Improving the Quality of Cancer Clinical Trials: Workshop Summary

Margie Patlak and Sharyl Nass, Rapporteurs

ISBN: 0-309-11669-4, 136 pages, 6 x 9, (2008)

IMPROVING THE QUALITY OF CANCER CLINICAL TRIALS

WORKSHOP SUMMARY

National Cancer Policy Forum

Margie Patlak and Sharyl Nass, *Rapporteurs*

INSTITUTE OF MEDICINE
OF THE NATIONAL ACADEMIES

THE NATIONAL ACADEMIES PRESS
Washington, D.C.
www.nap.edu

THE NATIONAL ACADEMIES PRESS 500 Fifth Street, N.W. Washington, DC 20001

NOTICE: The project that is the subject of this report was approved by the Governing Board of the National Research Council, whose members are drawn from the councils of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine.

This study was supported by Contract Nos. HHSN261200611002C, 200-2005-13434 TO #1, HHSP233200700373P, and 223-01-2460 TO #27, between the National Academy of Sciences and the National Cancer Institute, the Centers for Disease Control, the Agency for Healthcare Research and Quality, and the Food and Drug Administration, respectively. This study was also supported by the American Cancer Society, the American Society for Clinical Oncology, C-Change and the Association of American Cancer Institutes. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the view of the organizations or agencies that provided support for this project.

International Standard Book Number-13: 978-0-309-11668-8
International Standard Book Number-10: 0-309-11668-6

Additional copies of this report are available from the National Academies Press, 500 Fifth Street, N.W., Lockbox 285, Washington, DC 20055; (800) 624-6242 or (202) 334-3313 (in the Washington metropolitan area); Internet, http://www.nap.edu.

For more information about the Institute of Medicine, visit the IOM home page at: **www.iom.edu.**

Printed in the United States of America

The serpent has been a symbol of long life, healing, and knowledge among almost all cultures and religions since the beginning of recorded history. The serpent adopted as a logotype by the Institute of Medicine is a relief carving from ancient Greece, now held by the Staatliche Museen in Berlin.

Suggested citation: IOM (Institute of Medicine). 2008. *Improving the quality of cancer clinical trials: Workshop summary.* Washington, DC: The National Academies Press.

"Knowing is not enough; we must apply.
Willing is not enough; we must do."
—Goethe

INSTITUTE OF MEDICINE
OF THE NATIONAL ACADEMIES

Advising the Nation. Improving Health.

THE NATIONAL ACADEMIES

Advisers to the Nation on Science, Engineering, and Medicine

The **National Academy of Sciences** is a private, nonprofit, self-perpetuating society of distinguished scholars engaged in scientific and engineering research, dedicated to the furtherance of science and technology and to their use for the general welfare. Upon the authority of the charter granted to it by the Congress in 1863, the Academy has a mandate that requires it to advise the federal government on scientific and technical matters. Dr. Ralph J. Cicerone is president of the National Academy of Sciences.

The **National Academy of Engineering** was established in 1964, under the charter of the National Academy of Sciences, as a parallel organization of outstanding engineers. It is autonomous in its administration and in the selection of its members, sharing with the National Academy of Sciences the responsibility for advising the federal government. The National Academy of Engineering also sponsors engineering programs aimed at meeting national needs, encourages education and research, and recognizes the superior achievements of engineers. Dr. Charles M. Vest is president of the National Academy of Engineering.

The **Institute of Medicine** was established in 1970 by the National Academy of Sciences to secure the services of eminent members of appropriate professions in the examination of policy matters pertaining to the health of the public. The Institute acts under the responsibility given to the National Academy of Sciences by its congressional charter to be an adviser to the federal government and, upon its own initiative, to identify issues of medical care, research, and education. Dr. Harvey V. Fineberg is president of the Institute of Medicine.

The **National Research Council** was organized by the National Academy of Sciences in 1916 to associate the broad community of science and technology with the Academy's purposes of furthering knowledge and advising the federal government. Functioning in accordance with general policies determined by the Academy, the Council has become the principal operating agency of both the National Academy of Sciences and the National Academy of Engineering in providing services to the government, the public, and the scientific and engineering communities. The Council is administered jointly by both Academies and the Institute of Medicine. Dr. Ralph J. Cicerone and Dr. Charles M. Vest are chair and vice chair, respectively, of the National Research Council.

www.national-academies.org

WORKSHOP PLANNING COMMITTEE

JOHN MENDELSOHN, M.D. (Co-Chair), President, MD Anderson Cancer Center

HAROLD MOSES, M.D. (Co-Chair), Professor of Molecular Oncology, Cancer Biology, Medicine, and Pathology; Director Emeritus, Vanderbilt-Ingram Cancer Center, Vanderbilt University Medical Center

JAMES DOROSHOW, M.D., Director, Division of Cancer Treatment and Diagnosis, National Cancer Institute

ROY HERBST, M.D., Ph.D., Professor and Chief, Thoracic/Head and Neck Medical Oncology, MD Anderson Cancer Center

HEDVIG HRICAK, M.D., Professor and Chair, Department of Radiology, Memorial Sloan-Kettering Cancer Center

GILBERT OMENN, M.D., Ph.D., Professor, Internal Medicine, Human Genetics, and Public Health, University of Michigan Medical School

DAVID PIWNICA-WORMS, M.D., Ph.D., Professor, Department of Radiology, Washington University School of Medicine

JOHN WAGNER, M.D., Ph.D., Executive Director, Department of Clinical Pharmacology, Merck, Inc.

JANET WOODCOCK, M.D., Deputy Commissioner and Chief Medical Officer, Food and Drug Administration

---

IOM planning committees are solely responsible for organizing the workshop, identifying topics, and choosing speakers. The responsibility for the published workshop summary rests with the workshop rapporteurs and the institution.

NATIONAL CANCER POLICY FORUM

MARGARET SPITZ, M.D., Chair of Epidemiology, MD Anderson Cancer Center
ELLEN STOVALL, President and Chief Executive Officer, National Coalition for Cancer Survivorship
BARRY STRAUBE, M.D., Director, Office of Clinical Standards and Quality, Centers for Medicare & Medicaid Services
JANET WOODCOCK, M.D., Deputy Commissioner and Chief Medical Officer, Food and Drug Administration

Staff

SHARYL NASS, Ph.D., Study Director
ROGER HERDMAN, M.D., Director, National Cancer Policy Forum
LAURA LEVIT, J.D., Research Associate
CHRISTINE MICHEEL, Ph.D., Christine Mirzayan Science and Technology Policy Graduate Fellow
MICHAEL PARK, Senior Program Assistant
ALEX REPACE, Senior Program Assistant
MARGIE PATLAK, Science Writer

IOM forums and roundtables do not issue, review, or approve individual documents. The responsibility for the published workshop summary rests with the workshop rapporteurs and the institution.

Reviewers

This report has been reviewed in draft form by individuals chosen for their diverse perspectives and technical expertise, in accordance with procedures approved by the National Research Council's Report Review Committee. The purpose of this independent review is to provide candid and critical comments that will assist the institution in making its published report as sound as possible and to ensure that the report meets institutional standards for objectivity, evidence, and responsiveness to the study charge. The review comments and draft manuscript remain confidential to protect the integrity of the deliberative process. We wish to thank the following individuals for their review of this report:

Ray DuBois, M.D., Ph.D., University of Texas MD Anderson Cancer Center
Doug Hanahan, Ph.D., University of California, San Francisco
Hedvig Hricak, M.D., Ph.D., Memorial Sloan-Kettering Cancer Center
Susan M. Jerian, M.D., OncoRD, Inc.
Daniel C. Sullivan, M.D., Duke Cancer Center

Although the reviewers listed above have provided many constructive comments and suggestions, they were not asked to endorse the final draft of the report before its release. The review of this report was overseen

by **Melvin Worth, M.D.** Appointed by the Institute of Medicine, he was responsible for making certain that an independent examination of this report was carried out in accordance with institutional procedures and that all review comments were carefully considered. Responsibility for the final content of this report rests entirely with the authors and the institution.

Contents

Improving the Quality of Cancer Clinical Trials: Workshop Summary[1]

INTRODUCTION

The science underpinning cancer drug development has been changing rapidly in recent years because of a more mechanistic understanding of cancer. Today, hundreds of cancer therapeutics are in development, and many target specific molecules, genes, or pathways. To be most effective, preclinical studies indicate that many of these drug candidates need to be combined with other targeted agents, reflecting the complexity of multistep carcinogenesis.

Clinical trials must receive regulatory approval from the Food and Drug Administration (FDA) for these innovative drug candidates before bringing them into clinical use. Not only are these trials expensive and lengthy, but they are extremely prone to failure because prediction of efficacy and toxicity in humans from findings in animal models has often proved unreliable, as has early testing in humans. Only a small percentage of drug candidates ultimately become useful therapies. The novel and more mechanistically based cancer drugs may be even more inclined to fail traditional clinical trials, which are not tailored to the combination testing that may be required, or to the different standards or procedures needed when

[1]The planning committee's role was limited to planning the workshop, and the workshop summary has been prepared by the workshop rapporteurs as a factual summary of what occurred at the workshop.

agents are effective only in small subpopulations. For these and other reasons, scientists and clinicians seek a new paradigm that could improve the efficiency, cost-effectiveness, and overall success rate of cancer clinical trials, while maintaining the highest standards of quality. To explore innovative paradigms for cancer clinical trials and other ways to improve their quality, the National Cancer Policy Forum held a workshop, "Improving the Quality of Cancer Clinical Trials," in Washington, DC, on October 4 and 5, 2007. As Dr. John Mendelsohn explained, the main goals of the workshop were to examine new approaches to clinical trial design and execution that would (1) better inform decisions and plans of those responsible for developing new cancer therapies, (2) more rapidly move new diagnostic tests and treatments toward regulatory approval and use in the clinic, and (3) be less costly than current trials. At the workshop, experts gave presentations in one of five sessions:

- New clinical trial designs, including exploratory investigational new drug (IND) and Phase 0 trials, adaptive trials, trials that target multiple pathways with multiple drugs, and preclinical model systems that better inform clinical studies
- Molecular imaging, including molecular imaging strategies in drug development and how they can facilitate clinical trials
- Screening for predictive markers
- Collaborations among academia, the pharmaceutical or diagnostics industries, and government, including ways to reduce the costs and regulatory burdens of clinical trials and increase patient accruals
- Regulatory issues, including the laws, regulations, and policies that help or hinder improvements in cancer clinical trials

In addition, participants in five small-group discussions explored the following topics:

- Phase 0 trials
- Adaptive trial design
- Imaging
- Use of proteomics/genomics to assign therapy in lung cancer
- Use of genetics/genomics to assign therapy

This document is a summary of the conference proceedings, which will serve as input to the deliberations of an Institute of Medicine committee

that will develop consensus-based recommendations for moving the field of cancer clinical trials forward. The views expressed in this summary are those of the speakers and discussants, as attributed to them, and are not the consensus views of workshop participants or members of the National Cancer Policy Forum.

NEW CLINICAL TRIAL DESIGNS

Phase 0 Trials

The first session of the workshop was on new clinical trial designs. Dr. David Jacobson-Kram of the FDA began this session by giving his overview of the exploratory investigational new drug study and how it differs from the traditional IND study. The main purpose of the exploratory IND is to assess the likely therapeutic effectiveness of a compound, based on whether it affects its target in people and how long it is active in the body. An exploratory IND study tests a new experimental drug on human subjects prior to a Phase I clinical trial, which is the first traditional test of a compound in humans to assess safety and the dosing of subsequent trials. For that reason, the exploratory IND study is also called a Phase 0 trial.[2] Dr. Jacobson-Kram discussed the current problems in drug development and testing and how various types of exploratory IND studies might help assuage some of those problems.

As Dr. Jacobson-Kram noted, we currently face a crisis in drug development with the number of drugs in the pipeline declining, the number of drug failures increasing, and the costs of developing drugs rising. The FDA finds that less than 20 percent of new molecular entities progress through clinical trials to the point where approval for them is sought so they can enter the market as drugs. Currently about half of drugs in Phase III clinical trials fail because of toxicity or a lack of efficacy, Dr. Jacobson-Kram reported using FDA data. "That is really a disaster, because by the time you are in a Phase III trial you have invested an enormous amount of money, resources, and time," he said. The cost of developing a new molecular entity that makes it to the market is estimated to be nearly a billion dollars.

[2]To receive FDA approval for market, most drugs have to undergo three phases of clinical testing. Phase I testing determines safety and dose on a small number of individuals. Phase II testing is done on a larger group of volunteers to assess safety and effectiveness. If those tests are promising, a large-scale Phase III is usually done to confirm safety and effectiveness.

To help abate this crisis in drug development, the FDA published its guidance on exploratory INDs in January 2006. According to this guidance, the goals of an exploratory IND are to

- gain an understanding of the drug's mechanism of action and whether it affects a target relevant to a disease process (pharmacodynamics),
- provide information on how the drug is broken down by the body and how long it remains active (pharmacokinetics),
- indicate the most promising lead product from a group of candidate drugs designed to interact with a particular therapeutic target, and/or
- reveal where the drug is distributed in the body using various imaging technologies.

An exploratory IND study is done in a very small number of human subjects, with dosing up to 7 days, and is not designed to be therapeutic or assess the effectiveness of the experimental drug. "This is really important to keep in mind," said Dr. Jacobson-Kram. "These trials are not designed to treat patients. These are simply experiments that are being done in human beings."

The FDA's existing regulations are flexible in the amount of preclinical data it requires investigators to submit before conducting an exploratory IND. That data depends on the goals of the investigation, the testing being proposed, and the expected risks. For example, an exploratory IND that tests a single subpharmacologic drug dose would require a minimal dataset from a single animal species. More extensive data would be needed to conduct a repeated-dose clinical study designed to induce pharmacologic effects, but this expanded dataset would still be less than that required to initiate a traditional IND.

Exploratory INDs allow sponsors to evaluate up to five experimental drugs simultaneously in the clinic so as to better choose the most promising drug candidate to undergo traditional drug development and testing. Exploratory INDs can help to reduce the resources involved in drug development, including the amount of time and drug product needed to select promising drugs, and help to eliminate those that lack promise. The FDA guidance gives examples of several types of exploratory IND studies, including the microdose study, a study design developed by Pharmaceutical Research and Manufacturers of America (PhRMA), and a study design

proposed by the National Cancer Institute (NCI) to specifically study experimental cancer drugs.

The sole aim of a microdose study is to use imaging or other means to assess where in the body a compound is distributed and for how long it remains in these sites. A microdose is defined as less than 1/100th of the dose calculated to yield pharmacological effects, and less than 100 micrograms. A microdose study is designed *not* to induce pharmacologic effects; rather, it can indicate whether an experimental drug reaches its target.

The FDA assumes the risks of a microdose study are small and thus only requires a single study in a mammalian species, usually a rat, to assess safety prior to granting approval for a microdose exploratory IND study. The animals in this preclinical study would only have to be dosed a single time via the same route of administration that investigators would use in the exploratory IND study. The animal study must show a minimally toxic dose or show that the doses used in the microdose study would be well outside a toxic range. Genetic toxicology testing on the animals is not routinely needed. (The European Medicines Agency, in contrast, asks for additional safety data, including general toxicity studies using two ways of administering the compound, orally and intravenously, as well as in vitro genotoxicity studies.)

Another example of an exploratory IND study is the paradigm proposed to the FDA in 2004 by PhRMA. This study tests, in healthy subjects or minimally ill patients, up to five compounds that have a common biological target, but might not be structurally related. These compounds are given in up to seven repeated doses to assess pharmacological response, but not a maximum tolerated dose, as is determined in traditional Phase I studies. The risks in the PhRMA paradigm are greater than in the microdose study, so it requires genetic toxicity studies, as well as a repeated-dose toxicity study in rodents and another mammal, usually a dog. If the dog shows toxicity at a dose level that does not cause toxicity in the rat, the compound is not included in the exploratory IND, under the assumption that its toxicity had not been adequately evaluated to be tested in humans.

PhRMA used a database of 106 drugs tested in two species and in Phase I clinical trials to support the safety of its proposed exploratory IND using an analysis that assumed certain starting and stopping doses. That analysis revealed the trials would have been safe under the exploratory IND paradigm. In a presentation to the FDA, PhRMA discussed the advantages of its proposed exploratory IND versus a traditional IND (Table 1). The exploratory IND would accelerate discovery and development of new drugs,

TABLE 1 Comparison of the PhRMA Exploratory IND and the Conventional IND

	Conventional IND	PhRMA Exploratory IND
Active Pharmaceutical Ingredient (API)	• 1–3 kg	• 10–300 g
Preclinical Resources	• 9–12 studies • 220 rodent and 38 non-rodent • 9–18 months	• 5–6 studies • 170 rodent and 6 non-rodent • 3–6 months
Benefits	• Full toxicology profile • Escalation to maximum tolerated dose (MTD) in clinical trials • Progression directly to Phase II	• Predictable API requirement • Faster progression to clinical trials • Capability to evaluate candidates based on target activity • Better development decisions made more quickly • Early and less costly attrition
Disadvantages	• Larger quantity of API • Slower decisions • Late and costly attrition	• MTD not established • Potential delayed progression to Phase II

SOURCE: Jacobson-Kram presentation (October 4, 2007).

PhRMA claimed, because it would require a smaller number of animal studies that would take about one-third less time to perform using much less of the tested active ingredient of the drug. There would be a significant savings in non-rodent experimental animals, Dr. Jacobson-Kram pointed out. In addition, the exploratory IND would enable better development decisions to be made more quickly so there is early and less costly attrition of drugs that lack promise. This innovative IND would also give sponsors the ability to evaluate drug candidates based on target activity, and should enable faster progression to clinical trials. The only disadvantages cited for the exploratory IND were that it did not determine the maximum toler-

ated dose and thus could potentially delay the progression to Phase II trials, which entails closing the exploratory IND and opening a new, traditional IND, with the standard requirements for toxicology.

The FDA used its own data to analyze clinical studies that were preceded by 2-week or 4-week toxicology studies in two animal species and found that the PhRMA paradigm succeeded in identifying safe starting and stopping doses, but in many cases, the dogs or monkeys had lower no-observed-adverse-effect levels. In addition, an analysis of NCI data found that the toxicology data from the nonrodent species more closely resembled what is seen in humans than the data from rodents (Tomaszewski, 2004). "What this is saying is, at least for some cases, the dog better predicts the maximum tolerated dose in the clinical trial, so the exploratory IND might not be then as viable an option," said Dr. Jacobson-Kram.

He noted that the NCI developed its own version of an exploratory IND for oncology drugs. For what the agency termed "first-in-man" studies, researchers should aim to assess the blood levels of the drug needed to induce the desired effect, instead of focusing on toxicity and basing doses for future studies on such toxicology findings. The NCI exploratory IND is used to select promising drugs for life-threatening diseases, primarily cancers, with up to 3 days of dosing in the clinic. Participants for these tests are terminally ill patients without therapeutic options; but because there is no therapeutic intent in the studies, the safety bar is the same as it would be for healthy volunteers. "The thinking is if you are just doing an experiment, why would you make sick people sicker?" said Dr. Jacobson-Kram. In a later presentation, Dr. James Doroshow of NCI added that researchers need to address the ethical issues linked to an exploratory IND by consulting with their research oversight committees, Institutional Review Boards, to develop a process to obtain the appropriate informed consent from patient volunteers in these studies. According to Dr. Doroshow, because the exploratory IND is not considered therapy, participation in a Phase 0 trial should not preclude patient volunteers from then proceeding immediately to another clinical trial; the usual 3- to 4-week period between studies is not required in these cases.

Despite these various Phase 0 study options, the FDA has received only a handful of exploratory INDs, Dr. Jacobson-Kram reported (although it was added later during the discussion that the agency's recordkeeping of this may not be complete). "Although PhRMA was very excited about this possibility, in the 2 years that this tool has been available it has been used very sparingly," he said. He offered several reasons for why exploratory INDs are

not being done more often by drug sponsors, including the slowness of the established drug industry to adopt a new paradigm and the potential that microdose studies do not accurately predict what is likely to be seen with doses in the therapeutic range. But perhaps the biggest stumbling block to more widespread use of an exploratory IND, according to Dr. Jacobson-Kram, is excessive optimism on the part of a drug development team. "No development team thinks that their drug is a loser. So they don't want to use a tool that is going to kill their drug early, because they are convinced it is going to be a winner," he said.

In the discussion that followed Dr. Jacobson-Kram's presentation, participants voiced more reasons for hesitancy to adopt exploratory INDs. Oncology researcher Dr. Giulio Draetta of Merck noted that although the exploratory IND is an excellent concept that he and his colleagues welcome, "no established clinical oncologist inside or outside the company would think of these Phase 0 trials as being important for reaching a go or no-go decision about a drug," he said, implying that more knowledge is needed for such a decision.

Dr. Jacobson-Kram countered that exploratory INDs offer more than decisions on whether to move a drug forward in the clinical testing hierarchy. "With the current paradigm, from the tens of thousands of different structures you synthesize every day, you only take one into the clinic, and that is a big decision. But if you could take a handful of them in people and find the one that looks the most promising, based on clinical data, I think you have a much better chance of succeeding than just choosing that single one based on preclinical data," he said.

Another participant from Merck, Dr. John Wagner, concurred with Dr. Draetta that all of Merck's exploratory INDs have been in oncology, and asked what can be done to improve the usefulness of an exploratory IND for oncology purposes. Dr. George Mills, who, when he was at the FDA, helped develop the agency's guidance on exploratory INDs, responded by stressing the usefulness of an exploratory IND that uses imaging to determine which drug in a pool of candidates is the *most* promising. "All drugs will be promising at some level," he said. Further, Dr. Mills commented that clarification of which drug to focus on and thus accelerate the decision-making process for the group of drug candidates comes when rates and routes of clearances and target and non-target organ distribution are analyzed. Dr. Jerry Collins of NCI added that another advantage of the exploratory IND is that "it is an open invitation to a dialogue with the FDA." He added that an exploratory IND reduces the number of toxicology studies needed, which is a distinct

advantage for academic researchers, many of whom lack the expertise or resources to conduct such testing. Dr. Jacobson-Kram summed up his talk by stressing the FDA's commitment to improving the "critical path" to new medical products and sees the implementation of exploratory INDs as an important means for carrying out that commitment.

Dr. Mills, Vice President of Perceptive Informatics, Inc., expanded on some of the points Dr. Jacobson-Kram made, but narrowed the focus of his talk to the use of molecular imaging and linked nanotechnology techniques in exploratory INDs. He noted that small pharmaceutical companies and biotechnology companies developing biologic drugs are particularly keen on using exploratory INDs that employ imaging because this approach literally enables investors to visualize the likely effectiveness of a potential drug compound by showing if it hits targets such as tumors, abscesses, or the amyloid plaques in Alzheimer disease patients, and whether it is relatively absent in the liver, kidney, or other organs where it could pose toxicity problems (Figure 1). "These companies have limited amounts of funds and need rapid proof-of-concept for the investment community," Dr. Mills said. "You can take an image and show it to the investment banking industry person, who doesn't understand our world, but understands from the image that this drug does go to colorectal cancer and the other ones don't."

But particularly for oncology applications, it is not sufficient for a drug to just reach its target and be concentrated there. Its effectiveness or toxicity also depends on its duration in the target tissues as well as other parts or the body. An exploratory IND that uses imaging can show this effectively, he said. Radiation dosimetry studies can reveal rates and routes of clearance much more quickly and simply than the standard techniques used to determine these endpoints in Phase I studies, he claimed. After his presentation, audience participant Dr. Tim McCarthy from Pfizer pointed out that although an exploratory IND that uses imaging can reveal distribution data to compare compounds, it does not provide information about specificity of the target. Dr. Mills responded that he expected new software and perhaps combination products that might provide that specificity information in the future.

Dr. Mills added that the reduction in pharmacology and toxicology studies that an exploratory IND offers, especially one with an imaging component, is another incentive to drug companies. Many companies, he said, have several preclinically developed drug candidates, but are unwilling or unable to devote the financial resources to do the pharmacology and toxicology studies needed to take them to the next step. "The exploratory

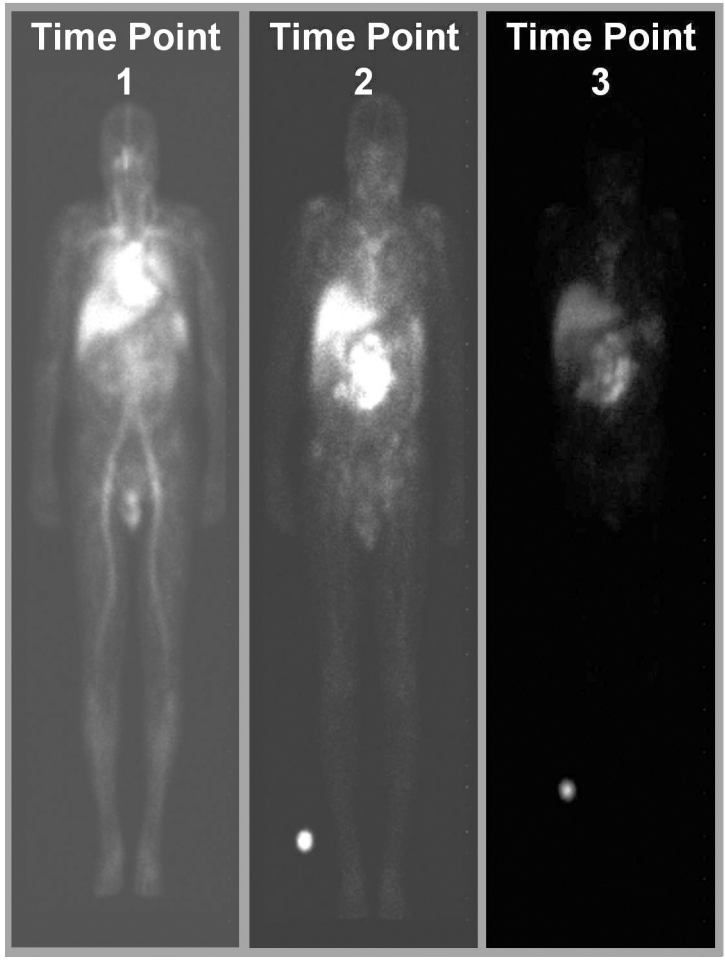

FIGURE 1 Whole-body biodistribution imaging. Time points 1, 2, and 3 show a radiolabeled bio-distribution study of a therapeutic agent as it targets an abdominal tumor. Time point 1 shows no tumor localization in the mid-abdomen; time point 2 shows localization in the abdomen; and time point 3 demonstrates routine clearance of the labeled agent from the body. Brightness of signal corresponds to density of therapeutic agent.
SOURCE: Mills presentation (October 4, 2007).

IND allows those products to come off the shelf and come into human experience to be able to determine if they are going to be promising," he said. A pre-IND development teleconference with the FDA can determine the exploratory IND's minimum pharmacology, toxicology, chemistry, and manufacturing and control information needed for all products evaluated, Dr. Mills said.

Dr. Mills also stressed the advantages of being able to make simultaneous comparative assessments of competitive drug compounds in a single study. Drug developers can also do imaging studies to see how new drug compounds compare to standard therapies. Those that do not perform better than the standard treatment, in terms of distribution and persistence in various regions of the body, are not developed further. He pointed out that sequential assessments of competitive drug compounds can also be done with a series of exploratory INDs so that the first to perform adequately moves on to Phase I trials, and no further testing is done on other similar compounds. Smaller drug companies tend to pursue this vertical approach to Phase 0 testing because it is more cost- and time-effective than the horizontal approach where compounds are compared simultaneously, Dr. Mills said (Figure 2).

Exploratory INDs can also address the concern recently raised by those pursuing nanotechnology that, when particle size is changed, the potential safety profile is changed as well. "With an exploratory IND, you can do comparative imaging to determine if particle size change will alter the distribution. It is very straightforward and immediate," he said. Some companies are developing nanoparticles to carry both a therapeutic and an imaging marker, he added.

Dr. Mills summarized his talk by concluding, "Exploratory INDs that use imaging can, in 5, 10, or 15 subjects, effectively let you make those business decisions that are so necessary and cost-effective in drug development."

Dr. Mills' talk was followed by a presentation on how best to use Phase 0 clinical trials in cancer drug development, given by Dr. James Doroshow of the NCI. Dr. Doroshow discussed the recent shift in cancer drug development from traditional cytotoxic chemotherapies for cancers to drugs that act on specific molecular signaling targets. This shift has created a need early on in drug development for reliable and sensitive tests that reveal if the drug is affecting its target, as well as confirmation of this in people before initiating large clinical trials to assess the drug's effectiveness. Phase 0 studies can address that need and establish standard operating procedures

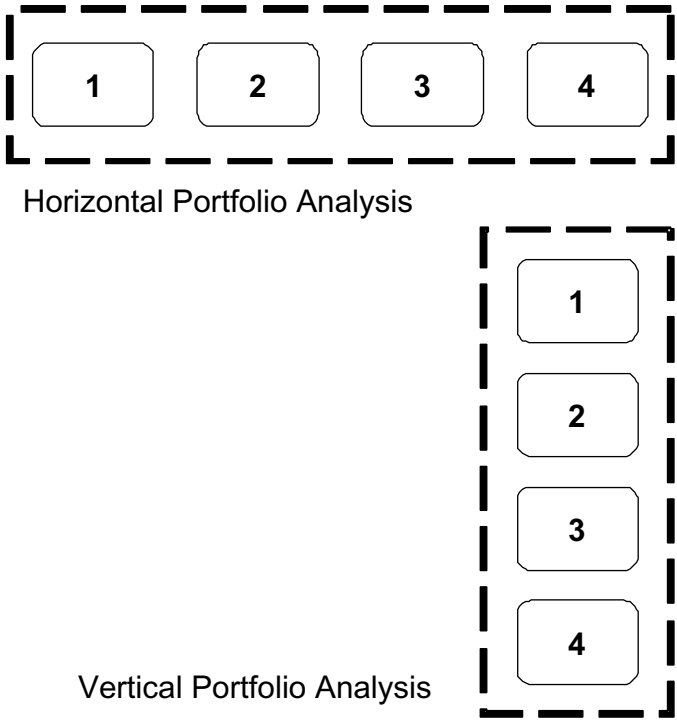

FIGURE 2 Exploratory IND assessment schemes of competitive drug compounds. The horizontal portfolio analysis assesses biodistribution of drug compounds simultaneously, regardless of performance and cost. The vertical portfolio analysis is a sequential, "top-down" assessment. In other words, "the first to perform, wins"; this is a cost-effective and time-effective approach. In the figure, 1–4 represent exploratory INDs dependent on chemistry, manufacturing, and controls (CMC) and pharmacology/toxicology.
SOURCE: Mills presentation (October 4, 2007).

needed to appropriately gather data in subsequent clinical studies, according to Dr. Doroshow. Researchers can also use findings from Phase 0 studies to closer approximate a safe, but potentially effective starting dose and limit the patient tissue sampling required in subsequent trials. "These are experiments that need to be performed to allow you to adequately inform the clinical trial, and even though they are not hypothesis driven, they are critical to the process," he said.

Dr. Doroshow pointed out that, for tests of a drug's effectiveness on tumor cells (or surrogate markers in the blood), clinical researchers often

do not concern themselves with accurately duplicating in people how those samples were acquired and processed in animals for the same tests. But variability in these standard operating procedures (SOPs) can affect the accuracy of the tests in clinical trials. Dr. Doroshow advocated using an exploratory IND to fine-tune SOPs for human subjects and create the appropriate bridge between what is done preclinically to what is done clinically.

Dr. Doroshow gave an example of an exploratory IND he and his colleagues conducted on the ability of a drug to inhibit the activity of the DNA repair enzyme poly(ADP-ribose) polymerase, known as PARP, in tumors, and how long that inhibition lasted. Before conducting this study, the researchers developed a sensitive test for PARP inhibition in tumor tissues and determined the SOPs for tissue removal, processing, and testing that were followed in the animal studies. "We tried to model the entire clinical experiment in a mouse—we had a veterinarian pretend that she was a radiologist and handle all the tissues the same way they would be handled in people," Dr. Doroshow said later in response to a question posed by an audience participant.

The exploratory IND study was done on only 13 patients, yet it gave the investigators the information needed to consider how best to combine the experimental drug with other cancer drugs in future clinical trials. Dr. Doroshow said this information was acquired much more quickly than in a traditional IND study that determines the maximum tolerated dose, yet lacks information on how long the drug affects its target.

Dr. Doroshow pointed out that Phase 0 studies, such as the example he gave, are best done on targeted drugs with a fairly wide therapeutic index, as opposed to traditional toxic chemotherapy drugs that have a much narrower range of doses in which they can be safely used. He also noted that his enthusiasm for conducting such studies would be dampened for experimental drugs that lack an accurate and reliable test for the drugs' effects on targets. "If you are going to go to the trouble of trying to do a proof-of-principle study, there has to be a principle to prove." He also reiterated the importance of researchers using Phase 0 studies to fine-tune their methods in people prior to progressing to large clinical trials. "It makes sense to take a small number of patients, ask them to volunteer, and to evaluate and develop your methodology prior to using them on a broader scale," he said.

Much of the discussion that followed the Phase 0 presentations focused on how to fund exploratory IND studies. Dr. Richard Schilsky of the University of Chicago pointed out that the Phase 0 study example that

Dr. Doroshow gave required an extensive research team, including surgeons who removed the tumor biopsies from patients and technicians who ran the tests on those samples (Figure 3). "An academic investigator who doesn't have a drug sponsor to work with may be facing some formidable challenges in putting this kind of team together," Dr. Schilsky said. Dr. Doroshow agreed and pointed out that because Phase 0 studies are done without therapeutic intent, health insurers are not likely to reimburse costs linked to the study, including computed tomography (CT) scans, biopsies, blood tests, etc. He estimated that, assuming a test to assess whether an experimental drug was affecting its target (pharmacodynamic assay) was already developed, the clinical costs of running an exploratory IND would be approximately $10,000 a patient.

One way to address the funding issue would be to use General Clinical Research Centers (GCRCs) or other government-funded institutions to conduct Phase 0 studies, said audience participant Dr. David Parkinson from Nodality, Inc. He pointed out that even drug sponsors might balk at the high hospital and surgeon expenses linked to doing the tumor biopsies as was done for Dr. Doroshow's exploratory IND study. "If you could take that into a GCRC-type mechanism, the surgeons become investigators, and it is out of the hospital billing system," he said. Dr. Doroshow concurred that using GCRCs for Phase 0 studies would be appropriate, as would using resources of the newly established Clinical and Translational Science Awards consortium, which is funded by the National Institutes of Health's (NIH's) National Center for Research Resources.[3]

Dr. Parkinson also noted that if Phase 0 studies were viewed as screening strategies to determine which patients could then benefit from a Phase I study of an experimental drug, they might be reimbursed by insurers. But Dr. Doroshow said later in the discussion that one cannot automatically proceed from a Phase 0 to a Phase I trial using the same patients for both without first having done the toxicology studies needed to proceed to a multidose investigation. "If you have already done that up front, that's fine, but if you haven't, you would have to now stop and do that as well. You can't just roll over from one to the other without having the toxicology base for safety for a traditional Phase I study," Dr. Doroshow said. He is actively involved in developing a molecule that will simultaneously be given in a

[3]For more information see http://www.ncrr.nih.gov/clinical_research_resources/clinical_and_translational_science_awards/.

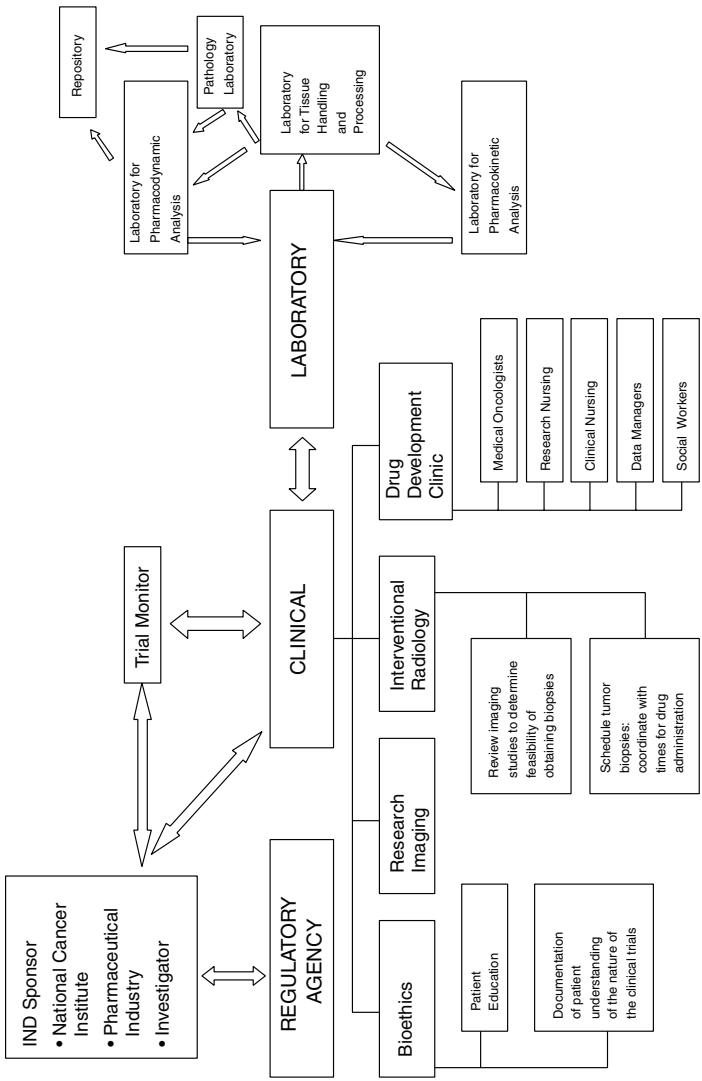

FIGURE 3 NCI integrated Phase 0 research team.
SOURCE: Doroshow presentation (October 4, 2007).

Phase 0–Phase I setting as a therapeutic, but with an imaging component to identify if the compound is hitting its target.

Another issue raised by discussant Steve Litwin of Biologics Consulting Group was how applicable exploratory INDs are to drugs produced using biotechnology, which are termed "biologics." His experience indicates that pharmacodynamic and pharmacokinetic data are not useful for biologics. Moreover, simultaneously testing a group of biologics aimed at the same substance may not work because there often are much larger differences between such compounds, he added. "The three licensed anti–tumor necrosis factor (anti-TNF) drugs have somewhat similar effects, but differ enormously and very importantly in the type of opportunistic infections the patients are prone to," he said. Dr. Jacobson-Kram countered that one could still use an exploratory IND to compare changes in the primary sequence or formulations of biologics that could affect how they are distributed in the body. "I think there really is a role for these types of studies in biologics," he said. Dr. Mills added that the horizontal drug testing model he presented was actually based on successful testing that was done on biologics.

Adaptive Trial Designs

In the next session of the conference, biostatisticians Dr. Donald Berry of the MD Anderson Cancer Center in Texas and Dr. Susan Ellenberg from the University of Pennsylvania gave presentations on adaptive trial designs. Dr. Ellenberg noted that "adaptive" simply means that one or more decision points are built into the trial design. How the trial proceeds following each decision point depends on the data observed up to that point. She and Dr. Berry described the many kinds of adaptive trials, including those which, during the course of the trial, adapt their stop date, range of doses tested, degree of randomization, and types of populations accrued and tested (Berry, 2005, 2006).

One of the more commonly used adaptive trial designs is one that stops early or continues based on results that indicate how effective the treatment under study is after a limited number of patients have been tested. The standard design for Phase II cancer trials has been of this type for many years, Dr. Ellenberg noted. It is also standard in Phase I cancer studies to have adaptive trials with dose escalation schemes dependent on observed toxicity at each stage, she said. Adaptive trial designs can also be used to make a seamless transition between phases in cancer clinical trials, Dr. Berry noted. The data from a Phase I portion of the trial, for example, determines the

design of the Phase II portion. The possibility of stopping a trial early on the basis of very positive results or very negative results is built into nearly every Phase III cancer trial. The so-called "stopping boundary" for superiority of an experimental agent is usually very conservative and therefore most Phase III trials accrue to their maximally targeted sample sizes.

Less commonly used adaptive trials are those that have adaptive borrowing, adaptive randomization, adaptive study populations, or adaptive accrual rates. Adaptive borrowing incorporates historical control data or data from other studies in the final study's conclusion. Dr. Berry noted that pharmaceutical companies may use adaptive borrowing to conduct, in a single trial, studies of a cancer drug in several different types of cancers. Dr. Berry mentioned several studies conducted at MD Anderson Cancer Center that have used an adaptive randomization design when testing various cancer drug combinations in a multiarm study. For example, in a study of treatments for acute myeloid leukemia, patients were initially randomized into three different treatment arms. But rather than maintaining an equal number of patients in each arm, the data were analyzed continually and patients were assigned to the better performing arms of the study with higher probabilities. After only five patients had been tested in a poorly performing arm of the study, the assignment probability to that arm became 0 and so it was effectively dropped (Figure 4 and Table 2). The trial ended after only 34 patients had been tested, about half the number that would have been tested in a standard randomized trial in which 25 patients would have been assigned to each treatment arm, Dr. Berry said. Not everyone sees the value to such a study design, he added. Although one journal rejected the study because only five patients had been tested in the one treatment arm, another journal published the study and the journal editor complimented the study design.

Dr. Berry noted that often with adaptive trials, investigators use mathematical modeling and simulations to determine the likely relationships among various factors in a trial and trial results. For example, the likely relationship between patient biomarkers and response to experimental treatments is predicted based on data collected during a Phase I trial. This information is then used to determine what types of patients to enroll in the various treatment arms of a subsequent Phase II trial, which then transitions seamlessly from the Phase I study.

Adaptive accrual ramps up the accrual rate of patients for a clinical study only after testing on an initial grouping of patients suggests the experimental therapy is likely to be effective and worth pursuing further. How-

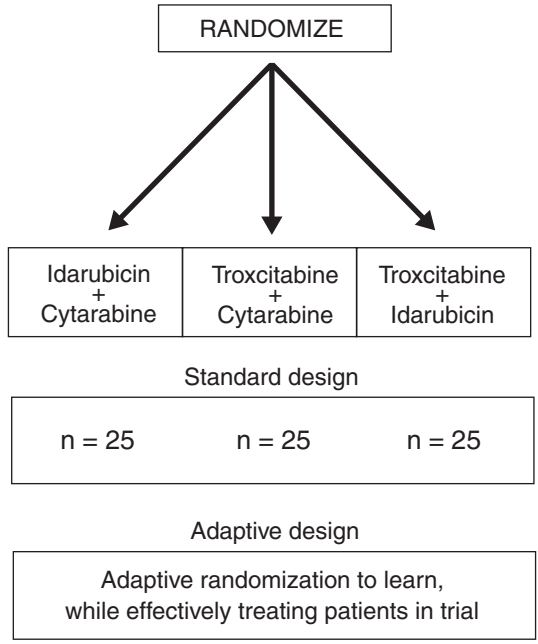

FIGURE 4 Study design of drug combinations in acute myeloid leukemia. In a standard design, patients would be assigned in equal numbers to each treatment combination. In the adaptive trial design, the data were analyzed continually and patients were assigned to the better performing arms of the study with higher probabilities.
SOURCES: Berry presentation (October 4, 2007) and Giles (2003).

TABLE 2 Study Results—Drug Combinations in Acute Myeloid Leukemia

Drug Combination	Complete Response by Day 50
Idarubicin/Cytarabine	10/18 = 56%
Troxicitabine/Cytarabine	3/11 = 27%
Troxicitabine/Idarubicin	0/5 = 0%

SOURCES: Berry presentation (October 4, 2007) and Giles (2003).

ever, Dr. Berry did not know of any adaptive accrual trials that had been conducted. He believed such an adaptive trial would be popular among drug sponsors, but noted that it is contrary to their traditional approach, which rewards fast patient accruals from the start.

Using adaptive clinical trials has several advantages. Dr. Ellenberg claimed that the primary rationale for doing adaptive designs has traditionally been ethical, especially for cancer studies. "You need to modify or terminate a study when interim data suggest that patients aren't being optimally treated," she said. Dr. Berry pointed out that adaptive trials that use information collected early in a trial to better segregate patients into treatment arms likely to be most favorable for them—personalized therapy—result in more patients in a trial assigned to better therapies. His experience with a number of patient groups suggests that adaptive trials will encourage more patients to enroll in cancer clinical studies because patients perceive such trials as offering them better treatment in addition to providing more efficient drug development.

Both Drs. Berry and Ellenberg noted that adaptive designs have practical advantages as well because they increase the likelihood that a study will be informative, and enable investigators to end studies early if they are not generating expected favorable results because the original design parameters were inaccurate. "If we see that things aren't going the way we thought, we are going to end up with data that are uninformative. We want to be able to stop studies early when it looks like they are going nowhere," Dr. Ellenberg said. Dr. Berry added that adaptive trials often enable faster, smaller, and more successful trials with substantial savings over nonadaptive trials.

Dr. Ellenberg concurred that there is consensus that adaptive approaches are appropriate in all phases of clinical research. Although not all proposed adaptive designs are uniformly favored, she added, the concept of adaptation is universally accepted. She added that although adaptive trial designs have been used since the 1960s, new types of adaptive designs have appeared in recent years. Improved computing power has stimulated use of Bayesian statistical methods, which are of increasing interest in clinical trials, particularly adaptive trials. "These [Bayesian] designs were impractical in the 1950s when these computer models weren't available," she said. "There is a lot of excitement now about seeing whether [such designs can improve efficiency], and more people are learning about how to apply Bayesian methods." Bayesian techniques are well suited to adaptive trials, Dr. Berry pointed out. They enable inferences based on observed data, continual updating, predictive probabilities, and longitudinal modeling. "The Bayesian approach

allows you to say, here I am today, this is what I know, where do I want to go, and what are the probabilities associated with going there," he said. Such analyses require prospective study designs. "It is a lot of work," he said. "You have to think about what you would want to do for the various kinds of things that happen in the course of the trial. You have to simulate to see what effects various factors have on operating characteristics, including duration of the trial and sample size," he said. He added that Bayesian analysis encourages modeling early and late endpoints. "One of the reasons for failure in Phase III trials is that in Phase II we use one endpoint, in Phase III we use another endpoint, and never the twain shall meet. We ought to be using both endpoints throughout and modeling the relationship," Dr. Berry said. Even when early and late endpoints lack traditional statistical power, they are still useful in Bayesian analyses, he pointed out. "Twenty percent power is better than zero percent power. Even though you don't learn a lot, you learn something," he said.

Dr. Berry pointed out that FDA Deputy Director Janet Woodcock stated at a recent conference that "improved utilization of adaptive and Bayesian methods" could help resolve the low success rate of and expense of Phase III clinical trials.[4] He added that in the past 7 years, MD Anderson has had more than 200 adaptive trials that used Bayesian techniques. "Adaptive design in the drug area has become what one pharmaceutical newsletter from Japan described as being a tsunami," he said. "Virtually every major company is getting involved because they are attracted to the possibility of building efficient trials."

But Dr. Ellenberg pointed out that adaptive trials are not free from controversy, particularly when they involve sample size reestimation in Phase III trials based on interim data. There is concern that such reestimation can bias the trial by revealing information about the accumulating data that in turn can change how the trial is being conducted. That concern is based on historical precedents. She noted that in the "old days," clinical trials were sometimes too adaptive. Often investigators and sponsors would review data as they came in and make decisions about stopping, continuing, or modifying a study based on emerging data. This led to studies that were inappropriately terminated early based on suggestive but not definitive data, and increased the chance of false-positive conclusions because investigators and sponsors would eye incoming data and "assume they had a winner"

[4]SPORE (Specialized Programs of Research Excellence) meeting, Baltimore, MD, July 18, 2006.

as soon as a *p* value crossed the "less than 0.05" threshold for determining statistical significance for study findings. At that point they would end the trial.

This led to the recognition that when changes are influenced by interim data, statistical tests may lose their meaning. Statisticians recognized some structure was needed to allow some mid-course changes that still permitted valid inferences to be made about final data. and developed sequential designs in the 1970s and 1980s that allowed regular looks at the accumulating data with the possibility of stopping the trial early, while maintaining the Type 1 error at an acceptably low level. Such designs are now routinely used in Phase 3 trials evaluating treatments for serious diseases.

In the mid-1990s, a new approach to adaptive clinical trial design emerged. Unlike the group sequential designs that specified a full sample size but with the possibility of stopping early if results were more impressive than expected, these new designs "started small" and enlarged the number of patients to be included only if the effect seen appeared large enough to be worthwhile, but likely to be smaller than anticipated or hoped for. The start-small approach was attractive to many sponsors because it did not require committing extensive resources to support a large clinical trial at the time the trial was initiated. These new adaptive study designs, like the earlier generation of group sequential designs, also preserve the low false positive rate, a necessary condition for an acceptable design. "You can still have a meaningful statistical test at the end of a study," said Dr. Ellenberg.

According to Dr. Ellenberg, the two main criticisms raised about these new adaptive trial designs is that they don't necessarily improve trial efficiency compared to standard designs, and that they create the potential for bias in trial conduct by providing information on emerging results to investigators and other interested parties. Because analyses of interim data in adaptive trial designs may reveal the need to enlarge a study population, they are not necessarily always more efficient. In fact, some investigators have shown that standard group sequential designs are always more efficient than the start-small type of adaptive design (Tsiatis, 2003). "This [does not necessarily mean there is no place for such a] design, but it is not so obvious that we are going to be able to do smaller studies on average if we go this route," said Dr. Ellenberg.

She is more concerned about the potential for bias in adaptive trials. When sample sizes are increased in adaptive trials, she said, sponsors, investigators, and even investment firms can back-calculate to figure out the interim data on which that change in size was based. This will

effectively "unblind" the trial so that investigators, for example, might be more inclined to notice favorable results in patients in the investigational treatment arm once they realize that the data are suggesting some benefit (albeit modest) for that treatment. Equally troubling, investigators may be uncomfortable continuing to enroll patients if there is evidence that one treatment is yielding better outcomes than the other. It is certainly true that use of stopping boundaries for more traditionally designed trials can also permit investigators and sponsors to make certain assumptions about the effectiveness of a treatment midstream in a trial. But those assumptions are much less precise than what can be inferred from an adaptive trial whose recipe for increasing sample size is prespecified, Dr. Ellenberg noted. A way to avoid this problem would be to keep confidential the aspects of the design relating to decisions to stop or enlarge the trial; but that would require sponsors to commit to (from their perspective) an open-ended trial, which is impractical, she said.

She concluded by stating that concerns about study integrity should be addressed before adaptive designs become more widely used to change sample size in Phase III trials. "I haven't seen any solution to this problem," she said. "You are not going to be able to keep it a secret if the trial enlarges, and you are not going to be able to keep secret what the study design was that led to that enlargement." In a later discussion, Dr. Berry countered that he thought solutions were possible for this problem. "I don't have a universal solution, but there are ways of preserving confidentiality in some circumstances," he said.

One discussant revisited response-adaptive randomization designs, in which patients are equally randomized into treatment arms at the beginning of the study, but then preferentially placed in specific treatment arms based on preliminary results. He noted that this design inaccurately assumes that patient characteristics do not change during the course of a trial.

Dr. Berry responded that patients *do* change during the course of a study. For example, at the beginning of a trial involving one or more intensive therapeutic regimens, investigators tend to present the trial to younger patients with more aggressive cancers than to older patients with less involved disease. However, as investigators become comfortable with the regimens and see that they can be given with a minimum of side effects they are more likely to offer the trial to older patients and to patients who have less aggressive disease.

But Dr. Berry noted that, although patients change over time during a clinical study, the treatment benefits usually do not. The use of controls

throughout the duration of the trial provides information about patient drift. These control patients have the same confounding factors as participants at every stage of the trial. "If you use the covariates and you have controls over time, you can at least partially resolve the issue," Dr. Berry said.

Targeting Multiple Pathways with Multiple Drugs

Research is increasingly finding that a specific cancer can be dependent on more than one altered biochemical pathway. Therefore, treatments that target multiple pathways are more likely to be effective than those that only target a single pathway. However, numerous challenges are involved in determining the best targeted cancer therapies to combine for different cancer types, and in conducting clinical trials of those combination treatments. These challenges and ways to overcome them were discussed by two speakers in the session focused on targeting multiple pathways with multiple drugs.

The first speaker in this session was Janet Dancey of the NCI's Investigational Drug Branch. She noted that combinations of more than one targeted therapy should be explored "earlier rather than later in the development of these agents," but added that combining investigational drugs prior to their receiving FDA approval for marketing presents multiple challenges. These challenges are not just scientific or medical. Other challenges include sharing data and intellectual property among companies and academic institutions, the greater risk of failure of combination therapies, and regulatory quandaries related to how best to show efficacy and safety for FDA approval of a combination therapy.

Agreements often have to be forged among different industry partners and academic institutions for the development of combination treatments that target multiple cancer pathways. To aid this process and foster early clinical trials of investigational drug combinations, the NCI turned to its Cancer Therapy Evaluation Program (CTEP). This program supports early "proof of principle" trials, which identify the appropriate molecular contexts for effectiveness. CTEP provides template agreement language among NCI, industry, and academic investigators concerning the sharing of data and intellectual property that stems from combination studies. CTEP currently has collaborative development agreements with more than 80 industry partners for more than 100 experimental drugs. CTEP also has clinical trial agreements with academic institutions, consortia, and cooperative groups. The program has sponsored more than 100 trials combining investigational

agents, as well as agreements for sharing resources in preclinical studies of 75 investigational agent combinations.[5]

Commenting on the success of CTEP in increasing the number of investigation drug combination trials that have been initiated, Dr. Dancey noted that "Our rate-limiting step is not necessarily being able to put together drug combinations. It is actually more to prioritize between possibilities." Such prioritization depends on overcoming scientific challenges such as determining

- which targets would be the most promising to combine,
- the best agents that can act on those targets simultaneously,
- the optimal types of patients likely to respond to the combination therapies,
- the most appropriate dosing regimens for combination treatments, and
- the best trial designs and endpoints that can reveal whether the combination is more effective than the treatments used alone.

Such determinations are difficult to make given the incomplete understanding of appropriate targets, agents, and likely responders, combined with the limits of what preclinical and clinical studies can reveal in that regard. Currently, nonclinical studies are the best means for determining mechanisms of action of investigational drugs and the development of useful biomarkers that can predict likely responders. These studies also can indicate which drugs have the most promising pharmacodynamics and work best in combination. These preclinical studies are usually done on animal models, but, Dr. Dancey said, "I think we would all agree that there are intrinsic differences between models and cancers in patients, and until we have models that look like patients or patients that look like models, we can't really have good predictive value from preclinical experimentation."

The limited number of models used to test drug combinations or look for biomarkers that predict response is unlikely to reflect the heterogeneity that occurs in cancer patients, she said. In addition, doses used in preclinical tests and endpoints may not be relevant for the clinical situation. Cancer clinical trial endpoints are usually patient survival or progression-free survival, but these endpoints are rarely used to evaluate for synergy in preclinical studies of drug combinations. The controls or standard treatments to

[5]For more information about CTEP, go to http://ctep.cancer.gov/industry/ipo.html.

which new drug combinations will be compared in a clinical trial also may not be appropriate for a preclinical study.

To improve these preclinical models, Dr. Dancey suggested that systematic efforts to molecularly characterize human tumors in nonclinical models "might help us get that much closer to that ideal of matching patients and models." She also advocated testing drug combinations in multiple tumor models to see if there is consistency in observed effects. Ideally, such testing should be done with a dosing regimen that mimics what is clinically achievable. This may require conducting nonclinical studies after acquiring preliminary human data to determine exposures likely to give a desired outcome in the clinic. All these preclinical results should be explained within the molecular context of the models used to better understand why synergy or antagonism occurs, and how that might be promoted or avoided, respectively, in clinical trials.

Another major scientific issue is which targets to aim for with combination treatments. Common strategies are to combine agents that target a pathway at the same point to maximize inhibition of that pathway, such as combining an agent that acts on the vascular endothelial growth factor (VEGF) with one that acts on the receptor for VEGF. Another approach is to combine an agent that targets a specific cancer growth factor, such as human epidermal growth factor receptor 2, or HER-2, in breast cancer with another compound that plays a critical role downstream from the activation of that target, such as mTOR (mammalian target of rapamycin). Combinations that block parallel pathways and different cellular processes that underlie a cancer and its progression can also be effective. These include combinations that target both the VEGF receptor and the epidermal growth factor (EGF) receptor, both of which are believed to play a key role in certain cancers. Other combinations include an agent aimed at a major cancer target and a second agent aimed at overcoming resistance to the first agent in the combination. (Dr. Gray expanded on this in his presentation, which is summarized below.)

Once appropriate targets are determined, researchers have to select agents that can collectively counter those targets without causing significant overlapping toxicities. Dr. Dancey pointed out that when a drug combination fails, it can fail for several reasons; perhaps the drugs individually or in combination did not effectively interact with their targets or the targets themselves singly or in combination were not relevant. "Therein lies the risk of evaluating combinations early on when you don't know a lot about

the targets in human cancer and you don't know a lot about the agents and their ability to effectively interact with those targets," she said.

Ideally, investigators should have biomarkers that predict response when testing targeted cancer drugs in clinical trials. But these biomarkers are often lacking, and without them, interpreting the results of clinical trials of combination targeted cancer drugs is especially problematic. To illustrate this point, Dr. Dancey showed a slide of possible outcomes with a three-drug combination therapy (Figure 5). If this combination of drugs

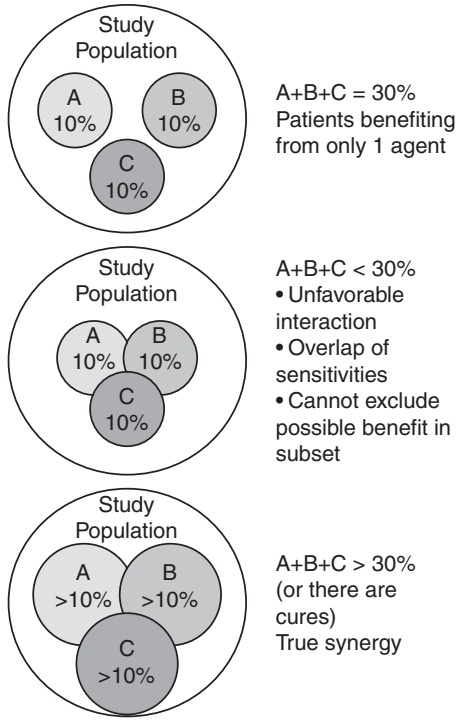

FIGURE 5 Possible outcomes with drug combinations in unselected patients. In the first example, patients respond to only one agent in the drug combination, so the overall response to the drug combination is additive. In the second example, combinations of agents result in fewer positive responses than if each patient was treated with only one agent. Possible reasons for the decrease in response are unfavorable interactions or that two agents target the same vulnerability. A positive response in a subset of the study population cannot be excluded, however. In the third example, a greater number of patients experience a positive therapeutic benefit or cure. This is due to synergistic activity of the combined therapeutic agents.
SOURCE: Dancey presentation (October 4, 2007).

A, B, and C is tested in a population in which 10 percent are responsive to each of the three drugs, then a total response rate of 30 percent will be seen even if there is no benefit to the combination. In other words, patients would respond the same to the combination as they would to the individual drug in the combination to which they are responsive. A response rate of greater than 30 percent would occur if the combination is more favorable than the individual agents. But without predictive markers for response to each of the three drugs, as well as predictors of response to the combination, researchers cannot conclude who is likely to benefit from the combination treatment and whether it is more beneficial than individual agents. "Looking for predictive markers in the context of developing the combination is probably going to be very difficult, and even more difficult than doing it with the individual agents," Dr. Dancey said.

The search for such markers requires multiple assessments of tumor response to assess markers for initial response as well as markers for the development of resistance, which occurs later in treatment when tumors are enriched with resistant clones. Such assessments are best done preclinically because such evaluations in the clinic are more difficult, complicated, and expensive, Dr. Dancey noted.

Determining the best dosing regimen for combination therapies also can be complex because of the need to consider multiple possibilities to discern the best dose and schedule. Often the optimal dose for individual agents within combination therapy is not the same as the dose for each drug when used alone. "This is one that we in particular have been wrestling with for a number of targeted agents that we have been testing in combinations—because individual single agents are well tolerated, but the combinations induce toxicity," Dr. Dancey said. For optimal effectiveness, the dosing schedule may need to be altered for the drugs used in combination versus alone. "So if you do have to modify the dosing schedule, that means you have even more potential combinations that you might have to test," Dr. Dancey said. An efficient approach for this is a multiarmed, controlled trial with an adaptive design, she noted.

Dr. Dancey did not discuss the regulatory challenges involved in evaluating combination therapies. But in one of her slides, she noted that the FDA may not require clinical toxicology data for the combination if the individual agents have been tested in the clinic and their toxicity is known. However, the FDA does require sponsors to show the contribution of each component of a fixed combination regimen to its total effectiveness. Testing

of drug combinations requires early discussion with regulatory authorities, Dr. Dancey noted in her slide.

In summary, Dr. Dancey asserted that the rapid emergence of hundreds of new agents on an expanding list of cancer-specific molecular targets offers tremendous hope to cancer patients, while presenting significant development challenges to the cancer research community. The major legal, regulatory, and scientific challenges involved in developing testing strategies for combination cancer treatments may be overcome with common agreements among industry and academic partners regarding intellectual property and data sharing; systematic evaluation of targets and agents in predictive nonclinical models; the development of biomarkers that predict response to individual agents or their combination; and controlled clinical trials that assess multiple combinations.

Dr. Dancey was followed by Dr. Joe Gray of Lawrence Berkeley National Laboratory (LBNL) and the University of California, San Francisco. He focused on how to model molecular heterogeneity to enhance multidrug clinical trial design. He summarized the efforts by the Greater Bay Area Consortium, which consists of investigators at the University of California, San Francisco, the University of California, Berkeley, LBNL, and SRI International, in collaboration with investigators at MD Anderson Cancer Center and pharmaceutical company GlaxoSmithKline. He described the immense challenges involved in dealing with the heterogeneous nature of cancer. Patients with the same type of cancer, or even with tumors that appear the same clinically, may differ in the molecular defects that underlie their cancers or fuel their growth.

Dr. Gray stressed that "it's not just the target, but everything that is going on in the tumor that is important." Researchers have amassed enough data to provide "a good catalog" of the ensemble of molecular abnormalities that play key roles in the progression and response to treatment for most major cancer types, he added. A slide of his data on breast cancers revealed several portions of cancer cells' genetic material (genome) that are abnormally activated due to duplications (Figure 6). These data indicate that at least 15 percent of the genes and the DNA that regulate their activity (the transcriptome) in breast tumors are activated abnormally. His research suggests that many, if not most, of these abnormally expressed genes play an important role in the progression of cancer and how well it responds to targeted therapeutics.

Some of these genes are part of the same molecular pathway and are activated when a linchpin molecule in the pathway, such as the ErbB2

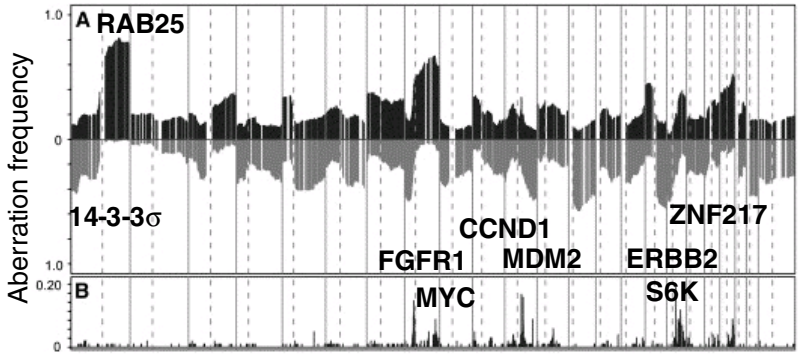

FIGURE 6 Recurrent copy number aberrations in breast cancer. Ten to fifteen percent of the transcriptome/proteome is deregulated by recurrent aberrations. Functional studies support the concept that many of these contribute to cancer pathophysiology. A: Frequencies of genome copy number gain and loss plotted as a function of genome location. Vertical lines indicate chromosome boundaries, and vertical dashed lines indicate centromere locations. Positive and negative values indicate frequencies of tumors showing copy number increases and decreases, respectively. B: Frequencies of tumors showing high-level amplification. Data are displayed as described in A.
ACRONYMS: 14-3-3σ (SFN, stratifin), CCND1 (cyclin D1), ERBB2 (v-erb-b2 erythroblastic leukemia viral oncogene homolog), FGFR1 (fibroblast growth factor receptor 1), MDM2 (transformed 3T3 cell double minute 2), MYC (v-myc myelocytomatosis viral oncogene homolog), RAB25 (member RAS oncogene family), S6K (ribosomal protein S6 kinase), ZNF217 (zinc finger protein 217).
SOURCE: Gray presentation (October 4, 2007), reprinted from *Cancer Cell*, Volume 10, Chin, K., S. DeVries, J. Fridlyand, P.T. Spellman, R. Roydasgupta, W.-L. Kuo, A. Lapuk, R.M. Neve, Z. Qian, T. Ryder, F. Chen, H. Feiler, T. Tokuyasu, C. Kingsley, S. Dairkee, Z. Meng, K. Chew, D. Pinkel, A. Jain, B.M. Ljung, L. Esserman, D.G. Albertson, F.M. Waldman, and J.W. Gray, Genomic and transcriptional aberrations linked to breast cancer pathophysiologies, pp. 529-541, Copyright 2006, with permission from Elsevier.

receptor, which the drug Herceptin targets, is activated farther upstream. But even when ErbB2 is overexpressed, downstream genes are activated to different degrees in different tumors, according to Dr. Gray's slide of gene expression in three different breast cancers (Figure 7). "We have to understand how these ancillary aberrations, co-acting with the target, condition response," Dr. Gray said.

Fortunately, recent large-scale "omics" technologies that enable simultaneous assessment of all expressed genes or proteins with automated devices

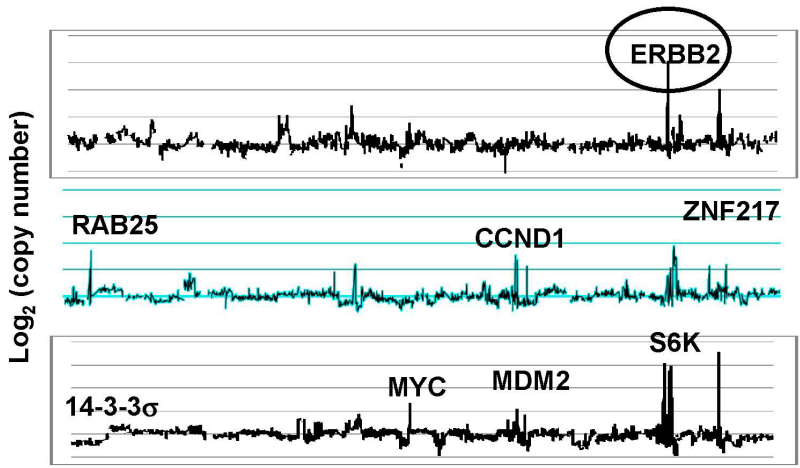

FIGURE 7 Aberration combinations in the same pathway vary considerably among tumors—even in *subsets* having the same therapeutic target.
ACRONYMS: 14-3-3σ (SFN, stratifin), CCND1 (cyclin D1), ERBB2 (v-erb-b2 erythroblastic leukemia viral oncogene homolog), MDM2 (transformed 3T3 cell double minute 2), MYC (v-myc myelocytomatosis viral oncogene homolog), RAB25 (member RAS oncogene family), S6K (ribosomal protein S6 kinase), ZNF217 (zinc finger protein 217).
SOURCE: Gray presentation (October 4, 2007).

can reveal telltale molecular patterns relevant to specific cancers and how they are likely to respond to various targeted treatments. This information can be used to identify markers that indicate which drug combinations are most likely to be effective for individual cancer patients. But these markers are not usually available until late in the drug development process, so they are not often used to guide early trials or to prioritize which drug combinations should be tested preclinically based on the likelihood that they will have synergistic effects.

Adding to the complexity is the fact that there are about 100 FDA-approved cancer drugs and more than 400 experimental cancer drugs in Phase II or III trials. The target specificities for most of these drugs are not well known, Dr. Gray pointed out, and clinical tests of these agents are not coordinated or guided by biomarkers. Unfortunately, the cost of molecularly characterizing all available cancer drugs and their effects on genes known to play a role in cancers would be enormous.

Another approach that Dr. Gray's consortium and others are taking is to develop preclinical models for the molecular heterogeneity found in tumors that can be used to determine which drug combinations are the best to test clinically and which patients are likely to respond to these treatments. He and his colleagues have collected and characterized about 50 breast cancer cell lines that have enough molecular diversity to enable the detection of molecular abnormalities linked to response. These cell lines also seem to adequately mirror clinical findings. For example, the cell lines have the same patterns of gene expression (genetic signatures) that are seen in primary tumors (Figure 8). Even when the cell lines are broken down by type of breast cancer (e.g., luminal versus basal), they closely mimic the gene expression of the primary tumors for each type.

Consortium investigators are using these cell lines to test large numbers of drug combinations in an automated fashion. Researchers can currently

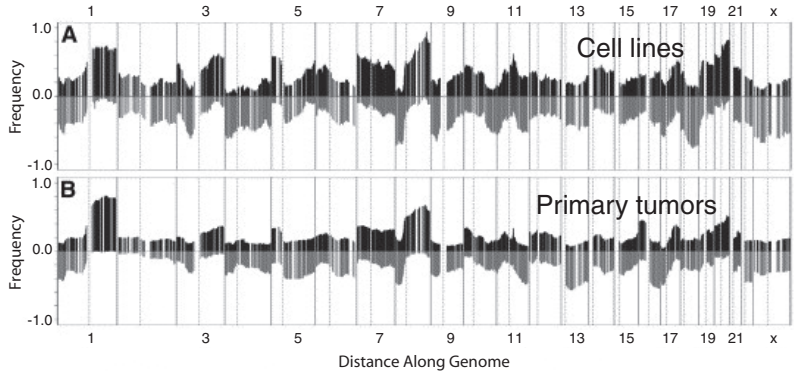

FIGURE 8 Cell lines retain the recurrent genomic characteristics of primary tumors. A and B: Frequencies of significant increases or decreases in genome copy number are plotted as a function of genome location for 51 cell lines (A) and 145 primary tumors (B). Positive values indicate frequencies of samples showing copy number increases [Log_2(copy number) > 0.3], and negative values indicate frequencies of samples showing copy number decreases [Log_2(copy number) < –0.3].
SOURCE: Gray presentation (October 4, 2007), reprinted from *Cancer Cell*, Volume 10, Neve, R.M., K. Chin, J. Fridlyand, J. Yeh, F.L. Baehner, T. Fevr, L. Clark, N. Bayani, J.-P. Coppe, F. Tong, T. Speed, P.T. Spellman, S. DeVries, A. Lapuk, N.J. Wang, W.-L. Kuo, J.L. Stilwell, D. Pinkel, D.G. Albertson, F.M. Waldman, F. McCormick, R.B. Dickson, M.D. Johnson, M. Lippman, S. Ethier, A. Gazdar, and J.W. Gray, A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes, pp. 515-527, Copyright 2006, with permission from Elsevier.

test hundreds of drugs and drug combinations simultaneously, and expect to use improved automation techniques to eventually boost such simultaneous testing to as many as 10,000 drugs or drug combinations. This testing has already revealed various cancer drugs' target specificities. For example, an AKT inhibitor appears to affect genes abnormally activated in luminal tumors, but not basal tumors. Some of these findings have been confirmed in clinical studies. For example, tests of lapatinib indicated it would only be effective in tumors that overexpress or phosphorylate ErbB2, and this was shown to be true when the drug was tested clinically (Di Leo et al., 2007).

Dr. Gray's studies have also revealed basic information about cancer pathways that will help to optimize targeted cancer treatments. There are two parallel molecular pathways relevant to breast cancer that are activated when ErbB2 is activated—the AKT pathway and the Raf-MAP kinase pathway (Figure 9). Research on the breast cancer cell lines reveals that luminal cancers have an activated Raf-MAP kinase pathway, whereas basal tumors have an activated AKT pathway. This suggests that using drug combinations that block the primary pathway activated by a mutation as well as the alternate "bypass" pathway could be effective.

Dr. Gray summarized the strength of this modeling approach by pointing out that cell lines can be characterized in exhaustive molecular detail, unlike patients or their tumor samples, and automated testing techniques can quickly indicate the most effective drug combinations to test clinically. In addition, the mechanism of action of an experimental drug can be easily assessed. For example, if it appears that the AKT pathway is important to a drug's effects, it can be tested by altering the activity of that pathway in a cell line and seeing if it correspondingly affects the drug's activity. The in vitro studies can also reveal promising new targets. Researchers have identified only about 20 percent of the genes in the abnormally duplicated regions of the breast cancer cell lines, Dr. Gray said in the discussion following his presentation.

The main weakness of his cell line model is that more cell lines, including resistant cell lines, are needed to more completely represent the molecular heterogeneity of breast cancer. In addition, better modeling of the in vivo microenvironment is needed, and some culture-specific aberrations may accumulate over time such that the cell lines eventually may not adequately mimic what is seen clinically. Despite that potential problem, Dr. Gray is "fairly confident that this is at least a way forward of helping us

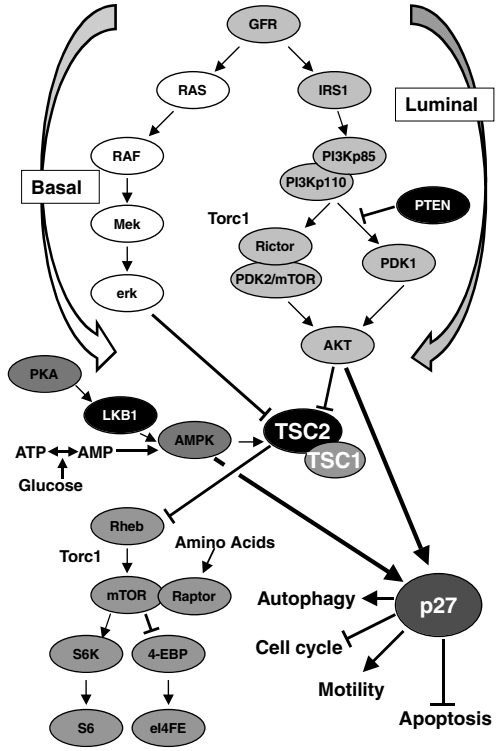

FIGURE 9 Basal and luminal tumors may use different parts of the growth factor signaling network. Drug combinations can be selected to block activating mutations and alternate bypass pathways. The signaling pathways shown impact cell motility, growth, and survival.

ACRONYMS: 4-EBP (translational repressor eukaryotic initiation factor 4E-binding protein), AKT (v-akt murine thymoma viral oncogene homolog), AMP (Adenosine Monophosphate), AMPK (AMP-Activated Protein Kinase), ATP (adenosine triphosphate), eI4FE (messenger RNA 5-cap binding protein), erk (extracellular-signal-regulated kinase), GFR (Rap guanine nucleotide exchange factor 5), IRS1 (insulin receptor substrate 1), LKB1 (serine/threonine kinase), Mek (mitogen-activated protein kinase), mTor (Mammalian target of rapamycin), p27 (SSSCA1, Sjögren syndrome/scleroderma autoantigen 1), PDK1 (pyruvate dehydrogenase kinase, isozyme 1), PDK2/mTOR (pyruvate dehydrogenase kinase, isozyme 2), pI3kp110 (phosphatidylinositol 3-Kinase p110 subunit), pI3kp85 (Phosphatidylinositol 3-Kinase p85 subunit), PKA (Protein Kinase A), PTEN (phosphatase and tensin homolog), RAF (a protein kinase), raptor (regulatory associated protein of mTOR), RAS (GTP-activated protein involved in cell growth regulation), Rheb (Ras homolog enriched in brain), rictor (rapamycin-insensitive companion of mTOR), s6 (ribosomal protein involved in translation), Torc1 (Target of rapamycin complex 1), TSC1 (tuberous sclerosis 1), TSC2 (tuberous sclerosis 2).

SOURCE: Gray presentation (October 4, 2007); pathways courtesy of Gordon Mills.

to prioritize these drugs and drug combinations for introduction into the clinic," he said.

In addition to using cell line models to indicate the most optimal drug combinations to test clinically, researchers can use them to select in vitro response biomarkers that are likely to work in a clinical setting. According to Dr. Gray, ideally, a clinically useful in vitro biomarker would be a genome aberration whose detection does not vary with culture condition. The marker should also be the same in both the cell cultures and the primary tumors. Dr. Gray gives higher priority to transcriptional markers than to protein markers because the former are currently easier to measure, although he acknowledged that genomic markers won't necessarily inform the biology as well as protein markers. Thus both approaches are ultimately needed.

Dr. Gray and his colleagues are currently pursuing an innovative marker-intensive clinical trial of breast cancer treatment that uses a series of core breast biopsies and magnetic resonance imaging (MRI) to determine before-and-during-treatment responses, and the effectiveness of markers in predicting such responses. "We think this is a reasonable way of taking the drugs and markers that come out of our in vitro system and quickly evaluating them just for general efficacy in the neoadjuvant environment. Then for those things that seem to be behaving the way that we expect them to, we will introduce them into a later phase clinical trial to assess long-term outcome," Dr. Gray said.

This approach should lead to more efficient clinical trials, he noted, because the early trials would target patient subpopulations most likely to respond, and would be less likely to miss drugs effective against small subpopulations. The model system would also provide a rationale for use of drug combinations that may not show independent efficacy. In addition, patients would be more likely to participate in such trials because they would be given treatments tailored to be effective against their specific type of cancer. The end results would be lower costs due to testing in patients more likely to respond first, and increased patient participation. Trials that have a primary focus on biomarker development would also provide material to assess not just target response, but the presence of other molecular aberrations that affect treatment effectiveness, including those that contribute to the development of resistance.

In a discussion following Dr. Gray's presentation, Dr. Roy Herbst from MD Anderson Cancer Center asked Dr. Gray about the role that animal models might play in modeling molecular heterogeneity to enhance multi-

drug clinical trial designs. Dr. Gray responded that the cell line model is just the first stage in the process, but that animal models can provide information that cell lines lack. "What our studies do is identify interacting aberrations that look like they condition response to drugs. But we will never get to the point in vitro where we model all of the nuances of the microenvironment, so the next logical step is to go into the mouse models and complement it there," he said. He added that the NCI's Mouse Models of Human Cancers Consortium[6] is developing a robust set of models that are genetically engineered to have many of the same molecular abnormalities linked to cancer progression or response to treatment that are seen in cell lines. (Dr. Anderson also discussed how to model the microenvironment in his presentation, which is summarized below.)

Another discussant, patient advocate Kathy Needham, raised the question of whether cancers should be grouped according to their underlying genetic abnormalities rather than by the type of organ in which they occur when assessing drug effectiveness. Dr. Gray noted that the genetic abnormalities in ovarian, prostate, and breast cancer are remarkably similar. But he added that the molecular conditioning abnormalities that affect response to treatment differ by organ site and tumor subtype. "So you have to pay attention to both. Clearly people are already pursuing targets, not organ types. But it is by organ site that the drugs get introduced into the clinic," he said. Dr. Dancey added that although there hasn't been a test case yet, the NCI has developed new clinical trial designs that enroll patients according to the molecular abnormalities in their tumors and not necessarily by where the tumors appear.

Discussant Dr. Steven Shak from Genomic Health then raised the issue that solid tumors often have tens of thousands of mutations, probably many of which are silent or biologically insignificant. But the large number of mutations makes it difficult to discern those mutations that do play a major role in the tumor. "Yes, there are a lot of mutations out there," Dr. Gray responded, "but they tend not to be recurrent. What we need to do is identify the ones that are recurrently present." He noted a recent journal article from researchers at Johns Hopkins University in which they catalogued mutations in 13,000 genes in breast and colorectal cancer. The researchers narrowed this list down to a few hundred genes that might be candidates for mutations that play an important role in the progression of these cancers (Sjöblom et al., 2006).

[6]See http://emice.nci.nih.gov/emice/MMHCC/mmhcc_organization.

Preclinical Model Systems

The discussion was followed by a presentation on translation from preclinical model systems to the bedside (and back) in multiple myeloma by Dr. Kenneth Anderson of the Dana-Farber Cancer Institute at the Harvard Medical School. Dr. Anderson developed laboratory models for myeloma that researchers used to predict the effectiveness of several new cancer therapies, most of which are now FDA-approved and widely used in the treatment of myeloma. The thrust of his talk was that researchers can use preclinical modeling to collect the information needed to choose which drugs should be developed and to design clinical trials for those drugs. Although this concept was explored by previous speakers, Dr. Anderson went a step further by showing how best to model the tissue microenvironment in which myeloma tumors form so as to gather more clinically relevant information from preclinical studies. This microenvironment determines the expression of the genes that foster myeloma tumors or enable their resistance to treatment. "If one is going to make a preclinical model of cancer that is valid, one needs very strongly to reflect the microenvironment," he said.

Because of the recent extraordinary explosion of genetic findings, Dr. Anderson said, myeloma is now classified into seven groups based on the genes expressed in the tumors. Although researchers have detected hundreds of genes that are abnormally expressed in such tumors, studies to systematically assess the effects of overexpression or deletion of these genes reveal a much smaller number of genes believed to play a major role in myeloma. But additional genes that strongly affect survival or metastasis of the tumor, or its resistance to treatment, are only expressed when myeloma cancer cells attach to particular bone marrow cells called stromal cells. Such attachment requires specific adhesion molecules. Some of the genes activated by attachment to the bone marrow stromal cells trigger the activity of a complex of proteins in the cells called proteasomes. By breaking down key proteins, proteasomes block normal cell death and enable cancer cells to live for a long time and actively divide.

This understanding of the microenvironment of myeloma tumors explains why a proteasome inhibitor drug such as bortezomib is more effective against myeloma cells with the preserved microenvironment of bone marrow stromal cells and adhesion molecules than in cell lines that lack this crucial microenvironment, Dr. Anderson pointed out. The microenvironment also explains why conventional myeloma therapies are not effective:

they are susceptible to cell adhesion-mediated drug resistance. "So testing the drug in the microenvironment is critical," Dr. Anderson said.

Dr. Anderson and his colleagues have developed both in vitro and animal models that mimic the microenvironment of myeloma tumors. In their in vitro models, myeloma cell lines or patient tumor cells are bound to bone marrow stromal cells grown in the laboratory. They also have in vivo mouse models, including a mouse with a transplanted human bone chip in which fluorescent human myeloma cells have been injected. Researchers use this model to test drugs and assess the genes that confer resistance or sensitivity to them, and how well that correlates with what is found from in vitro studies. "It is critical to look and see whether what you have proposed and observed qualitatively in vitro is reflected in vivo," Dr. Anderson said.

He also noted that the interplay between laboratory and clinical studies can be bidirectional. For example, his genomic studies in patients with myeloma revealed a gene, XBP-1, which is overexpressed in all the patients. He used this finding to develop a mouse model in which the mice are genetically engineered to overexpress XBP-1 and have bone destruction and other features similar to that seen in patients with myeloma. "This is a genetic model of multiple myeloma which came from an observation made in patients by the new genomics," he said. "We always think of bench-to-bedside research, but we can do it the other way around."

Dr. Anderson and his colleagues have used their preclinical models to screen many classes of drugs. They found that some drugs target the tumor and the microenvironment, while others target just one or the other. But whatever the mechanism, a drug must cause tumor cell death, even when the tumor is bound to the bone marrow stromal cells, in order to proceed further in the drug development and testing pathway. These studies led to four highly effective drugs receiving FDA approval in the past 3 years for the treatment of multiple myeloma, as well as several promising experimental drugs currently in clinical trials (Figure 10). Two of the approved drugs, bortezomib[7] and lenalidomide,[8] when used along with a steroid drug or, in the case of bortezomib, a steroid drug and various chemotherapy agents,

[7]Bortezomib (Velcade) received accelerated FDA approval as a single agent for relapsed, refractory multiple myeloma in 2003 (see http://www.fda.gov/bbs/topics/NEWS/2003/NEW00905.html).

[8]FDA granted approval in 2006 to lenalidomide (Revlimid) for use in combination with dexamethasone in patients with multiple myeloma who have received one prior therapy (see http://www.fda.gov/cder/Offices/OODP/whatsnew/lenalidomide.htm).

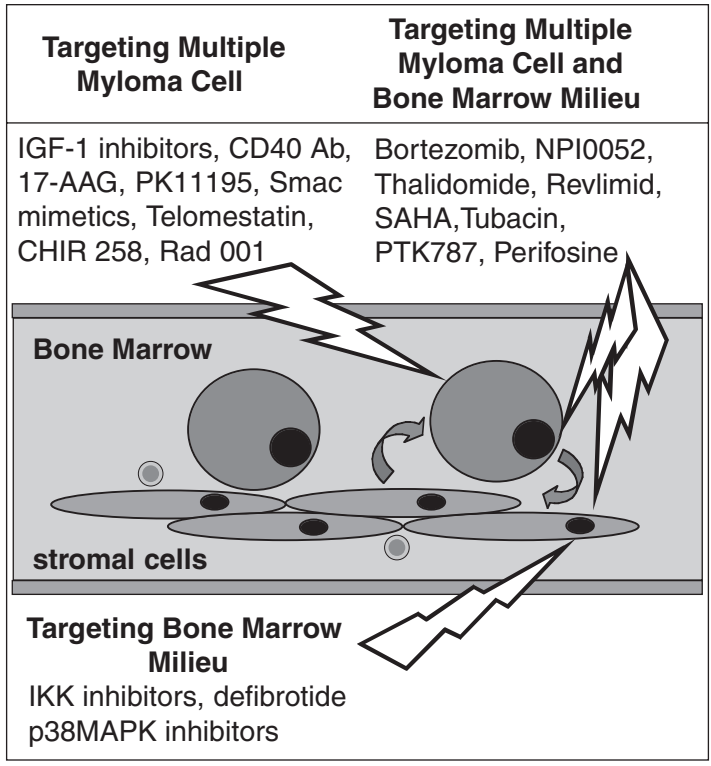

FIGURE 10 Novel agents targeting multiple myeloma (MM) cells and/or the bone marrow microenvironment.
ACRONYMS: 17-AAG (17-(Allylamino)-17-demethoxygeldanamycin), CD40 ab (antibody to the CD40 integral membrane protein), CHIR258 (Tyrosine Kinase Inhibitor), IGF-1 (Insulin-like growth factor 1), IKK (conserved helix-loop-helix ubiquitous kinase), NPI0052 (proteasome inhibitor), p38 MAPK (mitogen activated protein kinase 14), PK11195 (peripheral benzodiazepine receptor (PBR) ligand), PTK787 (multi-VEGF receptor inhibitor), Rad001 (serine-threonine kinase inhibitor of mTOR), SAHA (Suberoylanilide hydroxamic acid), Smac (Second mitochondria-derived activator of caspase).
SOURCE: Anderson presentation (October 4, 2007).

each produce remarkable and unprecedented response rates of 80 to 90 percent in newly diagnosed myeloma patients and about a 50 percent complete or near-complete response rate in some studies. Both bortezomib and lenalidomide take advantage of and overcome the growth, survival, and drug resistance potential that is conferred by the microenvironment,

Dr. Anderson noted. His preclinical models led to the bench-to-bedside development of lenalidomide in just 6 years—about half the typical amount of time needed for such development.

Researchers also used Dr. Anderson's preclinical models to determine the appropriate design of subsequent generations of myeloma drugs. Findings on the main proteasome activities that affect tumor cell growth and spread in the microenvironment led to the creation of a new type of proteasome inhibitor, NP10052, which inhibits a wider range of proteasome activities than bortezomib. Animal studies showed that about two-thirds of mice treated with NP10052 survived bortezomib-resistant myeloma, whereas all untreated mice died within 100 days. The drug is currently being tested in the clinic. "This drug came from preclinical lab and animal models that showed it was more effective to use broader inhibition of proteasome activities," Dr. Anderson said.

Dr. Anderson ended his talk by showing how his preclinical models help researchers discern which drugs to combine and how to test their combinations clinically. For example, these models revealed that proteasome inhibitors interfered with the ability of cultured myeloma cells to repair their DNA. This led the FDA to approve the use of the DNA-damaging agent doxorubicin combined with the proteasome inhibitor bortezomib for the treatment of multiple myeloma.[9] Doxorubicin is not FDA approved as a single agent for myeloma, Dr. Anderson noted, but its combination with bortezomib extended time to progression by about 3 months, and increased the response rate and overall survival, one clinical study found. "This combination would not have gone forward if it were not for preclinical modeling, which showed that this inhibitor of the proteasome has another feature inhibiting DNA repair," he said.

"The explosion in genetics and the ability to study the biology better than we have ever had before allows us to target the tumor directly, but as I hope we have illustrated for you, indirectly as well," Dr. Anderson concluded. The kinds of genetic studies that have been mentioned and the modeling that I have stressed not only define targets, but also define and inform the design of clinical trials." This new paradigm targeting the tumor cell in its microenvironment has great promise not only to change the natural history of multiple myeloma, but also to serve as a model for

[9]Approved in May 2007. See http://www.cancer.gov/cancertopics/druginfo/fda-doxorubicin-HCL-liposome.

targeted therapeutics directed to improve the outcome for patients with other types of cancers, he added.

MOLECULAR IMAGING

Recent advances in biomedical imaging provide potential opportunities to improve the discovery, development, and validation of novel therapeutics. Imaging applications offer the possibility to reduce the time, cost, and workload in drug development. In the session devoted to molecular imaging, five speakers addressed the current and near-horizon opportunities in molecular imaging, particularly how it can be used to detect biomarkers for assessing cancer treatment effectiveness, and the advantages it has over standard imaging. These speakers gave several promising examples of such molecular imaging biomarkers, and showed how they can be used throughout the drug development process. Also addressed in this session were current challenges involved with molecular imaging, suggestions for how to meet those challenges, and a discussion of how the current or potential pitfalls of molecular imaging compare to those of standard imaging.

"In the last 50 years, there have been tremendous advances of imaging," pointed out Dr. Hedvig Hricak of Memorial Sloan-Kettering Cancer Center (MSKCC). Those advances include the development and modification of positron emission tomography (PET), MRI, and ultrasound, as well as various genetic engineering techniques that enable imaging of specific molecules or molecular processes. All this innovative imaging can reveal, in a dynamic manner, key biological functions within the body related to cancer progression and response to treatment.

Current and Developing Methods

The development of PET and the use of the radioactive tracer FDG ([18F]-2-fluoro-2-deoxy-D-glucose) as a biomarker for the heightened metabolism that occurs in cancer cells paved the way for using an imaging probe to examine a biologic process. "This was a milestone in which a tracer was recognized as showing a particular biologic function," said Dr. Steven Larson of MSKCC, who noted that there are now more than 900 articles in medical journals related to PET imaging of tumor response. Such imaging is highly sensitive, with resolution down to 1 to 2 mm, and can indicate a response to treatment before standard CT imaging can. As Dr. Larson pointed out, PET imaging can show after just one or two treatments that a tumor is responding, even if it is not yet shrinking in size. The 1997 Food

and Drug Administration Modernization Act (FDAMA) and subsequent key FDA approvals of the PET FDG tracer spurred major technologic development over the past decade, according to Dr. Larson. This development includes combined PET and CT imaging devices at many cancer treatment facilities, as well as the development of new molecular probes that could be imaged with PET or other modalities.

Many of the newer cancer drugs target specific growth factors or their receptors that play a key role in the growth and spread of tumors. Consequently, researchers have developed several PET imaging probes for these compounds, including the receptors for HER-2, EGF, VEGF, estrogen, and androgen. Unlike standard imaging and non-imaging diagnostic techniques, molecular imaging of these growth factors or their receptors can reveal the heterogeneity of tumors and metastases. Dr. Hricak's slides showed that imaging with a probe for the estrogen receptor could reveal in breast cancer patients which metastases are estrogen receptor (ER) positive and thus likely to respond to hormonal therapy (Figure 11).

Researchers can also now use high-resolution MRI systems to create in reasonable time an anatomical map of the distribution of key metabolites relevant to cancer, pointed out Dr. John Gore of Vanderbilt University. Diffusion-rated MRI, which measures the degree to which water molecules are free to move around within tissue, can be used to measure cell density, which changes rapidly and early after particular cancer treatments, he said. He added that dynamic contrast MRI, by showing changes in blood volume and blood flow into tissues over time, is useful for detecting the abnormal proliferation and leaking blood vessels that typify malignancies. Farther in the future is the possibility of researchers using hyperpolarized carbon-13 as a radioactive label for tracers that can improve the sensitivity of MRI and enable the detection of specific metabolic pathways, as opposed to the heightened overall metabolism that is seen in PET with FDG.

Even ultrasound has been adapted to image molecular functions. For example, investigators are experimenting with adding molecular probes to the surface of the microbubbles of air used as contrast agents in ultrasound imaging. One company has labeled these microbubbles with an antibody that will bind to VEGF, whose expression is elevated in various tumors. Such ultrasound imaging may someday provide a low-cost alternative to more expensive techniques such as PET and MRI, Dr. Gore noted.

Genetic engineering techniques have enabled researchers to create probes for imaging the RNA or proteins expressed from specific genes in tumors, as well as to create innovations in optical imaging. Dr. David Piwnica-Worms of Washington University in St. Louis noted

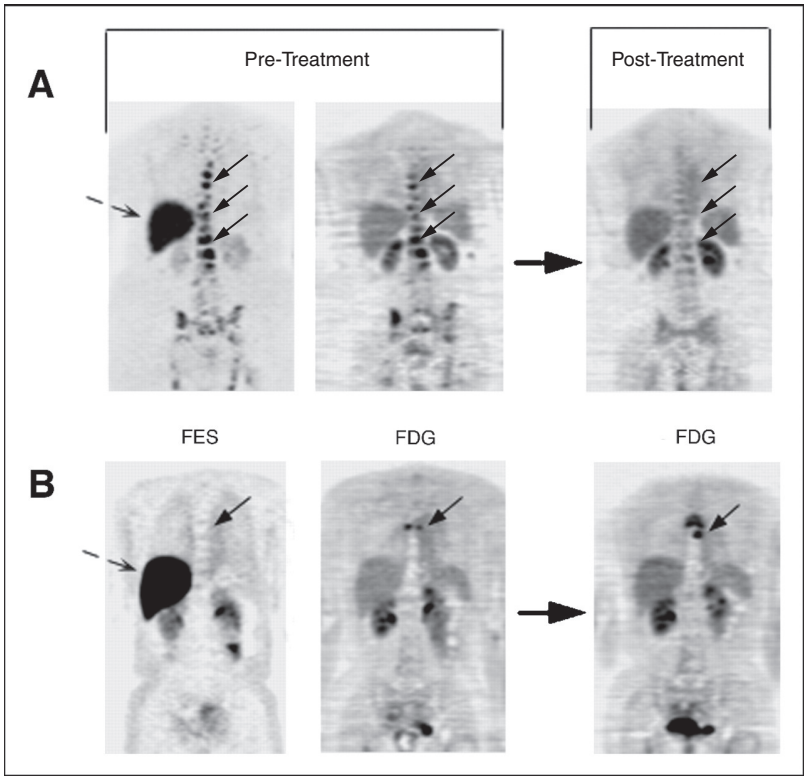

FIGURE 11 Targeted treatment selection: [18F]-fluoroestradiol (FES) in predicting response to hormonal therapy. Both patient A and patient B have ER+ primary tumors and bone metastases. Patient A was strongly FES positive and showed excellent response (size and SUV) after 3 months on letrozole. Patient B was FES negative and showed progression at 6 months. Solid arrows point out tumor locations, and lower posttreatment signal density corresponds to positive therapeutic response. Dashed arrows show normal liver FES uptake.
ACRONYMS: FDG ([18F]-2-fluoro-2-deoxy-D-glucose).
SOURCE: Hricak presentation (October 4, 2007) and Linden, H.M., S.A. Stekhova, J.M. Link, J.R. Gralow, R.B. Livingston, G.K. Ellis, P.H. Petra, L.M. Peterson, E.K. Schubert, L.K. Dunnwald, K.A. Krohn, and D.A. Mankoff. 2006. Quantitative fluoroestradiol positron emission tomography imaging predicts response to endocrine treatment in breast cancer. *J Clin Oncol* 24(18):2793-2799, reprinted with permission from the American Society of Clinical Oncology.

that researchers can add genes for proteins that generate bioluminescent compounds in the tumor cells that are introduced into animals. These "reporter" genes can indicate how the tumors are responding to an experimental drug. For example, the firefly gene that codes for the luciferase

enzyme, which causes light to be emitted, is often inserted into tumor cells. Specialized cameras can sensitively detect light emitted from these tumors deep inside an animal's body in preclinical studies. This optical imaging can be used with automated high-throughput systems that enable as many as 250 mice a day to be imaged, according to Dr. Piwnica-Worms. The luciferase gene can also be linked to a gene for a protein of interest so that it is activated only when this protein is expressed by the gene. With this system, the degree of light emitted will be proportional to the amount of protein produced. This enables optical imaging of molecules related to key cancer pathways in the body. Dr. Gore added that "optical imaging is a tremendously important tool in preclinical models of mice," and noted that once animal studies show the usefulness of an optical probe, researchers can then substitute a radiolabeled probe for the optical agent so that it can be imaged by PET in clinical trials.

Repetitive, non-invasive molecular imaging can provide the bridge between the genetic studies that are increasingly being done on tumor samples and the radiologic imaging done on patients, by revealing—in a dynamic in vivo fashion—the presence of key genetic biomarkers and pathways in the context of the whole organism over time, Dr. Piwnica-Worms said. In both preclinical drug development as well as in patients, molecular imaging can provide the added fourth dimension of time, which can reveal dynamic processes in the body, he stressed. For example, in one of Dr. Piwnica-Worms' studies, a luciferase reporter gene was used to reveal tumor production of a key protein targeted by an experimental drug. Repetitive optical imaging essentially provided "a real-time in vivo Western blot of protein content over time" that gave enough pharmacodynamic and pharmacokinetic data to determine an appropriate dosing regimen for a subsequent clinical trial, he said.

Researchers have also developed reporter gene imaging probes for key drug metabolizing enzymes. These probes can reveal how experimental drugs are metabolized and, in one animal study, indicated gender differences in drug metabolism (Zhang et al., 2003). In summary, Dr. Piwnica-Worms noted that these innovative probes and molecular imaging can aid both preclinical and clinical studies of experimental cancer therapies by validating mechanisms of action; providing pharmacodynamic and pharmacokinetic information; and humanizing the models so they better predict what will happen in patients. "You can get direct analysis of target-specific pharmacodynamics uncoupled from maximum tolerated dose that can guide the therapeutic clinical trial designs," he said.

Innovative imaging probes can reveal a number of physiological features of tumors that might provide earlier and more predictive measures of response to treatment than standard measures of tumor size currently used to determine response rates to treatments in clinical trials of experimental cancer drugs. Several speakers noted the promise of radiolabeled FLT (3′-Deoxy-3′-[F18] fluorothymidine) as a PET imaging probe for increased cell division. Retention of FLT in tumor cells reflects heightened activity of the thymidine kinase enzyme, which is linked to cell proliferation. Another indicator of cell proliferation is heightened production of compounds such as choline, which comprise cell membranes and can be imaged with PET or with higher magnetic field MRI systems, according to Dr. Gore. He added that there are also optical, PET, or single photon emission computed tomography (SPECT) imaging probes for the compound annexin-V, which can indicate cells are undergoing programmed cell death (apoptosis). In addition to FDG, there are other PET probes for the heightened metabolism of tumors, including those for amino acids and pH. Researchers have also developed imaging probes for how oxygenated tumors are, which radiologists can use to determine radiation dose escalation, Dr. Hricak noted.

Dr. Gore said imaging of tumor oxygenation can reveal complementary information to that conveyed by FDG imaging of tumor metabolism as to whether a treatment is having a major biological effect. "We tend to get obsessed with using single biomarkers, but there is tremendous potential in combining different kinds of imaging biomarkers. No single biomarker may be adequate," he said. This sentiment was echoed by most of the molecular imaging speakers. "We have to image a system of targets, not only one," said Dr. Hricak.

To illustrate the value of combining imaging biomarkers, Dr. Gore showed a slide of animals treated with a drug that targets the epidermal growth factor receptor, which has been linked to several solid tumors. Optical imaging revealed that the EGF receptor density was reduced by the treatment and apoptosis was increased, as indicated by an imaging probe for annexin V. But proliferation was not significantly decreased. "So we have hit the target, we have had some biological effect, but we haven't had the end result we hoped for," he said (Figure 12).

In addition to combining various biomarker molecular imaging, researchers can also combine different imaging modalities to gain more information on treatment response, Dr. Gore said. Dr. Larson noted that a PET/CT scan can show that a tumor hasn't changed in size in response to treatment, but has changed markedly in terms of its metabolism. Compa-

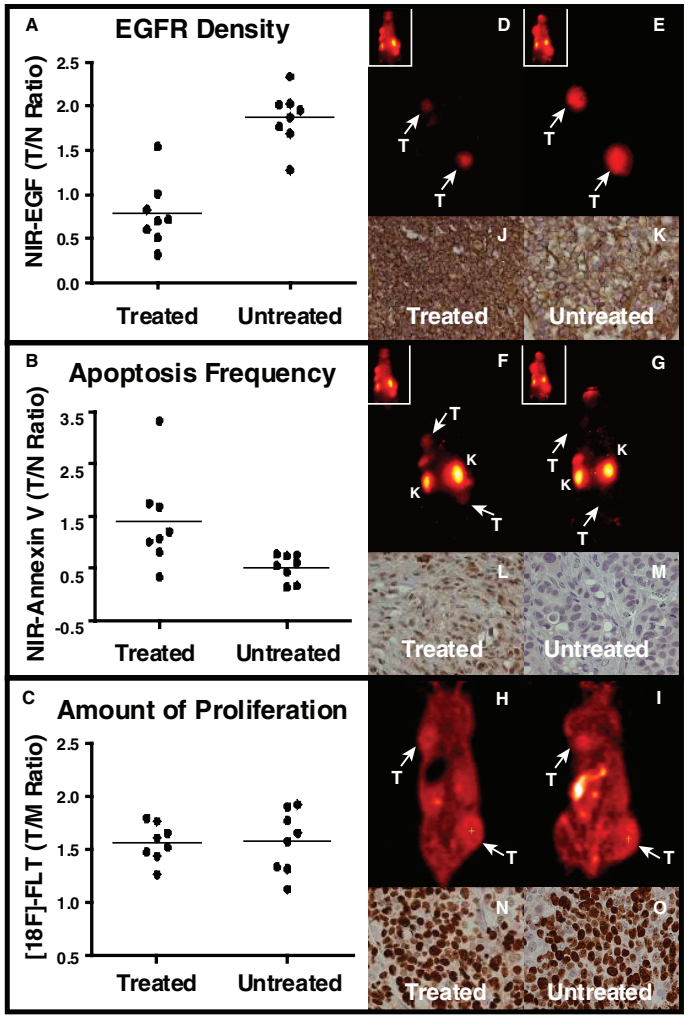

FIGURE 12 Optical imaging of animals treated with a drug that targets the epidermal growth factor receptor (EGFR), which has been linked to several solid tumors. EGFR density was reduced by the treatment (A) and apoptosis was increased (B), as indicated by an imaging probe for annexin V. But, proliferation was not significantly decreased (C). D-I show representative optical images reflecting data quantified in A-C. Brighter signal corresponds to higher target density. Tumors (T) and kidneys (K) are pointed out with arrows. J-O show representative immunohistochemistry staining in tumor tissue slices of the same markers imaged in D-I.
ACRONYMS: FLT (3′-Deoxy-3′-[F18] fluorothymidine), T/M ratio (tumor-uptake to muscle-uptake ratio), T/N ratio (tumor tissue-uptake to normal tissue-uptake ratio).
SOURCE: Gore presentation (October 4, 2007).

nies are already making hybrid scanners that combine CT with PET, and MRI–PET scanners are on the horizon, Dr. Gore noted.

Challenges of Molecular Imaging

Despite their promise, there are several challenges to ensuring that many of these molecular imaging probes meet the basic requirements for imaging biomarkers, several speakers noted. These basic requirements are that they be quantifiable, objective, accurate, sensitive to relevant biological changes (especially tumor-relevant processes), reproducible, validated, and standardized, Dr. Gore explained. But he also said imaging biomarkers have to be adequate, not perfect. "I think many people in the field criticize themselves for not having a better biomarker; but so long as it is better than the ones we are already using, it has already proven to be somewhat useful," he said. Another speaker, Dr. Larry Schwartz from MSKCC, noted that current standard imaging also has several shortcomings. He showed how standard imaging of tumor size is not reliable or standardized, and often lacks biological relevancy. For example, there is quite a bit of variability in the measurement of tumor size between tumors measured only by diameter and not bidimensionally. There can be a 3-month difference in time to progression when tumors are measured bidimensionally as opposed to unidimensionally, Dr. Schwartz pointed out (Schwartz et al., 2003).

A lack of standardization of image acquisition guidelines in clinical trials in regard to whether MRI or CT is used, the timing of contrast administration, and image slice thickness also can create discrepancies in the assessment of tumor response. For example, the size and number of cancer metastases visualized can vary greatly depending on when contrast is administered, and there can be a threefold increase in detection between images acquired at a 10-mm slice thickness versus a 2.5-mm slice thickness, Dr. Schwartz showed. He added that some of the endpoints used in clinical trials, such as a 20- or 30-percent response rate, are rather arbitrary and may not correlate with survival. "Conventional imaging uses poor surrogates as endpoints, and quite frankly biologically irrelevant response parameters," Dr. Schwartz said.

These same issues also pertain to molecular imaging. "Very often, we jump to the modalities that are not mature. When PET came, everybody wanted to use PET in clinical trials, and there were many failures because it was used before the modality was validated, standardized, and reproducible," Dr. Hricak said. Dr. Schwartz pointed out that different image set-

tings on a PET scanner can affect the results seen with FDG probes, and molecular imaging probes need to be validated as being clinically relevant. Dr. Hricak cautioned against using a contrast agent in molecular imaging before it is preclinically validated and shown to be sensitive and specific. She also questioned the use of bioluminescence imaging because it is not as quantitative as other molecular imaging techniques, such as PET. But Dr. Piwnica-Worms pointed out that once a target is validated by bioluminescence, then it can be quantified using a PET probe. "Although bioluminescence is semi-quantitative in terms of absolute photon output, it can be absolutely quantitative in terms of the biochemistry," he said, because changes over time in bioluminescence can reveal pharmacokinetics and pharmacodynamics.

Often there is a lack of standardization of imaging protocols in clinical trials, as well as a lack of harmonization between techniques used in different centers, Dr. Hricak noted. Imaging protocols are cancer type– and site-specific, she said. For example, MRI can adequately image local breast cancer, but bone metastases are best imaged with PET and an appropriate probe, yet the modality used during a study should not change. She suggested including an imaging expert when designing a clinical trial to help ensure the trial's success.

Ideally, tumors should be visualized in volumetric displays, Dr. Hricak and Dr. Schwartz said. A slide of Dr. Schwartz's showed a nearly 40-fold difference in tumor percentage change following treatment, depending on whether just the tumor diameter was measured versus whether the volume of the tumor was measured. Dr. Schwartz optimistically summed up the discussion of the challenges involved in molecular imaging by saying, "We in imaging view many of these as challenges that are readily achievable by obtaining appropriate image acquisition guidelines which could be standardized in a rational manner."

One additional challenge mentioned by Dr. Larson is how to disperse the bioimaging tracers being developed at individual laboratories to the wider research community and into clinical trials. In a later discussion, Dr. Hricak noted that at MSKCC alone, "there is a menu of radiotracers that have been around for at least 5 to 15 years that are not FDA approved and widely distributed." Dr. Larson advocated using the nuclear pharmacies that are scattered throughout the world to better distribute molecular imaging probes. These pharmacies use automated chemistry to make and ship labeled molecular PET probes.

In the panel discussion that followed the molecular imaging presenta-

tions, Dr. Mills echoed Dr. Hricak by stressing the lack of harmonization in bioimaging protocols between centers, which may not have the same hardware or software platforms, and which believe they have already optimized the components of their own protocols. "Until we establish harmonization on top of standardization, we are still going to have limitations for applications in clinical trials," he said. Panel member Dr. Jeff Evelhoch of Amgen agreed that "that is probably the biggest challenge that we have in using imaging in clinical trials," and that this challenge varies with the imaging modality. He noted that because bioimaging methods for dynamic contrast enhanced (DCE) MRI are continually developing and progressing, harmonization is more difficult to achieve than with more standard CT imaging. But there is some consensus on an appropriate acquisition protocol and other standards needed across centers in clinical trials using this technology, he added. In contrast, FDG PET, even though it is used more often in the clinic than DCE MRI, has fewer agreed-on standards even within the same institution.

Pfizer's Dr. McCarthy noted that a number of initiatives have been made to standardize or harmonize molecular imaging biomarkers, including one by the National Institute of Standards and Technology, and the Oncology Biomarker Qualification Initiative sponsored by the FDA, the NCI, and Centers for Medicare & Medicaid Services (CMS). But these efforts at standardization and harmonization are going on in parallel with each other without coordination and agreement. Genomic Health's Dr. Shak pointed out that genomic assays faced the same issues on standardization and harmonization, and successfully met them with financial investment in resources needed to ensure the technologies and the procedures were in place to provide quality control and harmonization. Dr. Mendelsohn added that such standardization and harmonization may not be as critical in clinical trials that measure tumor response or other variables over time as long as there is reproducibility at each participating institution. "If the tumor went down 50 percent in a reproducible way at that institution, it might be just as important as in another institution where they might have picked up four other lesions because [their measurements are] more sensitive. But you have still got to go down 50 percent," he said. "Where you need perfect standardization and harmonization is when it is a one-shot thing—either the androgen receptor is present or absent, and there you would need it."

SCREENING FOR PREDICTIVE MARKERS

Biomarkers may possibly be used to predict a number of factors relevant to cancer treatment, including aggressiveness of a tumor and the need for treatment, likelihood of responding to specific treatments, likelihood of developing adverse reactions to treatment, and prognosis. On the second day of the conference, the first session focused on the progress and challenges linked to identifying and validating such predictive markers, as well as applying them in a clinical setting. Drs. Pierre Massion of Vanderbilt Ingram Cancer Center, James Heath of California Institute of Technology, Dan Sullivan of Duke University, and Daniel Von Hoff of Translational Genomics Research Institute addressed these issues in the presentations, and provided answers to specific questions posed by the National Cancer Policy Forum at a panel discussion that followed the presentations.

Several presenters stressed the clinical need for predictive biomarkers in oncology. Using lung cancer as an example, Dr. Massion pointed out that such biomarkers are needed for every step in patient management. Markers that can predict a person's risk of developing lung cancer are needed for people who smoke and might benefit from heightened screening or participating in various chemoprevention trials. Blood, sputum, or other non-invasive biomarker tests are needed to improve diagnosis once physicians detect a suspicious lesion in a patient's lungs. Currently, diagnosis can only be done reliably with invasive surgery or bronchoscopies.

Also needed are biomarkers that can predict the likelihood that a small early lesion in the lungs will progress to a deadly cancer. A suspicious lesion that is less than 2 centimeters cannot be accurately diagnosed as malignant using a PET scan, and may not be accessible via bronchoscopy. A needle biopsy poses the risk of lung collapse, with the only other proactive option—surgical removal—being even more invasive and risky. Physicians can take the "wait-and-see" approach to such lesions, of which 20 percent may be malignant. But with that approach, one may miss a chance of curing an aggressive lung cancer.

Because only 30 percent of patients with lung cancer respond to radiation or chemotherapy, and most of them will experience some toxic reactions to those treatments, there is a great need for biomarkers that can predict response to treatment, recurrence, and prognosis. "We are overtreating cancer in general, and in some cases we undertreat the proper subset of patients," Dr. Massion said, adding that predictive biomarkers will "allow us to eventually provide adjuvant therapy to the appropriate population, narrow down the patient selection, and decrease costs of therapy."

Drs. Massion and Heath gave examples of biomarker tests that are already being used in a clinical setting. These include the ER and HER-2/neu biomarker tumor tests that predict response to various breast cancer treatments, gene signature tumor tests based on the activation patterns of 21 or 70 genes that predict breast cancer recurrence or survival (Paik et al., 2004), and a tumor genetic signature test based on 100 genes that can distinguish between two types of similarly appearing lymphomas, and predict survival following treatment (Dave et al., 2006).

Many other predictive biomarkers are in the early stages of clinical testing, including one that uses the patterns of 40 proteins in a blood sample to distinguish between prostate inflammation and prostate cancer,[10] an eight-gene signature blood test that predicts response to lung cancer treatment (Taguchi et al., 2007), and those that use the patterns of proteins in the blood, or genes activated in airway epithelial cells to predict lung cancer (Yildiz et al., 2007) (airway epithelial cells can be easily brushed off and collected noninvasively for such diagnostic testing). Initial testing of the blood biomarker test for predicting lung cancer suggests it has a sensitivity of 58 percent and a specificity of 85.7 percent. "Although this beats any biomarkers that are currently available in the blood for patients with lung cancer, it may not have the sensitivity you wish for in the early detection approach and the specificity may not be optimal," Dr. Massion noted.

Despite the need for biomarkers in oncology, and the thousands of cancer-related biomarker publications over the past 10 years or so, only about 20 cancer biomarkers have been approved by the FDA (Figure 13), and many of these are not used routinely in clinical practice. (Ludwig, 2005). "The world of cancer biomarkers is a very humbling one, and not a successful one to my eyes," Dr. Massion said. Dr. Heath added that "the world of biomarker discoveries has been advancing over the past decades in leaps and bounds, but it is still remarkably immature."

The Challenges of Clinical Validation

A major hurdle that needs to be overcome for more predictive biomarker tests to enter the clinic is the validation of the diagnostic accuracy and usefulness of existing candidates, according to Dr. Massion. Such validation should be done via several large, independent studies at multiple

[10]Personal Communication, J.R. Heath, E.W. Gilloon Professor of Chemistry, California Institute of Technology, October 5, 2007.

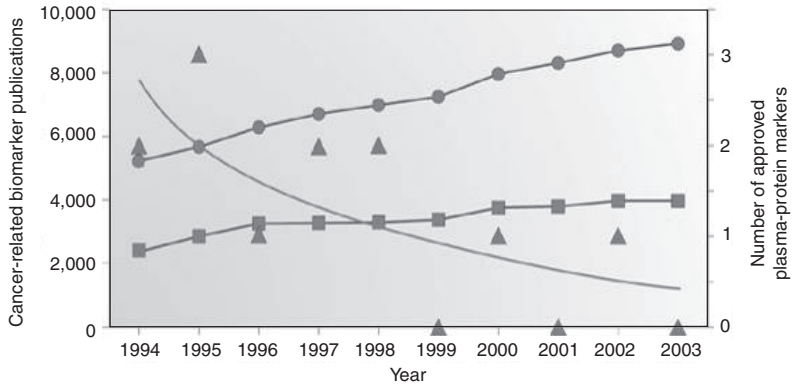

FIGURE 13 Publications and FDA-approved biomarkers. Despite the increasing rates of publications on biomarkers, the number of FDA-approved plasma-protein tests is decreasing. Triangles and the associated trend line represent the number of FDA-approved plasma-protein markers per year. Squares and circles indicate publications under the Medline medical subject heading "biomarker" and text word "biomarker," respectively.
SOURCES: Massion presentation (October 5, 2007) and Ludwig (2005). Reprinted by permission from Macmillan Publishers Ltd: *Nature Reviews Cancer* 5(11):845-856, Copyright 2005.

institutions that use different testing platforms and well-annotated patient samples. This validation remains challenging and needs to keep pace with the rapid progress in assay development. "You see a myriad of publications looking at biomarker discovery and first-phase validation, but very few are putting them within a clinical context. This is where we need to go—what we need to do," said Dr. Massion. Under the auspices of the NCI and the Specialized Programs of Research Excellence (SPORE) program, he and other researchers have created a group called the Lung Cancer Biomarkers Group, which aims to provide several academic institutions with access to four different sets of patient sample materials held at an NCI repository for the purpose of testing the accuracy of lung cancer biomarker tests. Such testing will not only access accuracy of the tests, but also their reproducibility within and between institutions and how they can provide clinically useful information. Later in the discussion, Dr. Massion added that "we need to establish repositories that allow us to validate biomarkers within institutions and across institutions and across platforms. This is unfortunately very time consuming and also requires centralized repositories of prospectively acquired samples and a great deal of collaboration between institutions."

Dr. Massion gave several examples of clinical trial designs for studies aimed at assessing the clinical utility of predictive biomarker test (Figures 14-16). The simplest study design is to compare outcomes of patients who test positive for the biomarker with those of historical controls (Figure 14). Another design that is much more costly to run is to randomize patients as to whether they undergo the biomarker test or not (Figure 15). Of those tested for the biomarker, patients with positive results receive the intervention the marker indicates is warranted, while those with negative test results receive standard care. This study is designed to determine whether the predictive test improves patient outcomes when compared with unselected patient treatment. This study requires a large number of patients. Another study design does not address the quality or value of the biomarker itself, but rather compares outcomes for two different interventions in those who test positive for the biomarker and as well as those who test negative for

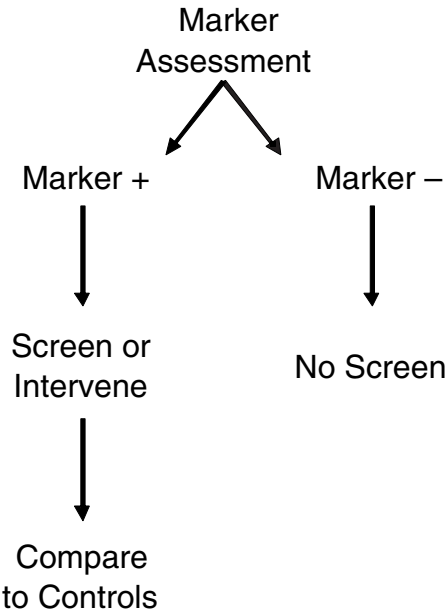

FIGURE 14 Clinical utility of predictive markers, study design 1: A single-arm validation study using historical controls for comparison. In this study, all patients receive the biomarker test and outcomes of patients who test positive for the biomarker are compared with those of historical controls.
SOURCE: Massion presentation (October 5, 2007).

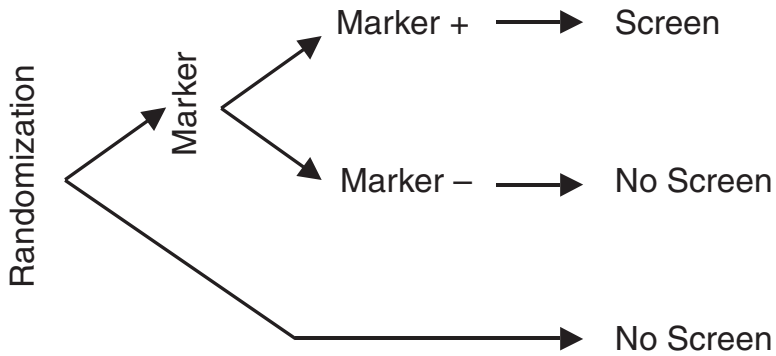

FIGURE 15 Clinical utility of predictive markers, study design 2: Randomization to receive or not receive the biomarker test. This design determines whether the predictive test improves patient outcomes when compared with unselected patient management. In this study, patients are randomized, and one group is given the biomarker test, the results of which influence the therapies received by the patients. The therapeutic outcomes of the tested and untested groups are compared.
SOURCE: Massion presentation (October 5, 2007).

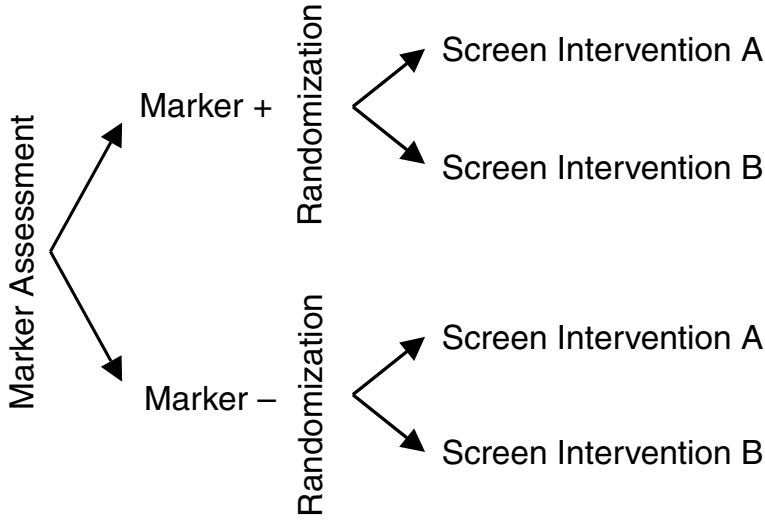

FIGURE 16 Clinical utility of predictive markers, study design 3: Randomization of treatment irrespective of biomarker test results. This design compares two interventions in both marker-positive and marker-negative groups. In this study, all patients receive the biomarker test, and are randomized into either of the two treatment arms.
SOURCE: Massion presentation (October 5, 2007).

the same biomarker (Figure 16). This study design randomizes between the two interventions.

Dr. Heath noted that often biomarker tests show promise when initially tested on a population of interest, such as men with prostate cancer, versus a healthy control population. "But as soon as you look at them across the general population, they tend to fall apart, and that is one of the reasons why the approval of biomarkers has been relatively slow. What you would like to be able to do is take your top 1,000 candidates and just measure them broadly across all population bases and do it cheaply and rapidly," he said.

One reason Dr. Massion gave for the lack of clinically successful biomarker tests in cancer are their lack of competitiveness in terms of costs and reimbursement. Dr. Heath expanded on this in his presentation on lowering the cost of in vitro diagnostics. He offered several suggestions for making biomarker tests more rapid, inexpensive, sensitive, and clinically relevant; the first and foremost is to ensure the tests are based on accurate and appropriate biology. "The first thing is to get the biology right because obviously if you are measuring the wrong thing, who cares," he said.

The reagents and materials used in the test are also critical. "You will see a lot of interesting devices these days to measure proteins or messenger RNAs or whatever. Many of these tests look exotic. But if it looks exotic, it is probably not going to be something you are going to use in the clinic," he said, adding that the test should use inexpensive and scalable technology, as well as inexpensive reagents and equipment, such as glass and plastic. "This is a huge issue, and it is probably the limiting issue in antibodies," Dr. Heath said.

Biomarker tests should also require very small amounts of tissue or blood, such as a finger prick of blood, yet be highly sensitive and quantitative because many of the compounds of interest are present in exquisitely minute amounts in patient samples. To improve the sensitivity and quantitativeness of such tests, Dr. Heath suggested using fluorescent markers and a scattering microscope, which has an aperture in front of a light microscope that enables automated counting of trace compounds of interest. It is about 10,000 times more sensitive than standard protein assays and can detect compounds at 100 attomolar concentrations.

Given the complexity of the molecular pathway networks that underlie various cancers, biomarker tests should be multiparameter tests that can be automated and done rapidly because, as Dr. Heath pointed out, time equals money. Time can be decreased by not making tests diffusion dependent, as are standard ELISA antibody-based assays. These tests require a few hours

for proteins of interest to diffuse to the labeled antibodies with which they bind. Tests can be based on the kinetics of proteins binding to antibodies without requiring such diffusion, according to Dr. Heath, and be completed in 5 to 30 minutes with a cost of 5 to 20 cents per measurement, depending how many measurements are made simultaneously.

"The idea of this kind of technology is to put every single possible biomarker you could imagine on the chip so we can capture the diurnal and dietary variation, and all the other fluctuations that tend to foul up the validation of a biomarker, but then also correlating it with traditional pathology and disease," Dr. Heath said. He also gave examples of what he called "PET on a chip," which is a microarray test developed by researchers at the University of California, Los Angeles, that partitions cells from a patient's tumor biopsy into 100 different wells that contain metabolic markers for response to drugs and can indicate—within an hour of when the patient was sampled—to which drug regimen the patient is likely to respond.

Bioimaging Predictive Markers

Following Dr. Heath's presentation, Dr. Sullivan gave examples of how bioimaging can be used to predict clinically relevant variables in oncology; the advantages and disadvantages of using imaging biomarkers; and the technical and regulatory challenges of making those biomarkers clinically useful. As previous speakers noted, bioimaging can predict where a cancer drug will concentrate in the body, and various physiologic states such as a lack of oxygen (hypoxia), or diffusivity that can affect drug response. Two small studies suggest DCE MRI might be useful as a predictor of survival following treatment for osteosarcoma or renal cell cancer (Reddick et al., 2001; Flaherty et al., 2008). Two clinical studies found PET imaging of hypoxia predictive of response to drug treatment, and larger multi-institutional trials have been planned to assess the effectiveness of such bioimaging (Rischin et al., 2006; Dehdashti et al., 2003). Two studies also indicate that PET imaging of labeled estradiol predicts response to hormonal therapy in advanced breast cancer (Linden et al., 2006; Mortimer et al., 2001). Diffusion imaging was found to predict response to treatment in brain cancer (Hamstra et al., 2005), and encouraging results from a study of magnetic resonance spectroscopy in non–Hodgkin lymphoma patients have led to a multisite trial to test prospectively whether such imaging can identify patients who would respond to conventional therapy versus patients who should receive more aggressive treatment, such as a bone marrow trans-

plant (Arias-Mendoza et al., 2004). Researchers are also starting to conduct clinical studies of the usefulness of assessing multiple imaging biomarkers to predict indolent disease (Shukla-Dave et al., 2007).

Although bioimaging is well suited for revealing physiologic measures such as hypoxia, diffusion, or tumor metabolism, such measurements may be strongly influenced by other interfering systemic conditions and result in a misleading reading, Dr. Sullivan pointed out. For example, for FDG imaging of tumor metabolism, "there can be other things going on in the body and the brain, the heart and the skeletal muscle that could suck up all the glucose and give a spuriously low value in the tumor, so these have to be controlled for," Dr. Sullivan said. Although imaging biomarkers lack the diversity and large number of parameters that can be simultaneously discerned compared to genomic or proteomic tests, he said, imaging biomarkers do provide the spatial and temporal context for the markers, unlike in vitro tests. "Given that cancer is a heterogeneous disorder and a systemic disorder, that contextual information might be important or useful in making predictions of response to therapy," he noted. For example, drugs such as tirapazamine (SR-4233) are most active in tissue lacking oxygen. PET scanning with a marker for oxygenation could be an especially useful predictor of which patients will respond to this drug.

Bioimaging can also preserve some physiological information, such as pH or oxygenation status, that might otherwise become disrupted or lost in the sample preparation involved in the in vitro tests, he added. Such imaging is more likely to accurately reflect the true physiological state of normal or cancer cells. Another advantage of molecular imaging over some genetic or proteomic biomarker tests is that it can be less invasive because it does not require tumor samples. But several technical challenges are involved in developing imaging biomarkers, particularly if they use radiolabels and small molecules. Although it is relatively easy to label antibodies with a radioactive probe for detection in an imaging system, it often is difficult to label small molecules with the radioactive carbon, oxygen, fluorine, or nitrogen atoms used in PET imaging. For example, although the cancer drug gemcitabine has a few fluorines as well as carbon atoms in its structure, several attempts to replace these with radioactive fluorine or carbon atoms for PET imaging failed after much trial and error. "So this has been years in development and they are still not there yet. It is not a straightforward process in many cases," Dr. Sullivan said.

One problem in this regard is that the drug kinetics may not be appropriate for imaging of a labeled drug. Imaging agents are usually better if they

have irreversible binding, but some drugs have reversible binding. One also has to develop a rapid synthetic pathway for the radiolabeled drug that can be accomplished within the time constraints of the half-life of the isotope used to label it, which may only be a few hours. It can also be challenging to produce a compound that is sterile, doesn't induce a fever, and can be immediately injected into a patient. "It can take years to get these bioimaging probes ready for use in patients, and meanwhile the drug development is moving along," Dr. Sullivan said.

An alternative to labeling the drug for bioimaging is to label another ligand for the target of interest, such as a drug analog or a growth factor that binds to the same receptor as the drug. However, although this may enable easier labeling synthesis, the labeled ligand may not reveal drug localization as accurately as the labeled drug and requires more validation. The time course for development of such labeled ligands can be as long or longer than that for labeled drugs, so that also may be out of synch with corresponding drug development.

In addition to the technical challenges of developing bioimaging biomarkers, there are substantial regulatory hurdles. Currently the regulatory process for these markers is the same as that for drugs: Studies must show their clinical benefit and safety for patients. "Right now there is no commercial pull for the industry to develop these agents and there is a lack of resources being devoted to their development partially because of this regulatory process," Dr. Sullivan said. "Many people believe that for these imaging tracers that have no pharmacologic effect, the target or benchmark for efficacy should be the same as it is for devices approved by the FDA; that is, the agent should provide the information which the producer or vendor claims that it provides, and that it would not necessarily provide clinical benefit," he added.

A final challenge that Dr. Sullivan discussed is the validation of imaging biomarkers. He noted many sources of variability in bioimaging that need to be considered and controlled for, including such physical sources of variability as scanner calibration, machine variation, different image acquisition parameters, and different algorithms for data processing. There are also physiological sources of variability, including intra- and interpatient variation and reader variability. Performing the necessary repeatability tests for imaging methods is especially difficult and costly because they are performed on people, not specimens.

If imaging is to become a reliable in vivo assay, Dr. Sullivan said, several factors must be in place: uniformity of instrumentation, protocol-specified

acquisition of images, independent quality control, reliability and independence of reader interpretation and provenance, auditability, and accessible storage of results.

In a presentation in the following session, Dr. Gwen Fyfe of Genentech also noted some of the potential pitfalls of bioimaging biomarkers and the need to properly validate them. "Expensive techniques save time and improve outcome only if carefully validated by clinical outcomes," she said. She gave an example of a small bioimaging study of a VEGF receptor inhibitor that showed that colorectal tumor vascularity and permeability decreased rapidly following treatment with the inhibitor and seemed to correlate with clinical benefit (Morgan et al., 2003, and Steward et al., 2004). But when a larger randomized study was done, the drug had no impact on overall progression (Hecht et al., 2005). She noted that a bioimaging study that predicts response after only one cycle of treatment might not predict a more durable response. "A biomarker probably has a good negative predictive value—if you don't see an impact it is unlikely to be useful—but the positive predictive value is really subject to interpretation," Dr. Fyfe said. "We need to validate these biomarkers very carefully with clinical outcomes before we assume that we can go from Phase I to Phase III based on an imaging result."

But it is debatable what the appropriate validation is, Dr. Fyfe added. "Is it response, [is it] durable response, or is it progression-free survival?" she asked. Not only does there have to be validation of technique, such as reproducibility that considers site and patient variability and timing of analyses, but there needs to be validation of patient benefit. Such validation may be disease-, pathway-, or drug-specific. Even within a pathway, agents may differ significantly in their mechanism of action, she pointed out. Follow-up exploratory trials that assess the relationship of the biomarker effect to clinical outcome are needed to avoid large negative trials.

Dr. Fyfe concluded by noting the need for information sharing among academia and pharmaceutical and biotechnology companies to validate biomarkers. The Biomarker Consortium[11] is a good start, she said, but

[11]The Biomarkers Consortium is a public–private biomedical research partnership managed by the Foundation for the National Institutes of Health that includes government, industry, patient advocacy groups, and other non-profit, private-sector organizations. In addition to the Foundation for NIH, founding members include the NIH, the FDA, and PhRMA. Other partners in the consortium include CMS and the Biotechnology Industry Organization. The Consortium aims to "rapidly identify and qualify biomarkers to support basic and translational research, guide clinical practice and, ultimately, support the develop-

she called for more government investment in such validation efforts and stressed that all biomarker validation results should be published, both negative and positive. "Each company has its own interests so for us to really work on this there needs to be an NCI-directed effort because I think pharma and biotech will help, but fundamentally that is going to be a very splintered effort, and we are going to learn best by doing careful studies that are in the public domain," Dr. Fyfe said. In the discussion that followed, Dr. Steven Larson added that the FDA and the U.S. Pharmacopeia, in addition to the NCI, should do more to aid efforts at validating and standardizing imaging biomarkers. "The FDA could develop a path which would allow for qualification of these individual biomarkers for the specific biochemical or pathway purpose for which they are intended," he said.

Clinical Translation

Given the complexity of the molecular mechanisms that underlie specific cancers and the challenges involved in developing and validating biomarker tests predictive of those mechanisms, translating the research findings on predictive biomarkers into tests with clinical usefulness can appear to be an especially difficult hurdle to overcome. But Dr. Von Hoff, the last speaker in this session, described a simplified approach to such translation that has been used at the Translational Genomics Research Institute (TGen). "It is an understatement to say that work on biomarkers is complicated and that screening for predictive markers is going to take a while. But we need to help patients who are sitting in front of us right now with refractory cancer. So our clinical research teams are focused on applying what we already know about mutations, translations, and deletions. We feel this is an important policy because there is a lot you can actually do right now," he said.

To apply that knowledge, Dr. Von Hoff proposed that oncologists assess the "clinical and molecular contexts of vulnerability" of their patients' tumors— which he referred to as a "sixth vital sign"—and use those contexts to help select therapy. As an example of clinical context of vulnerability, he described an 81-year-old patient with lung cancer who smoked for 72 years. Dr. Von Hoff said this patient's history suggests that his tumor can repair

ment of safe and effective medicines and treatments." The Consortium also plans to "harmonize approaches to identify viable biomarkers, verify their individual value, and formalize their use in research and regulatory approval" (http://www.biomarkersconsortium.org).

TABLE 3 Contexts of Vulnerability

Tumor Type	Vulnerability	Agent(s)
Ewing sarcoma	Growth factor receptor (IGFR1)	AMG479; CP751, 871
Ewing sarcoma	Phosphoinositide 3-kinase (enzyme)	SF1126
Synovial cell sarcoma	Gene translocation, Growth factor receptor	Iressa/Tarceva
Chondrosarcoma	Tumor necrosis factor-related apoptosis-inducing ligand (TRAIL)	TRAIL interactive agent
Alveolar soft part sarcoma	Oncogene—gene fusion	C-met inhibitor
Osteogenic sarcoma	C-met (Mesenchymal epithelial transition factor) abnormalities	C-met inhibitor
Small blue round cell tumors—Ewing, osteosarcoma, neuroblastoma, desmoplastic small round cell, synovial	Platelet-derived growth factor receptor	Platelet-derived growth factor receptor inhibitor
Chronic Myelogenous Leukemia	Oncogene—fusion protein (Brc-abl)	Gleevec

almost any DNA damage, so giving him chemotherapy that induces such damage is not likely to be effective, and other options are warranted.

The molecular or genomic context of vulnerability refers to what Dr. Von Hoff called the molecular addiction of the tumors. Many breast cancers, for example, are "addicted" to estrogen and need this growth factor to survive. Over the past few decades, researchers have noted numerous other growth factors, enzymes, and other compounds that tumors need to survive (Table 3). Drugs have already been developed that target these vulnerabilities, Dr. Von Hoff noted, so knowing the tumor addiction can help with treatment selection and improve treatment effectiveness.

He has found this approach to be remarkably effective in some cases, even in patients with advanced cancer that has not responded to conventional treatment. For an example, he described metastatic myxoid liposarcoma, a particularly aggressive cancer of the connective tissue whose hall-

TABLE 3 Continued

Tumor Type	Vulnerability	Agent(s)
Acute Lymphoblastic Leukemia	Oncogene—fusion protein (Brc-abl)	Gleevec
Chronic Neutrophilic Leukemia	Oncogene—fusion protein (Brc-abl)	Gleevec
Hypereosinophilic syndrome	Growth factor receptor mutations	Gleevec
Medulloblastoma	Growth factor receptor mutations	Gleevec, hedgehog
Gastrointestinal Stromal Tumor	Growth factor receptor mutation	Gleevec, sunitinib
Prostate cancer	Oncogene—fusion protein	HDAC inhibitor reversing the phenotype
Castleman disease	Increased interleukin-6 (growth factor)	CNTO 328

NOTE: This table shows the types of vulnerabilities and the therapeutic agents that have been developed to target them in specific tumor types.
ACRONYMS: IGFR1 (insulin-like growth factor receptor 1), AMG479 (fully human antibody against IGFR1), CP751,871 (monoclonal human antibody against IGFR1), SF1126 (Vascular Targeted pan-PI3K Inhibitor), TRAIL (Tumor necrosis factor–related apoptosis-inducing ligand), C-met (Mesenchymal epithelial transition factor), HDAC (Histone Deacetylase), CNTO 328 (human-mouse chimeric monoclonal antibody to interleukin-6).
SOURCE: Adapted from Von Hoff presentation (October 5, 2007).

mark is a specific and relatively simple genetic abnormality (translocation between the 12th and 16th chromosomes). A recent study showed that a drug in development, ET-743, is effective in only 7 percent of all patients with connective tissue cancers, but 97 percent effective in myxoid liposarcomas (Grosso et al., 2007).

According to Dr. Von Hoff, a tremendous number of tumors' genomic contexts of vulnerability are deletions, translocations, or other simple genetic abnormalities. For example, about 64,000 U.S. breast cancer patients have mutations in their BRCA1 or BRCA2 genes, which a recent study showed are likely to respond to drugs known as PARP inhibitors that target the abnormal DNA repair associated with these mutated genes. "Biomarker patterns are great, but they are going to take longer to be useful," Dr. Von Hoff said. In the later discussion he added that deletions and mutations are easier to measure and are more likely to be reproducibly measured than

biomarker pattern assays. Deletion, translocation, and mutation assays are reliable and don't have the "fudginess" of microarray assays, he said.

Dr. Von Hoff reported that researchers at TGen are using small interference RNA techniques to design drugs that target the specific genetic abnormalities that cause tumor addictions. "Instead of treating patients with a drug and then finding some pattern that indicates what the genetic deletion is that makes the drug effective for some patients, we design a drug that targets only that deletion," Dr. Von Hoff said. "We get the marker for the genetic abnormality and then design the drug to take out cells with the marker."

But determining the addictions of patients' tumors requires researchers to identify and catalogue the contexts of vulnerability as rapidly as possible. This suggests the need for a centralized clearinghouse for such characterization of patients' tumors, which would avoid the need to send tumor samples to several different facilities, each with the capacity to evaluate tumors for only one or two genetic abnormalities. To meet that clearinghouse need, Dr. Von Hoff and his colleagues created the Tissue Banking Analysis Center (TBAC). TBAC, located in Phoenix, assays tumors for all molecular targets for which there are therapeutics. TBAC is sponsored by U.S. Oncology and the Molecular Profiling Institute, and is the only one of its kind, according to Dr. Von Hoff. More such clearinghouses like TBAC are needed throughout the world, he said.

Patients who have their tumors analyzed at TBAC have the opportunity to participate in Phase I or II clinical trials enriched with patients whose tumors have specific molecular abnormalities. Such clinical trials are models for the approval of a new agent aimed at a specific molecular target in a patient's tumor rather than designated for a particular histologic type of cancer, Dr. Von Hoff noted. He reported on a new clinical trial design for evaluating an agent against a target rather than against a tumor type. With this design, patients whose tumors have the specific molecular target are treated with an agent aimed at that target. Patients who experience a complete or partial remission continue taking the agent, while patients who progress are taken off the study (Figure 17). Researchers at TGen have instituted five such trials, and using TBAC was key to enabling these trials, Dr. Von Hoff pointed out.

To see how commonly researchers would be able to discern a genetic abnormality in patients' tumors for which there are already agents that target them, Dr. Von Hoff and his colleagues conducted a pilot trial called Target Now (Von Hoff et al., 2006). This study of 112 cancer patients found

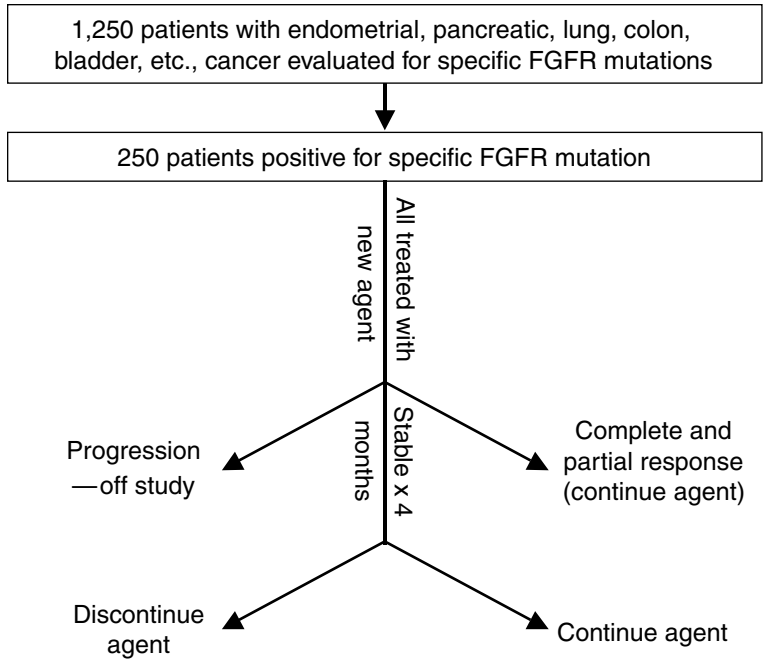

FIGURE 17 New clinical trial design for evaluating an agent against a target rather than against a tumor type. With this design, all patients receive a test for the FGFR mutation. Those with the mutation receive treatment designed to target the mutation. If the patient experiences a positive response, the therapy is continued.
ACRONYMS: FGFR (fibroblast growth factor receptor).
SOURCE: Von Hoff presentation (October 5, 2007).

that standard immunohistochemistry assays for 13 possible targets found at least one potential target in about three-quarters of the patients, with an average of 1.6 targets per patient for which a conventional therapeutic agent was available. Microarray analyses found an average of 11 targets per patient for which there was a potential therapeutic agent, and virtually all patients had at least one potential target identified with this analysis.

The physicians of the patients in this study provided abundant anecdotal evidence that this approach has been remarkably effective in some cases. For example, a patient with advanced ovarian cancer, who progressed on four prior regimens, responded to tamoxifen after estrogen receptors were found as a target. However, he said, there have some dramatic anec-

dotal failures of the approach as well. But overall, Dr. Von Hoff estimates that with this targeted approach to treatment, response rates range from 26 to 30 percent, and the response rates are even higher for patients with rare tumors. Meanwhile the average response rate for patients in Phase I clinical trials is about 4 percent. He and his colleagues are currently conducting a prospective clinical trial called the Bisgrove trial to assess this more accurately. The endpoint measured in this trial is time on the therapy selected by molecular profiling, versus time on the therapy the patient had just received prior to the study. The rationale for this endpoint is that the length of time a patient responds to a therapy usually gets progressively shorter with successive therapies as disease progresses. Thus, if the time on therapy increases, it suggests that the profiling-selected therapy has changed the natural history of the disease (Box 1). "If the results from this trial are promising, we will have to rethink whether or not Phase I trials should be done in patients who are not profiled," Dr. Von Hoff said.

"Discovery and use of biomarkers is tough in drug development—it takes a long time," he concluded. "But we shouldn't be paralyzed by that and instead should focus on the contexts of vulnerability that are deletions, mutations, or translocations. These are easier to find and will probably foster more dramatic results within smaller clinical trials." He added, "It is

BOX 1
Details on the Endpoint for the Bisgrove Trial

1. Usually the period of time a patient is on successive therapies is progressively shorter.

2.

Period A-TOTAL	Period B-TOTAL
Previous therapy	Therapy selected by molecular profiling

3. If period B is greater than period A, the profiling-selected therapy has changed the natural history of the patient's disease.

4. If 30% of patients on this Bisgrove trial have period B longer than period A, then molecular profiling helps.

SOURCE: Von Hoff presentation (October 5, 2007).

very feasible to molecularly profile nearly all patients' tumors [for patients] who are candidates for Phase I trials, and there is a great need for clearinghouses where patient tumors can be sent to be assayed for their context of vulnerability."

Panel Discussion

In the panel discussion following the presentations, Dr. Massion was asked what is needed to improve biomarker technologies to make them more useful for developing better and more efficient trials. He responded that "our ability to mine these datasets that we obtain from genomics and proteomics is falling behind our ability to generate these data. There is a great need for analytical tools and ways to analyze these data. We are trying to reconstitute a puzzle that is extremely complex."

He added that biomarkers also have to be discovered within specific clinical contexts that address the heterogeneity of cancers rather than testing them within broad populations. "If you address lung cancer as a whole, for example, you are limiting yourself and then you are actually going to face a lot of difficulty in applying biomarkers to specific subgroups. We should consult with our clinicians and epidemiologists to think about biomarker discovery in the specific clinical context and then rapidly test the biomarker in that context, and our preclinical models should mimic that clinical context."

Later in the discussion, Dr. Joe Gray from the University of California, San Francisco, added that microarrays generate an enormous volume of data, with a nearly infinite number of marker combinations that might be predictive. Rather than abstractly analyze these data and search for any pattern that might be predictive, he suggested analyzing the data with the awareness that the data are informative about the underlying biology of specific molecular pathways or networks. "We need to organize the data in the context of the biological process that is deregulated so that any of the following 27 markers actually inform you that it is [for example] the BRCA DNA repair pathway that is actually deregulated and all of these markers ought to point you to that. Until we start thinking about interpreting the data in that context, we are going to be lost in this chaos of marker space," he said.

Dr. Massion also reiterated the need for rigorously validating biomarkers. Such validation can be aided by establishing centralized repositories of prospectively acquired patients' samples, and by collaboration among institutions so biomarkers can be validated across institutions and platforms.

Dr. Heath was asked how the use of predictive markers would affect trial design and implementation. He replied that, although predictive markers would have great value in stratifying patients for clinical trials, the challenge still remains to generate for these trials predictive markers that are noninvasive, such as those found in the blood or another readily accessible body fluid as opposed to the tumor, which requires an invasive biopsy. He pointed out that such noninvasive markers cannot directly detect a translocation or other genetic abnormality, but rather reflect the status of the tumor. Such reflection requires multiple-parameter measurements. Yet physicians and diagnostic companies are more familiar with single-parameter tests such as the prostate-specific antigen (PSA) test, according to Dr. Heath. "They are very uncomfortable looking at a panel of markers that goes through some computation program to give them back an answer, and I think there is a significant amount of physician retraining that has to be done to counter this," he said.

Dr. Von Hoff asserted that Dr. Heath underestimates physicians. "We would love to work with you on this because medicine is really very pattern oriented and physicians are used to putting all those patterns into their decision-making process every day," he said, adding the example that antibiotic sensitivity testing "is never just black and white"; physicians have to consider many parameters when selecting the appropriate antibiotic for their patients. Dr. Heath noted, however, that "diagnostic companies have been reluctant to go down this pathway, and without commercialization, you can't get it into people's hands." Dr. Mills added that proper validation of predictive biomarkers through clinical trials would be a way "to reassure the practicing physician that the pattern is reproducible and does mean something effectively."

Dr. Mills then asked Dr. Sullivan if the use of predictive markers would increase the number of patients willing to participate in cancer clinical trials. Noting he had no data to back up his opinion, Dr. Sullivan said he suspected that predictive markers would increase the number of such patient volunteers because "if patients had some sense the tests are being used to intelligently sort them out to where there might be some benefit to them and less harm, I suspect that they would find that appealing." Patient advocate Kathy Meade of the Virginia Prostate Cancer Coalition added that predictive markers would appeal to many of the prostate cancer patients she deals with who are often interested in mechanistic explanations for the tests and treatments they receive and appreciate biomedical thinking that is "outside the box."

Further discussion centered on standardizing biopsy procedures so they are useful for biomarker analyses and whether health insurers will reimburse for predictive marker tests. Dr. Herbst of MD Anderson Cancer Center pointed out the need for reliability in the tumor tissue used to determine the predictive biomarker. He raised the point that biopsy tissue should be prepared in a way that is compatible with biomarker assays, and he questioned whether image guidance is needed to biopsy a "hot spot" in a tumor or to consider tumor heterogeneity. Dr. Heath responded that, ideally, measurements are made immediately on fresh tissue and that those measurements are available later to researchers who use the stored tissue. "We found that when we looked at stored tissues, including blood, one of the largest fluctuations was the protocol that was used to store it," he said.

As for addressing tumor heterogeneity in sampling and testing tumors, Dr. Von Hoff noted that one his colleagues, Michael Barrett, has developed a technique in which he separates cancer cells into those that have the normal number of chromosomes and those that do not. He then conducts biomarker assays on both populations as a way of addressing tumor heterogeneity. Dr. Sullivan added that he endorsed the idea of image-guided biopsies and suggested this could be done with a variety of imaging modalities. He added that "it is possible to reach any place in the body with a needle under image guidance now." He called for more collaboration among radiologists conducting image-guided biopsies, and oncologists and pathologists.

Dr. Mendelsohn raised the issue of the costs of doing biomarker assaying as part of clinical trials and who will pay those costs. Dr. Von Hoff reported that the cost of doing the biomarker assay in his Bisgrove clinical trial is $6,800 per patient and BlueCross BlueShield of Arizona agreed to reimburse that cost because they were interested in the results of the trial.

COSTS OF CLINICAL TRIALS

Clinical trials of new cancer drugs are expensive, and with implementation of better, more sophisticated studies, costs are projected to rise. Although the use of predictive biomarkers to enrich study populations might make smaller trials more likely to succeed, they pose the added costs of the biomarker tests and raise the issue of who will pay for those added costs. Currently, costs are only partially borne by NIH grants or by contracts with pharmaceutical companies. In some situations, third-party payers contribute for basic hospital or clinic services. However, considerable costs fall on the academic institutions at which studies of this type are being piloted.

There is a need to better understand the average cost of these sophisticated and multifaceted clinical trials, the factors driving such costs, and the drug development costs potentially saved by obtaining answers (positive or negative) more rapidly. More exploration of alternative funding approaches is needed, including more public–private collaborations among academia, pharmaceutical and biotechnology companies, and government agencies. The fourth session of the conference explored these issues with presentations by Robert Comis of the Coalition of National Cancer Cooperative Groups, Kevin Schulman of Duke University Medical School, and Gwen Fyfe of Genentech.

Dr. Comis began by noting that for Cooperative Group-funded oncology studies,[12] the overall per-patient cost is about $6,000, of which only $2,000 is reimbursed by government grants (The Lewin Group, 2005). In contrast, one study of four companies found that the per-patient costs for industry-sponsored studies ranged from $60,000 to $85,000 for Phase III studies (of which about $15,000 to $18,000 is reimbursed), and from $46,000 to $85,000 for Phase II studies (of which $20,000 to $25,000 is reimbursed). It is generally much less expensive to conduct a clinical trial in a foreign country, with the costs of certain cancer clinical trials in Western Europe being nearly half the cost of the same trial conducted in the United States.[13]

Regulatory Costs

Prepatient start-up costs account for a significant portion of clinical trial costs, Dr. Comis said. For publicly sponsored Phase II or III studies, these costs are about $5,000, and about $8,000 for privately sponsored studies (The Lewin Group, 2005). Much of this cost is due to addressing regulatory requirements of various institutional review boards (IRBs) and government agencies, Dr. Comis and Dr. Schulman pointed out. Of nine functional steps identified for the conduct of high-quality trials, six include elements

[12]The Clinical Trials Cooperative Group Program, which is sponsored by the NCI, is designed to promote and support clinical trials of new cancer treatments, explore methods of cancer prevention and early detection, and study quality-of-life issues and rehabilitation during and after treatment. Cooperative groups include researchers, cancer centers, and community physicians throughout the United States, Canada, and Europe. They work with the NCI to identify important questions in cancer research and to design clinical trials to answer these questions.

[13]TrialSpace Grants Manager (see http://www.fasttracksystems.net/grantsmanager_products.php).

BOX 2
Regulations Govern Most Functional Steps Required for Conducting Studies

- Of nine functional steps identified for the conduct of high-quality trials, six include elements related to federal regulations:
  - Institutional Review Board Submission
  - Site Approval
  - Preparation for Study Execution
  - Study Execution
  - Data Review
  - Study Closeout
- An average of 35% of clinical research costs is spent on compliance.

SOURCES: Comis presentation (October 5, 2007) and The Lewin Group (2005).

related to federal regulations (Box 2). "There are numerous regulatory functions that are included in bringing a study up, including the IRB costs, the approvals with the FDA and with the NCI, etc., and our estimate from working with those sites is that about 35 percent of the costs that accrue for a clinical trial relate to regulatory issues and regulatory compliance," Dr. Comis said. Dr. Schulman added that there is often "protocol creep"—after a protocol has undergone the regulatory review process with various agencies and internal boards, so many additional research steps are required that the actual cost for the study far exceeds the amount budgeted.

In the discussion following Dr. Comis's talk, conferee Dr. Birch pointed out that much of the paperwork required for regulatory approvals has been done in previous trials, and both time and money would be saved through a national database for this type of information. Dr. Schilsky added that there is unnecessary redundancy and a lack of harmonization among the multiple organizations that review clinical trials. For example, in addition to being reviewed by the NCI's central IRB, trials have to be reviewed by their own institution's review boards, in part because the Department of Health and Human Services Office for Human Research Protections also

allows institutions to exercise local control over trials. Dr. Schulman agreed, adding, "Having huge burdens related to the regulatory process that aren't adding value is costing us." Dr. Comis suggested that "We have to eliminate the bureaucracies in order to get advances for cancer patients."

Because the fixed start-up costs are independent of the number of subjects enrolled in a clinical trial and are so large, having higher patient accrual in fewer studies is more economically efficient than having lower accrual in more studies, Dr. Comis noted. Yet only 56 percent and 63 percent, respectively, of open government trials and open industry trials had subjects enrolled, one study found (The Lewin Group, 2005). Later in the discussion, Dr. Doroshow reported on a recent NCI study of four larger NCI-funded Comprehensive Cancer Centers. In those centers 25 percent of their trials accrued no patients, and 26 percent accrued four or fewer patients. A review of those four cancer centers along with two large cooperative groups and CTEP revealed that the amount of time it takes to start up a study is nearly 3 years, Dr. Doroshow noted. Start-up costs are not separately covered by government-sponsored studies. "There is a lot of inefficiency in the system," Dr. Comis said. "No site should ever bring up a study for which they don't have the patient or research resources. By the time you enter one patient in a study, you have already spent 5 or 6 thousand dollars that neither government nor industry reimbursement will probably make up."

But Dr. Doroshow voiced some optimism by noting that in 2005, NCI-supported cancer centers accrued 41,000 patients to treatment trials, with only a quarter of those supported by pharmaceutical companies. "That is an enormous amount of accrual—of those 50 centers, that is roughly about 20 percent of their patient population," he said. "Comprehensive cancer care centers do a very good job accruing patients to clinical trials." However, he added, a recent evaluation for the cooperative group system between 2000 and 2005 showed that while 463 sites accrued only 1 to 5 patients in 5 years, at 12 sites accrual was more than 500 patients for the same time period. "It is very clear that we do extraordinarily well at certain places. I think we can compete very effectively with Europe if we focus on enhancing the efficiency of our system," Dr. Doroshow said.

Patient Accrual

Much of the rest of this discussion focused on the problems associated with patient accrual and how to address them. Dr. Fyfe pointed out that

Genentech has had enormous difficulties recruiting patients for their cancer drug clinical trials because most patients in this country are not willing to agree to randomization for drugs. This forces many companies to move their studies to other countries where it is easier to accrue patients. "The recruitment percentages for patients in the United States is only about 2 or 3 percent, which is unfortunate because we are going to lose the opportunity to understand how drugs that will be marketed in the United States actually perform in the milieu of care in the United States as opposed to Ukraine or Russia or places where the standard of care is so decidedly different," Dr. Fyfe said. Dr. Comis agreed and added that more than half of patient accruals in the United States come from community-based practices, which are increasingly being endangered by financial pressures.

Margo Michaels from the Education Network to Advance Clinical Trials asked what the cooperative groups are doing at the local level in terms of working with community and advocacy groups to inform people about the types of trials offered and to increase accrual. Dr. Comis responded that the Coalition of Cancer Cooperative Groups[14] has been working closely with the American Society of Clinical Oncology, the American Cancer Society (ACS), and other major cancer-related organizations to inform physicians and patients about cancer clinical trials currently accruing patients. For example, the Coalition helped the ACS establish a search engine on its website, called Trial Check, which can be used to search for cancer clinical trials.

Dr. Von Hoff suggested using the "Just-in-Time" approach to improve patient accrual. With this approach, rather than having sites activated prior to screening for patients, a centralized IRB-approved protocol is activated only at sites that have enough potential patients to participate in the trial. One recent study found this approach improved patient accrual and reduced trial-related costs in a pancreatic cancer trial (Wiener et al., 2007).

Dr. Comis noted that complicated trials tend not to accrue as well as simpler studies, and the biggest draws for patient volunteers are clinical trials that offer a new drug as opposed to a new use of an established drug. This is in contrast to many cancer trials, which are done on "me too" drugs or on approved drugs being studied for new indications. Discussant Dr.

[14]The Coalition of Cancer Cooperative Groups is a nonprofit organization dedicated to improving patient awareness of clinical trials, facilitating access, and promoting participation. It is composed of members from 10 NCI-sponsored Cooperative groups, the country's leading patient advocacy organizations, and thousands of oncology and cancer research specialists.

Schilsky said studies show that the main influence on patient accrual is the physician, and that a minority of physicians in any practice setting refer the majority of patients to clinical trials. "It is anywhere between 20 to 30 percent of the doctors who enroll 70 to 80 percent of the patients who are in clinical studies," he said. "There is a core group of committed physicians out there who do most of the accrual and there are a lot of physicians who give lip service to it and do nothing."

Dr. Schilsky stressed that there is no incentive for physicians to enter their patients into clinical trials, but rather disincentives because such entry requires more time and effort on the physician's part. He estimated that for the average oncologist, the time needed to discuss a clinical trial with a patient and gain consent is probably at least three or four times the amount of time needed to discuss the standard chemotherapy the patient will be receiving. Dr. Schilsky suggested prompting more physicians to refer patients to clinical trials through a reimbursement mechanism, or by making referral a requirement for maintaining their credentials. "If there was a billing code that allowed a doctor to bill at a higher rate for managing a patient on a clinical trial than for giving off-protocol care, that might give an incentive to doctors to actually put patients on trials," Dr. Schilsky said. Conferee Dr. Patricia Ganz concurred and added that enrolling patients in clinical trials "should be the norm and it is only if you don't qualify or the patient doesn't want to participate that this doesn't occur. This would enable a much more efficient drug evaluation process as well as high-quality care," she said.

Global Outsourcing

Dr. Fyfe described the cost savings that could be gained if patient recruitment occurred more quickly and more robustly, as the length of a trial is often extended because of delays in patient accrual. If 25 to 50 percent of all eligible cancer patients enrolled in clinical trials, she said, "the cost of trials would decrease dramatically because you would have fewer sites, trials would take less [time] to recruit, and you would have your answer much more quickly. The reality is a lot of drug costs come from the fact that 20 percent of the physicians give you 70 percent of the patients and at this point in cancer, people have to go to the rest of the world because they can't get the patients here."

Clinical trials are increasingly being conducted outside the United States. Dr. Comis noted with a graph (Figure 18) showing that between

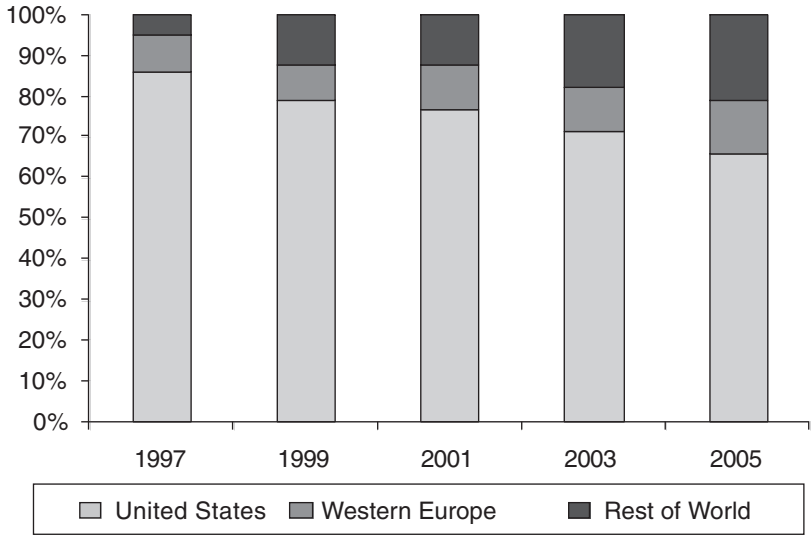

FIGURE 18 Clinical investigations are going global, as shown by the distribution of 1572 forms by location of investigative site. NOTE: A 1572 form must be submitted to the FDA by a clinical investigator prior to initiating a study in human subjects. SOURCE: Comis presentation (October 5, 2007), reprinted, with permission, from *Outlook 2007*, 2007. Copyright 2007 by Tufts Center for the Study of Drug Development.

1997 and 2005, the percentage of clinical trials being conducted in this country fell from 85 to 65 percent (Tufts Center for the Study of Drug Development, 2007). The United States ranked sixth in the countries conducting the most clinical trials, Dr. Comis said. Discussant Dr. Samir Khleif from the NCI pointed out that the majority of the countries that rank above the United States, such as France and Canada, have socialized medicine in which physicians are paid salaries. Therefore, spending more time enrolling patients in clinical trials does not have a negative impact on their income.

But Dr. Schulman urged caution in conducting clinical trials overseas. "It is cheaper to have a patient enrolled in your study in Eastern Europe at a per-patient level, but if you are not there, at the study site, making sure that they are not fabricating the data, it may not be less expensive at the end of the day, and there have been plenty of trials where an entire country's worth of data became lost as a result of that [data fabrication]," Dr. Schulman said.

Dr. Doroshow suggested that the electronic data capture system that is currently being finalized at the FDA and the NCI will improve the efficiency of clinical trials and prevent industry from globally outsourcing such studies. "It will enable the most efficient system for cancer clinical trials in the world, and studies will get done at a level of quality that just are not going to be comparable because we are spending so much time developing the infrastructure that is going to be unique," he said. Dr. Schulman countered that electronic data capture will not sufficiently speed up the time it takes to start up a clinical trial because much of that time is due to "onerous" regulations.

In his presentation, Dr. Schulman delineated the other costs in addition to start-up and patient accrual costs that explain why clinical trials are so expensive. These include the costs of patients, procedures, trial materials and distribution, as well as the costs of site and data management and statistical analyses. Obviously, the more patients a trial requires, and the more procedures done on those patients, the more costly the trial will be. Although the use of predictive biomarkers might reduce the number of patients required for a trial, they add the costs of conducting those tests on every patient who enters the trial, Dr. Schulman pointed out. They also raise the possibility that findings from studies that use them to enrich study volunteers may not be relevant to lower risk populations, effectively reducing the potential market size for the drug.

Time Is Money

Time is also a big cost driver, Dr. Schulman noted. The less time a new drug takes to make it to the market, the more time that is left on the drug's patent and the more likely investors will financially support the drug's development. A short time frame is also needed to ensure timely cash flow from sales of the drug that can offset its development costs. "Time is critical and is a huge cost that actually is probably larger than the difference in the cost per patient. Everyone says that every day's delay in making it to market is about a million dollars in the life cycle of a product," Dr. Schulman said. A trial with long patient accrual or follow-up times consequently is going to be costly to run.

Biomarkers that can predict survival and lessen follow-up time may reduce the duration of clinical trials, but as Dr. Schulman noted, "The question is whether the information is going to be as predictive as the survival data." Such biomarker tests also are an added cost in the study that may

not be trivial. "The more different measures we have in the study, the more expensive the study is," he said. "The more information we collect, the more we have to validate that information, and this adds costly complexity." Biomarker tests also increase the data analysis time, so they are not used as effectively as they could be. "We are rushing forward without actually using all the information we have, which makes it harder to put in more complex analyses in terms of some early-stage biomarkers," Dr. Schulman said.

Data management and analysis for large biomarker/genomic studies will be increasingly difficult and complex. "There are obviously not enough statisticians in the world to figure out how 23,000 genes interact with each other in cancer," he said. Electronic data capture and health records might help to alleviate some of the current costs linked to data management and analysis, he added; but the more complex the data are, the more costly it is not only to collect the data but also to make sure the dataset is accurate and to sift through in response to specific queries. The cost for each resolved data query may average between $50 and $90, Dr. Schulman noted.

Site management also has costs, which are associated with ensuring that site investigators are following the protocol. "Every different sponsor has a different way they want things done and that increases some of the complexity and costs," Dr. Schulman said.

One hidden cost is the cost of "loser" drugs. The estimate that each drug approved costs about a billion dollars to develop includes the $4 or $5 million cost of failed drugs, Dr. Schulman noted. To deal with those losses, the drug industry is increasingly relying on biotechnology companies to take on the financial risks of early drug development so "private investors are the ones that soak up a large number of losers right now," he said. Later in discussion, Dr. Mendelsohn pointed out that "if we could figure out ways to ditch products that are not going to work better, we could save a whole lot of money." An alternative, Dr. Schulman added, would be to devise different schemes for paying for failed drugs. Public funding for Phase I trials would reduce much of that cost, he said.

Public–Private Collaborations

Other suggestions were made on how to alleviate the costs of cancer clinical trials. Such trials are funded by either industry or the NCI. Industry tends to do early-phase testing and the NCI tends to fund mainly Phase III trials, most of which are to extend the indications of already approved drugs. "There is a symbiosis in the country between the public side of the system

and the private side of the system, which, in the end, benefits all patients and all cancer patients," Dr. Comis said. But an average of 29 percent of a clinical trial site's clinical research revenue originates from nontrial sources, he said (The Lewin Group, 2005). "This takes a tremendous amount of institutional commitment, which is true for both academic and community practices," Dr. Comis added.

Dr. Schilsky pointed out that there is a tremendous lack of public resources for cancer clinical trials. The budget for the cooperative groups is about $150 million annually, and that is used to support about 500 active clinical trials, including about 80 or 90 Phase III trials, he said. "Tell me any pharmaceutical company that can operate 80 or 90 Phase III clinical trials on a $150 million budget. It doesn't happen and so we are relying enormously on the contribution of our investigator population and the institutions that participate in the cooperative groups to make up the difference. It is getting more and more difficult for them to be able to do that," Dr. Schilsky said.

Dr. Comis warned that the U.S. public system that girds cancer clinical trials is grossly underfunded and endangered, and the U.S. private system is challenged by foreign competitors. A serious effort to combine resources, increase efficiency, and decrease regulatory burden will be required for the United States to continue to be on the forefront of cancer clinical research, Dr. Comis said. A few discussants and speakers suggested that CMS take a more active role in funding cancer clinical research because the large majority of people who develop cancer are Medicare recipients who would benefit from such studies. "I would be happy to propose no cancer patient would get paid for therapy within Medicare unless they were in a registry, unless we had some access to their tissue because this is all an experiment. We don't know all the answers and the quicker we can get some resolution, then the less money it will cost Medicare," said Dr. Schulman.

When a clinical trial done by an NCI-funded cooperative group has regulatory implications (e.g., if it will be a registration trial for a drug), the additional costs linked to that registration increasingly are paid for by the drug's sponsor, Dr. Comis said. Using this as a model, with judicious negotiating and planning with industry, the cooperative groups might be able to double their budget so that half comes from drug companies and half comes from the NCI, Dr. Mendelsohn pointed out. Dr. Comis agreed and noted that a similar model is already in use in Canada for funding clinical trials.

REGULATORY ISSUES

With its regulations aimed at ensuring safety and efficacy of therapeutics, the FDA controls not only the entry of drugs and diagnostics into the market, but affects the design of the clinical trials of these medical products. Consequently, FDA regulations have the potential to stimulate or hamper the implementation of biomarkers and other innovations aimed at improving the quality of cancer clinical trials. The third session of the second day explored current regulations and how they affect cancer clinical trials.

Regulatory Barriers to Innovation

The first speaker was former FDA cancer drug regulator Dr. Susan Jerian of OncoRD, Inc. Dr. Jerian discussed the interplay of current laws, regulations, and policies that inadvertently deter innovative clinical trial designs and foster what is expedient for business instead of providing the public with clinical trials that answer important questions. The first regulation she discussed was accelerated approval, which was codified in 1992, and is intended to expedite marketing of drugs for patients suffering from serious or life-threatening illnesses when the new drugs provide meaningful therapeutic advantage over existing treatment.[15] The acceleration is enabled by using surrogate endpoints for efficacy, such as response rate or time to progression. The regulations also allow for accelerated approval based on restricted use of a new treatment when "a drug, effective for the treatment of a disease can be used safely only if distribution or use is modified or restricted"; this has rarely, if ever been applied in the oncology setting.

The accelerated approval regulation stipulates that product approvals will be withdrawn if confirmatory studies fail to show clinical benefits, or if the drug sponsor fails to conduct the confirmatory study. But such product withdrawals have not occurred within oncology, most likely because "it is untenable and difficult to think of withdrawing a product when there may be patients who are benefiting," Dr. Jerian said.

Accelerated approval has benefited many cancer patients, Dr. Jerian noted. But it has also fostered new cancer treatments being predominantly tested in advanced, refractory cancer patients because the testing and approval process with such patients via the accelerated approval pathway

[15]Final Rule: New Drug, Antibiotic, and Biological Drug Product Regulations; Accelerated Approval, 57 FR 58942 (December 11, 1992).

occurs faster than it would for drugs tested in patients with early-stage cancers. "We have inadvertently incentivized sponsors to do their more rigorous studies in the end stages of disease because the sponsors see this as a quick path to market," Dr. Jerian said. "The public wants cancer prevention and cures, but more often gets end-stage care."

Furthering this trend is the increasing use of compendia to support reimbursement of off-label use of cancer drugs. This was facilitated by congressional legislation passed in the early 1990s that made it acceptable for Medicare to reimburse for some off-label use of products if that use is supported by clinical research that appears in peer-reviewed medical literature, or if it is listed in one of two compendia; either the American Hospital Formulary Service Drug Information or the U.S. Pharmacopoeia DrugPoints.[16] For the former, physicians are responsible for acquiring the supporting medical literature, but this can be challenging in a busy practice, Dr. Jerian said. Pharmaceutical companies may not routinely disseminate publications of off label use, but FDAMA regulation[17] and a landmark case in 1998[18] made it easier for them to provide such publications when physicians query the companies for additional information.

As for using the compendia route to support off-label uses, there is a lack of transparency and third-party oversight on the requirements for listing a drug for an off-label use in these compendia, Dr. Jerian pointed out. The use of compendia by CMS and other insurers to evaluate the appropriateness of off-label uses of drugs requires clinical and scientific expertise that often is lacking by third-party payers, she added. The compendia set a different and less rigorous standard for safety and effectiveness of drugs than that of the FDA. Dr. Jerian noted that CMS convened the Medicare Payment Advisory Commission (MedPAC), which developed the criteria for a desirable compendium. But none of the compendia that they reviewed met all of their criteria.

"So what we have now is a nontransparent system that we can't really access, and the ability to influence the system. The physicians who are busy in their practices are really stuck in the middle," Dr. Jerian said. Although

[16]See http://www.ashp.org and http://www.micromedex.com/products/drugpoints/.

[17]Food and Drug Administration Modernization Act, 21 U.S.C. §§ 351 et seq. (1997) amended by 21 U.S.C. § 401 (1998).

[18]Washington Legal Foundation v. Friedman, 13 F. Supp. 2d 51, (D.D.C. 1998), *aff'd mem.*, 36 F. Supp. 2d 16 (1999), *aff'd mem.* 36 F. Supp. 2d 418 (1999), *aff'd mem. sub nom.* Washington Legal Foundation v. Henney, 56 F. Supp. 2d 81 (1999).

regulations supporting reimbursement of off-label use of cancer treatments have improved overall cancer care in the United States, she said, they also pose logistical challenges for conducting rigorous cancer clinical trials. Early-stage cancer patients are being enrolled in studies for compendial listings rather than in more rigorous scientific studies intended for full-market approval of drugs. Accelerated approval and enhanced use of off-label drugs via the compendia route has "created a disincentive to companies to conduct rigorous registration studies in less refractory or earlier stage patients. Conduct of these studies is hampered by competition for patients, many of whom enroll in 'compendial' studies," Dr. Jerian said.

Another FDA process that can be problematic for the adoption of biomarkers in clinical trials, according to Jerian, is the special protocol assessment (SPA), developed in 1997.[19] On request, the FDA will evaluate within 45 days certain protocols and issues relating to the protocols to assess whether they are adequate to meet scientific and regulatory requirements. The SPA process led to improvement of clinical study designs for trials intended to support efficacy claims, and improved the likelihood of success for such regulatory approval.

But Dr. Jerian's experience with the SPA is that the more innovative or complex the study design is, and the more it employs the use of complex elements such as biomarkers or adaptive trial designs, the more difficult it is for the trial to successfully complete the SPA agreement process within one 45-day review cycle. It might take as long as a year for the sponsor to reach agreement with the FDA under the SPA process because they must go through multiple cycles; the 45-day review clock is reset each time a protocol change is made. Consequently, to have their trials up and running quicker, sponsors often avoid these innovations in their study designs, Dr. Jerian said. With sponsors much less inclined to submit novel trial designs, the promise of personalized medicine becomes more challenging, she noted. "The SPA process is not user friendly for complex study designs that the Critical Path wants to see more of," Dr. Jerian said.

The final regulation Dr. Jerian discussed was the 2005 FDA ruling on

[19]PDUFA Reauthorization Performance Goals and Procedures, an enclosure to a letter from Donna E. Shalala, the Secretary of Health and Human Services, to Senator James M. Jeffords (November 12, 1997) (on file at http://www.fda.gov/CBER/genadmin/pdufago111297.htm). See also, Center for Biologics Evaluation and Research, Guidance for Industry: Special Protocol Assessment (May 17, 2002) (on file at http://www.fda.gov/cder/guidance/3764fnl.htm).

the definition of a combination product.[20] This ruling states that in cases where a diagnostic is used to select patients for treatment with a therapeutic, the diagnostic and therapeutic together are considered a combination product. For example, if trastuzumab (Herceptin) were approved by the FDA today, it would be considered a combination product only with the IHC test for HER-2, the HercepTest. This regulatory link between a therapeutic and biomarker test can lead to an undesirable business outcome between the therapeutic's sponsor and the device's sponsor, she said. Screening using a biomarker can result in many ineligible patients, which can extend enrollment time lines. The combination product ruling also makes it more difficult to introduce scientific advances by not allowing for flexibility as the science advances. The understanding of how best to employ biomarkers is rarely present at the onset of clinical trials, Dr. Jerian noted; complex molecular pathways take time to delineate and thus need a flexible regulatory system.

Dr. Jerian proposed ways to improve the regulatory environment to foster more innovative and rigorous, high-quality cancer clinical trials. She suggested "cleaning up" the compendial process so it is more transparent and congruent with FDA standards. Scientific and medical expertise should be required and documented for those individuals who provided input for product monographs in compendia and for those individuals making reimbursement decisions. She also stressed the need to improve the depth and quality of data review. These changes should reduce the number of non-rigorous "frivolous" clinical studies done to gain compendial listings for off-label uses of drugs.

Dr. Jerian also suggested that the FDA could consider using the "restricted use" pathway for accelerated approval of products whose use would be restricted based on an in vitro diagnostic assay or other testing procedure. The restricted use pathway allows for accelerated approval by restricting the use of a drug to certain facilities or physicians with special training or experience, or is conditioned on the performance of specified medical procedures, which could be biomarker tests. Accelerated approval could be granted to a drug if the use of that drug was restricted to those patients predicted to respond to it via biomarker tests. "I think as oncologists, if we saw compelling efficacy data in a biomarker-selected group of patients, that would be more meaningful to us than a 10 percent response

[20]Final Rule: Definition of Primary Mode of Action of a Combination Product, 70 FR 49848 (August 25, 2005).

rate in end-stage disease, and I propose that this rises to the same level of importance," Dr. Jerian said. "This is one way to give an incentive to a commercial sponsor to do a biomarker-based study. It also is a way to try to get the most exciting scientific advances to patients earlier in their state of disease and really answer the important questions that we want to answer."

Finally, Dr. Jerian proposed a number of revisions to the SPA process, including hiring, training, and retaining more reviewers who are expert in complex trial design issues, and making it easier to modify study protocols without extending the FDA review process. "There are clear changes that a sponsor can make to a protocol that are easily addressed and can be reviewed efficiently within a reasonable amount of time. To have to restart the 45-day review clock for that type of situation adds to the regulatory burden," she said.

There already exist examples of frequent communication used in other regulatory interactions between sponsors and the FDA. Those same principles of frequent and open communication could be applied to the SPA process, but it would require an increase in the number of review staff at the FDA. The FDA currently does not have the capacity to provide this level of communication for the SPA process. Implementation of increased communication and allowing for protocol revisions during the 45-day review process would likely result in an increase in the submission of novel trials designs, an improvement in the quality of clinical trials and an improvement in clinical development timelines. A strong and well-staffed FDA is good for product development, Dr. Jerian said.

In the discussion following the presentations, Dr. George Mills suggested the initial 45-day review for SPA be followed 2 weeks later by a teleconference to address any needed adjustments to the SPA. After that, an annotated follow-up review should be completed and submitted to the agency in a timely manner, Dr. Mills said. Dr. Jerian said she thought this was a good suggestion, but noted that it would depend on staffing at the FDA to carry it out.

Patient Advocacy Perspective

The next speaker, Ellen Stovall, a cancer survivor and President and Chief Executive Officer of the National Coalition for Cancer Survivorship (NCCS), gave a patient advocate's perspective on regulations and what is needed to improve the quality and patient participation in cancer clinical trials. Ms. Stovall began her presentation by noting the diversity of patient

advocate perspectives and pointing out that views about patient participation in clinical research are not universal. But studies and anecdotal experience suggest that clinical trial design is not a major barrier for patients participating in clinical trials. "Few cancer patients would say 'if it weren't for the design of the trial, I would have definitely wanted to participate,'" she said. One survey of 1,000 adults noted the primary reasons that patients do not participate in clinical trials were unavailability of an appropriate trial, not meeting eligibility criteria, and reluctance of their physicians to even raise the issue of a clinical trial with them (Comis et al., 2003). Another study confirmed these main reasons for lack of participation and added other reasons, such as unwillingness to be randomized as to whether they receive an innovative treatment, time constraints, excessive distance from a treatment center, insurance denial, and distrust of the medical establishment (Metz et al., 2005).

Ms. Stovall noted that people with life-threatening diseases such as cancer are less likely to be concerned about the safety of new drugs than patients who take drugs for other conditions. "No medical intervention is entirely free from risk, and pharmaceutical innovation may be deterred if decision makers focus too much on risk avoidance," she said. But Ms. Stovall added that NCCS, while advocating for maximum access to new therapies for people with cancer, also believes that access should be based on sound and reliable medical evidence. This is in contrast with other advocacy groups, such as the Abigail Alliance, which, along with the Washington Legal Foundation, has raised a lawsuit against the FDA that claims individual access to unapproved therapies by those who can pay for them is a constitutional right.

The regulatory and financing issues that Ms. Stovall identified as most pressing in the cancer patient advocate community include the recent revisions to CMS requirements for coverage of routine patient care costs within clinical trials, the FDA's ability to follow up on supplemental labeling issues once a drug has been approved under the accelerated approval process, the status of the FDA's issuance of its final guidance on expanded access programs, and proper oversight that ensures adherence to FDA policies in the review of new cancer indications by the Oncologic Drugs Advisory Committee. "Some new drugs have gone to groups to review that don't have the complement of oncology reviewers on them that we would like to see, nor people who understand how these drugs are actually going to be used in clinical practice," Ms. Stovall said. She also proposed creating incentives that reward individual physicians for engaging in clinical research. "People

value what they get paid to do and we need to put the rewards in place" to promote more physician participation in clinical trials, she said.

Ms. Stovall suggested that patient advocates work with the FDA and clinical trial sponsors to design drug development programs and clinical trials, noting that such cooperation would facilitate greater patient accruals to clinical trials and lead to expanded access to innovative cancer treatments. After Genentech developed a working relationship with the National Breast Cancer Coalition, she said, patient accrual for their clinical trial of Herceptin rapidly jumped from 16 to 40 women a month. "Articulate, educated advocates who understand the science and can come to the table can really transform the way clinical trials accrue, the way they are designed, and the way they are monitored. There is a fourth wheel to the bus at times—it is not just academia, government, and industry, but patients. We are the end user," she concluded. Dr. Janet Woodcock concurred that "it has been the patient groups that have driven many of the major policy changes that have been made over the past two decades."

Regulation of In Vitro Diagnostics

Following Ms. Stovall's talk, Dr. Steve Gutman of the FDA gave a presentation on the regulation of in vitro diagnostics, otherwise known as lab tests. The FDA has been regulating these tests since 1976, when the Medical Device Amendments gave the agency the mandate to conduct premarket reviews of in vitro diagnostics, and to monitor good manufacturing practices and postmarket reporting of adverse events. For such premarket reviews, tests considered high risk are reviewed primarily via two pathways. The premarket approval application pathway is the most rigorous because it often requires clinical studies and is reserved for diagnostics that pose the most risk to patients. Those diagnostics considered less risky are reviewed via the premarket notification, or 510k pathway, which involves showing that the test is similar to tests already on the market and performs adequately. For both pathways, the FDA uses the same core science or standards to evaluate tests. The evaluation process is well established and transparent, and "although it is rapidly evolving, we have a very rich base on which to ground our evaluation of new diagnostics—there is rich literature and numerous standards," Dr. Gutman said (Box 3).

Dr. Gutman did note, however, that for innovative biomarker tests where there is not an intuitive relationship between the analytical and clinical signal, the agency might require the sponsor to conduct a feasibil-

BOX 3
Model for Evaluation of Diagnostics

- **Literature**
- **Standards**
  - Clinical and Laboratory Standards Institute (CLSI)
  - International Organization for Standardization (ISO)
  - Standards and Reporting of Diagnostic Accuracy (STARD)
- **Guidances**
- **FDA template**

SOURCE: Gutman presentation (October 5, 2007).

ity study that specifies a model for a subsequent clinical study. This model would indicate the intended use and targeted performance of the test, what the test population and type of test sites will be, and what the cut-offs are for the test parameters. This feasibility study "works out all the bugs" for the subsequent validation clinical study, Dr. Gutman said.

In that study, sponsors have to demonstrate that "whatever signal is generated stands the test of time with independent validation of the data and demonstrates that the hypothesis works and is linked to the claim," Dr. Gutman said. For that demonstration, the clinical sensitivity and specificity of the test usually are required. Dr. Gutman noted that the agency tends to shy away from reviewing predictive values of positive and negative tests because these values vary according to the prevalence of the condition in the population being studied. He added that the FDA is "very concerned with the endpoint against which the new diagnostic is being measured, whether that is a drug effect or the presence of disease, or the risk of future disease, and the weaker the endpoint then the more semantically interested we become in cautionary labeling."

The "bad news," as Dr. Gutman put it, is that for cutting-edge science with clinical implications, such as biomarker tests, there is a lack of material and method standards, complex bioinformatics for the agency to wade through, and a lack of gold standards. "Often as we look at a new diagnostic, we will look for a silver or bronze standard. Sometimes we will settle for

a lead standard, but we do look for a yardstick of truth," Dr. Gutman said. As is true for other diagnostics, the agency will continue to be on guard for biases in sampling, selection, verification, and spectrum. The agency will pay close attention to the impact of missing datasets because the impact of even small sets of missing data can sometimes be magnified, he noted.

The "good news," according to Dr. Gutman, is that the agency has interest in and understanding of adaptive designs and Bayesian statistics, and will let sponsors take regulatory shortcuts by using cautionary labeling and by initially narrowing their claims to bring a test quickly to market, with expanded claims dependent on the collection of future data after the product is on the market. The Critical Path plan for modernizing the agency is "alive and well" and being executed on many fronts within the FDA's diagnostics arena, Dr. Gutman stressed.

The agency also has a flexible regulatory toolbox, he added, including a process for expedited reviews, de novo classifications, niche submissions, and preinvestigation device exemptions (pre-IDEs). Expedited reviews enable cutting-edge new products with real public health impact to move to the head of the queue of products to be reviewed at the FDA. Products with a modest amount of data may undergo a niche submission, which is reviewed within 30 days. The de novo classification enables the agency to down-classify new devices or tests that are not similar to those already on the market, but not likely to pose much risk. The pre-IDE is a free service of the FDA in which it they will review a sponsor's protocol within 60 days. "It is sort of the antithesis of a pop quiz because, when we get the protocol, we tell you all the questions we are going to ask when the study comes in, and you get to sort of negotiate. We will argue and talk and learn so that there is a decrease in uncertainty when the product actually hits the decks in my shop," Dr. Gutman said. This results in more well-developed submissions to the FDA that the agency can review more quickly.

Once a test is approved based on its performance in small studies, there is still uncertainty whether it will perform similarly in the hundreds or thousands of labs that will use it on millions of patients. "My center is undergoing a transformation in which it is deliberately looking at what kinds of mechanisms it has for tracking real-world use of products, and we have a variety of programs looking at active surveillance, better integration of signals," Dr. Gutman said. In the discussion following the presentations, a conferee pointed out the variability in the accuracy of HER-2/neu testing and noted "it is not only the test that matters, but how it is performed in the laboratories and how CLIA [the Clinical Laboratory Improvement

Amendments] regulates the performance of laboratory testing," he said. Dr. Gutman agreed and also stressed the importance of standardizing the preanalytical phase of testing, that is, the process for procuring and handling the sample being tested. Dr. Woodcock added that "we can develop the best drugs and diagnostics in the world, but if the health-care system is going to be error prone, we still are not going to be delivering quality care to the patients."

Dr. Gutman concluded his talk by saying, "FDA has a dual mission across the board—we are trying to promote public health by getting new diagnostic devices out more quickly, and we are trying to protect public health by keeping bad ones off the market. There is a clear tension which we try to address by applying good science, by asking the right questions."

Regulatory Issues in Improving Cancer Clinical Trials

In the final talk of the session, Dr. Woodcock of the FDA gave her regulator's perspective on what is needed to improve cancer clinical trials. The need for such improvement is indicated by the average success rate of only 5 percent for oncology drugs, from first in human to registration (Kola, 2004). Dr. Woodcock pointed out that most experts in drug development agree that better evaluation of candidates earlier in the process will increase the success rate and decrease the amount of residual uncertainty about the performance of a product. This requires changing the traditional cancer trial process, which is more empiric than mechanistic.

Dr. Woodcock gave suggestions for improving every phase of cancer clinical studies. Although she recognized the value of Phase 0 trials for revealing proof of mechanism in people early in the testing phase of drug development, they require personnel and approaches that are not traditional and thus less likely to be adopted. "It's going to take a lot to change traditional drug development because there is a tremendous amount of inertia," she said. Phase I is typically a dose tolerance study where there is a dose escalation to maximal tolerated dose, "and often very little else is learned except how much people can take and that is not very informative from a scientific perspective," she said. More information could be gained by using "biomarkers, pharmacokinetic analyses, and other scientific lenses into the performance of the product as early as possible," she said. But, like other speakers, Dr. Woodcock stressed the need for proper validation of biomarkers.

As for Phase II trials, Dr. Woodcock advocated using adaptive trial

designs. These designs are especially suited to finding the optimum dose, which is the main goal of Phase II trials. "Better dose finding means studying more dose strata and rapidly trying to converge on appropriate doses, and nothing would be better than using adaptive designs to do that. They are tailor-made for this," she said, adding that there is nothing controversial in their use for this purpose. Dr. Woodcock also proposed mathematically modeling the relationships among dose, toxicity, and effectiveness, using Phase II data. "This is extremely labor intensive, but it is extremely informative. It turns what is basically an observational descriptive exercise of Phase I and II into something that is quantitative and understandable," she said. She hopes to reinstitute a highly popular FDA pilot program that did such modeling for drug sponsors.

Dr. Woodcock encouraged the use of composite endpoints in Phase III trials, noting that this is not a novel concept; the use of several outcome measures in one composite is accepted for many other conditions, such as arthritis and cardiovascular disease. The validity of the composite needs to be established, but this need not entail conducting large numbers of validation trials, and instead could be done by gathering expert agreement on how best to combine already individually validated outcome measures.

Dr. Woodcock also discussed the paired development of investigational drugs and diagnostics. Such development does not bypass the need to do a normal safety workup of the drug, or the need to establish analytical validity of the diagnostic, but rather means that the same trial(s) can determine the clinical utility of both the drug and the diagnostic. The most controversial aspect of this codevelopment pathway, according to Dr. Woodcock, is the need to demonstrate that the test contributes some information of value. "The test needs to actually discriminate two populations and you need to measure that at some level—to understand the predictive value of a negative test," she said.

As for combining several investigational drugs in a single development program, Dr. Woodcock explained that this is not traditionally done, but may be needed for the innovative targeted cancer drugs that must be combined to be most effective. In addition to commercial concerns, such as the ability of the drug sponsors to work together, a combination treatment trial has to show that each agent makes a contribution to the clinical effect, and has to measure the toxicity of each agent individually, as well as in combination. "We want to be sure that we are not adding one of the agents that adds no benefit and adds significant toxicity to the regimen," Dr. Woodcock said. "Smart" Phase I trials that measure pharmacokinetics

can indicate individual toxicities, she noted, and a factorial design can demonstrate the individual effectiveness of each agent being studied in combination.

In the discussion following the presentations, one conferee asked for more specific guidance from the FDA on how to study early in development two combined investigational agents, such as how many patients the combination needs to be tested in for the various phases of clinical testing. "The sad aspect is that it is just much easier to add [a new investigational drug] onto standard chemotherapy because we know that toxicity profile," he said. Dr. Woodcock responded, "This is an action item for the FDA to take home and think about because this is coming."

Dr. Woodcock pointed out that many surrogate endpoints are used now in oncology trials, including the Response Evaluation Criteria in Solid Tumors (RECIST), but these surrogates do not always correlate with "something a patient values such as prolonged survival or progression-free survival," she said. Such surrogates need to be refined, and the use of functional imaging might provide such refinements without requiring complex endeavors, Dr. Woodcock noted. "For the purposes of drug development, it is often better just to have a reliable response measure. It doesn't have to be a surrogate," she said, but instead could be the use of FDG PET, for example, to reveal whether a drug is having an effect on tumor metabolism. "It won't be a surrogate endpoint, but it will tell you that it looks good or bad, which can be an extremely helpful type of measure to have in the Phase I stage of clinical testing. In many cases we need a good predictor of ultimate success more than we need a surrogate endpoint," Dr. Woodcock said.

She added, however, that the field of cancer drug development would benefit from rigorous identification of some candidate surrogate markers and an assessment of what is needed to qualify them. Qualification could be carried out by a consortium of interested parties, such as the Biomarker Consortium. "There are some low-hanging fruits—markers that could be developed [relatively easily] and used as surrogate endpoints in the future," Dr. Woodcock said. Surrogate endpoints for prevention trials are especially needed because the time frame for outcomes is so long, yet "that is where we have the greatest risk if we are wrong and exposing large numbers of people to an ineffective or maybe harmful intervention," she said.

Dr. Woodcock's final suggestion for improving the quality of cancer clinical trials is to standardize all aspects of trial design and execution because the mechanics of trial conduct and execution across all disease areas are extremely suboptimal, she said. "Quality of cancer trials also includes

the quality and efficiency of execution, something that adds to the cost and is holding this field down," Dr. Woodcock said. She mentioned that the FDA plans to form a public–private partnership to drive standardization of study execution.

Dr. Woodcock ended her talk by noting that the purpose of trials is to develop the evidence that products will save lives and improve health, as well as provide access to investigational products. Quality trials within this context mean those that meet the needs of patients and health-care providers, who are the "ultimate customers," she said. "Regulatory agencies are supposed to be surrogates for what the patients and the providers actually want, and are not supposed to stand in the way," Dr. Woodcock said. She urged the conferee to consider not only scientific problems to be addressed in improving the quality of cancer clinical trials, but also to think about whether the trials are meeting the needs of patients and providers—whether they are saving lives and improving the health of the population.

In the discussion following the presentations, Dr. Sullivan pointed out that the imaging agent industry believes radiotracers used in imaging studies have a risk/benefit profile similar to devices and should be regulated accordingly. Dr. Woodcock agreed that it was important for the FDA to consider this as a way to streamline the approval of such tracers. Later during the discussion, Dr. Parkinson noted that overcoming regulatory hurdles and gaining FDA approval for a drug or diagnostic does not guarantee insurance reimbursement for the product, especially when the product is used abroad. "A common problem globally is this disconnect between regulatory approval and reimbursement," he said.

REPORTS FROM THE CASE STUDY DISCUSSION GROUPS

Adaptive Trial Design

Dr. Herbst presented a summary of his group's breakout session on adaptive trial design. His group agreed on a number of issues, including the variability of adaptive trial designs and the notion that adaptive trials are not a new phenomenon—for a long time, many investigators have been using a few of these adaptive designs in their clinical trials, most notably sequential monitoring with early stopping boundaries. But what is new is the creation and expanded use of other forms of adaptive designs, such as sample size reestimation, adaptive dropping and adding of trial arms, as well as the other types delineated by Dr. Donald Berry in his presentation

earlier in the day. The group agreed that although Bayesian methods are particularly well suited for building adaptive designs, traditional methods can also be adaptive. The group also agreed that it can be important to monitor relationships among early and late endpoints in adaptive studies, and to calculate operating characteristics, such as sample size distribution and frequency of incorrect conclusions. Such calculations usually require simulation.

Dr. Herbst stressed the importance of noting that although adaptive designs can sometimes result in smaller trials, that is not always the case. Hopefully, adaptive trials provide more accurate conclusions, he said. Additional possible benefits of such designs are that more questions can be considered in a single trial; they can foster faster, more efficient drug development; they can lower costs of medical care; and they can lead to better treatment of participants. Because of its flexibility, another advantage of adaptive trials is the ability to work additional study parameters into an ongoing study as more knowledge is gained. For example, data collection on the predictive value of a newly discovered biomarker can be added into an ongoing trial of the predictive value of other biomarkers.

Dr. Herbst also described the challenges involved in conducting adaptive trials, such as additional logistical infrastructure needs. Adaptive designs must be prospective, which requires extra work before trials begin. Extra costs and potential delays in setting up adaptive trials may arise due to the need for increased communication with government regulators and local oversight boards, and the potential need for a larger drug supply. Data-flow needs increase the logistical burden, and it can be challenging to prepare for different alternatives—for managing a greater variety of treatment regimens that may change over time within a single trial.

During the discussion, Dr. Richard Chappell pointed out the potential bias of response-adaptive randomization trials due to changes in patient characteristics during the accrual that can confound treatment effect with time effects. Quoting Peto (1985), he noted, for example, that patients entering the European Coronary Bypass Trial showed a statistically significant trend toward better prognosis at baseline as accrual continued. If there had also been a trend in allocation proportions toward the apparently better treatment, an appreciable bias might have been engendered. Although models can be designed to address these issues, models tend to be subjective, based on sparse trend information, and, even when used, trials may be left with ambiguous and multiple adjusted answers—difficulties that should be prevented by randomization, Dr. Chappell stressed. A stratified group-

sequential approach (Karrison et al., 2003) can address this bias, he added. Although he recognized that results from response-adaptive randomized Phase II studies may be superior to those produced from nonrandomized designs, Dr. Chappell still thought fears of bias from confounding with time trends warrant not using response-adaptive randomization in confirmatory trials.

During his summary, Dr. Herbst reiterated this concern about potential bias in response-adaptive randomized trials, but added that "although this needs to be looked at in any trial and analyzed and is a real concern, the belief was that we could go forward despite this." He also expressed concern about the potential bias being introduced by revealing interim trial results to investigators, patients, and the investment community, as Dr. Ellenberg discussed in her presentation. (See page 16.) But all agreed that this potential bias could be addressed, although it may limit the use of adaptive methods in some applications. Group members agreed that adaptive methods should be considered in designing clinical trials because they offer some benefit, although their use still presents challenges.

Phase 0 Trials

Dr. Giulio Draetta presented a summary of his group's discussion of Phase 0/exploratory IND trials. This group discussed the advantages, disadvantages, ethics, and costs of Phase 0 trials. The group agreed that Phase 0 trials were especially useful in doing compound triage when multiple compounds are being considered because they can indicate the compound with the most favorable pharmacokinetics or pharmacodynamics. But as Dr. Schilsky pointed out, generally only the large pharmaceutical companies, as opposed to academic investigators or researchers at small biotechnology companies, simultaneously have available several compounds that target the same pathway or disease. He also noted that Phase 0 trials add an extra step in the testing process for individual compounds. "At the end of the day, if you're going to bring your drug through the full clinical development plan, you're still going to need the full package of data. This is not a shortcut," he said. But Dr. Collins pointed out the time savings by running Phase 0 and animal studies simultaneously. Dr. Schilsky expressed additional reservations that the dose ranges in a Phase 0 study are more limited than those in a Phase I study and therefore may not reveal as much. He added that Phase 0 studies may not be appropriate for a drug with multiple targets or with an unknown target. They might also be unsuitable for a drug that is metabo-

lized very differently in people than in animals, or for a drug that affects patients in a way that is substantially different from the effects observed in blood samples or other surrogate tissues used in the Phase 0 study.

In addition to aiding compound triage, another advantage of Phase 0 trials pointed out by several discussants is that it can reveal the clinical pharmacodynamics of a compound early in the drug development process, so that compounds that appear worthwhile in preclinical studies—but are clinically irrelevant—are essentially thrown out early in the process. The group agreed that Phase 0 studies have unquestioned benefits in addressing early biological endpoints in patients, and impacts on target and potential downstream biology. But such studies require extensive and time-consuming efforts to develop and validate pharmacodynamic assays in animals and then in human tissues. Dr. Draetta said the group believed such effort is worthwhile, noting that, "You need to spend the time to know what you are doing. You cannot wait until Phase II to know whether you are hitting the tumor target."

Given the large time and expense that may be involved in developing and validating a pharmacodynamic or biomarker assay for a Phase 0 trial, there was some discussion over who would be willing to pay for such assay development. Dr. Schilsky suggested the NIH Clinical Center provide assay development and transfer the technology to interested parties. Dr. Doroshow agreed it would be a wise use of the Center's resources and that this has already occurred for some assays, which are now publicly available. For academic researchers without drug sponsors, acquiring the resources to run a Phase 0 trial is also a problem because these studies are essentially being conducted outside the realm of routine medical care and therefore are not likely to be reimbursed by insurers, Dr. Schilsky pointed out.

The group also discussed the ethical implications of a Phase 0 trial and noted that they are equivalent to what is seen in normal volunteer studies. But healthy volunteers are often paid to participate in clinical trials from which they personally receive no benefits, so Donna Przepiorka of the FDA raised the question of whether volunteers in Phase 0 trials should be paid as well. Dr. Mills noted that his long-term experience with imaging studies in volunteers indicates their broad acceptance of the concept that their participation will not benefit them personally.

Dr. Darlene Rosario of Mannkind Corporation added that even many patients participating in placebo-controlled randomized studies do not receive a personal benefit from their participation. All agreed that informed consent was essential in Phase 0 trials since they are for other clinical trials,

and that volunteers in Phase 0 trials should be made aware that they are not likely to personally gain anything from their participation. Dr. Doroshow pointed out that when the IRB reviewed his study of PARP inhibition in tumor biopsies, the IRB noted that because of the extensive preclinical work done by the investigators, the data generated by their tumor analyses would be more accurate than data typically garnered from a Phase I study. The low doses used in the study also meant that the risks would be lower to participating patients. As a result, the IRB considered the overall risk/benefit ratio to be lower than what might be seen in a typical Phase I trial, and the study was considered ethically sound as long as patients were made aware in writing that they personally may not benefit from the study. Katherine Meade, a patient advocate who volunteers for Us TOO International Prostate Cancer Education and Support Network, stressed the importance of involving patient advocates early in the process to gain feedback on the trial design and communication to volunteers.

Imaging

Drs. Hricak and Piwnica-Worms summarized their group's discussion on imaging. Dr. Piwnica-Worms noted that imaging is used most often in preclinical or Phase I studies to determine if a drug is hitting its target (pharmacodynamics), to confirm the drug's mechanism of action, to evaluate the clinical response, and to do pharmacokinetic analyses when labeled drugs are used as the imaged agent. The group agreed that there were opportunities to use imaging in Phase II or III trials to optimize patient selection and conduct treatment follow-up. "We have tools that we are not using," said Hricak.

The group spent some time discussing the differences between targeted diagnostic imaging and imaging of therapeutic response. Diagnostic imaging often involves imaging a single target at a single time point. Given that it can take between 100–200 million dollars to develop such a targeted diagnostic imaging agent (Nunn, 2006), which has a very limited use, this application of imaging may be too expensive to develop relative to the size of the market that would use it, Dr. Piwnica-Worms said. The group agreed that imaging targets that can be generalized to several different tumor types, such as imaging of apoptosis, proliferation, and tumor metabolism, are more broadly useful than tumor-type-specific targets.

Discussants Drs. Tim McCarthy from Pfizer and Jeff Evelhoch from Amgen pointed out that imaging has and continues to have a growing

impact on drug development. Pfizer has used PET assessment of whether a drug is reaching its target to make decisions about whether to stop a trial. They also used FDG and FLT PET imaging and DCE MRI measures to confirm mechanisms of action for novel therapeutics and to help make "go/no go" decisions regarding further drug development. Dr. Evelhoch noted that PET FDG measures of tumor metabolism demonstrated a metabolic effect even in the absence of a traditional clinical response, and enabled some of Amgen's drugs to progress further in the clinical trial hierarchy than they would have if only standard radiologic measures of tumor response were used. CT volume imaging of tumors was also noted to be more accurate than RECIST. "This is just low-hanging fruit of today's technology. It can have a significant impact on how you interpret the data, and is very accessible, straightforward, and can be executed today," said Dr. Piwnica-Worms.

By "killing failures early and fast," imaging saves substantial drug development costs, Drs. McCarthy and Evelhoch noted. Some savings are due to avoiding Phase II or III trials of drugs likely to fail or repeated Phase II failures, while others were due to shortening the drug development time line. Even a delay of a few months can mean millions of dollars lost.

Despite the recent explosion in U.S. clinical PET centers, as indicated by a slide Dr. Piwnica-Worms showed (Figure 19), there is a lack of

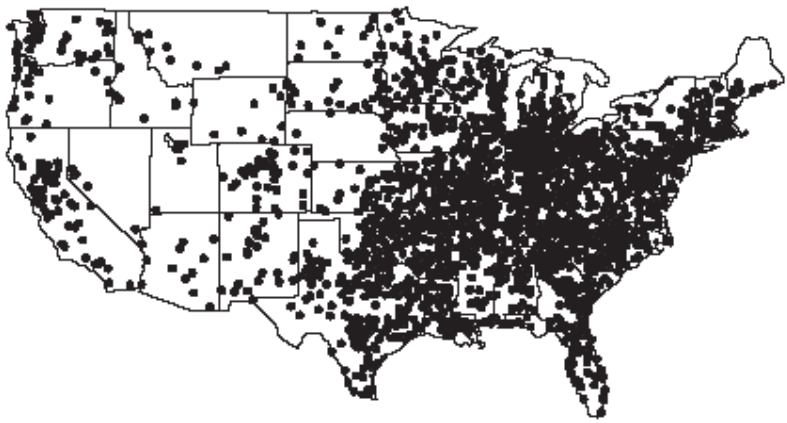

FIGURE 19 Clinical PET centers in the United States in 2007.
SOURCE: Piwnica-Worms presentation (October 4, 2007), reprinted, with permission, from *AMI Winter 2007 News*, 2007. Copyright 2007 by Academy of Molecular Imaging.

standardization, harmonization, and training for these and other imaging centers that can create inaccuracies in multisite clinical trials, the group pointed out. They suggested NCI-designated Comprehensive Cancer Centers should have imaging and image analysis core laboratories for clinical trials. These core labs would be especially useful in the conduct of Phase I trials and would save money in the future by improving the quality of data, Dr. Piwnica-Worms said. Key to these core labs would be research protocol assistants, who ensure proper protocol execution, and Image Response Assessment Teams (IRATs), which ensure consistent interpretation of imaging. IRATs "are not widely distributed, but their value is integral, and with a modest investment could be integrated into the overall process and probably increase quality control substantially," Dr. Piwnica-Worms said. For multisite trials, uniform central reading of images is key to forging an efficient path forward, the group agreed. The group also suggested that academic imagers (radiologists) be engaged at the starting point of the trial design process to produce better trial results.

The group also discussed validation of imaging biomarkers. Such validation requires precisely defining the question the biomarker is expected to answer, and using positive and negative controls in the protocol design, Dr. Piwnica-Worms said. The group requested that the FDA establish a pathway for qualifying imaging biomarkers akin to that used to qualify FDG PET—one that is mechanism-based and not limited to use on a specific organ or a similarly narrowed application. The group also suggested that the validation process for tracers used as research tools be differentiated from that needed for imaging agents used for clinical diagnostic purposes.

Although some imaging technology, such as PET, is expensive, it can offer information that saves patients from undergoing surgeries or other invasive procedures, which are even more expensive, Dr. Piwnica-Worms explained. "PET was approved as a cheap alternative for surgery in the work-up of single pulmonary nodules 8 years ago," he added. "The insurance companies loved it. They liked the chance of spending $2,000 to have a one-in-four chance of avoiding a $25,000 surgery. Can that same kind of logic optimize the use of more expensive bioimaging tests as long as the information they provide has value?"

Other suggestions made by the group included that genomic and proteomic correlation studies in image-guided biopsies be standardized, and that researchers establish how serum, urine, and tumor biomarkers complement the information gained with imaging biomarkers. The group also suggested changing the culture of the imaging community so they are

more inclined to participate in clinical research. This could be done by developing practice environments that encourage and reward imagers to engage in research.

Dr. Piwnica-Worms concluded by reiterating the caveats of imaging that were described by other speakers. These caveats include that imaging may demonstrate that a target is being hit or confirm the expected mechanism of action, but it alone does not imply clinical benefit. He acknowledged the regulatory and financial barriers linked to imaging biomarker validation. These barriers hinder sponsors from running clinical trials in this country. Many imaging trials are moving overseas, he noted, and this poses a threat to the U.S. trial infrastructure.

Use of Proteomics/Genomics to Assign Therapy in Lung Cancer

Dr. Mendelsohn summarized his group's discussion on the use of proteomics and genomics to assign therapy in lung cancer. This group's discussion was focused on the presentations by Drs. David Carbone from Vanderbilt University and Mark Kris and William Pao of MSKCC. Dr. Carbone used the MALDI (matrix assisted laser desorption/ionization) mass spectrometry system to detect protein signatures associated with longer survival in 139 advanced lung cancer patients following treatment with tyrosine kinase inhibitors gefitinib or erlotinib. In this retrospective study, he found the elevated production of eight key proteins in blood serum linked to longer survival. A second retrospective study in a different group of lung cancer patients found that the eight proteins did not correlate with longer survival in patients treated with standard chemotherapy, or surgery and radiation. This suggests that the eight-protein signature specifically predicts longer survival following treatment with a tyrosine kinase inhibitor and not merely in those patients likely to survive longer no matter what treatment they receive. He plans to do a prospective study on the usefulness of the protein signature in predicting lung cancer patients who will respond best to tyrosine kinase inhibitor treatments. Dr. Carbone also has used two-dimensional gel electrophoresis to find a more complex protein pattern, encompassing more than 1,000 proteins, that is present in the lung biopsies of cancer patients, but not in normal lung tissue biopsies. He is currently looking for candidate diagnostic markers among these proteins.

Drs. Kris and Pao reported that EGFR is overexpressed in 45 percent of non-small cell lung cancers (NSCLCs), as measured by immunohistochemistry. Studies done by Dr. Pao and others (Lynch et al., 2004; Paez et

TABLE 4 Factors Predicting Sensitivity to Gefitinib (Iressa)

Overall Response Rate	11%
Women	18%
Men	5%
Never Smokers	29%
Current/Former Smokers	5%
Adenocarcinoma	12%
Other Non-Small Cell Lung Cancer	7%

SOURCE: Kris/Pao presentation (October 4, 2007).

al., 2004; and Pao et al., 2004) detected four EGFR mutations associated with sensitivity to gefitinib or erlotinib, and indicated that the KRAS mutation predicts a lack of response to the same drugs. Seventy-five percent of lung cancer patients with these EGFR mutations responded to tyrosine kinase inhibitors, and only 1 percent of those with the KRAS mutations responded. By comparison, clinical predictors are not as informative in predicting the likelihood of response. Nonsmoker patients, for example, have the highest response rate, but only 30 percent of nonsmokers with NSCLC respond to gefitinib or erlotinib (Table 4). There are also molecular predictors for acquired resistance to tyrosine kinase inhibitor treatment, including second-site EGFR mutations in about 50 percent of the cases, and MET amplification in about 20 percent of patients. This suggests that MET inhibitor drugs may have a role in treating patients with acquired resistance to erlotinib or gefitinib. MSKCC is starting to screen lung cancer patients for KRAS or EGFR mutations and using this information, which can be acquired from the tissue removed in a needle biopsy, to decide which patients to treat with tyrosine kinase inhibitors. In conjunction with Dr. Varmus, Drs. Kris and Pao also developed a transgenic mouse lung cancer model, with the same EGFR mutations that cause lung cancer in humans, to screen for more effective tyrosine kinase inhibitors. "This is a nice example of a mouse model using transgenic technology in order to do preclinical studies," Dr. Mendelsohn said.

The cost of doing tumor biopsies and genomic and proteomic tests on the biopsied tissues was discussed. Drs. Kris and Pao noted that in their studies, this cost was about $5,000 per lung cancer patient. The cost was covered through grants or philanthropy and not by third-party payers. The New York researchers continue to do studies aimed at confirming the useful-

ness of these predictive markers. "We still have work to do to convince the [insurance] companies that this is something that they should foot the bill for," Dr. Mendelsohn pointed out.

Dr. Kris stressed the need for tumor biopsies and molecular analyses on such biopsies in clinical trials, pointing out that EGFR and KRAS mutations play a critical role in one-fifth of lung adenocarcinomas. But Dr. Parkinson noted that patient subsets are a major problem for drug developers, especially because so much redefining of those subsets is done each year. Discussant Dr. Sam Hanash of the Fred Hutchinson Cancer Center added that gene defects may not be so definitive, with the many overlapping and changing molecular pathways to cancer. One can target a pathway, he noted, but at some point it might not be the critical path. Dr. Carbone raised the question of when a biomarker is good enough to be predictive and prognostic, and said how good a biomarker is depends on what other alternatives there are, and how bad the outcome may be if a biomarker is not used. He said it might be appropriate to take risks when there are bad outcomes and poor alternatives.

There also was some discussion on which regulatory pathways cancer biomarkers should follow to enter the clinical market. The lengthy FDA approval process is not necessarily required for some predictive biomarker tests performed in laboratories, for which only the laboratory is subject to CMS scrutiny under CLIA. The costs of such tests are likely to be reimbursed if there is enough evidence for their usefulness in compendia or other published data, Dr. Mendelsohn noted. "It may be more important to have good data than to worry about all the criteria that would require the FDA to actually put this on the drug sheet in the package saying that this test is certified for this particular treatment," he said.

Dr. Mendelsohn summarized the discussion by saying, "We are finally taking what has been experimental and moving it into the clinic, and getting to the point of proof of principle. But it is painfully hard to prove the principle. These are 5-year projects, and they are still very much investigative projects rather than in-clinical-practice projects."

Use of Genetics/Genomics to Assign Therapy

Dr. Pierre Massion summarized his group's discussion on the use of genetics/genomics to assign therapy. At this discussion, Drs. Sparano, Shak, Kucherlapati, and Chang presented their study results, which suggest that at least for breast, lung, and colon cancer, genetic or genomic tests for predict-

ing treatment response or prognosis are showing evidence that they could be useful for reducing the overtreatment of patients that is so common. "These still are primarily investigator-initiated trials and validation is on the way, but I think we have got strong evidence that this will be beneficial," Dr. Massion said.

Drs. Joseph Sparano of Albert Einstein College of Medicine and Steven Shak of Genomic Health noted that patients with ER-positive, lymph node–negative breast cancer comprise about half of all newly diagnosed breast cancers. More than three-quarters of these patients can be adequately treated with surgery and hormonal therapy, with or without radiation. Adding chemotherapy to these women's treatment regimens provides an absolute benefit of only 5 percent or less, while adding significant toxicity. To better assess that risk of recurrence and aid the decision of whether to add chemotherapy to the treatment, the Oncotype DX 21-gene test was developed. This test, which is already on the market, has been studied in patients treated with the hormonal therapy tamoxifen as well as in patients treated with both chemotherapy and hormonal therapy, and in patients who received no therapy. These studies show that the greater the "recurrence score" in the Oncotype DX test, the greater the likelihood of recurrence and death from breast cancer within 10 years. With the NCI-sponsored TAILORx study, researchers at 900 sites will be assessing prospectively in 4,000 women the benefit of adjuvant chemotherapy for those women with an intermediate Oncotype DX recurrence score. The women all have ER-positive, lymph node–negative, HER2/neu-negative breast cancers.

Dr. Raju Kucherlapati from Harvard University described his much smaller iTarget trial that assessed the usefulness of EGFR mutation as a biomarker test to predict response to gefitinib in first-line therapy of advanced NSCLC. "The results are quite impressive," Dr. Massion said. "They move an overall 20 percent response rate for all NSCLCs, which is probably an optimistic estimate, to a 75 percent response rate for those patients with EGFR-mutated positive tumors." The cost of the mutation analysis was absorbed by Harvard and not reimbursed by third-party payers, he added.

Dr. David Chang reported the results of his prospective analysis of more than 400 archival tissues, which indicated that a wild-type KRAS in colon cancers predicted prolonged survival following treatment with an EGFR antibody, whereas patients with tumors that had KRAS mutations did not benefit at all from that therapy.

"Genomic biomarkers have evolved in clinical decision making," Dr. Massion summed up. "They are not on the horizon anymore but rather

right in front of us." He stressed the need to continue to systematically discover and select across biological materials and molecules, new diagnostic biomarkers for cancer, including those that indicate posttranslational modifications, or other new classes of genetic biomarkers. Dr. Massion noted that the newly launched Cancer Genome Atlas project,[21] which is sponsored by the NCI and the National Human Genome Research Institute, aims to find new classes of tumors and will focus on biomarker discovery and validation of those that are specifically related to the function of a target of interest. This project will involve comprehensive molecular analyses of a large series of tumor specimens at multiple sites. Lung, brain, and ovarian cancer will be the first three cancers that will be studied in the pilot phase. The goal is to test the feasibility of using large-scale genome analysis technologies to determine all of the important genomic changes involved in cancer.

The group also recommended carefully accessing existing cohorts that are valuable resources for testing biomarkers. An example of such a resource is the blood samples from the Women's Health Initiative study. "These studies need to be accessed for testing the performance of prevalidated biomarkers, and we also need guidelines as to how we can access them in a rational way," Dr. Massion said. A major point raised in the discussion was the lack of funding to manage these clinical trial tissue repositories, which could be so useful for the discovery and validation of biomarkers. The group suggested finding a new funding mechanism to support the management of these tissue repositories.

The group also discussed how to move new diagnostic tests and treatments more rapidly into clinical practice. Suggestions included incorporating biomarker tests early in drug development, even at the level of preclinical models, Dr. Massion said. He also noted the need for both professional societies and regulatory agencies to establish guidelines for the evaluation and standardization of specific biomarkers, as well as for the standardization of the clinical elements associated with the biomarkers. "The biological nature of the study, the stage of the study, the performance of the tests, the mode of action, the stage validation, and the technologies—all these need to be clearly standardized, and guidelines would help the field move forward," Dr. Massion said. Such standardization and guidelines, as well as incorporating biomarker tests early in the development process, should help to reduce the costs of biomarker development, provide stronger evidence for improved outcomes, and thus make biomarker test reimbursement by third-

[21] See http://cancergenome.nih.gov/about/index.asp.

party payors more likely, he added. Finally, he pointed out that biomarker-based trial designs will narrow down the population most likely to benefit and therefore reduce sample size and toxicity, and provide opportunities for enrollment in other trials.

CONCLUDING REMARKS

In final closing comments before adjourning the meeting, planning committee co-chairs Drs. Moses and Mendelsohn both commended the quality of the presentations and subsequent discussions. Dr. Mendelsohn remarked that a valuable outcome from the meeting was "educating each other in this room." For example, "clearly there was tremendous cross communication between the people that are in [the molecular bioimaging] field and the people who want to use [the technologies of] that field but are not in it," he said.

Dr. Mendelsohn also noted one important omission in the workshop. The planning committee had hoped to address the topic of sharing intellectual property in greater detail, but the expert invited to speak about intellectual property issues unfortunately had to cancel the week of the meeting. Noting that this would be an essential topic for a consensus committee to address, Dr. Mendelsohn asserted that "we have got to figure out a way to incentivize." "You have to show [the stakeholders] that it is to their advantage [to share intellectual property]," he said.

Drs. Mendelsohn and Moses both noted that important policy issues had been identified and explored in each session, and that these issues would benefit from further study. Accordingly, this summary of the conference proceedings will serve as input to the deliberations of an Institute of Medicine committee that will develop consensus-based recommendations for moving the field of cancer clinical trials forward.

Numerous suggestions were put forth by speakers and discussion groups, including the following:

- Consider adaptive methods more often in designing clinical trials.
- Consider increased use of Phase 0 trials for addressing early biological endpoints in patients and for compound triage when multiple compounds are being considered.
- Increase efforts to standardize and harmonize imaging methodologies used in clinical trials.

- Encourage incorporation of biomarker tests early in drug development, even at the level of preclinical models.
- Devote greater efforts to establishing guidelines for the evaluation, standardization, validation, and qualification of biomarkers, especially those used in clinical decision-making.
- Develop a new funding mechanism to support the management of existing tissue repositories, and to support rational access to these valuable resources for testing biomarkers.

Addressing some issues will require a great deal of work and research, Dr. Mendelsohn noted, but there was also some "low-hanging fruit" that could be accomplished in the near future as well. In particular, he noted that developing partnerships among Federal agencies like the NCI and the FDA in exploratory areas of research were likely to help move the field forward.

REFERENCES

Arias-Mendoza, F., M.R. Smith, and T.R. Brown. 2004. Predicting treatment response in non-Hodgkin's lymphoma from the pretreatment tumor content of phosphoethanolamine plus phosphocholine. *Acad Radiol* 11(4):368-376.

Berry, D.A. 2005. Statistical innovations in cancer research. In *Cancer Medicine* e.7: 411-425. London: BC Decker.

Berry, D.A. 2006. Bayesian clinical trials. *Nature Reviews Drug Discovery* 5:27-36.

Chin, K., S. DeVries, J. Fridlyand, P.T. Spellman, R. Roydasgupta, W.-L. Kuo, A. Lapuk, R.M. Neve, Z. Qian, T. Ryder, F. Chen, H. Feiler, T. Tokuyasu, C. Kingsley, S. Dairkee, Z. Meng, K. Chew, D. Pinkel, A. Jain, B.M. Ljung, L. Esserman, D.G. Albertson, F.M. Waldman, and J.W. Gray. 2006. Genomic and transcriptional aberrations linked to breast cancer pathophysiologies. *Cancer Cell* 10:529-541.

Comis, R.L., J.D. Miller, C.R. Aldigé, L. Krebs, and E. Stoval. 2003. Public attitudes toward participation in cancer clinical trials. *J Clin Oncol* 21(5):830-835.

Dave, S.S., K. Fu, G.W. Wright, L.T. Lam, P. Kluin, E.J. Boerma, T.C. Greiner, D.D. Weisenburger, A. Rosenwald, G. Ott, H.K. Müller-Hermelink, R.D. Gascoyne, J. Delabie, L.M. Rimsza, R.M. Braziel, T.M. Grogan, E. Campo, E.S. Jaffe, B.J. Dave, W. Sanger, M. Bast, J.M. Vose, J.O. Armitage, J.M. Connors, E.B. Smeland, S. Kvaloy, H. Holte, R.I. Fisher, T.P. Miller, E. Montserrat, W.H. Wilson, M. Bahl, H. Zhao, L. Yang, J. Powell, R. Simon, W.C. Chan, L.M. Staudt, and Lymphoma/Leukemia Molecular Profiling Project. 2006. Molecular diagnosis of Burkitt's lymphoma. *N Engl J Med* 354(23):2431-2442.

Dehdashti, F., M.A. Mintun, J.S. Lewis, J. Bradley, R. Govindan, R. Laforest, M.J. Welch, and B.A. Siegel. 2003. In vivo assessment of tumor hypoxia in lung cancer with 60Cu-ATSM. *Eur J Nucl Med Mol Imaging* 30(6):844-850.

Di Leo, A., H. Gomez, Z. Aziz, Z. Zvirbule, M. Arbushites, C.R. Oliva, M. Koehler, L.S. Williams, J. Dering, and R.S. Finn. 2007. Lapatinib (L) with paclitaxel compared to paclitaxel as first-line treatment for patients with metastatic breast cancer: A phase III randomized, double-blind study of 580 patients. *J Clin Oncol 2007 ASCO Annual Meeting Proceedings Part I* 25(18S):1011.

Flaherty, K.T., M.A. Rosen, D.F. Heitjan, M.L. Gallagher, B. Schwart, M.D. Schnall, and P.J. Dwyer. 2008. Pilot study of DCE-MRI to predict progression-free survival with sorafenib therapy in renal cell carcinoma. *Cancer Biol Ther* 7(4):Epub ahead of print.

Fridlyand, J., A.M. Snijders, B. Ylstra, H. Li, A. Olshen, R. Segraves, S. Dairkee, T. Tokuyasu, B.M. Ljung, A.N. Jain, J. McLennan, J. Ziegler, K. Chin, S. Devries, H. Feiler, J.W. Gray, F. Waldman, D. Pinkel, and D.G. Albertson. 2006. Breast tumor copy number aberration phenotypes and genomic instability. *BMC Cancer* 6:96.

Giles, F.J., H.M. Kantarjian, J.E. Cortes, G. Garcia-Manero, S. Verstovsek, S. Faderl, D.A. Thomas, A. Ferrajoli, S. O'Brien, J.K. Wathen, L.C. Xiao, D.A. Berry, and E.H. Estey. 2003. Adaptive randomized study of idarubicin and cytarabine versus troxacitabine and cytarabine versus troxacitabine and idarubicin in untreated patients 50 years or older with adverse karyotype acute myeloid leukemia. *J Clin Oncol* 21(9):1722-1727.

Grosso, F., R.L. Jones, G.D. Demetri, I.R. Judson, J.Y. Blay, A. Le Cesne, R. Sanfilippo, P. Casieri, P. Collini, P. Dileo, C. Spreafico, S. Stacchiotti, E. Tamborini, J.C. Tercero, J. Jimeno, M. D'Incalci, A. Gronchi, J.A. Fletcher, S. Pilotti, and P.G. Casali. 2007. Efficacy of trabectedin (ecteinascidin-743) in advanced pretreated myxoid liposarcomas: A retrospective study. *Lancet Oncol* 8(7):595-602.

Hamstra, D.A., T.L. Chenevert, B.A. Moffat, T.D. Johnson, C.R. Meyer, S.K. Mukherji, D.J. Quint, S.S. Gebarski, X. Fan, C.I. Tsien, T.S. Lawrence, L. Junck, A. Rehemtulla, and B.D. Ross. 2005. Evaluation of the functional diffusion map as an early biomarker of time-to-progression and overall survival in high-grade glioma. *Proc Natl Acad Sci USA* 102(46):16759-16764.

Hecht, J.R., T. Trarbach, E. Jaeger, J. Hainsworth, R. Wolff, K. Lloyd, G. Bodoky, M. Borner, D. Laurent, and C. Jacques. 2005. A randomized, double-blind, placebo-controlled, phase III study in patients (Pts) with metastatic adenocarcinoma of the colon or rectum receiving first-line chemotherapy with oxaliplatin/5-fluorouracil/leucovorin and PTK787/ZK 222584 or placebo (CONFIRM-1). *J Clin Oncol 2005 ASCO Annual Meeting Proceedings Part I of II* 23(16S):3.

Karrison, T.G., D. Huo, and R. Chappell. 2003. A group sequential, response-adaptive design for randomized clinical trials. Control Clin Trials 24(5):506-522.

Kola, I., and J. Landis. 2004. Can the pharmaceutical industry reduce attrition rates? *Nat Rev Drug Discov* 3(8):711-715.

The Lewin Group. 2005. Guidance document for implementing effective cancer clinical trials. http://www.c-changetogether.org/pubs/default.asp (accessed Jan. 15, 2008).

Linden, H.M., S.A. Stekhova, J.M. Link, J.R. Gralow, R.B. Livingston, G.K. Ellis, P.H. Petra, L.M. Peterson, E.K. Schubert, L.K. Dunnwald, K.A. Krohn, and D.A. Mankoff. 2006. Quantitative fluoroestradiol positron emission tomography imaging predicts response to endocrine treatment in breast cancer. *J Clin Oncol* 24(18):2793-2799.

Ludwig, J.A., and J.N. Weinstein. 2005. Biomarkers in cancer staging, prognosis and treatment selection. *Nat Rev Cancer* 5(11):845-856.

Lynch, T.J., D.W. Bell, R. Sordella, S. Gurubhagavatula, R.A. Okimoto, B.W. Brannigan, P.L. Harris, S.M. Haserlat, J.G. Supko, F.G. Haluska, D.N. Louis, D.C. Christiani, J. Settleman, and D.A. Haber. 2004. Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. *N Engl J Med* 350(21):2129-2139.

Metz, J.M., C. Coyle, C. Hudson, and M. Hampshire. 2005. An Internet-based cancer clinical trials matching resource. *J Med Internet Res* 7(3):e24.

Morgan, B., A.L. Thomas, J. Drevs, J. Hennig, M. Buchert, A. Jivan, M.A. Horsfield, K. Mross, H.A. Ball, L. Lee, W. Mietlowski, S. Fuxuis, C. Unger, K. O'Byrne, A. Henry, G.R. Cherryman, D. Laurent, M. Dugan, D. Marmé, and W.P. Steward. 2003. Dynamic contrast-enhanced magnetic resonance imaging as a biomarker for the pharmacological response of PTK787/ZK 222584, an inhibitor of the vascular endothelial growth factor receptor tyrosine kinases, in patients with advanced colorectal cancer and liver metastases: Results from two phase I studies. *J Clin Oncol* 21(21):3955-3964.

Mortimer, J.E., F. Dehdashti, B.A. Siegel, K. Trinkaus, J.A. Katzenellenbogen, and M.J. Welch. 2001. Metabolic flare: indicator of hormone responsiveness in advanced breast cancer. *J Clin Oncol* 19(11):2797-2803.

Neve, R.M., K. Chin, J. Fridlyand, J. Yeh, F.L. Baehner, T. Fevr, L. Clark, N. Bayani, J.-P. Coppe, F. Tong, T. Speed, P.T. Spellman, S. DeVries, A. Lapuk, N.J. Wang, W.-L. Kuo, J.L. Stilwell, D. Pinkel, D.G. Albertson, F.M. Waldman, F. McCormick, R.B. Dickson, M.D. Johnson, M. Lippman, S. Ethier, A. Gazdar, and J.W. Gray. 2006. A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. *Cancer Cell* 10(6):515-527.

Nunn, A.D. 2006. The cost of developing imaging agents for routine clinical use. *Invest Radiol* 41:206-212.

Paez, J.G., P.A. Jänne, J.C. Lee, S. Tracy, H. Greulich, S. Gabriel, P. Herman, F.J. Kaye, N. Lindeman, T.J. Boggon, K. Naoki, H. Sasaki, Y. Fujii, M.J. Eck, W.R. Sellers, B.E. Johnson, and M. Meyerson. 2004. EGFR mutations in lung cancer: Correlation with clinical response to gefitinib therapy. *Science* 304(5676):1497-1500.

Paik, S., S. Shak, G. Tang, C. Kim, J. Baker, M. Cronin, F.L. Baehner, M.G. Walker, D. Watson, T. Park, W. Hiller, E.R. Fisher, D.L. Wickerham, J. Bryant, and N. Wolmark. 2004. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N Engl J Med* 351(27):2817-2826.

Pao, W., V. Miller, M. Zakowski, J. Doherty, K. Politi, I. Sarkaria, B. Singh, R. Heelan, V. Rusch, L. Fulton, E. Mardis, D. Kupfer, R. Wilson, M. Kris, H. Varmus. 2004. EGF receptor gene mutations are common in lung cancers from "never smokers" and are associated with sensitivity of tumors to gefitinib and erlotinib. *Proc Natl Acad Sci USA* 101(36):13306-13311.

Peto, R. 1985. Discussion of "On the Allocation of Treatments in Sequential Medical Trials" by J.A. Bather and "The Search for Optimality in Clinical Trials." *Int Stat Rev* 53(1):31-34.

Reddick, W.E., S. Wang, X. Xiong, J.O. Glass, S. Wu, S.C. Kaste, C.B. Pratt, W.H. Meyer, and B.D. Fletcher. 2001. Dynamic magnetic resonance imaging of regional contrast access as an additional prognostic factor in pediatric osteosarcoma. *Cancer* 91(12):2230-2237.

Rischin, D., R.J. Hicks, R. Fisher, D. Binns, J. Corry, S. Porceddu, and L.J. Peters. 2006. Prognostic significance of [18F]-misonidazole positron emission tomography-detected tumor hypoxia in patients with advanced head and neck cancer randomly assigned to chemoradiation with or without tirapazamine: A substudy of Trans-Tasman Radiation Oncology Group Study 98.02. *J Clin Oncol* 24(13):2098-2104.

Schwartz, L.H., M. Mazumdar, L. Wang, A. Smith, S. Marion, D.M. Panicek, and R.J. Motzer. 2003. Response assessment classification in patients with advanced renal cell carcinoma treated on clinical trials. *Cancer* 98(8):1611-1619.

Shukla-Dave, A., H. Hricak, M.W. Kattan, D. Pucar, K. Kuroiwa, H.N. Chen, J. Spector, J.A. Koutcher, K.L. Zakian, and P.T. Scardino. 2007. The utility of magnetic resonance imaging and spectroscopy for predicting insignificant prostate cancer: an initial analysis. *BJU Int* 99(4):786-793.

Sjöblom, T., S. Jones, L.D. Wood, D.W. Parsons, J. Lin, T.D. Barber, D. Mandelker, R.J. Leary, J. Ptak, N. Silliman, S. Szabo, P. Buckhaults, C. Farrell, P. Meeh, S.D. Markowitz, J. Willis, D. Dawson, J.K. Willson, A.F. Gazdar, J. Hartigan, L. Wu, C. Liu, G. Parmigiani, B.H. Park, K.E. Bachman, N. Papadopoulos, B. Vogelstein, K.W. Kinzler, and V.E. Velculescu. 2006. The consensus coding sequences of human breast and colorectal cancers. *Science* 314(5797):268-274.

Steward, W.P., A. Thomas, B. Morgan, B. Wiedenmann, C. Bartel, U. Vanhoefer, T. Trarbach, U. Junker, D. Laurent, and D. Lebwohl. 2004. Expanded phase I/II study of PTK787/ZK 222584 (PTK/ZK), a novel, oral angiogenesis inhibitor, in combination with FOLFOX-4 as first-line treatment for patients with metastatic colorectal cancer. *J Clin Oncol, 2004 ASCO Annual Meeting Proceedings (Post-Meeting Edition)* 22(14S):3556.

Taguchi, F., B. Solomon, V. Gregorc, H. Roder, R. Gray, K. Kasahara, M. Nishio, J. Brahmer, A. Spreafico, V. Ludovini, P.P. Massion, R. Dziadziuszko, J. Schiller, J. Grigorieva, M. Tsypin, S.W. Hunsucker, R. Caprioli, M.W. Duncan, F.R. Hirsch, P.A. Bunn, Jr., and D.P. Carbone. 2007. Mass spectrometry to classify non-small-cell lung cancer patients for clinical outcome after treatment with epidermal growth factor receptor tyrosine kinase inhibitors: a multicohort cross-institutional study. *J Natl Cancer Inst* 99(11):838-846.

Tomaszewski, J.E. 2004. Multi-species toxicology approaches for oncology drugs: The U.S. perspective. *Eur J Cancer* 40(6):907-913.

Tsiatis, A.A. and C. Mehta. 2003. On the inefficiency of the adaptive design for monitoring clinical trials. *Biometrika* 90(2):367-378.

Tufts Center for the Study of Drug Development. 2007. *Outlook 2007*. http://csdd.tufts.edu/InfoServices/OutlookPDFs/Outlook2007.pdf (accessed Jan. 15, 2008).

Von Hoff, D.D., R. Penny, S. Shack, E. Campbell, D. Taverna, M. Borad, D. Love, J. Trent, and M. Bittner. 2006. Frequency of potential therapeutic targets identified by immunohistochemistry (IHC) and DNA microarray (DMA) in tumors from patients who have progressed on multiple therapeutic agents. *J Clin Oncol, 2006 ASCO Annual Meeting Proceedings Part I* 24(18S):3071.

Wiener, M.B., H.M. Newman, and E.A. Spradley. 2007. Revolutionizing oncology patient enrollment in clinical trials: Just-in-time approach. *J Clin Oncol, 2007 ASCO Annual Meeting Proceedings Part I* 25(18S):6577.

Yildiz, P.B., Y. Shyr, J.S. Rahman, N.R. Wardwell, L.J. Zimmerman, B. Shakhtour, W.H. Gray, S. Chen, M. Li, H. Roder, D.C. Liebler, W.L. Bigbee, J.M. Siegfried, J.L. Weissfeld, A.L. Gonzalez, M. Ninan, D.H. Johnson, D.P. Carbone, R.M. Caprioli, and P.P. Massion. 2007. Diagnostic accuracy of MALDI mass spectrometric analysis of unfractionated serum in lung cancer. *J Thorac Oncol* 2(10):893-901.

Zhang, W., A.F. Purchio, K. Chen, J. Wu, L. Lu, R. Coffee, P.R. Contag, and D.B. West. 2003. A transgenic mouse model with a luciferase reporter for studying in vivo transcriptional regulation of the human CYP3A4 gene. *Drug Metab Dispos* 31(8):1054-1064.

Acronyms

ACS	American Cancer Society
CLIA	Clinical Laboratory Improvement Amendments
CMS	Centers for Medicare & Medicaid Services
CT	computed tomography
CTEP	Cancer Therapy Evaluation Program
DCE	dynamic contrast enhanced
EGF	epidermal growth factor
ER	estrogen receptor
FDA	Food and Drug Administration
FDAMA	Food and Drug Administration Modernization Act
FDG	[18F]-2-fluoro-2-deoxy-D-glucose
FR	Federal Register
GCRC	General Clinical Research Center
HER-2	human epidermal growth factor receptor 2
IND	investigational new drug

IRAT	Image Response Assessment Team
IRB	Institutional Review Board
LBNL	Lawrence Berkeley National Laboratory
MALDI	matrix assisted laser desorption/ionization
MedPAC	Medicare Payment Advisory Commission
MRI	magnetic resonance imaging
MSKCC	Memorial Sloan-Kettering Cancer Center
mTOR	mammalian target of rapamycin
NCCS	National Coalition for Cancer Survivorship
NCI	National Cancer Institute
NIH	National Institutes of Health
NSCLC	non-small cell lung cancer
PARP	poly(ADP-ribose) polymerase
PET	positron emission tomography
PhRMA	Pharmaceutical Research and Manufacturers of America
pre-IDE	preinvestigation device exemption
PSA	prostate-specific antigen
RECIST	Response Evaluation Criteria in Solid Tumors
SOP	standard operating procedure
SPA	special protocol assessment
SPECT	single photon emission computed tomography
SPORE	Specialized Program of Research Excellence
TBAC	Tissue Banking Analysis Center
TGen	Translational Genomics Research Institute
TNF	tumor necrosis factor
VEGF	vascular endothelial growth factor

Glossary

Adaptive trial design—a trial with one or more decision points built into the trial design. How the trial proceeds following each decision point depends on the data observed up to that point. One of the more commonly used adaptive trial designs is one that stops early or continues later than expected based on results that indicate how effective the treatment under study is after a limited number of patients have been tested.

Amplification—a process resulting in an increase in the number of copies of specific genetic sequences within a genome. Amplification of some genes can cause some types of cancer.

Analytical validity—the accuracy of a test in detecting the specific entity that it was designed to detect. This accuracy does not imply any clinical significance, such as diagnosis.

Apoptosis—programmed cell death.

Bias—the systematic but unintentional erroneous association of some characteristics within a group in a way that distorts a comparison with another group.

BRCA—a gene that when mutated increases risk of developing breast cancer and/or ovarian cancer. Two BRCA genes have been identified and are known as BRCA1 and BRCA2. The acronym BRCA stands for "breast cancer gene."

Clinical endpoint—characteristics or variables expected at the end of a clinical trial, reflecting how patients feel, function, or survive. After achievement of a clinical endpoint, a patient's participation in the clinical trial ends.

Clinical trial—a formal study carried out according to a prospectively defined protocol that is intended to discover or verify the safety and effectiveness of medical procedures or interventions in humans.

Clinical utility—the physical and psychological benefits and risks of a given technique or test.

Computed tomography (CT)—a radiographic technique that uses a computer to assimilate multiple X-ray images into two-dimensional, cross-sectional images or a three-dimensional image. Use of this technique can reveal many soft-tissue structures not shown by conventional radiography.

De novo classification—a Food and Drug Administration classification of a device or diagnostic that is not equivalent to a legally marketed product. De novo classification is a way for a low-risk medical device to bypass the premarket approval process.

Deletion—the loss of genetic material. Some cancers are triggered by the deletion of key genes, portions of genes, or their regulatory sequences.

Diagnostic—an investigative tool or technique used in biological studies or to identify or determine the presence of a disease or other condition.

Epidermal growth factor receptor (EGFR)—a receptor that is overproduced in several solid tumors, including breast and lung cancers. Its overproduction is often linked to a poorer prognosis because it enables cell proliferation, cell migration, and blood vessel development. Several new drugs recently approved by the Food and Drug Administration specifically target EGFR.

Estrogen-receptor positive (ER+)—a tumor, either a primary tumor or a metastasis, that tests positive for estrogen receptors. Such tumors, often found in cancers such as breast cancer or uterine sarcoma, may be treated by hormonal therapy that decreases or blocks estrogen to prevent or slow tumor growth. Some tumors may also be progesterone-receptor positive (PR+) and may be treated with a different type of hormonal therapy.

Genome—an organism's entire complement of DNA, which determines its genetic characteristics.

Genomics—the study of all of the nucleotide sequences, including structural genes, regulatory sequences, and noncoding DNA segments, in the chromosomes of an organism or tissue sample. One example of the application of genomics in oncology is the use of microarray or other techniques to uncover the genetic "fingerprint" of a tissue sample. This genetic fingerprint is the pattern that stems from the variable expression of different genes in normal and cancer tissues.

Global outsourcing—conducting a clinical trial outside the United States in an effort to save money.

High-throughput system—any approach using robotics, automated machines, and computers to process many samples at once.

Human epidermal growth factor receptor 2 (HER-2/neu)—a growth factor receptor that is used as a breast cancer biomarker for prognosis and treatment with the drug trastuzumab (Herceptin), which targets the protein. The HER-2/neu protein is overexpressed in approximately 25 percent of breast cancer patients, due to amplification of the gene.

Magnetic resonance imaging (MRI)—a method by which images are created by recording signals generated from the excitation (the gain and loss of energy) of hydrogen atoms in tissue when placed within a powerful magnetic field and pulsed with radio frequencies.

Mass spectrometry—a method for separating ionized molecular particles according to mass by applying a combination of electrical and magnetic fields to deflect ions passing in a beam through the instrument.

Microarray—a high-throughput tool for biological assays in which many different probes (sometimes 10,000 or more) are deposited on a chip surface (glass or silicon) for analysis. DNA microarrays are the most commonly used microarrays.

Microdose study—a study employed in phase 0 clinical trials that uses imaging or other means to assess where in the body a compound is distributed and for how long it remains in these sites. A microdose is defined as less than 1/100th of the dose predicted to yield pharmacological effects, and less than 100 micrograms. A microdose study is designed *not* to induce pharmacologic effects; rather, it can indicate whether an experimental drug reaches its target.

Off-label use—the doctor-prescribed use of a drug for a condition or dis-

ease for which it has not been approved by the Food and Drug Administration, or the use of a drug by a non-approved method.

Pharmacodynamics—the study of the biochemical and physiological effects of drugs, the mechanisms of drug action, and the relationship between drug concentration and effect. Pharmacodynamics is the study of what a drug does to the body, as opposed to pharmacokinetics, which is the study of what a body does to a drug.

Pharmacokinetics—the study of the metabolism of, or chemical changes experienced by, substances in an organism over time, such as drugs. Pharmacokinetics is used to determine how quickly and for how long a drug acts on its target.

Phase 0 trial—an exploratory investigational new drug study (IND). The main purpose is to assess the likely therapeutic effectiveness of a compound, based on whether it reaches its target in people and how long it is active in the body. An exploratory IND study tests a new experimental drug on human subjects prior to a phase I clinical trial.

Phase I trial—a clinical trial in a small number of patients in which the toxicity and dosing of an intervention are assessed.

Phase II trial—a clinical trial in which the safety and preliminary efficacy of an intervention are assessed in patients.

Phase III trial—a large-scale clinical trial in which the safety and efficacy of an intervention are assessed in a large number of patients. The Food and Drug Administration generally requires new drugs to be tested in Phase III trials before they can be put on the market.

Positive predictive value—the probability that an individual with a positive test has or will develop a particular disease or characteristic that the test is designed to detect. It is a measure of the ratio of true positives to the sum of true and false positives.

Positron emission tomography (PET)—a highly sensitive technique that uses radioactive probes to image *in vivo* tumors, receptors, enzymes, DNA replication, gene expression, antibodies, hormones, drugs, and other compounds and processes.

Premarket approval—a Food and Drug Administration approval process for a new test or device that enables it to be marketed for clinical use. To

receive this approval, the manufacturer of the product must submit clinical data showing the product is safe and effective for its intended use.

Premarket notification or 510(k)—a Food and Drug Administration review process that enables a new test or device to be marketed for clinical use without undergoing the premarket approval process. To qualify for the 510(k), manufacturers must provide documentation supporting the claim that their product is substantially equivalent to one already on the market, in terms of safety and efficacy.

Proteomics—the study of the structure, function, and interactions of the proteins produced by the genes of a particular cell, tissue, or organism. The application of proteomics in oncology may involve mass spectrometry, two-dimensional polyacrylamide gel electrophoresis, protein chips, and other techniques to uncover the protein "fingerprint" of a tissue sample. This protein fingerprint is the pattern that stems from the various amounts and types of all the proteins in the sample.

PSA test—a blood test that detects prostate-specific antigen (PSA). The PSA test was approved by the Food and Drug Administration in 1985 for prostate cancer recurrence, but it is now widely used as a screening test for prostate cancer.

Qualification—the evidentiary process of linking an assay with biological and clinical endpoints that is dependent on the intended application.

Randomization—randomly placing trial participants in different arms of a trial; for example, one arm may use a standard treatment while another uses a standard treatment plus a new drug.

Sensitivity (clinical)—a measure of how often a test correctly identifies patients with a specific diagnosis. It is calculated as the number of true-positive results divided by the sum of true-positive and false-negative results.

Specificity (clinical)—a measurement of how often a test correctly identifies the proportion of persons without a previous diagnosis. It is calculated as the number of true-negative results divided by the sum of true negatives and false positives.

Surrogate endpoint—a biomarker that is intended to substitute for a clinical endpoint in a therapeutic clinical trial and is expected to predict

clinical benefit, harm, or lack thereof, based on epidemiologic, therapeutic, pathophysiologic, or other scientific evidence.

Two-dimensional gel electrophoresis—a technique used to separate molecules from one another based on their isoelectric point, charge, and size. One-dimensional electrophoresis, in contrast, has fewer molecule-distinguishing capabilities, as it only separates molecules on the basis of their charge and size.

Type 1 error—In statistics, the error of the "false positive." In other words, concluding that a test result is positive when it is, in fact, negative is a Type 1 error.

Validation—the process of assessing an assay or measurement performance characteristics.

Appendix A

Workshop Agenda

National Cancer Policy Forum
Workshop on
Improving the Quality of Cancer Clinical Trials

The Keck Center of The National Academies
Room 100
500 Fifth Street, NW
Washington, DC 20001

Agenda
October 4–5, 2007

Day 1: October 4, 2007

Welcome and Opening Remarks
8:00 am – 8:15 am

John Mendelsohn, M.D., MD Anderson Cancer Center

Session 1: New Clinical Trial Designs
8:15 am – 10:45 am

A) Exploratory INDs and Phase 0 Trials
8:15 am – 9:45 am

Moderator: James Doroshow, M.D., NCI

David Jacobson-Kram, Ph.D., FDA
"Overview of the Exploratory IND: Differences in the Traditional IND"
George Mills, M.D., Parexel International Corporation
"Molecular Imaging and Nanotechnology: Strategic Implementation of the 2006 Exploratory IND Guidance"
James Doroshow, M.D., NCI
"Phase 0 Clinical Trials in Cancer Drug Development: From Concept to Practice"

B) Adaptive Trial Designs
9:45 am – 10:45 am

Moderator: John Wagner, M.D., Ph.D., Merck

Don Berry, Ph.D., MD Anderson Cancer Center
"Adaptive Designs for Cancer Trials"
Susan Ellenberg, Ph.D., University of Pennsylvania
"Adaptive Designs in Cancer Trials: Consensus and Debate"

BREAK
10:45 am – 11:00 am

SESSION 1 RESUMES
11:00 am – 12:30 pm

C) Targeting Multiple Pathways with Multiple Drugs
11:00 am – 12:00 pm

Moderator: Roy Herbst, M.D., Ph.D., MD Anderson Cancer Center

Janet E. Dancey, M.D., NCI
"Strategies to Develop Combinations of Investigational Agents"
Joe Gray, Ph.D., UCSF Comprehensive Cancer Center
"Modeling Molecular Heterogeneity to Enhance Multidrug Clinical Trial Design"

D) Preclinical Model Systems
12:00 pm – 12:30 pm

Ken Anderson, M.D., Dana-Farber Cancer Institute
"Translation from Preclinical Model Systems to the Bedside in Multiple Myeloma"

LUNCH BREAK
12:30 pm – 1:15 pm

SESSION II: MOLECULAR IMAGING
1:15 pm – 3:15 pm

Moderators: Hedvig Hricak, M.D., Ph.D., Memorial Sloan-Kettering Cancer Center, and David Piwnica-Worms, M.D., Ph.D., Washington University School of Medicine

Introduction and Mission Statement
Hedvig Hricak, M.D., Ph.D., Memorial Sloan-Kettering Cancer Center

Lawrence Schwartz, M.D., Memorial Sloan-Kettering Cancer Center
"Imaging Studies That Facilitate Clinical Trials Today"
Steven Larson, M.D., Memorial Sloan-Kettering Cancer Center
"Molecular Imaging: Biomarkers for Oncology"
John Gore, Ph.D., Vanderbilt University
"Imaging Biomarkers on the Near Horizon"
David Piwnica-Worms, M.D., Ph.D., Washington University School of Medicine
"Molecular Imaging Strategies in Drug Development"

Panel Discussion
Tim McCarthy, Ph.D., Pfizer
Jeff Evelhoch, Ph.D., Amgen, Inc.
Jerry Collins, Ph.D., NCI

BREAK
3:15 pm – 3:30 pm

Breakout Discussions: Case Studies of Clinical Trial Designs
3:30 pm – 5:30 pm

1) Phase 0 Trials

Moderator/Reporter: Giulio Draetta , M.D., Ph.D., Merck

James Doroshow, M.D., NCI

Invited Discussant: Richard Schilsky, M.D., University of Chicago

2) Adaptive Trial Design

Moderator/Reporter: John Wagner, M.D., Ph.D., Merck

Lung Cancer Personalized Therapy
Jack Lee, Ph.D., MD Anderson Cancer Center
"Design for Targeted Therapies in Lung Cancer: Statistical Considerations"
Roy Herbst, M.D., Ph.D., MD Anderson Cancer Center
"Toward Personalized Therapy for Lung Cancer"

Invited Discussant: Rick Chappell, Ph.D., University of Wisconsin
"Comments on the Controversy Over Response-Adaptive Randomization"

3) Imaging

Moderators/Reporters: David Piwnica-Worms, M.D., Ph.D., Washington University School of Medicine, and Hedvig Hricak, M.D., Ph.D., Memorial Sloan-Kettering Cancer Center

Tim McCarthy, Ph.D., Pfizer
Jeff Evelhoch, Ph.D., Amgen, Inc.
"Incorporating Imaging Biomarkers in Phase I Oncology Trials"

4) Use of Proteomics/Genomics to Assign Therapy in Lung Cancer

Moderator/Reporter: John Mendelsohn, M.D., MD Anderson Cancer Center

Proteomics and Lung Cancer Prediction (EGFR)

David Carbone, M.D., Ph.D., Vanderbilt University

"Molecular Signatures to Guide Selection of Lung Cancer Patient Therapy"

Use of Genomics to Assign Therapy in Lung Cancer (genetic targets for Iressa and Tarceva)

William Pao, M.D., Ph.D., Memorial Sloan-Kettering Cancer Center

Mark Kris, M.D., Memorial Sloan-Kettering Cancer Center

"The Lung Cancer Oncogenome Group: Bedside to Bench and Beyond"

5) Use of Genetics/Genomics to Assign Therapy

Moderator/Reporter: Pierre Massion, M.D., Vanderbilt University

Breast Cancer Personalized Therapy: The TAILORx Trial

Joseph Sparano, M.D., Albert Einstein Comprehensive Cancer Center

"Rationale for and Design of TAILORx"

Steven Shak, M.D., Genomic Health

"The 21 Gene Oncotype DX Assay and the NCI-Sponsored TAILORx"

Using Genetic and Genomic Technologies in Design and Execution of Cancer Clinical Trials

Raju Kucherlapati, Ph.D., Harvard Partners Center for Genetics and Genomics

Invited Discussant: David Chang, M.D., Ph.D., Amgen, Inc.

ADJOURN DAY 1
5:30 pm

Day 2: October 5, 2007

WELCOME AND OPENING REMARKS
8:00 am – 8:15 am

Hal Moses, M.D., Vanderbilt University

SESSION III: SCREENING FOR PREDICTIVE MARKERS
8:15 am – 10:15 am

Moderator: George Mills, M.D., Parexel International Corporation

Pierre Massion, M.D., Vanderbilt University
"Are Genomics and Proteomics Biomarkers Ready for Prime Time?"
James Heath, Ph.D., California Institute of Technology
"Lowering the Cost of In Vitro Diagnostics Measurements Associated with Clinical Trials by a Factor of 10 (or more)"
Daniel Sullivan, M.D., Duke University
"Is There a Role for Imaging as a Predictive Biomarker?"
Daniel Von Hoff, M.D., Translational Genomics Research Institute
"Improving a Patient's Chance of Benefiting from Early Clinical Trials"

BREAK
10:15 am –10:30 am

SESSION IV: COLLABORATIONS AMONG ACADEMIA, PHARMA, BIOTECH, AND GOVERNMENT
10:30 am – 12:30 pm

Moderator: John Mendelsohn, M.D., MD Anderson Cancer Center

Robert Comis, M.D., Coalition of National Cancer Cooperative Groups
"The Public Sector Perspective"
Kevin Schulman, M.D., Duke University Medical School
"Cost of Clinical Trials"
Gwen Fyfe, M.D., Genentech
"The Industry Perspective"

LUNCH BREAK
12:30 pm – 1:15 pm

SESSION V: REGULATORY ISSUES
1:15 pm – 3:15 pm

Moderator: Janet Woodcock, M.D., FDA

Susan Jerian, M.D., OncoRD, Inc.
"The Interplay of Laws, Regulations, and Policies: Moving Cancer Therapeutics Development Out of the Quagmire"
Ellen Stovall, National Coalition for Cancer Survivorship
"Clinical Trial Design, Drug Development, and Policy Issues of Importance to Cancer Advocates"
Janet Woodcock, M.D., FDA
"Issues in Cancer Drug Development of the Future"
Steven Gutman, M.D., FDA
"Regulation of Biomarkers"

REPORTS FROM THE CASE STUDY DISCUSSION GROUPS
3:15 pm – 4:15 pm

Giulio Draetta, M.D., Ph.D., Merck Research Laboratories
"Phase 0 Trials"
Roy Herbst, M.D., Ph.D., MD Anderson Cancer Center
"Adaptive Trial Design"
David Pinwica-Worms, M.D., Ph.D., Washington University School of Medicine, and Hedvig Hricak, M.D., Ph.D., Memorial Sloan-Kettering Cancer Center
"Imaging"
John Mendelsohn, M.D., MD Anderson Cancer Center
"Use of Proteomics/Genomics to Assign Therapy in Lung Cancer"
Pierre Massion, M.D., Vanderbilt University
"Use of Genetics/Genomics to Assign Therapy"

WRAP-UP/SUMMARY
4:15 pm – 4:30 pm

John Mendelsohn, M.D., MD Anderson Cancer Center, and Hal Moses, M.D., Vanderbilt-Ingram Cancer Center

ADJOURN DAY 2
4:30 pm

Appendix B

WORKSHOP SPEAKERS,* MODERATORS,† AND INVITED DISCUSSANTS‡

Ken Anderson, M.D., Dana-Farber Cancer Institute*
Don Berry, Ph.D., MD Anderson Cancer Center*
David Carbone, M.D., Ph.D., Vanderbilt University*
David Chang, M.D., Ph.D., Amgen, Inc.‡
Rick Chappell, Ph.D., University of Wisconsin‡
Jerry Collins, Ph.D., National Cancer Institute‡
Robert Comis, M.D., Coalition of National Cancer Cooperative Groups*
Janet Dancey, M.D., National Cancer Institute*
James Doroshow, M.D., National Cancer Institute*†
Giulio Draetta, M.D., Ph.D., Merck Research Laboratories†‡
Susan Ellenberg, Ph.D., University of Pennsylvania*
Jeff Evelhoch, Ph.D., Amgen, Inc.*‡
Gwen Fyfe, M.D., Genentech, Inc.*
John Gore, Ph.D., Vanderbilt University*
Joe Gray, Ph.D., University of California–San Francisco Comprehensive Cancer Center*
Steven Gutman, M.D., Food and Drug Administration*
Sam Hanash, M.D., Ph.D., Fred Hutchinson Cancer Research Center‡

James Heath, Ph.D., California Institute of Technology*
Roy Herbst, M.D., Ph.D., MD Anderson Cancer Center*†
Hedvig Hricak, M.D., Ph.D., Memorial Sloan-Kettering Cancer Center*†
David Jacobson-Kram, Ph.D., Food and Drug Administration*
Susan Jerian, M.D., OncoRD, Inc.*
Mark Kris, M.D., Memorial Sloan-Kettering Cancer Center*
Raju Kucherlapati, Ph.D., Harvard Partners Center for Genetics and Genomics*
Steven Larson, M.D., Memorial Sloan-Kettering Cancer Center*
Jack Lee, Ph.D., MD Anderson Cancer Center*
Pierre Massion, M.D., Vanderbilt University*†
Tim McCarthy, Ph.D., Pfizer, Inc.*‡
John Mendelsohn, M.D., MD Anderson Cancer Center*†
George Mills, M.D., Parexel International Corporation*†
Hal Moses, M.D., Vanderbilt-Ingram Cancer Center*
William Pao, M.D., Ph.D., Memorial Sloan-Kettering Cancer Center*
David Piwnica-Worms, M.D., Ph.D., Washington University School of Medicine*†
Richard Schilsky, M.D., University of Chicago‡
Kevin Schulman, M.D., Duke University Medical School*
Lawrence Schwartz, M.D., Memorial Sloan-Kettering Cancer Center*
Steven Shak, M.D., Genomic Health, Inc.*
Joseph Sparano, M.D., Albert Einstein Comprehensive Cancer Center*
Ellen Stovall, National Coalition for Cancer Survivorship*
Daniel Sullivan, M.D., Duke University*
Daniel Von Hoff, M.D., Translational Genomics Research Institute*
John Wagner, M.D., Ph.D., Merck Research Laboratories†
Janet Woodcock, M.D., Food and Drug Administration *†