection, while retaining the potential for future analyses building upon the collected data, which is not the case with other methods (e.g., total anonymization).

Patient-initiated study participation efforts should allow controlled access to patient-level data, and researchers seeking such data would have to make appropriate commitments including the following: (i) use only for approved

research; (ii) no sharing of data/samples with others without such sharing having been referenced in the consent form; (iii) no effort to reidentify; (iv) return of unused biospecimens to the repository; and (v) the repository would be obligated to confirm that the proposed research is consistent with the scope of the consent forms and to track the disposition of all specimens.

Governance

Governance policies are required to establish oversight of data collection and use. This is essential to maintain aspects of compliance, privacy, and access to data and models within the project. Existing projects involving clinical/genomic data set generation by structures such as the Cancer Biomedical Informatics Grid (caBIG), The Cancer Genome Atlas, and trials such as the BATTLE trials and the I-SPY trial network provide precedents for establishing these governance rules and processes. Relevant policies may pertain to the use of data collected and how to ensure that the effort uses a sustainable funding model, among other topics.

Regulatory submissions

One potentially important issue involves the types of entities that might bring data forward to a regulatory agency as a result of patient-initiated study participation. Given that such efforts may be spearheaded by nonprofit as well as commercial organizations, it is possible that a nonprofit organization, not affiliated with a commercial product sponsor, might develop and submit data on molecular markers associated with response/nonresponse to an approved drug for review by a regulatory agency. It is unclear whether there is a pathway for evidence to be brought to the FDA by these nontraditional sponsors. If the evidence pertained to a biomarker for treatment response in general and without reference to a particular drug under development, one potential pathway might be through the FDA's recently proposed qualification process for drug development tools. In such cases, if the biomarker is qualified, it could be incorporated into any future drug development based on the qualified context of use. However, the process for translation of evidence from patient-initiated study participation into labeling changes may still require clarification and consideration regarding whether data arose from a specific drug development program as opposed to a postmarketing study. Such changes might occur independently of the product sponsor and possibly without the sponsor's agreement. Arguably, these changes would likely be in the interest of the patient community and society in general, but they might not always be in the interest of product sponsors.

Principles for Effective Management of Patient-Initiated Data Collection

As with any clinical research, it is essential that data be of a standard form for analysis. In traditional research settings, standardization of multiple data sources is accomplished

through use of common data collection forms and adherence to common practices in form completion. Years of practice in the oncology community have produced a large library of these common data elements using terminologies and ontologies that are national and international standards. In a partnership among academia, industry, and the FDA, these elements and ontologies have been used to create a common information model that supports electronic regulatory submission. Wherever possible, data collection should leverage these and other standard information representations.

Data generated through patient-initiated study participation is unlikely to arise solely from the clinical research arena. Instead, data will arise from health care encounters in settings using a variety of information representation standards. In addition, clinical information represents only a single dimension among the multiple diverse types of data that must be captured, managed, and interconnected. Similar considerations exist for biospecimens, imaging data, and the molecular data that will be used to characterize the individual participants. This information must have common representation across the diverse organizations in multiple disparate locations acquiring and sharing variant dialects of data often captured in unstructured (narrative) form. The NCI's caBIG program has created such representations and a collection of tools, accessible as Web tools, which utilize them. However, as is the case for clinical information, the caBIG Integration Hub permits disparate types of information to be cross-mapped to a common representation. Researcher-generated data can then be collected in a standardized manner and captured in an infrastructure that can support reuse by other investigators as authorized by patients.

Aggregation and analysis of the complex, multidimensional data also requires novel infrastructure. The caBIG community has created data mart/data warehouse tools that facilitate the collection and effective use of the multidimensional clinical and molecular data through its caIntegrator capabilities. These tools effectively manage the large volume and complexity of data for projects such as The Cancer Genome Atlas.

Next Steps

Patient-initiated study participation is a potentially promising way of rapidly generating evidence to support better targeting of previously approved cancer therapies. Cancer patients, caregivers, and their advocates have demonstrated strong enthusiasm for improving the efficiency of clinical research. It is possible that better patient education will enhance the quality of data collected through patient-initiated participation in clinical studies. However, some limitations should be acknowledged and addressed, including potential issues in selection of patients for these types of studies. For example, it is possible that certain highly motivated patients may be disproportionately represented, which could influence results obtained.

We have proposed a model that leverages the motivation and commitment of cancer patients to overcome some of the challenges in the collection of data and biospecimens that can be used to identify biomarkers predictive of nonresponse to previously approved chemotherapeutic agents. Care should be taken to ensure that such studies are designed with broad-based input from all stakeholders so that patients are informed appropriately, the correct types of data and biospecimens are collected, information is compiled and managed efficiently, the resulting database is made available to researchers with appropriate protections and security features in place, and that the data are analyzed in a way that yields evidence of sufficient quality to inform regulatory decisions and clinical practice.

In order to determine the true potential of patient-initiated study participation, pilot efforts are an important first step. These pilot studies will necessarily be informed by the ongoing data collection efforts of advocacy organizations. To inform pilot studies, a Guidance from FDA regarding what data will be considered actionable for labeling changes would be helpful. This information could be gathered unofficially as part of meetings that convene regulatory

authorities, industry representatives, patient groups, and academia around this issue.

Disclosure of Potential Conflicts of Interest

Dr. Schilsky reports having served as a member of the Board of Directors for Universal Oncology, Inc. and as a member of the Scientific Advisory Board of Foundation Medicine, Inc. No potential conflicts of interest were disclosed by the other authors.

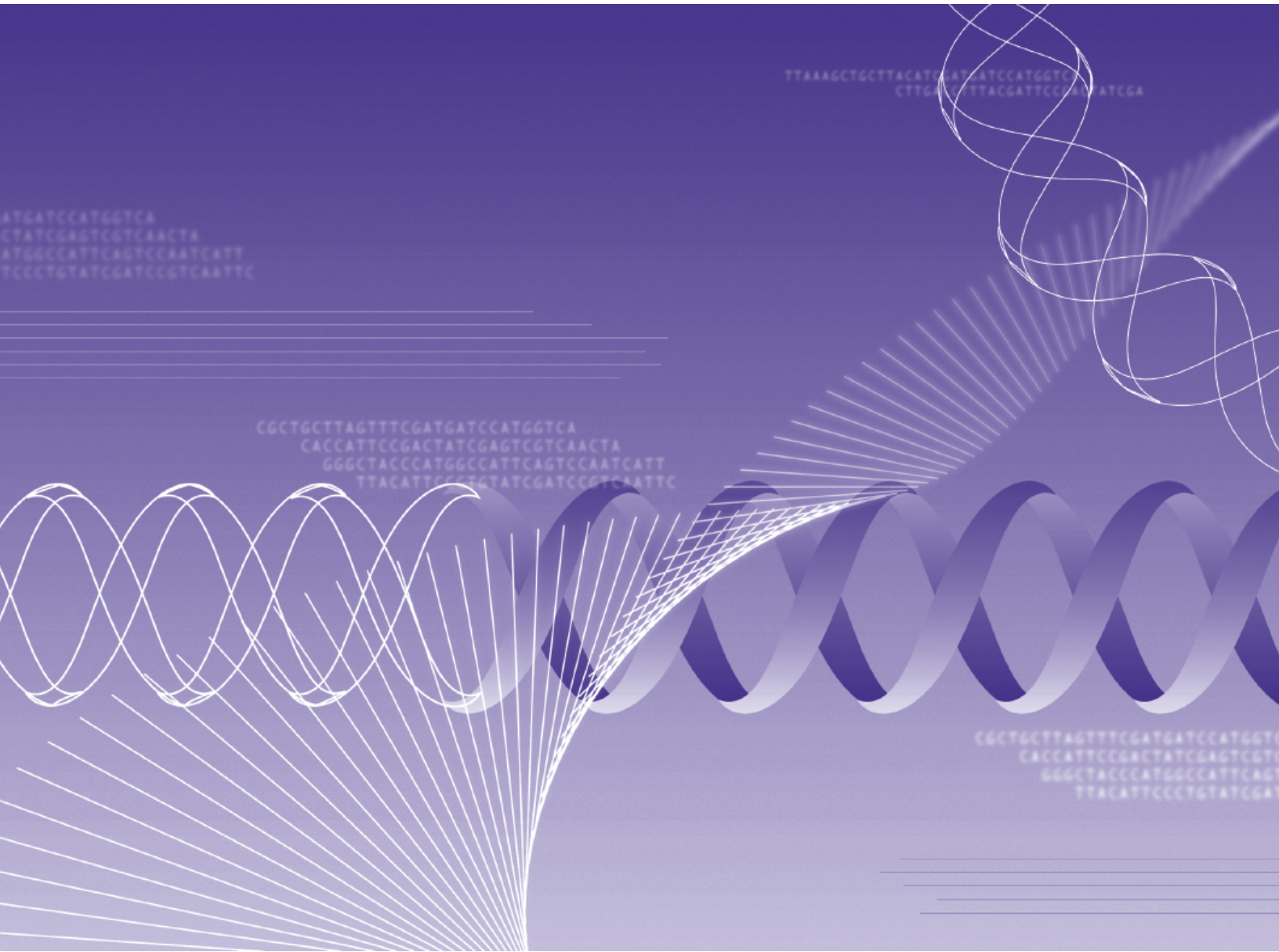
Acknowledgments

The authors gratefully acknowledge the contributions of several individuals. Sue-Jane Wang (U.S. Food and Drug Administration) offered valuable insights during the development of the concept for this paper and participated in the panel presentation at the 2010 Conference on Clinical Cancer Research. Roy Herbst (Yale School of Medicine), Annie Martin (GlaxoSmithKline), and Ignacio I. Wistuba (The University of Texas MD Anderson Cancer Center) provided targeted advice as the authors refined the concept. Joshua S. Benner, Erin Karnes, and Ed Walters (all from the Engelberg Center for Health Care Reform, The Brookings Institution) as well as Elaine Khoong (medical student at Washington University School of Medicine) helped to coordinate panel discussions and provided research and drafting assistance throughout the process.

Received May 6, 2011; revised July 5, 2011; accepted August 25, 2011; published online November 1, 2011.

References

1. Roberts SA, Karnes EK, Allen JD, Benner JS, Sigal EV, McClellan M. Achieving the goals of effective, safe, and individualized cancer care. *Clin Cancer Res* 2011;17:6632-3.
2. Scher HI, Fuld Nasso S, Rubin EH, Simon R. Adaptive clinical trial designs for simultaneous testing of matched diagnostics and therapeutics. *Clin Cancer Res* 2011;17:6634-40.
3. Turteltaub KW, Davis MA, Burns-Naas LA, Lawton MP, Clark AM, Reynolds JA. Identification and elucidation of the biology of adverse events: the challenges of safety assessment and translational medicine. *Clin Cancer Res* 2011;17:6641-5.
4. Cleeland CS, O'Mara A, Zagari M, Baas C. Integrating pain metrics into oncology clinical trials. *Clin Cancer Res* 2011;17:6646-50.
5. Friend SH, Golub T, Radich J, Sawyers C, Schilsky R. Identification of non-responders to approved cancer drugs through patient-oriented sample and data collection: Strategic Summary and AML Pilot. Overview for NCI Meeting, Sep 23, 2010. [cited 2011 Oct 13]. Available from: http://sagebase.org/downloads/NRP_Appendices.pdf.
6. Friend SH, Schilsky RL. Sage Project: finding non-responders to approved drugs through patient oriented sample/data collection and hosted in the public domain. Mar 2010.
7. Cetuximab (Erbix) and panitumumab (Vectibix) [database on the Internet]. Silver Spring (MD): U.S. Food and Drug Administration. [cited 2011 Mar 21]. Available from: <http://www.fda.gov/AboutFDA/CentersOffices/CDER/ucm172905.htm>.
8. Comis RL, Colaizzi DD, Miller JD. Barriers to cancer clinical trials (CCT) participation: "We have met the enemy and he is us." ASCO Annual Meeting Proceedings (Post-Meeting Edition). *J Clin Oncol* 2007; 25(18S): abstr 6567.
9. Wang SJ, Cohen N, Katz DA, Ruano G, Shaw P, Spear B. Retrospective validation of genomic biomarkers—what are the questions, challenges and strategies for developing useful relationships to clinical outcomes—Workshop Summary. *Pharmacogenomics J* 2006;6:82-8.
10. Wang SJ, O'Neill RT, Hung HMJ. Some statistical considerations in evaluating pharmacogenomics confirmatory clinical trials. *Clin Trials* 2010;7:525-36.
11. Love / Avon army of women [homepage on the Internet]. [cited 2010 Sep 23]. Available from: <http://www.armyofwomen.org/>.
12. Health of women study [database on the internet]. [cited 2010 Sep 23]. Available from: http://www.armyofwomen.org/HOW_Study.
13. Lung Cancer Alliance. What is Give a Scan? [cited 2010 Sep 23]. Available from: <http://www.lungcanceralliance.org/giveascan>.
14. The Clarity Foundation. [cited 2010 Sep 23]. Available from: <http://www.clarityfoundation.org>.
15. About the MMRC [homepage on the Internet]. [cited 2010 Sep 23]. Available from: http://www.themmr.org/about_mmr.php.
16. Reimer S. Dr. Love builds an anti-cancer army [monograph on the Internet]. Baltimore (MD): The Baltimore Sun; [updated 2011 May 12; cited 2011 Jun 29]. Available from: http://articles.baltimoresun.com/2011-05-12/health/bs-hs-love-profile-20110512_1_breast-cancer-cervical-cancer-susan-love.
17. Scagliotti GV, De Marinis F, Rinaldi M, Crinò L, Gridelli C, Ricci S, et al. Phase III randomized trial comparing three platinum-based doublets in advanced non-small-cell lung cancer. *J Clin Oncol* 2002;20:4285-91.
18. National Cancer Institute best practices for biospecimen resources. Bethesda, MD: National Cancer Institute; 2007. [cited 2011 Jun 20]. Available from: http://biospecimens.cancer.gov/global/pdfs/NCI_Best_Practices_060507.pdf.



ATGATCCATGGTCA
CTATCGAGTCGTCAACTA
ATGGCCATTTCAGTCCAATCATT
TCCCTGTATCGATCCGTCAATTC

TTAAAGCTGCTTACATGATATCCATGGTC
CTTGAATTACGATTCCCAATCGA

CGCTGCTTAGTTTCGATGATCCATGGTCA
CACCATTCCGACTATCGAGTCGTCAACTA
GGGCTACCCATGGCCATTTCAGTCCAATCATT
TTACATTCCTGTATCGATCCGTCAATTC

CGCTGCTTAGTTTCGATGATCCATGGTC
CACCATTCCGACTATCGAGTCGTCAACTA
GGGCTACCCATGGCCATTTCAGTCCAATCATT
TTACATTCCTGTATCGATCCGTCAATTC