Genome-Based Diagnostics: Demonstrating Clinical Utility in Oncology: Workshop Summary

ISBN
978-0-309-26959-9

120 pages
6 x 9
PAPERBACK (2013)

Adam C. Berger and Steve Olson, Rapporteurs; Roundtable on Translating Genomic-Based Research for Health; Board on Health Sciences Policy; Institute of Medicine; Center for Medical Technology Policy

GENOME-BASED DIAGNOSTICS

Demonstrating Clinical Utility in Oncology

WORKSHOP SUMMARY

Adam C. Berger and Steve Olson, *Rapporteurs*

Roundtable on Translating Genomic-Based Research for Health

Board on Health Sciences Policy

INSTITUTE OF MEDICINE
OF THE NATIONAL ACADEMIES

and

Center for Medical Technology Policy

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

THE NATIONAL ACADEMIES PRESS 500 Fifth Street, NW Washington, DC 20001

NOTICE: The workshop that is the subject of this workshop summary 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 project was supported by contracts between the National Academy of Sciences and the American Academy of Nursing (unnumbered contract); American College of Medical Genetics and Genomics (unnumbered contract); American Heart Association (unnumbered contract); American Medical Association (unnumbered contract); American Society of Human Genetics (unnumbered contract); Association for Molecular Pathology (unnumbered contract); Blue Cross and Blue Shield Association (unnumbered contract); Centers for Disease Control and Prevention (Contract No. 200-2011-38807); College of American Pathologists (unnumbered contract); Department of the Air Force (Contract No. FA7014-10-P-0072); Department of Veterans Affairs (Contract No. VA248-P-1528); Eli Lilly and Company (unnumbered contract); Genetic Alliance (unnumbered contract); Health Resources and Services Administration (Contract No. HHSH250201100119P and Contract No. HHSH25034017T); International Society for Cardiovascular Translational Research (unnumbered contract); Johnson & Johnson (unnumbered contract); The Kaiser Permanente Program Offices Community Benefit II at the East Bay Community Foundation (Contract No. 20121257); Life Technologies (unnumbered contract); National Cancer Institute (Contract No. N01-OD-4-2139, TO#189 and Contract No. HHSN263201200074I, TO#5); National Coalition for Health Professional Education in Genetics (unnumbered contract); National Heart, Lung, and Blood Institute (Contract No. N01-OD-4-2139, TO#275); National Human Genome Research Institute (Contract No. N01-OD-4-2139, TO#264 and Contract No. HHSN263201200074I, TO#5); National Institute of Mental Health (Contract No. N01-OD-4-2139, TO#275 and Contract No. HHSN263201200074I, TO#5); National Institute of Nursing Research (Contract No. HHSN263201200074I, TO#5); National Institute on Aging (Contract No. N01-OD-4-2139, TO#275 and Contract No. HHSN263201200074I, TO#5); National Society of Genetic Counselors (unnumbered contract); Northrop Grumman Health IT (unnumbered contract); Office of Rare Diseases Research (Contract No. N01-OD-4-2139, TO#275); Pfizer Inc. (unnumbered contract); and P*h*RMA (unnumbered contract). The views presented in this publication do not necessarily reflect the views of the organizations or agencies that provided support for the activity.

International Standard Book Number-13: 978-0-309-26959-9
International Standard Book Number-10: 0-309-26959-8

Additional copies of this workshop summary are available for sale from the National Academies Press, 500 Fifth Street, NW, Keck 360, Washington, DC 20001; (800) 624-6242 or (202) 334-3313; 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). 2013. *Genome-based diagnostics: Demonstrating clinical utility in oncology: 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. C. D. Mote, Jr., 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. C. D. Mote, Jr., are chair and vice chair, respectively, of the National Research Council.

www.national-academies.org

PLANNING COMMITTEE[1]

WYLIE BURKE (*Co-Chair*), Professor and Chair of the Department of Bioethics and Humanities, University of Washington, Seattle
ROBERT McCORMACK (*Co-Chair*), Head, Technology Innovation and Strategy, Veridex, LLC, Raritan, NJ
MARIAN BIRKELAND, Oncology Scientist, Molecular Medicine, National Comprehensive Cancer Network, Fort Washington, PA
PATRICIA DEVERKA, Senior Research Director, Center for Medical Technology Policy, Baltimore, MD
ANDREW FREEDMAN, Branch Chief, Clinical and Translational Epidemiology Branch, Division of Cancer Control and Population Sciences, National Cancer Institute, National Institutes of Health, Rockville, MD
USMAN IQBAL, Director, Oncology–Global Evidence and Value Development, Sanofi US, Cambridge, MA
MUIN KHOURY, Director, Office of Public Health Genomics, Centers for Disease Control and Prevention, Atlanta, GA
GARY LYMAN, Professor of Medicine, Director of Comparative Effectiveness and Outcomes Research–Oncology, Duke University School of Medicine and the Duke Comprehensive Cancer Center, Durham, NC
ELIZABETH MANSFIELD, Director of the Personalized Medicine Staff, Office of In Vitro Diagnostics and Radiological Health, Center for Devices and Radiological Health, U.S. Food and Drug Administration, Silver Spring, MD
ROBERT S. McDONOUGH, Head of Clinical Policy Research and Development, Aetna, Hartford, CT
MARGARET PIPER, Director of Genomics Resources, Blue Cross and Blue Shield Association Technology Evaluation Center, Atlanta, GA

IOM Staff

ADAM C. BERGER, Project Director
CLAIRE F. GIAMMARIA, Research Associate (*until July 2012*)
TONIA E. DICKERSON, Senior Program Assistant

[1] Institute of Medicine 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.

ROUNDTABLE ON TRANSLATING GENOMIC-BASED RESEARCH FOR HEALTH[1]

WYLIE BURKE (*Co-Chair until June 2013*), Professor and Chair, Department of Bioethics and Humanities, University of Washington, Seattle

GEOFFREY GINSBURG (*Co-Chair from June 2013*), Director, Center for Genomic Medicine, Institute for Genomic Sciences and Policy, Duke University, Durham, NC

SHARON TERRY (*Co-Chair*), President and Chief Executive Officer, Genetic Alliance, Washington, DC

NAOMI ARONSON, Executive Director, Technology Evaluation Center, Blue Cross and Blue Shield Association, Chicago, IL

EUAN ANGUS ASHLEY, Representative of the American Heart Association; Director, Center for Inherited Cardiovascular Disease, Stanford University School of Medicine, Palo Alto, CA

PAUL R. BILLINGS, Chief Medical Officer, Life Technologies, Carlsbad, CA

BRUCE BLUMBERG, Institutional Director of Graduate Medical Education, Northern California Kaiser Permanente, The Permanente Medical Group, Oakland, CA

DENISE E. BONDS, Medical Officer, Division of Prevention and Population Sciences, National Heart, Lung, and Blood Institute, Bethesda, MD

JOANN A. BOUGHMAN, Formerly Executive Vice President, American Society of Human Genetics, Bethesda, MD

PAMELA BRADLEY, Staff Fellow, Personalized Medicine Staff, Office of In Vitro Diagnostics and Radiological Health, Center for Devices and Radiological Health, U.S. Food and Drug Administration, Silver Spring, MD

PHILIP J. BROOKS, Health Scientist Administrator, Office of Rare Diseases Research, National Center for Advancing Translational Sciences, National Institutes of Health, Bethesda, MD

ANN CASHION, Acting Scientific Director, National Institute of Nursing Research, National Institutes of Health, Bethesda, MD

C. THOMAS CASKEY, Professor, Baylor College of Medicine, Houston, TX

SARA COPELAND, Former Acting Chief, Genetic Services Branch, Health Resources and Services Administration, Rockville, MD

[1] Institute of Medicine 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.

ROBERT B. DARNELL, President and Scientific Director, New York Genome Center; Investigator, Howard Hughes Medical Institute, Heilbrunn Cancer Professor and Senior Physician, Head, Laboratory of Molecular Neuro-Oncology, The Rockefeller University, New York, NY

MICHAEL J. DOUGHERTY, Director of Education, American Society of Human Genetics, Bethesda, MD

VICTOR DZAU, President and Chief Executive Officer, Duke University Health System; Chancellor for Health Affairs, Duke University, Durham, NC

W. GREGORY FEERO, Contributing Editor, *Journal of the American Medical Association*, Chicago, IL

ANDREW N. FREEDMAN, Branch Chief, Clinical and Translational Epidemiology Branch, Division of Cancer Control and Population Sciences, National Cancer Institute, Rockville, MD

JENNIFER L. HALL, Representative of the International Society for Cardiovascular Translational Research; Associate Professor of Medicine, University of Minnesota, Minneapolis

RICHARD J. HODES, Director, National Institute on Aging, Bethesda, MD

SHARON KARDIA, Professor and Chair of Epidemiology; Director, Public Health Genetics Program; Director, Life Science and Society Program; Codirector, Center for Public Health and Community Genomics, University of Michigan School of Public Health, Ann Arbor

MOHAMED KHAN, Representative of the American Medical Association; Leader of Radiation Oncology, Vancouver Cancer Centre, BC Cancer Agency, Vancouver, British Columbia, Canada

MUIN KHOURY, Director, National Office of Public Health Genomics, Centers for Disease Control and Prevention, Atlanta, GA

GABRIELA LAVEZZARI, Assistant Vice President, Scientific Affairs, P*h*RMA, Washington, DC

THOMAS LEHNER, Director, Office of Genomics Research Coordination, National Institute of Mental Health, Bethesda, MD

DEBRA G. B. LEONARD, Representative of the College of American Pathologists; Professor and Chair of Pathology and Physician Leader of Pathology and Laboratory Medicine, Fletcher Allen Health Care, University of Vermont College of Medicine, University of Vermont, Burlington

MICHELE A. LLOYD-PURYEAR, Representative of the Office of Rare Diseases Research; Former Senior Medical and Scientific Advisor, National Institute of Child Health & Human Development, Bethesda, MD

TERI A. MANOLIO, Director, Division of Genomic Medicine, National Human Genome Research Institute, Rockville, MD

ELIZABETH MANSFIELD, Director of the Personalized Medicine Staff, Office of In Vitro Diagnostics and Radiological Health, Center for Devices and Radiological Health, U.S. Food and Drug Administration, Silver Spring, MD

KATHRYN McLAUGHLIN, Program Officer, National Hemophilia Program, Genetic Services Branch, Health Resources and Services Administration, Rockville, MD

KELLY McVEARRY, Former Senior Scientific Advisor, Health Solutions, Northrop Grumman Information Systems, Rockville, MD

GARRY NEIL, Former Corporate Vice President, Corporate Office of Science and Technology, Johnson & Johnson, New Brunswick, NJ

ROBERT L. NUSSBAUM, Chief, Division of Medical Genetics, Department of Medicine and Institute of Human Genetics, University of California, San Francisco, School of Medicine

OLUFUNMILAYO F. OLOPADE, Walter L. Palmer Distinguished Service Professor of Medicine; Director, Center for Clinical Cancer Genetics; Associate Dean for Global Health, The University of Chicago, IL

MICHELLE A. PENNY, Senior Director, Translational Medicine Group, Eli Lilly and Company, Indianapolis, IN

AIDAN POWER, Vice President and Head PharmaTx Precision Medicine, Pfizer Inc., Groton, CT

VICTORIA M. PRATT, Representative of the Association for Molecular Pathology; Associate Professor of Clinical, Medical and Molecular Genetics and Director, Pharmacogenomics Diagnostic Laboratory, Department of Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis

RONALD PRZYGODZKI, Associate Director for Genomic Medicine and Acting Director of Biomedical Laboratory Research and Development, Department of Veterans Affairs, Washington, DC

ALLEN D. ROSES, President and Chief Operating Officer, Cabernet, Shiraz and Zinfandel Pharmaceuticals; and Jefferson–Pilot Professor of Neurobiology and Genetics, Professor of Medicine (Neurology); Director, Deane Drug Discovery Institute; Senior Scholar, Fuqua School of Business, R. David Thomas Executive Training Center, Duke University, Durham, NC

KEVIN A. SCHULMAN, Professor of Medicine and Business Administration; Director, Center for Clinical and Genetic Economics; Associate Director, Duke Clinical Research Institute, Duke University School of Medicine, Durham, NC

JOAN A. SCOTT, Formerly Executive Director, National Coalition for Health Professional Education in Genetics, Lutherville, MD

SAM SHEKAR, Chief Medical Officer, Health IT Program, Northrop Grumman Information Systems, McLean, VA

KATHERINE JOHANSEN TABER, Senior Scientist, Genetics and Molecular Medicine, American Medical Association, Chicago, IL
DAVID VEENSTRA, Professor, Pharmaceutical Outcomes Research and Policy Program, Department of Pharmacy, University of Washington, Seattle
MICHAEL S. WATSON, Executive Director, American College of Medical Genetics and Genomics, Bethesda, MD
DANIEL WATTENDORF, Deputy Chief, Medical Innovations, Department of the Air Force; Program Manager, Defense Advanced Research Projects Agency/Defense Sciences Office, Arlington, VA
CATHERINE A. WICKLUND, Past President, National Society of Genetic Counselors; Director, Graduate Program in Genetic Counseling; Associate Professor, Department of Obstetrics and Gynecology, Northwestern University, Chicago, IL
JANET K. WILLIAMS, Representative of the American Academy of Nursing; Professor of Nursing, University of Iowa College of Nursing, Iowa City

Fellows

SEAN P. DAVID, James C. Puffer, M.D./American Board of Family Medicine Fellow
SAMUEL G. JOHNSON, American Association of Colleges of Pharmacy/American College of Clinical Pharmacy Anniversary Fellow

IOM Staff

ADAM C. BERGER, Project Director
SARAH H. BEACHY, Associate Program Officer
CLAIRE F. GIAMMARIA, Research Associate (*until July 2012*)
TONIA E. DICKERSON, Senior Program Assistant

Board on Health Sciences Policy Staff

ANDREW M. POPE, Director
DONNA RANDALL, Administrative Assistant

Reviewers

This workshop summary 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 workshop summary as sound as possible and to ensure that the workshop summary 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 process. We wish to thank the following individuals for their review of this workshop summary:

Siraj Ali, Foundation Medicine, Inc.
Katrina Goddard, Center for Health Research, Kaiser Permanente Northwest
Lisa M. McShane, Biometric Research Branch, National Cancer Institute
Bruce Quinn, Foley Hoag LLP
David F. Ransohoff, University of North Carolina Schools of Medicine and Public Health

Although the reviewers listed above have provided many constructive comments and suggestions, they did not see the final draft of the workshop summary before its release. The review of this workshop summary was

overseen by **Dan G. Blazer,** Duke University Medical Center. Appointed by the Institute of Medicine, he was responsible for making certain that an independent examination of this workshop summary was carried out in accordance with institutional procedures and that all review comments were carefully considered. Responsibility for the final content of this workshop summary rests entirely with the rapporteurs and the institution.

Acknowledgments

The support of the sponsors of the Institute of Medicine Roundtable on Translating Genomic-Based Research for Health was crucial to the planning and conduct of the workshop Evidence for Clinical Utility of Molecular Diagnostics in Oncology and the development of the workshop summary report titled *Genome-Based Diagnostics: Demonstrating Clinical Utility in Oncology*. Federal sponsors are the Centers for Disease Control and Prevention; Department of the Air Force; Department of Veterans Affairs; Health Resources and Services Administration; National Cancer Institute; National Heart, Lung, and Blood Institute; National Human Genome Research Institute; National Institute of Mental Health; National Institute of Nursing Research; National Institute on Aging; and Office of Rare Diseases Research. Nonfederal sponsorship was provided by the American Academy of Nursing; American College of Medical Genetics and Genomics; American Heart Association; American Medical Association; American Society of Human Genetics; Association for Molecular Pathology; Blue Cross and Blue Shield Association; College of American Pathologists; Eli Lilly and Company; Genetic Alliance; International Society for Cardiovascular Translational Research; Johnson & Johnson; The Kaiser Permanente Program Offices Community Benefit II at the East Bay Community Foundation; Life Technologies; National Coalition for Health Professional Education in Genetics; National Society of Genetic Counselors; Northrop Grumman Health IT; Pfizer Inc.; and P*h*RMA.

The Roundtable would like to thank Sean Tunis and the Center for Medical Technology Policy staff for their support and partnership in bringing about the success of this workshop. The Roundtable would also like to

thank the members of the planning committee for their work in developing an outstanding workshop agenda. The Roundtable also wishes to express its gratitude to the expert speakers whose presentations helped outline the challenges in developing clinical utility data for molecular diagnostics. The project director would like to thank the project staff who worked diligently to develop both the workshop and the resulting summary.

Contents

Boxes, Figures, and Tables

BOXES

FIGURES

TABLES

Abbreviations and Acronyms

ASCO	American Society of Clinical Oncology
CALGB	Cancer and Leukemia Group B
CER	comparative-effectiveness research
CLIA	Clinical Laboratory Improvement Amendments
CMS	Centers for Medicare & Medicaid Services
CPT	current procedural terminology
EMR	electronic medical record
ER	estrogen receptor
FDA	U.S. Food and Drug Administration
IOM	Institute of Medicine
LDT	laboratory-developed test
NCCN	National Comprehensive Cancer Network
NCI	National Cancer Institute
NIH	National Institutes of Health
PCCTC	Prostate Cancer Clinical Trials Consortium
PSA	prostate specific antigen
QOPI	Quality Oncology Practice Initiative
RCT	randomized controlled trial

1

Introduction[1]

The sequencing of the human genome has greatly accelerated the process of linking specific genetic variants with disease. These findings have yielded a rapidly increasing number of molecular diagnostic tests designed to guide disease treatment and management. Many of these tests are aimed at determining the best treatments for specific forms of cancer, making oncology a valuable testing ground for the use of molecular diagnostic tests in medicine in general.

Nevertheless, many questions surround the clinical value of molecular diagnostic tests, and their acceptance by clinicians, payers, and patients has been unpredictable. A major limiting factor for the use of these tests has been the lack of clear evidence of clinical utility. Barriers to the generation of evidence of clinical utility include a lack of consensus among stakeholders regarding both the level of evidence needed to move a test into clinical practice and the acceptable methodologies to collect and validly demonstrate that evidence.

Capturing the benefits of molecular diagnostics will require stakeholders to help shape and define methodologies for efficiently generating reliable information about which tests will improve health outcomes for patients. Sustained dialogue among stakeholders is needed to help close the current

[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. Statements, recommendations, and opinions expressed are those of individual presenters and participants, and are not necessarily endorsed or verified by the Institute of Medicine, and they should not be construed as reflecting any group consensus.

evidence gaps and foster the development of clinically valuable tests that can inform both clinical and policy decision making.

PRIOR ROUNDTABLE WORKSHOPS

The Roundtable on Translating Genomic-Based Research for Health held two previous workshops examining barriers to the development and use of genomic-based diagnostic tests. At the first workshop, held in November 2010, it became clear, said Robert McCormack, head of technology, innovation, and strategy for Veridex LLC, that multiple stakeholders are involved and that their multiple needs are not always aligned (IOM, 2011a). To move forward, he said, these needs have to be combined and reconciled, which will require dialogue and coordination.

Participants at the 2010 workshop also emphasized the need for rules demonstrating that clinical utility has been achieved, both for future tests and for those in use today. Demonstrating clinical utility is a higher hurdle to meet for diagnostic tests than technical feasibility, analytic validation, or clinical validation, said McCormack. To overcome this hurdle, he argued, the concept of clinical utility needs to be better defined. In addition, the evidence generated and analyzed to demonstrate clinical utility needs to be adapted to the clinical setting, and clinical utility needs to be matched with the indication to make the task manageable.

Finally, McCormack said, the workshop demonstrated that the full picture is much bigger than most stakeholders imagine. The challenge of establishing clinical utility extends throughout the diagnostic discovery and development process, encompassing not only evidence development but also reimbursement and regulatory hurdles. "You cannot just fix one element," said McCormack. "It all has to come in line for the [system] to work."

The second workshop, held in November 2011, focused on this broadened landscape and emphasized the many changes needed to develop, regulate, and reimburse for genomic tests (IOM, 2012a). The involvement of providers is essential to creating short- and long-term change by enabling access and demonstrating value for genomic tests, McCormack said. At the same time, payers are the most important stakeholder in changing the overall climate for tests. The venture capital perspective was a "sobering" example at the workshop, according to McCormack. Venture investment in biotechnology has decreased precipitously over the last 10 years, and it is apparent, said McCormack, that in order for venture capitalists to continue investing their dollars in the United States, they will "need to see a seamless path from FDA [U.S. Food and Drug Administration] to CMS [Centers for Medicare & Medicaid Services] to the private payer. Until they see that, they are taking their money offshore." In addition, many participants at the

workshop called for greater oversight by the FDA of genomic test development, particularly for laboratory-developed tests (LDTs), though this call was not unanimous.

THE CURRENT EFFORT

On May 24, 2012, the Roundtable on Translating Genomic-Based Research for Health and the Center for Medical Technology Policy cohosted a workshop in Washington, DC, to foster the identified need for further sustained dialogue between stakeholders regarding the clinical utility of molecular diagnostics. Titled Evidence for Clinical Utility of Molecular Diagnostics in Oncology, the workshop brought together a wide range of stakeholders, including patients, health care providers, policy makers, payers, diagnostic test developers, researchers, and guideline developers, to identify the challenges and opportunities in advancing the development and use of molecular diagnostic tests designed to guide the treatment and management of patients with cancer.[2] Box 1-1 lists the goals of the workshop.

BOX 1-1
Workshop Goals

- Assess the evidentiary requirements for clinical utility of molecular diagnostics used to guide treatment decisions for patients with cancer.
- Discuss methodologies, including innovative models, related to demonstrating these evidentiary requirements that meet the needs of all stakeholders.
- Consider innovative, sustainable research collaborations for generating evidence of clinical utility involving multiple stakeholders.

This summary document describes the presentations and rich discussions that occurred at the workshop. Chapter 2 establishes a framework for the discussion by examining the history and broad issues associated with the development and use of molecular tests in oncology, the challenges that requiring clinical utility evidence generation for molecular tests poses, the potential impact molecular diagnostics could have on medicine, and the difficulties in establishing evidence standards that satisfy all stakeholders. Chapter 3 presents the perspectives of five different stakeholder groups—clinical guideline developers, health care providers, payers, aca-

[2] The workshop agenda, speaker biographical sketches, full statement of task, and registered attendees can be found in Appendixes A-D.

demic health systems, and patients—demonstrating both the commonalities and differences in their positions. Chapter 4 looks at the tools that have been and are being developed for use in assessing clinical utility, including a recommended framework for evaluating omics-based tests, comparative-effectiveness research, randomized controlled clinical trials, and observational studies. This chapter also provides considerations for determining whether a test is useful from statistical and economic perspectives. Chapter 5 examines possible paths forward to apply those tools to the many challenging issues associated with the use of molecular tests in medicine. The role of establishing successful partnerships to overcome challenges encountered during test development and the importance of the availability of biospecimens and data are discussed. This final chapter also provides a summary of the proposals suggested by individual speakers to advance the development of clinical utility measures for molecular diagnostics.

2

Setting the Context

Important Points Emphasized by Individual Speakers

- Test development may be stifled if creative solutions for demonstrating utility in a timely manner are not developed.
- Sustained dialogue between stakeholders is necessary to understand their views on clinical utility.
- The ability to discover potentially beneficial markers far exceeds the ability to translate them for patient use.
- A critical factor constraining the development and use of molecular diagnostics is a clear link between their use and improved patient outcomes.
- Clear, consistent, and predictable evidentiary expectations are essential to move forward in designing studies of clinical utility.
- Stakeholders need to collectively determine the optimal balance between access to new technologies and the need for certainty about risks and benefits associated with their use.

Bringing diagnostic tests to market previously required developing evidence of technical feasibility, analytic validity, and clinical validity, said Robert McCormack, co-chair of the workshop. Decision makers now need

TABLE 2-1 Genomic Predictive Markers of Cancer Treatment Efficacy and Safety

Test/Markers	Drugs	Cancer Outcomes
	In Clinical Use	
HER2/neu	Trastuzumab, Pertuzumab	Breast Cancer—Recurrence/ Survival
Oncotype Dx	Use of Adjuvant Chemotherapy	Breast Cancer—Recurrence/ Survival
EGFR mutation	Erlotinib	Lung Cancer—Progression/ Survival
K-ras	Cetuximab, Panitumumab	Colorectal Cancer—Progression/ Survival
EML4-ALK mutation	Crizotinib	Lung Cancer—Progression/ Survival
BRAF V600E	Vemurafenib	Melanoma Cancer—Progression/Survival
BCR-ABL	Imatinib, Dasatinib, Nilotinib	CML—Response
C-Kit	Imatinib	GIST—Response/Recurrence/ Progression
TPMT	6-MP, 6-TG	ALL, AML—Toxicity
DPD	5-FU	Toxicity
UGT1A1	Irinotecan	Toxicity
	Emerging Evidence	
MSI Status	5-FU	Colorectal Cancer—Prognosis/ Recurrence/Survival
Mammaprint	Treatment Regimen	Breast Cancer—Recurrence/ Survival
K-ras Mutation	Anti-EGFR Therapy	Lung Cancer—Recurrence/ Survival
ERCC1	Cisplatin-Based Therapy	Lung Cancer—Recurrence/ Survival

NOTE: ALL, acute lymphoblastic leukemia; AML, acute myeloid leukemia; BCR-ABL, breakpoint cluster region-abelson; BRAF, v-raf murine sarcoma viral oncogene homolog B1; CML, chronic myeloid leukemia; DPD, dihydropyrimidine dehydrogenase; EGFR, epidermal growth factor receptor; EML4-ALK, echinoderm microtubule-associated protein-like 4-anaplastic lymphoma kinase; ERCC1, excision repair cross-complementing rodent repair deficiency, complementation group 1; FU, fluorouracil; GIST, gastrointestinal stromal tumor; HER2, human epidermal growth factor receptor 2; MP, mercaptopurine; MSI, microsatellite instability; TG, thioguanine; TPMT, thiopurine methyltransferase; UGT1A1, UDP glucuronosyltransferase 1 family, polypeptide A1.
SOURCE: Adapted from Andrew Freedman, workshop presentation, May 24, 2012.

the answers to four key questions in considering the use of molecular diagnostics in oncology.[1]

1. Does the genomic application provide correct information? This question addresses the *analytic validity* of the test by assessing whether an application is measuring what it is supposed to measure.
2. Is there a significant association between the results of the genomic application and the clinical phenotype? This question of *clinical validity* assesses whether a relationship exists between the results of a test and a condition affecting health.
3. Does the genomic application provide clinically significant information? This measure of *clinical utility* determines whether the information from the application leads to a clinical decision that improves outcomes, taking into account the benefits and harms associated with those actions.
4. Finally, does the genomic application lead to improved patient outcomes as compared with the alternatives? This is an additional measure of clinical utility that relies on comparisons of utility or added clinical value.

The addition of clinical utility questions has changed the traditional path and altered the processes involved in developing diagnostic tests.

An increasing number of genomic predictive markers are either in clinical use or are undergoing testing to answer at least the first two questions listed above (see Table 2-1). Nevertheless, clinical utility is still largely unknown for most genomic applications. Demonstrating the clinical value of these technologies may potentially reduce the waste of health care resources from inconsistent or unnecessary use of tests and increase the quality of care received. Still, the evidentiary requirements to demonstrate clinical utility for genome-based diagnostics remains unclear, said McCormack. "We are hopeful that we can come to a point where we can understand the level of evidence that is required to get us to the next level of a seamless pathway for introducing these [diagnostic tests] into patient use."

THE POTENTIAL AND THE PROBLEMS

According to Sean Tunis, director of the Center for Medical Technology Policy, which was a cohost of the workshop, molecular diagnostics

[1] Andrew Freedman, chief of the Epidemiology and Genomics Research Program's Clinical and Translational Epidemiology Branch at the National Cancer Institute, cited these four questions during his workshop presentation, which is summarized in Chapter 4.

could have a transformational impact on medicine. Though molecular diagnostics currently apply to only about 2 percent of the population, that number could eventually rise to 60 percent (Ferreira-Gonzalez et al., 2008). PricewaterhouseCoopers (2009) has estimated that the diagnostic and therapeutic segment of the personalized medicine market will be $42 billion by 2015, with a 10 percent annual growth rate.

The current reality belies this vast potential, however. As Teutsch et al. (2009) have written, "Of most concern, the number and quality of studies [of genetic tests] are limited. Test applications are being proposed and marketed based on descriptive evidence and pathophysiologic reasoning, often lacking well-designed clinical trials or observational studies to establish validity and utility but advocated by industry and patient interest groups" (p. 3). This is a "serious indictment," said Tunis, and contrasts strongly with the expectations for the field.

The Centers for Medicare & Medicaid Services' (CMS's) guidelines for the evaluation of diagnostics tests center on two questions. First, is the evidence adequate to determine whether the test provides accurate diagnostic information? Second, if the test changes accuracy, is the evidence adequate to determine how the changed accuracy affects health outcomes? The clear message is that diagnostic accuracy by itself is not enough, said Tunis. The important factor is whether test results lead to changes in practice that can be linked to improved health outcomes.

Establishing this linkage to improved patient outcomes requires clear, predictable, and consistent standards of evidence by which diagnostic technologies will be judged, stated Tunis. These standards, in turn, will dictate the infrastructure and partnerships that are needed, and they will be essential for investors and entrepreneurs to judge accurately the risk and potential returns on investment.

In the past, coverage decisions have not necessarily been based on clear evidentiary standards, Tunis said. In some cases, coverage decisions have been dictated by legal challenges. Private payers tend to follow the lead of CMS in making their coverage decisions. For the field to move forward, Tunis said, a decision-making process among the payers needs to be more clearly tied to a clear and shared understanding of clinical utility.

FROM TUMOR MARKERS TO BIOMARKERS

Cancer diagnostics have progressed from an era of tumor markers to biomarkers, said McCormack. Tumor markers were helpful in making decisions, but they were also problematic. Many were based on serum tumor markers that were validated using sample banks that were not well pedigreed or well stored and were drawn from readily available patients.

The evidence generated for such markers "met the standard of the day," said McCormack, "but left a lot to be desired."

Today, far more biomarkers are available. From 1960 to 1989, fewer than 50,000 publications in the Library of Medicine mentioned biomarkers. In just the first decade of the 21st century, more than 250,000 did. During that period, medicine has moved from classifying cancers based on organ or tissue to classifying cancers based on pathway. Laboratory medicine also has evolved. An explosion of technology, especially at the genomic level, has resulted in procedures and results that require skilled specialists to acquire and analyze data. Yet the answers provided by those data "are not overly obvious," said McCormack, and the data are being interpreted by a generation of practitioners and researchers who were not necessarily schooled to fully understand that information. "It is apparent to everyone that our ability to discover potentially beneficial markers far exceeds our ability to translate them for patient use," he added.

A CHANGING LANDSCAPE

Highlighting how the need for clinical utility information has recently altered the development of molecular diagnostics, McCormack drew on his experience in having been involved in several major diagnostic development projects. Of four tests—the development of the prostate specific antigen (PSA) test in 1993, high-throughput hepatitis testing in 1997, testing for tumor cells in the blood in 2004, and a two-gene pathology lab test to detect cancer cells in women undergoing resection for primary breast cancer in 2007—only in the last case did regulatory approval require a demonstration of clinical utility, which was done in a postmarket context. In launching these tests, McCormack was repeatedly asked three questions by providers. First, has the test been validated? Second, is the test covered? And, third, how will this test change the way I practice medicine? "It was this last question that, of course, stumped me the most, because I did not have the clinical utility . . . they were looking for."

It was not until this most recent test that the clinical utility question drove a change in the development process, said McCormack. Veridex partnered with the Southwest Oncology Group to study the clinical utility of basing therapy on counts of circulating tumor cells 3 weeks after the start of therapy. The study opened in the fourth quarter of 2006, closed to enrollment in March 2012, and still has 2 years before outcomes can be determined. Nonetheless, noted McCormack, "it shouldn't take 6 years to [demonstrate the utility of a test]. We need to be creative to deliver the information that people want."

INHERENT TENSIONS

Tunis emphasized the inherent tension between the level of certainty about risks and benefits and innovation and early access to new technologies. The optimal balance to maximize long-term public health is not easy to determine and varies by stakeholder interest and perspective. Thus, said Tunis, "there is a critical need for stakeholders to come together and develop some common, shared understanding of what constitutes adequate evidence of clinical utility and a process for doing that."

Clear, consistent, and predictable evidentiary expectations are essential, said Tunis. "Until we translate that dialogue into specific methodological recommendations for designing studies of clinical utility, we can't really move the field forward in a meaningful and predictable way."

These methodological recommendations are as much a product of a social consensus as a scientific consensus, Tunis remarked. They reflect a collective social judgment about what is optimal, and they cannot be determined by specialists in a single field working in isolation. In balancing public policy objectives that compete with each other, the stakeholders need to arrive at a consensus that can be translated into scientific and technical terms.

3

Perspectives from Stakeholders

Important Points Emphasized by Individual Speakers

- The increasing number of molecular diagnostic tests calls for the integration of clinical practice guidelines into routine oncology practice.
- Incentives that focus on outcomes, rather than reimbursement, for diagnostics will enable providers to improve patient responses and reduce the high costs associated with cancer care.
- A provisional period of several years during which payers cover part of the costs of a molecular diagnostic's use would allow additional evidence to be gathered, after which the test could be accepted only if it produces substantial improvements in health outcomes.
- Situations may arise when a randomized controlled clinical trial is not needed for the approval, use, and reimbursement of a biomarker, but these situations must be chosen with great care.
- The term "clinical utility" does not mean much to most patients; a more relevant concept is "personal utility" or "personal guidance."

Multiple stakeholders are involved in the generation, analysis, and use of evidence for clinical utility of molecular diagnostics in oncology. Representatives of five stakeholder groups—guideline developers, health care providers, payers, academic health systems, and patients—offered perspectives on the challenges that need to be overcome to assess the value of molecular diagnostics.

CLINICAL GUIDELINE DEVELOPERS

Clinical guidelines are not prescriptions for care, said Al Benson, professor of medicine and associate director for clinical investigations at the Robert H. Lurie Comprehensive Cancer Center, Northwestern University. They are tools to inform decision making between the individual patient and clinician. In oncology, in particular, physicians are confronted with a growing list of biomarkers and treatment options across multiple diseases. As the number of biomarkers grows and understanding evolves, the integration of guidelines into routine oncology practice will be increasingly important.

Benson discussed three concepts related to the evaluation of medical technology, drawing on the analysis of Archie Cochrane (1972):

1. *Efficacy* is the extent to which an intervention does more good than harm under ideal circumstances—that is, in circumstances designed to maximize the effect of the intervention and eliminate confounding factors. Considerations of efficacy address the question, will it work?
2. *Effectiveness* is the extent to which an intervention does more good than harm when provided to real-world patients by physicians practicing in ordinary clinical settings. The relevant question here is, does it work in practice?
3. *Efficiency* measures the effect of an intervention in relation to the resources it consumes. In other words, is it worth it?

Benson also distinguished two principal types of evidence-based guidelines. The first category consists of integrated interventions over time, sometimes called a continuum-of-care approach. This category, which Benson covered in his presentation, includes the many hundreds of decision points reached in guiding treatment decisions. The second category consists of systematic reviews of single issues, which are described later in this chapter. The two are not mutually exclusive. Rather, the continuum-of-care guidelines help identify important areas for both systematic reviews and additional clinical research. Likewise, the results of systematic reviews of single issues can be integrated into continuum-of-care guidelines.

For a biomarker to progress to a clear clinical test, it should have significant and independent value and be validated by clinical testing, Benson said. Its use also needs to be feasible, reproducible, and widely available with quality control. Finally, use of the biomarker should benefit the patient. "Too often," said Benson, "tests are ordered without clear benefit or understanding of how the tests will be used to inform a decision-making process."

For an assay to have clinical utility, it must improve clinical decision making and patient outcomes. Measures of effectiveness include the probability of achieving a cure, the impact on survival, the impact on disease control, the impact on improving performance status, and the impact on disease-related symptom control. These outcomes often depend on the clinical situation, the availability of effective therapies, and the magnitude of the clinical benefit (or lack thereof) in one group versus another. They also depend on the relative values that patients, caregivers, and society place on the differences in the benefits and risks, including the benefits and risks that occur during continued surveillance of patients over time. These perceptions of benefits and risks can vary greatly and are often marked by a lack of understanding on the part of both patients and clinicians, Benson said.

Modern marker-generated clinical trial designs seek to answer complex questions regarding utility. For example, does the presence of a marker imply one kind of treatment while its absence implies another, which would provide potential clinical utility for that marker? Or does marker status make no difference on the effects of a given treatment, in which case the marker may not have clinical utility in that setting? A common situation, said Benson, is that a marker is prognostic and identifies risk but does not have a predictive correlation that an intervention will benefit the patient, in which case it has limited clinical utility.

Challenges in demonstrating clinical utility are already evident in oncology, noted Benson. As common cancers are broken down into smaller subsets, trials with smaller numbers of patients will become more common and evidence is likely to be limited. Decisions regarding evidentiary standards will become paramount in situations where large randomized controlled trials (RCTs) are no longer realistic. When dealing with a limited number of patients, the trials' researchers will have to rely on national databases to gain access to larger numbers of patients. An increase in patient numbers may result in the development of stronger evidence for an intervention, but will also lead to increased expenses for screening eligible populations. An increased reliance on tumor banks of appropriate tissues, whether metastatic or primary, will become greater as well, perhaps with serial biopsies and the use of multiple markers over time to deal with tumor heterogeneity.

TABLE 3-1 Disciplines Represented in National Comprehensive Cancer Network Clinical Practice Guidelines Panels

• Medical oncology	• Interventional radiology
• Surgery/surgical oncology	• Nursing
• Radiation oncology	• Cancer genetics
• Hematology/hematology oncology	• Psychiatry, psychology
• Bone marrow transplantation	• Pulmonary medicine
• Urology	• Pharmacology/pharmacy
• Neurology/neuro-oncology	• Infectious diseases
• Gynecologic oncology	• Allergy/immunology
• Otolaryngology	• Anesthesiology
• Orthopedics/orthopedic oncology	• Cardiology
• Pathology	• Geriatric medicine
• Dermatology	• Epidemiology
• Internal medicine	• Patient advocacy
• Gastroenterology	• Palliative, pain management
• Endocrinology	• Pastoral care
• Diagnostic radiology	• Oncology social work

SOURCE: Al Benson, workshop presentation, May 24, 2012. Derived from National Comprehensive Cancer Network (http://www.nccn.org/clinical.asp; accessed August 11, 2012).

Guideline Development in the National Comprehensive Cancer Network

The National Comprehensive Cancer Network (NCCN) seeks evidence-based consensus to allow for the development of comprehensive guidelines for treatment from prevention and screening to survivorship and hospice care, Benson stated. In many areas of treatment, high-level evidence exists, but in other areas gaps in evidence must be filled by expert consensus. Achieving this consensus requires the use of multidisciplinary panels representing a broad range of specialties (see Table 3-1). In examining the use of biomarkers, these panels evaluate the data demonstrating that the biomarker affects treatment decisions, the evidence that the biomarker can divide patients into clinically relevant subgroups, and the availability of reliable testing. They determine the levels of evidence using the results of tumor marker studies, taking into consideration whether the studies were prospective or retrospective, whether the studies used archived samples or were observational, and whether validation studies were available. The NCCN then classifies the test into one of three categories on the basis of the levels of evidence and consensus determination:

- Category 1: On the basis of high-level evidence, there is uniform NCCN consensus that the intervention is appropriate.
- Category 2A: On the basis of lower-level evidence, there is uniform NCCN consensus that the intervention is appropriate.

- Category 2B: On the basis of lower-level evidence, there is NCCN consensus that the intervention is appropriate.
- Category 3: On the basis of any level of evidence, there is a major NCCN disagreement that the intervention is appropriate.

"Unfortunately, in oncology the minority of decisions are based on a category 1 level of evidence," said Benson. "The majority of the guidelines represent category 2A, which is based on a lower level of evidence, but uniform consensus."

Another NCCN tool evaluates evidence derived from the use of archived tissues to determine the clinical validity of tumor markers. The tool considers such factors as the clinical trial design; the patients studied; specimen collection, processing, and archival; statistical design and analysis; and validation.

Using these and other tools, the NCCN can integrate markers into guidelines. For example, with colon cancer, KRAS or BRAF mutation testing has been integrated and linked with pathologic reviews so there is guidance about the testing and the methodology that is most appropriate. Similarly, with metastatic melanoma, recent targeted-therapy treatments have reached the category 1 level of evidence, with specific recommendation for use based on BRAF mutation testing.

The NCCN is also working on a biomarkers compendium. This collection is intended to ensure access to appropriate testing as recommended by NCCN guidelines. It seeks to identify the utility of a biomarker to screen, diagnose, monitor, or provide predictive or prognostic information. It is also meant to discriminate between clinically useful biomarkers and those that are not clinically indicated. More than 800 biomarkers are currently included in NCCN guidelines, and all will be integrated into the compendium, said Benson. The reference document will include the indication, molecular abnormality, test purpose, methodology, NCCN level of evidence, specimen types, and NCCN recommendation.

Today, people are making decisions on the basis of incomplete datasets, Benson said.[1] Many medical devices, not just molecular tests, enter the market with insufficient information. Still, it is an enormous challenge to test these devices adequately. Clinical trials become complicated as populations are continuously subdivided, adding such expenses as screening for marker positive and negative individuals and evaluating markers over time to judge whether tumor biology is changing. "That comes at enormous cost. Who is

[1] Although not discussed in this workshop, prior Roundtable workshops (IOM, 2011a, 2012a) have examined the significant challenges facing guideline development, including the inherent tension that exists between the need for greater certainty regarding benefits and risks and providing early access to innovative technologies.

going to pay for that?" asked Benson. No one group can cover this work. A major commitment of patients, insurers, government, public and private institutions, and clinicians will be needed to foster partnerships aimed at innovation and technology development, Benson concluded.

Development of Guidelines by the American Society of Clinical Oncology

Like Benson, Gary Lyman, professor of medicine and director of comparative effectiveness and outcomes research–oncology at the Duke University School of Medicine and the Duke Cancer Institute, observed that clinical practice guideline recommendations face particular challenges with molecular diagnostics. Many tests are already in existence; new tests and data are emerging rapidly; and analytic validity, clinical validity, and clinical utility all have to be established.

The American Society of Clinical Oncology (ASCO) has the goal of producing valid, reliable, and useful clinical practice guideline recommendations, Lyman said. In deciding whether to take on a guideline topic, ASCO asks several questions:

- Is the burden or the importance of the condition or intervention large enough to warrant guideline development?
- Is there uncertainty or controversy about the effectiveness or safety of available clinical strategies for the condition?
- Is there sufficient variation in practice in the management of a given condition or use of an intervention?
- Is there sufficient scientific evidence of good quality to allow guideline development?
- Is there potential for an impact on clinical decision making, clinical outcomes, or practice variation?

Once guideline development is initiated, ASCO bases its recommendations on exhaustive, systematic reviews overseen by a steering committee. Using well-defined inclusion and exclusion criteria for studies, it conducts quality appraisals and undertakes a formal data-abstraction process. It then places all the data before a guideline panel of content and methodology experts, patient representatives, and sometimes members of industry to generate its guidance. Draft guidance undergoes multiple internal and external reviews prior to being finalized. The recommendations are then disseminated through publication in the society's *Journal of Clinical Oncology* and the *Journal of Oncology Practice* and by various other means. Recommendations do not go out for public review before publication.

This process meets most of the recommendations for the development of clinical practice guidelines established by the Institute of Medicine (IOM)

(2011b,c), said Lyman. The process is transparent, conflicts of interest are disclosed and managed, expert panels are multidisciplinary, reviews are rigorous and systematic, the format for recommendations is standardized and clear, and external review takes place. The one major area where the process falls short of the IOM's recommendations is in the development of a formal rating of the strength of the evidence and the strength of the recommendation, which is "a controversial area within the field," according to Lyman.

Biomarkers

Biomarkers pose particular challenges to the guideline development process, said Lyman. Biomarkers are complex, as are the data describing them. The types of prognostic and predictive biomarker studies that have been done vary widely, and most biomarker studies are retrospective rather than prospective. Demonstrating clinical validity or clinical utility becomes difficult. All these factors create major obstacles for developing and updating evidence-based guidelines for biomarkers.

To date, ASCO's recommendations around biomarkers have been limited, Lyman noted. The focus has tended to be on the analytic validity of a number of tests that are currently used in practice, such as HER2 testing and immunohistochemical testing of estrogen/progesterone receptors in breast cancer. Other tumor biomarkers have been discussed by ASCO, but not recommended because panels have concluded that their clinical validity or utility was insufficient, said Lyman.

A major challenge, said Lyman, is to learn within the evidence-based structure established by the IOM and ASCO to appraise and update oncology biomarkers to enhance their trustworthiness and impact on clinical practice. As an example of how impact on practice can be measured, Lyman briefly described the Quality Oncology Practice Initiative (QOPI) (Neuss et al., 2005). QOPI is a program offered by ASCO to its members for assisting with the evaluation of the quality of care that hematology-oncology practices provide their patients. By sharing limited data about more than 150 quality measures, QOPI can identify gaps in care and the resources needed to improve practices. The ASCO panel that develops the guideline recommendations defines the quality measures which are then put into the QOPI library. Ultimately, many of these measures are incorporated into the QOPI measurement process, enabling practices at particular sites to be benchmarked against similar practices. Though currently voluntary, "there may come a time when this type of process, or something like it, will be fairly mandatory for oncology practices," said Lyman.

ASCO is also developing a decision-support system to provide real-time, point-of-care data and understanding that can be used in clinical

decision making. Known as Cancer Linq, the system was being piloted at the time of the workshop. ASCO is also looking at quality measures, rapid systematic reviews for guidelines panels, and a point-of-care guide on regimen benefits, toxicities, and costs. The goal, said Lyman, is "to bring the membership and practicing oncologists real-time data—updated, current, yet properly validated and assessed by an expert group—as they care for their patients."

Additional barriers involve the lack of awareness, slow dissemination of new recommendations into clinical practice, inadequate access to the guidelines, reluctance to accept guidelines, and lack of accountability. "It's a work in progress as far as biomarkers are concerned," Lyman concluded. However, Lyman added that "we cannot afford not to demonstrate clinical utility." The biggest challenge will be setting the bar where there is agreement about sufficient demonstration of clinical utility, he said. Well-defined outcome measures need to be accepted, and then the magnitude of the impact on those outcomes needs to be set to justify the adoption of a test into guidelines or regulatory approval.

In 2005, CMS stated: "Clinicians armed with appropriate assessments and the best evidence-based practice guidelines can reduce some of the unpleasant and frequent side-effects that often accompany cancer and chemotherapy treatment, obtain the best possible clinical outcomes, and avoid unnecessary costs" (CMS, 2005). It is an optimistic vision, said Lyman, and it is a goal that everyone working in the field shares.

HEALTH CARE PROVIDERS

As Lloyd Everson, vice chairman and founder of The U.S. Oncology Network, said, all of the great technological advances currently under way create a more promising situation for patients now than in the past. But providers are struggling with what he termed a "wild west" environment. Which tests will make a difference in the clinical care of patients? How will the use of tests influence costs? When should tests be moved into practice? Moreover, as cancers are divided into ever smaller subsets, how can evidence be developed to make such decisions? The process of validating outcomes according to marker status is in an "embryonic stage," he said. And with venture capital fleeing the field, the development of molecular tests is likely to be hindered.

The situation is even more dire with cancer treatments, which have costs rising much faster than health care costs in general (see Figure 3-1). The tension between the ongoing technological explosion and constrained resources will not go away, said Everson. To address this tension, the provider community needs incentives that focus on outcomes, not on per-unit reimbursements.

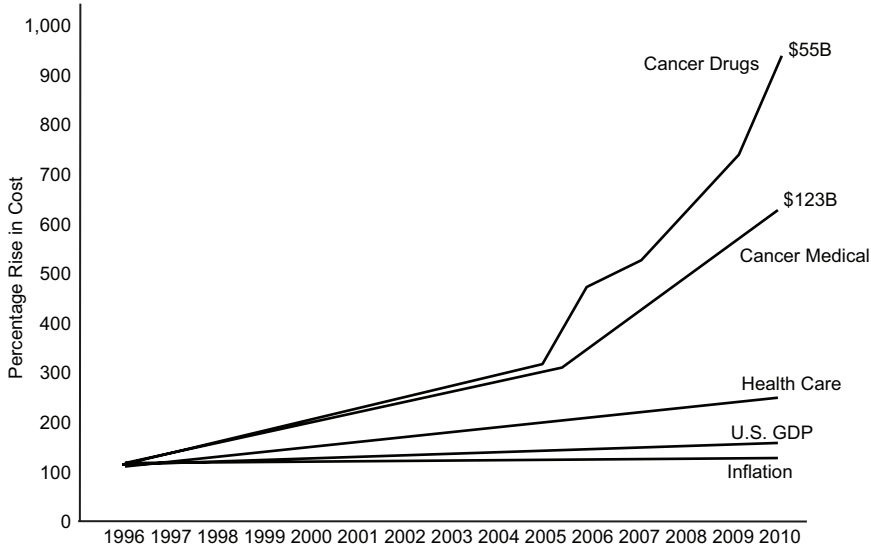

FIGURE 3-1 The costs of chemotherapy are rising faster than the costs of cancer medicine and health care in general.
NOTE: U.S. GDP, U.S. gross domestic product.
SOURCE: Lloyd Everson, workshop presentation, May 24, 2012.

Like other stakeholders in the system, The U.S. Oncology Network has embraced the evidence-based approach to treatment. The network encompasses more than 1,000 affiliated physicians and almost 2,000 affiliated nurses at more than 350 sites of care. According to Everson, about 30 percent of all cancer cases in the United States come through an affiliated practice or treatment facility.

The U.S. Oncology Network's pathways approach is similar to the approaches taken by the NCCN and ASCO. It develops evidence-based treatment guidelines that provide a precise, clinically proven approach to cancer care. In particular, level 1 pathways support physicians in making treatment decisions to provide a consistent platform for delivering, documenting, and reporting high-quality, evidence-based care. The goals of the evidence reviews during the guidelines development process are to permit flexibility of choice, find the balance point that maximizes patient benefits but maintains accountability for health care expenditures, ensures the ability of patients to participate in clinical trials, integrates cancer care with physicians' workloads, and remains current. Flexibility is particularly important, Everson emphasized. Because their cancers are uncommon or

because of comorbidities, 20 to 30 percent of patients do not fit into the pathways that have been established. As a result, physicians often have to be flexible in interpreting pathways in a clinical context.

Analyses have shown that the pathways can save money without adversely affecting outcomes. According to Everson, level 1 pathways can reduce variation in patient care, improve the predictability of costs for health plans, offer up-to-date clinical tools for documentation and reporting, prepare oncologists to succeed in pay-for-performance relationships, and demonstrate fiscal responsibility to patients and payers. In particular, the patient perspective is critical because this level is where change has to happen.

The approach developed by The U.S. Oncology Network has met with resistance in the past, said Everson, but it is now being embraced by physicians. They see the benefits it brings to their patients in terms of better outcomes and less toxicity. "The evidence-based approach is something that can work," concluded Everson. "It all depends on whether or not you can demonstrate in these smaller and smaller subsets of patients clinical utility. We have an enormous challenge, but if we can't do it, I don't know who is going to do it."

PAYERS

If health care costs continue to rise at the current rate, today's preschoolers will immediately have to earn the average U.S. salary when they graduate from high school just to pay their health care premiums, observed Lee Newcomer, senior vice president, oncology, at United HealthCare Corporation. At the same time, the mapping of the human genome has created phenomenal potential to better understand disease biology, target medicines to specific diseases, improve health, and advance the field of medicine dramatically. "The potential for the next decade is huge," Newcomer said. The problem, he added, is that "if you can't afford it, what difference does it make?"

Few people have come to terms with the unsustainable trajectory of rising health care costs. But unless new understandings from biomedical research dramatically improve the outcomes of care or markedly lower the cost, biomedical research will have relatively little impact, said Newcomer. The challenge, therefore, is learning how to pick from the genome the things that will make a difference.

A "Blue Sky" Proposal

Newcomer made what he called a "blue sky" proposal at the workshop wherein a new diagnostic or drug would need to lower the cost of care by

10 percent or improve an outcome by 10 percent to demonstrate its value. That bar is high, Newcomer noted, because most advances produce an improvement on the order of only 1 percent or so. To help facilitate the development of evidence to meet this goal, he proposed that a laboratory that has developed an analytically and clinically valid test could have the test covered by all payers for a 3-year period at a price that would cover some of the costs of using and continuing to develop the test. If the test achieves the 10 percent hurdle by the end of that 3-year period, it will be accepted. If it does not, it will not be accepted.

Newcomer also emphasized that the manufacturer would still need to provide and analyze the necessary data. The payers should not be trying to determine whether a test is useful. Payers could work with physicians, for example, to identify patients who have had particular responses, and the analysis could be conducted by a neutral third party, with protections for privacy. The manufacturer would work with that group to direct the study and bring out an unidentified or de-identified result.

This type of system would represent a major departure from current procedures. It would require that payers collaborate to offer provisional coverage, which would probably require an antitrust exemption. In the past, such collaboration has not been allowed, "but this may be a new world," said Newcomer. Also, the customers of the payers, most of which are self-funded businesses, would need to agree to such a system, because they would be the ultimate funders of such an approach. Finally, current health care legislation limits payers to using 15 percent of premium revenues for administrative costs,[2] and if the provisional funding were considered an administrative cost rather than a medical cost, it probably would not be a viable option.

These obstacles are substantial, said Newcomer, but they are all surmountable. And such a program would manage budgetary constraints while allowing biomedical advances to proceed. "We need to collaborate. We need to think about new models. I also think it is entirely possible," he said.

One of the reasons Newcomer made his proposal, he noted, is that "it's going to happen no matter what." Payers, providers, and patients are going to have to find the advances of highest value if health care is to continue to progress. "The more we can begin to find those things of highest value," he said, "the better off we will all be."

Newcomer concluded by stating that he is trying to bend the cost curve. "There are an awful lot of coded technologies whose value is quite uncertain, yet we pay for them," he said.

[2] 45 CFR Part 158.

The Evaluation of Evidence

Molecular diagnostics have raised particularly difficult challenges for Palmetto GBA, which administers Medicare health insurance for CMS, said Elaine Jeter, Palmetto's medical director. The coding system for diagnostics has been inadequate, and no process has been put in place to evaluate evidence. CMS requires that tests be "reasonable and necessary," which the agency defines as both being safe and effective and as demonstrating improvement in health outcomes. But assessing whether these standards have been met is difficult, and reimbursement issues are complex and contentious.

Jeter focused on the evaluation of evidence, which has been hindered by the inadequate coding system. The current procedural terminology (CPT) codes developed by the American Medical Association are insufficient, she said, and their descriptions are inadequate. Furthermore, there is no system to predetermine which tests qualify for payment. According to Medicare, a demonstration of clinical utility requires that a test or intervention improve patient outcomes by such measures as better functional status, improved quality of life, reduced disability, or changes in the physician's management of a patient. But published evidence for clinical utility or evidence-based decisions is lacking.

Few molecular assays are going through the FDA regulatory process, and thus many have not been evaluated for analytic or clinical validity. Most are LDTs for which no hard look has been taken at the science, Jeter said. Furthermore, many assays are so new that they have not undergone the reviews conducted by professional organizations such as ASCO or the NCCN. "I'm usually seeing [an assay] 3 years before it comes to any of the professional societies. I'm the one who is having to make a determination: Are we going to cover it or not?"

Deciding on Claims

To determine which claims to cover, Palmetto has created a system known as MolDx Solution, which it is implementing in CMS's J1 jurisdiction of California, Hawaii, and Nevada.[3] The system assigns a unique identifier known as either a Z code or a PTI (for Palmetto Test Identifier) code to a test. This code needs to be submitted on the claim in the comment narrative field; without this unique identifier, claims will be rejected. This is not a replacement for the CPT codes but a more specific way of identifying tests, Jeter said.

[3] More information about the MolDx system can be accessed at: http://www.palmettogba.com/palmetto/MolDX.nsf/DocsCatHome/MolDx (accessed August 11, 2012).

Once a code has been assigned, Palmetto requires a technical assessment by teams of experts in the subject matter; exceptions are made in cases where enough evidence already exists to make such an assessment unnecessary. Palmetto has identified approximately 50 of the 2,316 submitted applications that require technical assessments. These have either been completed already or are in the process of being done, noted Jeter. The Palmetto website has a short summary of the assessments. Once the assessment is completed, Palmetto makes a decision about coverage—to cover a test, cover it under certain situations, or not cover it at all. If an assay does not have evidence of clinical utility, the developer is notified in writing that the test will not be covered.

Some laboratories have resisted this system, saying that they have not had enough time to implement changes in their computer systems. As a result, implementation of the system was delayed until June 2012. The data in the system will be open to clinicians, patients, and the public, said Jeter, and coverage decisions will be published on the website.

ACADEMIC HEALTH SYSTEMS

Robert Bast, vice president for translational research and Harry Carothers Wiess Distinguished University Chair for Cancer Research at the University of Texas MD Anderson Cancer Center, took a different approach to the analysis of cancer biomarkers. He proceeded from insights into a particular disease to more general observations about where randomized controlled trials of clinical utility are needed.

About 22,280 new cases of epithelial ovarian cancer[4] occur annually in the United States, with about 15,500 deaths despite advances in surgery and chemotherapy. It is often diagnosed late, after it has spread throughout the abdominal cavity, often first seen as a pelvic mass that requires surgical removal. Surgery for ovarian cancer is complex and requires specific training.

Decades of experience indicate that even when ovarian cancer cannot be removed, prognosis is improved when residual metastases are decreased to less than 1 centimeter. Whether surgical expertise or biology is the most important factor in this observation is unknown, Bast said, and a pro-

[4] As of January 1, 2010, a reported 186,138 individuals were living with an ovarian cancer diagnosis (Howlader et al., 2013). The incidence rate for the disease is 12.5 per 100,000 women (ACS, 2012). Screening tests have limited accuracy for early detection and pelvic examination can generally only detect advanced ovarian cancer. Women at high risk may be referred for pelvic exam, transvaginal ultrasound, and testing for the tumor marker CA125. Treatment routinely includes surgery followed by chemotherapy. Bevacizumab and cediranib are currently being evaluated in clinical trials as targeted therapeutics for ovarian cancer treatment (ACS, 2012).

spective randomized trial of previously untreated patients is not feasible. Nevertheless, a retrospective meta-analysis of more than 50 nonrandomized studies involving almost 7,000 patients indicated that optimal versus nonoptimal cytoreduction is associated with 11 months of improved survival, which represents a 50 percent improvement (Bristow et al., 2002). For each 10 percent increase in cytoreduction, a 5.5 percent increase in survival results on average. Thus, Bast said, referral to gynecologic oncologists who are specifically trained in this kind of surgery improves outcomes for ovarian cancer patients.

Currently, however, only 30 to 50 percent of American women with ovarian cancer are referred to gynecologic oncologists. Those who are not referred tend to be poor, rural, and elderly, said Bast. The decision to refer is generally made not by oncologists but by general gynecologists, family practitioners, and internists (Goff et al., 2011).

Biomarkers of Malignancy

More than 200,000 exploratory operations for pelvic mass take place each year in the United States, and 13 to 22 percent of those lead to the diagnosis of cancer, said Bast. Biomarkers can help distinguish malignant from benign pelvic masses. A risk-of-malignancy index for ovarian cancer was developed in 1990 that incorporates the biomarker CA125, ultrasound, and menopausal status, providing a sensitivity of 71 to 88 percent and specificity of 74 to 97 percent (Jacobs et al., 1990). A more recently developed biomarker panel improves on CA125 and does not depend on ultrasound, which is observer specific, with better than 90 percent sensitivity and 75 percent specificity (Moore et al., 2010). A follow-up trial of this panel found that it had 100 percent sensitivity in premenopausal patients and had a negative predictive value of 99 percent (Moore et al., 2011a). This finding has prompted the referral within the last year of using the newer panel to triage patients, said Bast.

A second assay, developed by Vermillion, examines a panel of five biomarkers. It had better than 90 percent sensitivity, but specificity was 42 percent (Ueland et al., 2011). Although the difference in specificity should not affect patient outcomes—because a gynecologic oncologist can perform surgery on benign as well as malignant tumors—it could affect the distribution of medical resources. Neither of these is a screening test, Bast said, and should be used only for patients who are undergoing exploratory surgery. The real challenge, said Bast, is to encourage the use of either test.

Biomarkers can personalize the care of patients with epithelial ovarian cancer. When ovarian cancer is limited to the ovaries and has not metastasized, up to 90 percent of patients can be cured with the currently available chemotherapy and surgery, Bast said. Disease that has spread from

the pelvis is curable in less than 20 percent of patients. Currently, only a quarter of women with ovarian cancer are diagnosed in stage I. Detection of preclinical disease at an earlier stage could improve survival from 10 percent to 30 percent, predicted Bast.

Screening has stringent epidemiological requirements, however. The prevalence of the disease is 1 in 2,500 in the postmenopausal population, which is at greatest risk, requiring high sensitivity to detect early stage disease or, ideally, asymptomatic preclinical disease. But extraordinarily high specificity is also needed to avoid false positives—on the order of 99.6 percent to achieve a positive predictive value of 10 percent. In this context, these figures would imply 10 operations for each case of ovarian cancer detected.

Used alone, neither CA125 nor transvaginal ultrasound has adequate specificity. Ovarian cancer, however, is associated with rising CA125, whereas benign disease is not. Very high specificity and sensitivity can be attained when rising CA125 is used to trigger use of ultrasound in a two-stage strategy. A risk-of-ovarian-cancer algorithm was developed as a screening mechanism and uses each woman's own CA125 baseline to determine whether a significant increase has occurred (Skates, 2012). A randomized trial in the United Kingdom of 200,000 women that will conclude in 2015 has reported that 48 percent of cancers found by using this algorithm for screening were in stages I or II, doubling the detection of early stage disease, and up to 89 percent of all cancers were detected (Menon et al., 2009). Only about three operations were required per case of ovarian cancer when CA125 was followed by ultrasound, compared to 36 operations per case with an annual ultrasound alone. This study is consistent with earlier data indicating that ovarian cancers appear to develop 2 years before they are detected by conventional means, suggesting that annual screening might be effective (Menon et al., 2009).

With Karen Lu at MD Anderson and in collaboration with seven different sites, Bast has participated in a smaller trial evaluating the use of CA125 and transvaginal ultrasound in postmenopausal women at average risk for developing ovarian cancer. The study is powered to test the specificity and positive predictive value of the screen and to explore the feasibility of using this methodology to screen in the United States. During the past 10 years, it has obtained 15,000 samples from more than 4,000 postmenopausal women. Less than 1 percent have gone on to ultrasound over a single year, and less than 3 percent over multiple years. The risk-of-ovarian-cancer algorithm screen has prompted 10 operations to detect 6 cases of ovarian cancer—"a very small number," said Bast, "but encouraging." Two were borderline cases, and four were invasive, and all were in stages I or II. With a positive predictive value of 60 percent for all cancers and 40 percent for

invasive cancers, no more than three operations would be required to detect each case of ovarian cancer using this strategy.

Remaining Questions

One question raised by this research is whether multinational trials are acceptable forms of evidence for national decisions. "We're living in a global medical environment," said Bast, "but I'm not sure that there is complete comfort with that." In this case, a trial in the United States has shown the feasibility of the approach, but Bast questioned whether this was going to be considered adequate evidence.

Bast also raised the issue of LDTs, several of which have been applied previously to the detection of ovarian cancer. He concluded that FDA guidance needs to be applied to such tests. In particular, where significant risk is involved, the relevant question is whether LDTs should be held to the same standard as tests submitted for FDA approval.

Finally, Bast discussed current therapies for ovarian cancer. Most patients with ovarian cancer are treated routinely with both carboplatin and paclitaxel. However, only 70 percent of patients respond to platinum-based therapy, and 42 percent respond to paclitaxel as a single agent, and no synergy exists between the two (Muggia et al., 2000). Therefore, more than half of patients are treated with a drug that produces significant neurotoxicity without any obvious benefit. Biomarkers with high negative predictive value are clearly needed, Bast said.

Predicting responses to targeted therapy is an important issue for ovarian cancer, as it is for other diseases. Several potential candidate biomarkers exist, but given the potential toxicity and cost of treatment, a test with high negative predictive value would be very useful, Bast noted.

These treatment issues raise the more general question of whether accurate prediction of failure to respond to a toxic drug is adequate evidence of clinical utility, or are prospective, randomized controlled trials required to validate biomarkers or panels of biomarkers? Is a 90 percent negative predictive value an adequate benchmark? And is there a place for a test with positive predictive value? Would such a test simply have to demonstrate statistical significance, or would clinical utility be necessary? While some convergence has occurred around stakeholder evidentiary requirements (IOM, 2011a, 2012a), general agreement has not been reached on the evidence needed to garner approval, use, and reimbursement for tests.

In general, Bast concluded, situations may arise where prospective RCTs for utility are not needed to approve a biomarker for widespread use and reimbursement. But such situations need to be chosen carefully, he said, and early detection is not one of those situations.

PATIENTS

The field of molecular diagnostics has had some successes but has also generated many issues for patients, said Deborah Collyar, president of Patient Advocates in Research. Our research environment is not set up for the collaborative efforts that are critical to moving forward in the molecular age. False positives and false negatives can have serious effects on individuals. High costs are also a substantial problem and limit patient support for cancer research, she noted. Collyar said, "I regularly get the question, 'Why should we support cancer research if they are just going to produce drugs that nobody can afford?'" Finally, little information is available about the clinical utility of molecular tests even after a decade or more of use (Sparano and Solin, 2010).

Patients and people in general want true prevention of disease, but not at all costs, Collyar said. They want to reduce their risk of cancer or recurrence of cancer, but they also want to maintain their lifestyle as much as possible. They want cures as the word is customarily used and resist discussions of improved 5-year survival rates. They want a safe system, and those at high risk of developing a disease want to lower that risk for themselves, she said.

The word "diagnostics" has a variety of connotations for people, including hope; fear; anticipation of costs; vulnerability; and potential loss of self, family, culture, community, or privacy, Collyar said. The term that most people associate with that word is "risk," however, though the medical and common individual definitions are quite different. Individuals put more emphasis on their absolute risk as opposed to the relative risk in a population. For example, tamoxifen may produce a 50 percent relative reduction in risk (Vogel et al., 2006), but the decline in absolute risk is only from 2.6 to 1.3 percent (Howlader et al., 2012) for a drug with serious side effects, which is one reason why women have not rushed to their doctors to secure the drug, said Collyar.

People's lives can change dramatically after they are given the results of a diagnostic test, said Collyar. Yet, discussions of risk are complicated by a general lack of knowledge by the populace about the "world of the sick." They may seek information from the Internet, get multiple opinions from physicians, rely on their gut instincts, or turn to their families for advice. They may have questions about procedures, pain, and suffering; what the test results mean; how to lower their risks; the implications for relatives; the options that remain open or are closed; their work; their family and social life; and protection against misuses of their personal information. "We have to have knowledge before we can actually create results for people," Collyar said.

Clinical Utility

What does "clinical utility" mean to the common individual? Collyar asked. Most often, it means reliability—is a test going to predict how someone will react to a treatment, and what are the ramifications from the decisions that are made on the basis of the information gained? Most tests focus on a single marker, but the human body is an integrated circuit with many pathways that interact. Tests of single markers therefore need to fit into a diagnostic whole that includes imaging, clinical examinations, exploratory surgery, and so on. In general, however, "clinical utility" does not mean much to most people. A more relevant concept is "personal utility" or "personal guidance."

From this perspective, people want to get test results in a reasonable time frame and have results explained in clear language, Collyar said. Health care providers should be comfortable interpreting the test information and conveying what result is most important. In addition, test results need to be updated quickly as the test or a person's condition changes, with medical personnel receiving adequate training to keep up to date with rapidly changing technologies. Patients also need to have the choice to receive test results because some people will want to know a result even if no intervention is available, but others will not. Also, if a test has implications for family members, people want counseling because of the immense consequences for families that can follow the results.

Patients and advocates want the health care and research worlds to honor the true meaning of "patient-centered," according to Collyar. "Nothing about us without us" is a key message—"we need to be involved in the dialogue." Also, when groups are asked, they are all interested in being included in research, Collyar added, and they need to be so they are not excluded from the benefits of research.

From the perspective of the public and of patients, failures in developing molecular diagnostics can waste time and money, erode trust, and cost lives. "We have to get this right," Collyar said.

4

Tools for Generating and Synthesizing Evidence

Important Points Emphasized by Individual Speakers

- Test development needs to be rigorous, using meaningful and well-designed studies, proper statistical analysis, independent external validation, and interdisciplinary expertise.
- The identification and dissemination of best practices for the entire pathway of test development ensures that everyone understands the test development process.
- Clinical utility needs to receive earlier and more intense focus, with more education about how to interpret the results.
- Appropriate control groups are important to determine whether a biomarker distinguishes a group that benefits from a treatment.
- Studies of clinical utility should be conducted in settings that are relevant to more real-world clinical decisions.
- Focusing more on value than on cost-effectiveness in assessments of molecular diagnostics will enable analyses to be descriptive in addition to prescriptive and will allow consideration of the full context of care.
- The collection of blood and tissue from every cancer patient, including those who die, could greatly advance research.

Five speakers covered diverse aspects of the development of methodologies and tools that, as the statement of task put it, are related to demonstrating the evidentiary requirements for clinical validity and clinical utility that meet the needs of all stakeholders. Discussions included guidelines for test development, the role of comparative-effectiveness research (CER) in demonstrating clinical utility, statistical techniques, cost-utility analyses, and innovation mechanisms in small companies. Common themes included the need for clearly defined standards for analyses, the importance of context in determining clinical utility, and the importance of access to well-documented biospecimens.

DEVELOPING OMICS TESTS

Debra Leonard, professor and vice chair in the Department of Pathology and Laboratory Medicine and director of the Clinical Laboratories at Weill Cornell Medical Center, summarized the findings of a recent Institute of Medicine (IOM) report titled *Evolution of Translational Omics: Lessons Learned and the Path Forward* (2012b). The report was written by an IOM committee in response to the development of gene-expression array tests at Duke University that were said to predict sensitivity to chemotherapeutic agents. Papers written about the tests suggested that they represented a major advance and would better direct cancer therapy. Clinical trials were initiated in 2007, with the tests being used to select which chemotherapeutic agent patients would receive.

A paper by Baggerly and Coombes (2009), however, pointed to numerous errors and inconsistencies in the data and stated that the results could not be reproduced. Following a 2010 letter from more than 30 bioinformaticians and statisticians to the National Cancer Institute (NCI) urging the suspension of the clinical trials and an investigation of the test and computational models by the NCI, the clinical trials were stopped. The NCI then asked the IOM to review the situation and provide guidance for the field.

The IOM committee was charged with recommending an evaluation process to determine when omics-based tests are fit for use in a clinical trial. It also was asked to apply these criteria to omics-based tests used in the three cancer clinical trials conducted by the Duke investigators and to recommend ways to ensure adherence to the developed framework.

A Recommended Framework

An omics test is defined as being composed of or derived from multiple molecular measurements and interpreted by a fully specified computational model to produce a clinically actionable result (IOM, 2012b). The test can assess genomics, transcriptomics, proteomics, epigenetics, and

so on. Characteristics of an omics test include the use of complex, high-dimensional datasets and interpretation by a computational model, with a high risk that the computational model will overfit the data. The term is not meant to apply to a single gene test or to noncomplex testing, though Leonard added that she believes the committee's findings should apply to the development of any test.

The committee developed a recommended framework for the evaluation of omics tests from discovery to clinical use (see Figure 4-1). The framework begins with a discovery phase in which a candidate test is developed on a training set of data. Then the computational model is fully defined and locked down. The testing method and computational model are subsequently confirmed on a separate set of specimens or a subset of samples from the discovery set that were not used for training. If the test is intended for clinical development and eventual use, the data, computer code, and metadata should be made available to the public. The candidate test should be defined precisely, including not only the molecular measurements and computational model but also the intended clinical use for the test. This is standard for FDA, Leonard explained, but academic investigators tend not to think about how a test will be used in the clinic.

After the discovery phase, the committee concluded, test validation should be done under approval from an institutional review board (Jacobs et al., 1990) and in consultation with the FDA. The clinical testing method should be defined along with the analytical validation or confirmation of the analytical performance characteristics of a test. This can be done in a looped process, Leonard said, in which the test is modified to achieve the desired analytical performance. The defined test method should then be used on a validation sample set, with the intended use, assay, computational procedures, interpretation criteria, and target patient population all defined. The sample set can come from the discovery phase if the samples were from patients. If the discovery phase was done on cell lines or samples that were not like the patient samples, however, then validation needs to be done on a patient sample set. If the sample set is annotated with patient treatments and outcomes, it needs to be blinded to those doing the testing.

Once the test is defined and locked down, validation cannot be done iteratively. Rather, if the test needs to be changed, it must then begin the validation phase anew.

The committee recommended that the test be discussed with the FDA prior to the validation studies to learn what the FDA would want to see for approval of the test. The test development and validation should be performed in a Clinical Laboratory Improvement Amendments (CLIA)-accredited clinical laboratory if the test is intended to direct patient management. The CLIA laboratory should design, optimize, validate, and

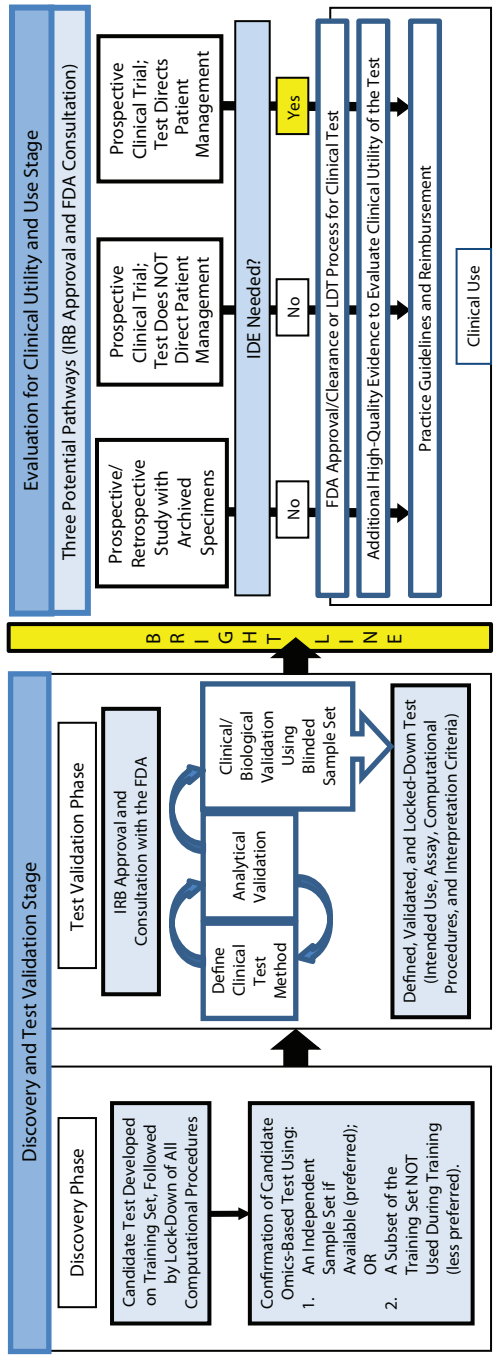

FIGURE 4-1 The recommended framework for the evaluation of omics tests extends from discovery to clinical use.
NOTE: FDA, U.S. Food and Drug Administration; IDE, investigational device exemption; IRB, institutional review board; LDT, laboratory-developed test.
SOURCE: IOM, 2012b.

implement the test under the current clinical laboratory standards, and CLIA requirements should be met by each laboratory in which the test will be performed for the clinical trial.

Pathways to Clinical Utility

During the final stage—which is separated by a bright line from discovery and validation—the test is evaluated for clinical utility and use. Clinical utility is not assessed by the FDA or in the LDT process, so the committee recommended that the process of gathering evidence to support clinical use begin before the test is introduced into clinical practice.

Three potential pathways are available for developing evidence of clinical utility, Leonard observed. Prospective or retrospective studies can be conducted using archived specimens from previously conducted clinical trials. Also, prospective clinical trials can be performed where either the test does or does not direct patient management. Whether test results direct patient management affects both the design of the prospective clinical trials and where the test is done. Regardless of the chosen method, the study or trial should receive approval from the institutional review board, and the FDA should be consulted. For investigators conducting a clinical trial in which the test will be used to manage patient care, the committee recommended that they communicate early with the FDA regarding the process and requirements of an investigational device exemption. Omics-based tests should not be changed during the clinical trial without a protocol amendment and discussion with the FDA. A substantive change to the test may require restarting the study, noted Leonard.

If supportive evidence is generated using these pathways, FDA approval or clearance can be sought or the test can be further developed as an LDT, said Leonard. Evidence can continue to be generated during this period to facilitate coverage and reimbursement discussions with payers and adoption into clinical practice guidelines.

Concluding Observations

The test development pathway is segmented, and different parts of it are done by different groups, especially in the academic environment, Leonard noted after describing the committee's recommendations. But oftentimes the groups do not fully comprehend the impact they have on each other. The IOM report, by describing the entire pathway of test development, defines best practices so that everyone can understand the interrelatedness of the test development process.

Unfortunately, the report does not look in depth at the barriers to

the recommended pathway, Leonard noted. In addition to the cost of the clinical trial, test validation and development is expensive, and how this expense will be covered is not clear. Also, there is a lack of availability and access to annotated specimens and datasets. The National Institutes of Health (NIH) does not routinely fund the maintenance and biobanking of specimens along with associated clinical data. Finally, there is no process for establishing whether a test will be covered or the level of payment that will be received for a test.

Leonard also suggested the idea of a clearinghouse that holds data gathered from various sources to determine whether a product has clinical utility. These data could be used both in guidelines development and in deciding whether to cover or not cover the clinical use of a molecular diagnostic.

THE ROLE OF COMPARATIVE-EFFECTIVENESS RESEARCH

One way to build the evidence base for decision making in cancer genomic medicine is through CER, said Andrew Freedman, chief of the Epidemiology and Genomics Research Program's Clinical and Translational Epidemiology Branch at the NCI. CER is intended to create evidence for decision making by finding out "what works" in health care. According to the IOM (2009, p. 13), "CER is the generation and synthesis of evidence that compares the benefits and harms of alternative methods to prevent, diagnose, treat, and monitor a clinical condition or to improve the delivery of care. The purpose of CER is to assist consumers, clinicians, purchasers, and policymakers to make informed decisions that will improve health care at both the individual and population levels."

One problem with CER is that it typically focuses on average treatment effects. Some treatments may have a significant effect across a broad population but may in reality only benefit particular patients and not others. Alternatively, a treatment may not reach significance and be considered ineffective when looked at across a large population, but in fact work for a certain subpopulation. The failure to recognize the heterogeneity of treatment effects can undermine the interpretation of the clinical trial results and the generalizability of those findings to patient populations. In cancer genomics, the goal is to figure out, for patients with a similar diagnosis, what tumor markers or genomic markers predict who will respond to treatment, who will not respond to treatment, and who will have adverse effects.

The methods used in CER are typically the same as in traditional genomic studies (see Table 4-1), but CER has a different orientation in that it encourages the stakeholders—including patients, clinicians, and payers—to prioritize research to help deliver the answers they need. The stakeholders identify the questions that would generate the needed evidence, and a

TABLE 4-1 Comparative-Effectiveness Research Versus Traditional Studies of Genomic Tests for Cancer

Feature of Research	Comparative-Effectiveness Research	Traditional Studies
Priority of study among alternatives	Determined by multiple stakeholders, using criteria such as disease burden or cost, lack of information, variability in care	Opportunity as dictated by expert assessment of emerging technology
Study design	Retrospective or prospective analysis	Retrospective analysis of existing tumor specimens; occasional prospective analysis of observational data
Comparisons	Direct comparisons of new therapy with usual care	Direct comparisons of competing therapies, often not considering usual care
Topics	Prevention, treatment, monitoring, and other broad topics	In most cases, prediction of narrow effects such as serious drug interactions, response to treatment, tumor recurrence
Perspectives	Multiple, including clinician, patient, purchaser, and policy maker	Clinician and patient
Study populations and settings	Representative of clinical practice	Highly selected
Data elements	Patient characteristics, quality of life, safety of treatment, resource use and costs, patients' preferences	Patient characteristics, clinical end points
Funding	"Coverage with evidence development" programs, public–private partnerships	Private investors, research grants from federal sources such as the National Institutes of Health

SOURCE: Ramsey et al., 2011.

synthesis of the evidence then informs clinical practice. CER emphasizes new therapies in usual care rather than in highly selected clinical trials, because the studies need to be relevant to clinical practice and not look just at clinical end points. Important considerations include quality of life, resources used, costs, and patient preferences, among other factors.

Types of Randomized Controlled Trials

In a randomized controlled trial (RCT), patients are randomized into two groups on the basis of either the treatment or the genomic test or marker. Freedman focused particularly on several varieties of RCTs as sources of evidence for both traditional genomic studies and CER. He divided RCTs into explanatory RCTs,[1] adaptive clinical trials, pragmatic clinical trials, and cluster randomized trials (Meyer, 2011). For example, the demonstration that HER2-positive patients benefited from treatment with Herceptin compared with others was the classic example in cancer genomics of an explanatory RCT (Smith, 2001).

Adaptive clinical trials are "learn-as-you-go trials" where the biomarker, the treatments, or both are changed as the trial progresses and more results become available, said Freedman. One or more decision points are built into the trial design for analysis of outcomes and associated patient or disease characteristics to identify subgroups that are responding favorably.

Pragmatic clinical trials—also called practical clinical trials, effectiveness trials, or large simple trials—are designed to help decision makers choose between options for care in routine clinical practice. These trials include a broad range of health outcomes, including morbidity, quality of life, symptom severity, and costs and are similar in many ways to explanatory trials. A purely pragmatic trial is not necessarily looking for regulatory approval of efficacy; rather, it is trying to figure out what works in the real world. It may not have an ideal experimental setting and is aimed more at normal practice. Generally, such trials have broader eligibility criteria that are not highly restrictive. An example is the ongoing 4,000-patient RxPONDER trial that is looking at the use of Oncotype DX to predict chemotherapeutic benefit for patients with estrogen receptor (ER)-positive/HER2-negative breast cancer where the cancer has been detected in one to three lymph nodes (Wong et al., 2012). Patients are being randomized to receive either chemotherapy and endocrine therapy or just endocrine therapy in order to identify the best cutoff point for use of the Oncotype DX recurrence score. It has elements of not only a traditional explanatory RCT but also a pragmatic trial. It convened a stakeholder group to identify

[1] Explanatory RCTs examine an intervention under ideal circumstances while pragmatic clinical trials assess interventions under real-world circumstances.

end points and the elements that would make the trial more relevant to current practice (Ramsey et al., 2013). The setting is reasonably representative of general clinical practice, as opposed to a more ideal clinical trial setting. A fairly simple intervention is being compared with usual care, and insurance firms are helping to pay for some of the tests. The outcomes are the cutoff point and disease-free survival, but quality of life, decision making, and cost-effectiveness are also being examined.

Cluster randomized controlled trials are a type of pragmatic trial in which social units or clusters rather than individuals are randomly allocated to use of a treatment. For example, such a trial can be randomized on clinical practices, where some practices deliver the intervention and some do not. Relatively few of these have occurred in oncology, especially with biomarkers.

Observational Trials Versus Randomized Controlled Trials

Observational studies are a valuable and complementary approach to generating evidence, said Freedman. These studies are nonrandomized and include retrospective analysis of biospecimens from RCTs, retrospective and prospective cohort designs, studies based on registries, and studies with case-control designs. The strength of the evidence increases in the progression from observational studies to pragmatic RCTs to explanatory RCTs. External validity generally gets stronger in the opposite direction, however. Trade-offs, therefore, have to be made in the decisions of which type of study to use.

Freedman proposed several sets of criteria that can be used to make these decisions. RCTs may be more suitable to determine comparative-effectiveness in genomic medicine in the following situations:

- When decisions require the highest level of certainty.
- When detecting small or modest differences in the results of treatment or testing.
- When ensuring high levels of internal validity (by controlling for selection bias, patient compliance, and other confounding factors).
- When accessible biospecimens are required for all participants.
- When detailed information on outcomes is needed.
- When genomic markers are incorporated in the design.
- When examining complex testing of multitherapy treatments.

Observational studies may be more suitable in the following situations:

- When study populations are not represented in RCTs, as with patients distinguished by age, comorbidities, or medications.

- When larger studies and diverse populations are needed, especially when looking at rare outcomes or the analysis of subgroups.
- When long-term follow-up is needed.
- When an RCT is not ethical or feasible.
- When testing or treatments are used off label.
- When comparing outcomes from multiple treatment regimens.
- When detecting larger differences in the results of treatment or testing.
- When confirming results from RCTs.
- When generating hypotheses to be tested in RCTs.
- When study results need to be generalizable.
- When study results are needed quickly.
- When treatment adherence differs.

An RCT is neither desirable nor feasible in every circumstance, said Freedman. High-quality observational study designs and evidence of underlying biological mechanisms can contribute to the evidentiary framework. For example, large prospective cohort studies with very large effects and compelling data can make a convincing argument for clinical utility even if the evidence was only generated with observational studies. The major concerns about using observational studies to inform clinical utility are that they can be poorly designed, their findings may be difficult to replicate, reliable outcome measures may be difficult to obtain, and they may be subject to bias and confounding through such factors as selection, response, adherence, attrition, or misclassification.

These limitations have not been lost on epidemiologists, and several groups are now working on guidelines to ensure high-quality studies that can be replicated. With bias and confounding, for example, new techniques such as instrumental variables or propensity score matching can adjust for the nonrandomization of subjects studied. Also, sensitivity analysis can assess the variability in results.

Considering the Context

Health policy decisions have to take into consideration the clinical context, the type of genomic application, the quality and availability of evidence to assess a marker's benefits and risks, and the risk to patients that a wrong decision could pose, Freedman said. Clinical utility is particularly difficult to determine because it encompasses the context in which the application is being used. Strong evidence is critical, but situations in which the benefits outweigh the risks often have to be weighed on a case-by-case basis.

A comprehensive approach is needed to resolve questions about the

clinical utility of genomic applications, said Freedman. He offered the following suggestions:

- Future research should consider more outcomes measures and be conducted in settings that are relevant to more real-world clinical decisions.
- A multitude of stakeholders having a role in evidence generation could lead to better studies.
- New strategies involving transformation of the research infrastructure to "learning systems" could allow continual additions to the knowledge base.
- Any changes to the evidentiary framework should uphold rigorous best-research practice standards.
- Collaborations among cancer centers are essential, particularly to investigate rare cancers.
- Clear priorities for CER could ensure that limited resources are used to resolve the most compelling questions.
- An evidentiary framework needs to articulate the minimal evidence necessary before clinical application is warranted.

DESIGNING STUDIES TO EVALUATE BIOMARKERS FOR CLINICAL APPLICATIONS

Lisa McShane, senior mathematical statistician in the Biometric Research Branch, Division of Cancer Treatment and Diagnosis, at the NCI, provided a statistician's perspective on the issues associated with the development of genomic-based tests. There are "widespread problems in the literature," she said, and everyone associated with the development of these tests needs to understand how to ensure the quality and interpretability of the studies evaluating them.

Molecular diagnostics have a range of potential roles in medicine (see Figure 4-2). McShane focused on two of those roles: as a prognostic indicator and as a predictive indicator. Prognostic indicators are molecular signatures associated with clinical outcomes in the absence of therapy or with a standard therapy that all patients are likely to receive. In the latter case, an untreated population generally will not exist.

Predictive indicators are molecular signatures associated with the benefit or lack of benefit—or potential harm—from a particular therapy relative to other available therapies. In the simplest case, a patient group with a particular biomarker may benefit from a therapy while a group without that biomarker may not benefit.

When is a prognostic test clinically useful? McShane asked. First, statistically significant does not mean clinically useful. "The literature is abso-

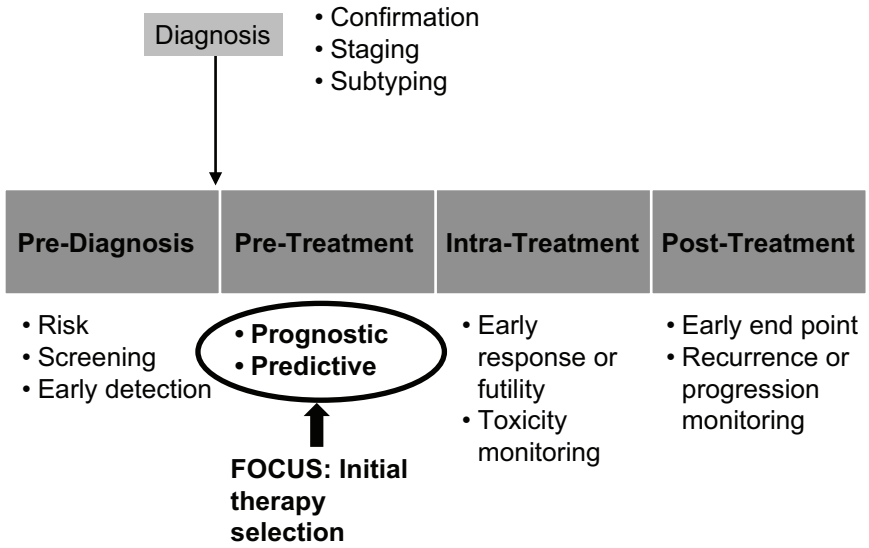

FIGURE 4-2 Potential roles for molecular diagnostics in medicine extend from prediagnosis to posttreatment.
SOURCE: Lisa McShane, workshop presentation, May 24, 2012.

lutely polluted with studies that show that a biomarker or a genomic test has some kind of prognostic information in it. But very, very few of those will ever make it into anything that is clinically useful," she said. If a prognostic marker can split a patient population into two groups, one of which has an outcome so much better than the other that different treatment decisions are made, then that might reach the necessary level of utility. An example is the Oncotype DX test, which can identify patients who have such good outcomes after surgery that they do not need chemotherapy. A test may also split a patient population into two distinguishable groups that both have bad outcomes, however. Unless the test can guide treatment or monitoring decisions, that information may not be clinically useful. Basing clinical decisions on results from such a test may rest on very little evidence.

Test developers also can confuse prognostic and predictive markers, said McShane. A marker may be prognostic of outcomes, but a new treatment may have the same effect in people with or without the marker. This dichotomy emphasizes the importance, said McShane, of appropriate control groups to determine whether a marker is predictive and distinguishes a group that benefits from a treatment.

To determine when a predictive test is clinically useful, a treatment-by-marker interaction can be assessed, but care needs to be taken in inter-

preting the results. One should not simply rely on statistically significant p-values, McShane said. For example, a new treatment may help only people with a particular biomarker, or it may help people regardless of biomarker status but to different extents. In the latter case, it might be beneficial to give everyone the new treatment, depending on such considerations as toxicity, cost, and patient preferences.

Prospective Versus Retrospective Studies

Prospective studies to establish the clinical utility of molecular tests are much easier to design than retrospective studies, McShane observed. They can use unbiased patient cohorts and adjust for standard variables, and several design options are available to answer specific questions (Freidlin et al., 2010; IOM, 2012b). The challenges with these studies tend to be their feasibility and cost.

Retrospective studies can provide a high level of evidence if they are performed properly. For example, a prospective-retrospective study can produce high levels of evidence for the value of a test if specimens were collected carefully in a clinical trial, if a sufficient number of representative specimens were collected, if the assay was analytically validated, if an analysis plan was prespecified, and if results were validated in one or more similar but separate studies (Simon et al., 2009).

Unfortunately, McShane said, many retrospective studies are poorly conducted. There may not be sufficient specimens available from completed trials or there might not be any suitable trials to address the question of interest. The retrospective study design may be flawed. Patient characteristics may be heterogeneous, and treatments may be unknown, nonrandomized, and not standardized. Specimens may be poorly characterized and accompanied by data of uncertain quality. And the studies may be subject to misinterpretation or deficient reporting. As an example of the problems that can arise, McShane quoted from the conclusion of an American Society of Clinical Oncology update of the recommendations for the use of tumor markers in breast cancer (Harris et al., 2007, pp. 5287–5289): "The primary literature is characterized by studies that included small patient numbers, that are retrospective, and that commonly perform multiple analyses until one reveals a statistically significant result. Furthermore, many tumor marker studies fail to include descriptions of how patients were treated or analyses of the marker in different treatment subgroups."

A major problem is that, as McShane said, "if you torture the data long enough, they will confess to anything." For example, multiple testing of many markers, many end points, many subgroups, and so on can produce false positives. Such testing is particularly problematic when there is no prespecified analysis plan and findings are selectively reported on the

basis of statistical significance. "People will do a zillion analyses, and what they will put into the paper are the ones that came up significant," she said.

Model overfitting is also a major problem, especially with high-dimensional marker data such as those generated by omics technologies. With overly complex models that have too many parameters or predictor variables, the model will describe random error or noise instead of the true underlying relationship. Similarly, if there are many more variables than independent subjects or if the data are sparse in a high-dimensional biomarker space, a model will generally have poor predictive performance on an independent dataset.

Model validation is essential, said McShane, yet many researchers make errors when attempting validations. In particular, as discussed by Leonard, they often use the same dataset to validate the model that they used to train it, even though such resubstitution estimates of model performance are useless (Subramanian and Simon, 2010). There are ways to do internal validation of models using data from the same sample set, but they require careful planning (e.g., Molinaro et al., 2005).

McShane also described several other common problems with evaluations of predictive tests. Randomized clinical trials adequately powered to detect treatment effects are often not sufficiently powered to establish predictive marker effects. For example, the nonsignificance of a treatment effect in a "marker negative" subgroup is often misinterpreted as no treatment effect, even though the test may not be adequately powered to exclude a treatment benefit. Also, sufficient information is sometimes not reported in studies to know whether an effect is meaningful. The p-value does not have much meaning without looking at an estimated effect size along with a measure of its uncertainty (e.g., a confidence interval).

Improving Assessments

McShane had several ideas about how to improve assessments of predictive tests. One is to place earlier and more intense focus on clinical utility, with more education about the proper interpretation. The test-development process and study design need to be rigorous, using meaningful and well-designed studies, proper statistical analysis, independent external validation, and interdisciplinary expertise.

A biomarker study registry, as suggested by Andre et al. (2011), could aid in identifying relevant biomarker studies for overviews and meta-analyses. It also could make study protocols available, including prespecified analysis plans, and help reduce nonpublication bias and selective reporting. Several sets of guidelines exist that could encourage more complete and transparent reporting (Altman et al., 2008, 2012; McShane et al., 2005; Moore et al., 2011b).

The Worldwide Innovative Networking consortium in personalized cancer medicine has provided seed funding to establish such a biomarker study registry. The aim is to create an entity similar to the website clinicaltrials.gov. Just as journals require that a clinical trial therapy protocol be registered with clinicaltrials.gov from the start, that would happen with biomarker studies. This would provide a placeholder for studies that do not result in publication. It also would allow the distribution of prespecified analysis plans, thus addressing not only nonpublication bias but also selective reporting in published papers.

Finally, expanded access to useful specimens, including alternative sources of specimens because trial specimens are optimal but limited, would be especially useful. Specimens should be well annotated with clinicopathologic data, treatments, and clinical outcomes. Health maintenance organizations and other large health care entities could be important partners in such efforts because people move around, which makes it difficult to piece together a patient's data for retrospective analyses.

McShane cautioned that we do not know the optimal way to collect a specimen and preserve it for every technology that might appear in the next several decades. Also, she has been involved with large biobanks of specimens that do not get used for a variety of reasons. But representative specimen collections from populations for which the tests might be used, along with data on treatment, outcome, the handling of specimens, and so on, are an important place to start. Adequate funding for collection and storage of annotated specimens from clinical trials and other carefully followed cohorts would also be very helpful, she added.

ASSESSING THE VALUE OF ONCOLOGY-BASED MOLECULAR DIAGNOSTICS

In theory, molecular diagnostics should save money or at least provide better care at lower cost. But some observers have called personalized medicine either a myth or hype, and others have declared it unaffordable, said Kathryn Phillips, professor of health economics and health services research at the University of California, San Francisco. Others have called more testing the medical equivalent of Moore's law in computing, claiming that testing causes more visits to the doctor, which results in exponentially more visits to the doctor. "There's a lot of debate regarding whether molecular diagnostics are really ever going to provide cost-effective care," she concluded.

Phillips conducted an informal study using data from the Tufts registry of cost-utility analyses to learn what has already been done on the

economics of molecular diagnostics for cancer.[2] More cost-utility analyses are available now than in the past, with about 14 percent focusing on cancer. The cost-effectiveness of diagnostics used to treat cancer is similar to that of other conditions. About half have a reasonable incremental cost-effectiveness ratio, but only about 10 percent of the analyses demonstrate that these interventions save money, Phillips said. Another 10 percent of the studies concluded that the interventions cost more and provide less health benefits than the standard of care, and these studies generally were done before the recent increase in high-cost diagnostics and cancer drugs. "In general, new health-care interventions do not save money," said Phillips. "They provide better health [benefits] at a reasonable cost." Furthermore, the cost-utility analyses of 64 molecular diagnostics for cancer were even less encouraging, with 20 percent costing more and resulting in less health benefits.

Challenges to Cost-Effectiveness Analyses

Phillips listed several well-known challenges to the use of cost-effectiveness analyses:

- The lack of data on effectiveness and costs.
- The need to consider the effect of the diagnostic on downstream decisions and outcomes.
- No or limited use of cost-effectiveness analyses by stakeholders.

She noted that diagnostics are complicated to analyze because of the evolving nature of the field, the complexity of the tests, the uncertainties surrounding them, and the nature of the disease. Also, because cancer can be an inherited condition, diagnostic tests may require consideration of family members. For example, a cost-effectiveness analysis of Lynch syndrome screening found that the screening is only cost-effective if family members change their behavior (Ladabaum et al., 2011). "And that's a big if," said Phillips. "If you are just looking at what happens to the proband, then you shouldn't be doing Lynch syndrome screening."

Whole-genome sequencing will be the next big dilemma, Phillips said. The complexity is far worse because of the huge amount of data available. Information may range from clinically actionable to not directly clinically actionable to unknown or no clinical significance, with various levels of risk and possible outcomes.

[2] The registry is available at https://research.tufts-nemc.org/cear4/Default.aspx (accessed August 10, 2012).

A Focus on Value

Phillips recommended that future analyses focus not only on cost-effectiveness but also on value. Cost-effectiveness analysis is "a hard sell," she said. Methodological concerns are common, such as defining and measuring quality-adjusted life years. Furthermore, in the United States, there is a lack of support for explicit consideration of cost.

Different frameworks for assessing value exist, said Phillips. Cost-effectiveness analysis can compare one alternative to another, but without the context. Alternatively, just the benefits and risks can be compared. Still, neither of these takes into account the larger impact of the technology on the health care system. For example, a technology may have a good cost-benefit ratio but have little impact because only a few individuals are being treated. A variety of methods can be used to capture the magnitude and scope of the technology being examined, said Phillips. Also, budget impact analysis can be used to assess whether a technology is affordable, for example, within a particular health plan.

As an example of a process for assessing value, Phillips laid out potential steps for conducting a multicriteria analysis:

1. Establish a decision context. What are the aims? Who are the decision makers?
2. Identify options.
3. Identify objectives and criteria that reflect the value associated with the consequences of each option.
4. Describe the expected performance of each option against the criteria and score options.
5. Assign weights for each criterion to reflect relative importance.
6. Combine weights and scores for each of the options to derive overall value.
7. Examine the results.
8. Conduct sensitivity analyses of the results.

She noted that this type of framework could be used to make systematic decisions about new health care interventions in a way that still captures costs and benefits but does not solely focus on these variables.

Phillips pointed out that cost-effectiveness analyses based on the ideal world may not adequately reflect actual implementation. Cost-effectiveness analyses are often normative—demonstrating what should be better—when they need to be descriptive, that is, taking into account the full context of care. For example, Elkin et al. (2011) demonstrated that few cost-effectiveness analyses of breast cancer diagnostics explicitly evaluated the relationships among the methods of targeting, the accuracy of the test, and

the outcomes of the intervention. The analyses tended to assume that the tests were perfect and did not consider the impact of test thresholds. As another example, Phillips (2008) found many real-world impacts on the cost-effectiveness of testing strategies. Data were missing on groups, especially the uninsured, Medicaid recipients, and minorities. Test results could be inaccurate, and some of the treatment courses did not correspond with what would be recommended by the test results. In addition, the claims and medical records for testing did not match 25 percent of the time.

Cost-effectiveness analyses are being applied to molecular diagnostics for cancer, Phillips concluded, but they raise methodological and political challenges. Focusing more on value than on cost-effectiveness will allow for headway to be made for molecular diagnostics but will also require changes in methods and public discourse. But shifting the focus to real-world analyses of value will allow decisions to be considered in the full context of care.

ADVANCING THE UTILITY OF ONCOLOGY DIAGNOSTICS

As Robert Bast did in the previous session, Noel Doheny, chief executive officer of Epigenomics, used a specific disease to make several general points about the tools available to assess the clinical utility of molecular diagnostics in oncology.

Colorectal cancer is the second largest cancer killer in the United States, causing 50,000 deaths and 140,000 new cases each year, or 15 to 20 deaths per 100,000 inhabitants (ACS, 2012). It is a disease of the developed world, with the highest prevalence in North America and Europe. Colorectal cancer is curable, though, if it is detected early enough. The 5-year survival rate for diagnosed and treated stage I or II colorectal cancer is 90 percent (ACS, 2011).

Colorectal cancer costs the U.S. health care system $17 billion per year—$7 billion in the initial year, $5 billion in the continuing-care years, and $5 billion in the last year of life (CDC, 2011). The key challenge to improving health and controlling costs, said Doheny, is changing current noncompliance with colonoscopy or stool-based screening. “If you [detect colorectal cancer] early, you never get to that last year where the costs go through the roof,” he said.

In the United States today, 100 million people are eligible for colorectal screening. Half of them get a colonoscopy; another 12 percent or so are tested by fecal tests; and the rest do not undergo any form of screening. Major impediments to screening are unpleasantness associated with stool tests; time constraints, risks, and fears associated with the colonoscopy preparation and anesthesia; and unreimbursed costs. Noncompliance is also greater among people who lack health insurance, have no other source of health care, and have not visited a doctor within the preceding year. In

addition, patients are often lulled into a false sense of security after one colonoscopy and fail to undergo subsequent testing.

Colonoscopy and fecal tests are not perfect. Colonoscopists differ in their ability to detect an adenoma, a polyp, or colorectal cancer in patients, and there are discrepancies in screening between the left and right side of the colon. Nevertheless, if everyone older than age 50 were screened regularly, as many as 60 percent of deaths from colorectal cancer could be prevented, suggested Doheny.

A Blood Test for Colorectal Cancer

The availability of a blood test could promote higher rates of screening by providing the ability to evaluate patients who would not otherwise have been screened, said Doheny. In a survey of more than 1,300 adults, 75 percent said they were more likely to get screened more frequently if a blood test were available.[3]

Epigenomics has been developing a blood test for colorectal screening based on a circulating marker called methylated Septin9. The test, which was going through the FDA's premarket approval process at the time of the workshop (Vogel et al., 2006), uses real-time polymerase chain reaction to detect free-circulating tumor DNA in blood. If one of three triplicate polymerase chain reaction tests is positive, the assay is considered to be positive, and patients are sent for a colonoscopy. The intended use of the test as filed in the premarket approval application is as a qualitative assay to aid in screening for patients with average risk.

During the test's technology development phase, Epigenomics conducted case-control studies using patients' samples that were both positive and negative for colorectal cancer to optimize the research protocol and increase the sensitivity and specificity of the test. It then gauged the performance of the test in two prospective studies. In one of the trials, of about 8,000 people, about two-thirds had no evidence of disease, and the test found 51 cancers among this group. A "very positive dialogue" with the FDA subsequently changed the number of patients considered to have no evidence of disease, which "clarified what we needed to do in this trial," said Doheny. At the time of the workshop, the company was getting ready to submit its clinical data later in 2012.

Epigenomics has been developing the test as an LDT, but on an interim basis. It has given licenses for the test to two laboratories, licenses that

[3] Results of the survey are available at http://www.prnewswire.com/news-releases/nearly-one-in-three-men-and-women-age-50-and-over-have-not-been-screened-for-colon-cancer-one-in-four-say-their-healthcare-provider-didnt-recommend-screening-124002559.html (accessed August 9, 2012).

will automatically end when the test receives regulatory approval and clearance. This procedure ensures that the laboratories switch to using the regulated product. About 1,000 tests per week are being done in the United States currently, which means that about five patients each week are being detected by a blood test as having colorectal cancer that would not have been detected if that test had not been done. "That, in our eyes, is a very positive position to be in," said Doheny.

Remaining Challenges

Doheny discussed some of the issues that "keep him awake at night." Imperfect standards are firmly entrenched, he said. Colonoscopies are not perfect, but they have become the gold standard. Also, incentives are misaligned for providers and payers, with current rewards disproportionately skewed toward chemotherapy.

Innovation is being driven by small companies in a difficult capital environment. Large companies tend to buy de-risked assets, and many small companies cannot afford the up-front costs of diagnostics development. Said Doheny, "It's very difficult to build an accurate, meaningful, and valid business model."

He also described several opportunities for improvement. Research partnerships are needed between pharmaceutical companies, diagnostic companies, and government. Although the incentives for these groups are different, they could be aligned if ways were found to combine complementary strengths, Doheny said.

Too few care approaches allow full cost clarity from first patient encounters to interventions and follow-up, Doheny observed. But several experiments, such as those being conducted by the Kaiser and Veterans Administration systems, could show a path forward.

He also suggested that the payment process could be linked through a visible mechanism to the regulatory process. In Japan, for example, when a company receives clearance on a project, a level of payment is established for each time the test is performed, which is an effective "closed loop mechanism," according to Doheny. Similarly, rewards could be differentiated on the basis of the regulatory credential of the offering.

A registry of patients with apparent false positives is needed to demonstrate clinical utility, he said. Also, a "platinum" standard to compensate for colonoscopy variability should be identified.

The collection of blood and tissue from patients with cancer, including those who die, could greatly advance research, Doheny said. He also recommended creating an accelerated review and publication format spe-

cifically for personalized medicine assays to overcome the extended and biased review cycles in traditional publications. Finally, real-world LDT performance should be linked to FDA filings, he said.

Doheny concluded by saying that "[waiting for] perfection takes too long. Why don't we just move ahead on some of these as appropriate?"

5

Advancing Molecular Diagnostics for Oncology

Important Points Emphasized by Individual Speakers

- A critical step to improve the generation of evidence for molecular diagnostics in oncology is to determine what unmet medical needs require prospective randomized trials to develop the evidence base.
- Other critical steps are to develop informative assays for use in research and practice and to overcome the segmentation within the provider community, including the divide between the medical benefit and the pharmacy benefit.
- The totality of evidence from interventional studies, observational studies, registries, and other sources needs to be combined to produce better outcomes for patients.

The final session of the workshop focused on pathways that could address both the opportunities and the challenges associated with the development of molecular diagnostics for oncology. Each of the five speakers emphasized that partnerships are an especially valuable way to accelerate evidence development. The availability of specimens and good quality clinical data, to name just one example, inevitably requires collaborative efforts among stakeholders. Also, collaborations are essential to overcome barriers imposed by costs, limited numbers of patients, and regulatory requirements.

Several speakers cited examples of successful collaborations, and all pointed toward the steps needed to replicate and extend such successes.

BIOMARKER STUDIES IN MULTICENTER CANCER CLINICAL TRIALS: THE ROLE OF COOPERATIVE GROUPS

As chair of the NCI-funded Cancer and Leukemia Group B (CALGB) for 15 years, Richard Schilsky, professor of medicine and chief of the Section of Hematology and Oncology at the University of Chicago, has had many experiences involving the kinds of partnerships that will be essential to accelerate the development and use of molecular diagnostics in oncology. He has participated in exploratory studies using clinically annotated biospecimens and research assays (commonly called correlative studies), prospective-retrospective studies using clinically annotated specimens with known clinical outcomes and using either research or analytically validated assays, prospective biomarker-drug codevelopment studies, prospective biomarker development studies, and prospective biomarker validation studies.

These studies require a large and expensive infrastructure, Schilsky said, and an advisory committee to help coordinate a range of activities involving many different disciplinary and interdisciplinary groups (see Figure 5-1). CALGB, like all of the NCI collaborative groups, has an NCI-funded biobank; it is located at The Ohio State University and is called the Pathology Coordinating Office. It also has a leukemia tissue bank that has collected frozen leukemia specimens and a lung cancer tissue bank that has collected frozen lung cancer specimens. All of the specimens, with the exception of those in the lung cancer bank, have been collected only from patients enrolled in clinical trials. "They were generally high-quality specimens collected in a uniform way from patients who met the eligibility criteria to participate in the study and for whom the outcomes were known," Schilsky said. In addition, CALGB established a number of reference laboratories that had specific analytical expertise. Finally, he pointed out that a collaboration with the Pharmacogenomics Research Network led to germline genotyping studies that were implemented in the group, which led to a collaboration with the Riken Institute in Japan that did much of the genotyping.

Examples of Biomarker Development

As examples of projects enabled by CALGB, Schilsky cited several exploratory biomarker studies. In an adjuvant chemotherapy study done in patients with node-positive colon cancer, treatment with irinotecan did not add any benefits to the previous standard of care (Saltz et al., 2007).

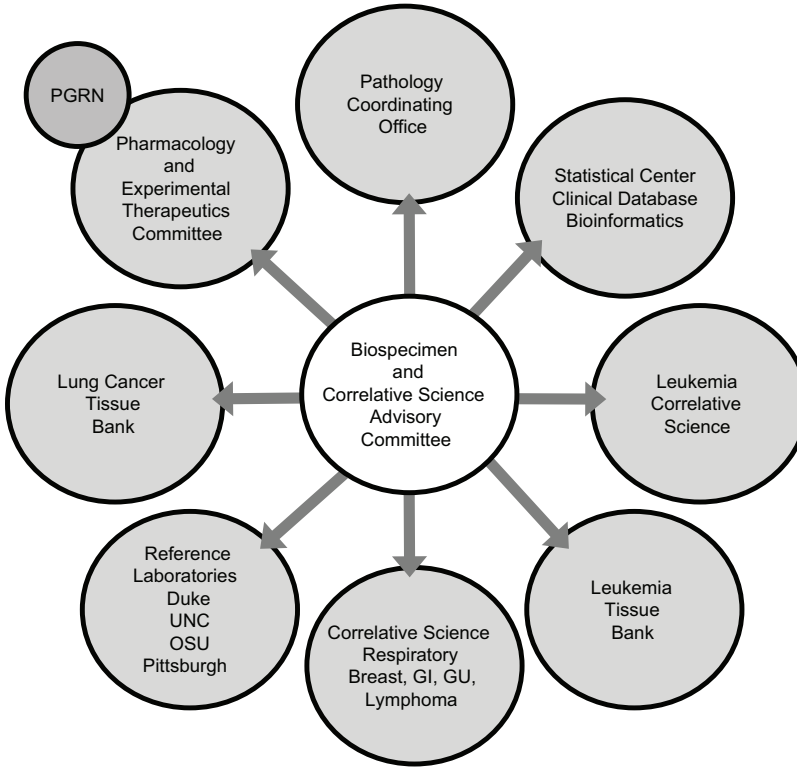

FIGURE 5-1 Many kinds of organizations interact in the translational science infrastructure.
NOTE: GI, gastrointestinal; GU, genitourinary; OSU, The Ohio State University; PGRN, pharmacogenomics research network; UNC, University of North Carolina at Chapel Hill.
SOURCE: Richard Schilsky, workshop presentation, May 24, 2012.

Primary tumors were collected from all 1,200 patients enrolled in the study, which afforded the ability to look at a variety of biomarkers rather than a treatment effect. This work found, for example, that KRAS mutation is not prognostic in stage III colon cancer, but BRAF mutation is prognostic of survival. A hypothesis emerging from the study is that irinotecan might be beneficial in patients with BRAF mutations, but that hypothesis needs further testing.

A breast cancer study done in the 1990s established that Taxol is a useful component of adjuvant chemotherapy in women with node-positive breast cancer (Henderson et al., 2003). Yet, as in many adjuvant studies,

the incremental benefit of Taxol is relatively small when examined over the entire trial population. Subsequent study of the tumors from trial participants revealed that the benefits were limited to women who had ER-negative and/or HER2-positive breast cancer (Hayes et al., 2007), thus targeting the therapeutic to those who will benefit while relieving those who would not benefit of deleterious side effects. The remaining biospecimens from this study are being used now to develop a taxane-sensitivity signature.

A randomized Phase II study of the drugs zileuton and celecoxib in non-small-cell lung cancer found that they did not produce better outcomes than standard chemotherapy (Edelman et al., 2008). But subsequent analysis of the biospecimens from the study participants suggested that celecoxib use in patients whose tumors express high levels of COX-2 might be beneficial. An ongoing prospective RCT is testing whether the use of celecoxib in this biomarker-selected population will produce a survival benefit.

As an example of biomarker-drug codevelopment, Schilsky cited a placebo-controlled, prospective randomized trial for patients with acute myeloid leukemia in which FLT3 is expressed at high levels to examine standard chemotherapy with the addition of a FLT3 inhibitor. This study was done in collaboration with Novartis and would have been impossible to do without the company, said Schilsky, because Novartis supplied the drug and had access to sufficient numbers of patients with the mutation to do the study. The study was done on three continents, Europe, North America, and South America, and involved eight reference laboratories in different regions of the world, all of which used the same reagents and procedures. Although the results from the trial were not available at the time of the workshop, the study is a good example of the strategies that may need to be used when doing assay-drug codevelopment.

As an example of a prospective-retrospective study, Schilsky described the use of specimens from a negative clinical trial of a monoclonal antibody therapy in early stage non-small-cell lung cancer to provide a validation of the Oncotype DX colon cancer test (O'Connell et al., 2010). The results, which were presented at an American Society of Clinical Oncology meeting, mirror the results developed by Genomic Health using other datasets.

With regard to prospective marker validation studies, Schilsky mentioned a "perfectly designed biomarker study that fell flat on its face." A collaboration among several cooperative groups sought to validate the utility of fluorescence in situ hybridization (FISH) testing to select patients to receive erlotinib as part of their therapy for non-small-cell lung cancer. The study was designed collaboratively with NCI and "essentially met the gold standard for the way a biomarker validation study would be designed," according to Schilsky. By the time the study got under way, however, the

lung cancer community had lost interest in the clinical question posed by the study.

Finally, Schilsky mentioned the TAILORx (Trial Assigning IndividuaLized Options for Treatment (Rx)) study, which was a collaborative effort across cooperative groups to validate the Oncotype DX test in breast cancer as a test that can be used to allow patients with an intermediate risk score to safely forgo receiving adjuvant chemotherapy. The study has completed accrual but does not have results yet.

General Observations

Cooperative groups have the capacity to conduct many types of biomarker studies, including formal validation trials, said Schilsky, but there are many challenges such collaborative efforts must overcome, such as the following:

- The adequacy of the biospecimen collection.
- Access to CLIA-certified laboratories that can conduct analytically validated assays in a reproducible way.
- Funding for biomarker studies, especially for large prospective studies.
- Regulatory requirements.
- Contractual agreements with commercial partners.

For some questions, according to Schilsky, stakeholders need to be willing to accept that less-than-gold-standard RCTs may need to serve as sufficient evidence to make regulatory, payment, and clinical decisions. Large clinical trials are not always necessary or possible because not enough resources, patients, time, and investigators are available to answer every question. Therefore, the most important step to improve the generation of evidence for molecular diagnostics in oncology, according to Schilsky, is to determine what unmet medical needs require prospective randomized trials to develop the evidence base.

Schilsky also noted, as did other people at the workshop, that most adult patients with cancer are not part of clinical trials. The primary determinant of whether a patient enrolls in a clinical trial is whether a physician recommends doing so. But in the United States, there are almost no incentives for physicians to recommend that patients participate in a trial, and there are many disincentives. Instead, physicians are likely to prescribe a drug off label. Countries that do not tolerate off-label prescribing are much more successful than the United States in enrolling patients in clinical trials. An interesting idea Schilsky mentioned is that of the "cancer information donor," where someone with a cancer diagnosis could volunteer to provide

information for cancer research even if that person is not participating in a clinical trial.

PARTNERING FOR THE CURE: AN INNOVATIVE ROLE FOR ACADEMIA IN ONCOLOGY DRUG AND DIAGNOSTIC DEVELOPMENT

As an example of an especially effective collaboration, Howard Scher, the D. Wayne Calloway Chair in Urologic Oncology and chief of the Genitourinary Oncology Service at Memorial Sloan-Kettering Cancer Center, described the Prostate Cancer Clinical Trials Consortium (PCCTC). It, too, has brought together a number of partners, with funding in part from the Department of Defense, to build the infrastructure to collaborate (see Figure 5-2). The mission of the collaboration is to design, implement, and complete hypothesis-driven Phase I and II trials of novel agents and combinations that could prolong the lives of patients with prostate cancer.[1] The 13 cancer centers involved in the collaboration each have scientific programs to support biomarker discovery and a translational clinical research enterprise.

The guiding principles of the collaboration are that centrally managed, harmonized, and comprehensive clinical trial processes will accelerate drug development and improve outcomes. This goal can be achieved by streamlining any process that can impede trial activation, conduct, completion, and analysis, Scher said.

A framework to conduct clinical trials was developed by consensus within the groups in order to synchronize clinical research with clinical practice (Scher et al., 2008). "The same way that a drug is focused on an indication," said Scher, "we're focusing on the context of use." Aligned to member-prescribed scientific priorities, teams of experts design trials in a sequence, each with "go–no go" metrics. Embedded in the collaboration is an extensive effort to discover and validate biomarkers analytically and clinically.

Since 2005, the PCCTC has submitted 152 letters of intent, with 118

[1] As of January 1, 2010, a reported 2,617,682 individuals were living with a prostate cancer diagnosis; 241,740 new cases and 28,170 deaths were reported in 2012 with an overall incidence of 152 per 100,000 men. Higher incidences have been found in African Americans than whites (228.5 versus 144.9 per 100,000 men) (ACS, 2012; Howlader et al., 2013). Digital rectal examination and prostate serum antigen (PSA) testing have been used for detection, though the U.S. Preventive Services Task Force recently recommended against use of PSA-based screening (USPSTF, 2012). Treatment options include active surveillance, surgery, external beam radiation, brachytherapy, hormonal therapy, chemotherapy, or a combination approach depending on disease advancement. Sipuleucel-T or Abiraterone may also be employed in cases where tumors are no longer responsive to traditional therapy (ACS, 2012).

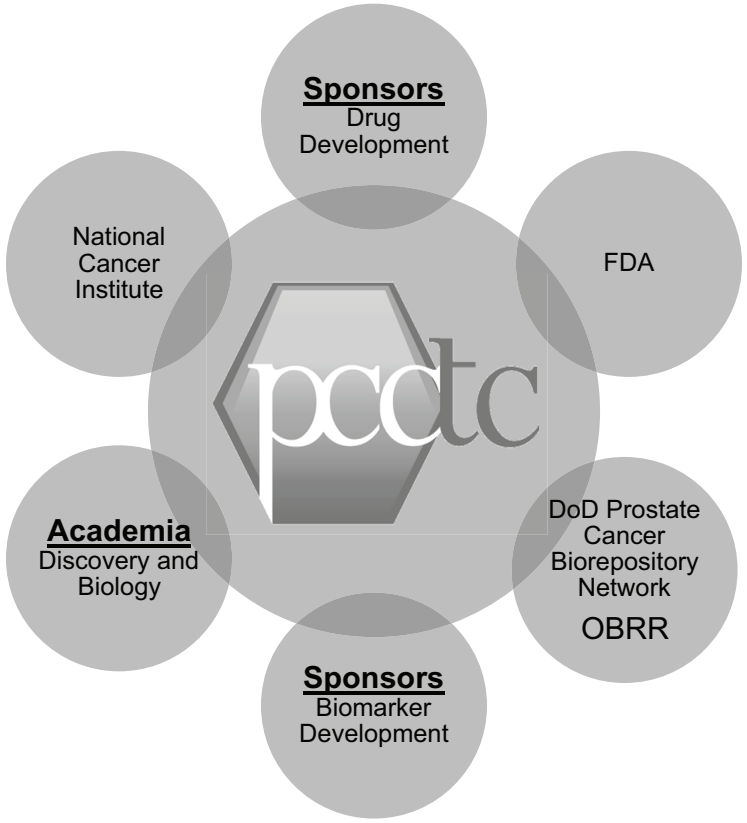

FIGURE 5-2 The Prostate Cancer Clinical Trials Consortium collaborates with critical stakeholders in drug and biomarker development.
NOTE: DoD, U.S. Department of Defense; FDA, U.S. Food and Drug Administration; OBRR, FDA Office of Blood Research and Review; PCCTC, Prostate Cancer Clinical Trials Consortium.
SOURCE: Howard Scher, workshop presentation, May 24, 2012.

protocols approved for activation. More than 3,200 men have been enrolled in Phase I and Phase II trials, and 8 therapeutic candidates have advanced to Phase III study.

Changing the Clinical Research Enterprise

In its first 6 years, the PCCTC has moved beyond its original charter in ways that have changed the clinical research enterprise, according to Scher. To illustrate this point, he focused on two new androgen axis inhibitors,

abiraterone and Medivation 3100. The development of these drugs was a great success, but that success also had a downside. Future trials would be more difficult because studies, in particular placebo-controlled studies, would need to be larger and longer; they would also be more costly because crossover to an effective treatment can confound a survival effect.

As a result, qualified surrogate biomarkers for survival are urgently needed that can be used for accelerated drug approvals. Also needed are qualified predictive biomarkers of sensitivity to better match drugs to an individual patient's tumor. "The era of 'all comers' trials will soon be ending," said Scher.

Both abiraterone and Medivation 3100 were studied in pre- and post-chemotherapy castration-resistant prostate cancer. Circulating tumor cell numbers were included as an end point, though before this could be used as a biomarker, the platform used to do the assay had to undergo analytical testing. This testing was designed to establish the minimum performance characteristics to justify the use of the assay in clinical testing and to achieve analytical validity across laboratories. Analysis of the first Phase III registration trial has led to positive results, Scher reported, "so we're in a very good position to look at both circulating tumor cells and other markers for their potential impact on survival." The development process led to FDA approval of abiraterone and submission of a New Drug Application for Medivation 3100 the same week as the workshop. Throughout this process, interactions with FDA "have been extremely favorable and extremely helpful," said Scher.

A rate-limiting factor has been the availability of analytically valid assays, Scher said. The collaboration has been looking at various putative predictive markers for patients who respond to treatment, do not respond to treatment, and develop resistance after treatment. Several markers have been postulated, but none has warranted testing in a large-scale trial. Meanwhile, the collaboration has been storing specimens for future analysis and has been working to develop assays of biomarkers that it thinks will be included in future panels.

Implementation of a Precision Medicine Paradigm

With a Stand Up To Cancer award, members of the collaboration are now pursuing precision therapy for advanced prostate cancer. The objective is to establish a "Rosetta stone" resource of mutation profiles of advanced prostate cancer for researchers and patients. This effort will establish advanced prostate cancer as a model tumor type for the precision medicine paradigm and facilitate the use of clinical sequencing for cancer management, Scher said. Specific goals include the following:

- Establishing the use of precision tumor boards to help guide the management of advanced prostate cancer.
- Identifying resistance mechanisms and sensitivity biomarkers for new prostate cancer therapies.
- Identifying rare "actionable" mutations in advanced prostate cancer and providing rational clinical trial options to patients.

The key to success, said Scher, will be the availability of analytically valid assays when the trials are ready to begin. When hypotheses about the contributors to a cancer cannot be explored because of a lack of an effective assay, "it's quite frustrating," he said. "You know what you want to do, but you can't." At the same time, it is essential that the assays provide correct information. "If we are not confident of the diagnostic, then what we may do clinically may actually harm patients, which is what no one wants to do," he added. Clinicians need to work closely with pathologists, he said, and pathologists need to be closely integrated into the development process.

Scher also said that less can be more with regard to data collection. Instead of trying to record everything, it may be better to capture the milestone events. "You want to get the key elements but not necessarily waste time on things that are not adding value to the patient, to the drug, or to the investigator," he noted.

Scher also said that physicians are busy and that the demands on their time are increasing. Asking them to provide information on patients may be too difficult unless they get something in return. If they get data that improves practice, they will use and support a system. The system needs to serve the provider rather than having the provider serve the system.

PATIENT APPROACHES TO COLLABORATION

Patient advocacy has many dimensions, said President of Patient Advocates in Research Deborah Collyar, who spoke for a second time in the final session of the workshop. Patient advocates do fundraising, political advocacy, direct patient support, watchdog advocacy, and research advocacy. "Many of us have done all of those different things," she said.

As a result, patient advocates tend to be involved in many different types of networks and partnerships, including networks of advocates. For example, patient advocates have been extensively involved in the cooperative groups described by Schilsky. As part of this work, they have helped to develop and design research concepts, protocols, consent forms, and results summaries. They also have been represented in advisory groups for biospecimen collections and correlative studies and have helped develop standards for consent processes.

Patients also have worked closely with translational research programs

such as the Specialized Programs of Research Excellence. For example, they have been involved in the process of tissue collection and tissue awareness programs within the different communities. In addition, they conduct grant reviews and serve on advisory boards for companies and government agencies, Collyar said.

Patient advocates work with such groups as the Army of Women, which gathers information from survivors and from women who do not have breast cancer to find out more about research. If patients could play a more active role during their cancer experience, they would be more willing to contribute biospecimens and clinical data, Collyar said. Studies have shown that patient-reported outcomes are accurate, and such data could be used in multiple ways. Patient advocates also get involved in specific issues, such as the reproducibility of studies, which is "integral to how good information is once it goes to people," she noted. In turn, that involvement produces opportunities to work with institutions to change policies and resolve barriers.

NOVEL PARTNERSHIP STRATEGIES TO DEVELOP EVIDENCE OF CLINICAL UTILITY

As Gabriela Lavezzari, director of development and diagnostics at Express Scripts, observed, the drug development pipeline is full of therapies accompanied by biomarkers (see Figure 5-3). But "the stars are not aligned," Lavezzari said, regarding what different stakeholders want from those drug-diagnostic combinations. The payers want to lower costs, offer safer and more effective treatments, and have consistent management across benefits. Patients want better health outcomes, fewer health and safety issues, and lower out-of-pocket expenses for medications. Physicians want the current buy-and-bill system to remain, the administrative burden to be reduced, and clinical information to be improved.

"Everyone is working in silos," said Lavezzari. "The patient has his own issues, trying to face a new disease and a new treatment. The physician is trying to find what the best treatment is for patients. The payer is trying to manage the cost of all these drugs. . . . In the end, everybody is frustrated."

Barriers and Opportunities for Diagnostic Companies

Lavezzari also stated that diagnostic companies face many barriers, including incomplete disease knowledge, lack of market education (which depends on a strong understanding of clinical utility), market segmentation, little intellectual property protection, an uncertain regulatory environment, and no guarantee of reimbursement. In particular, diagnostic companies

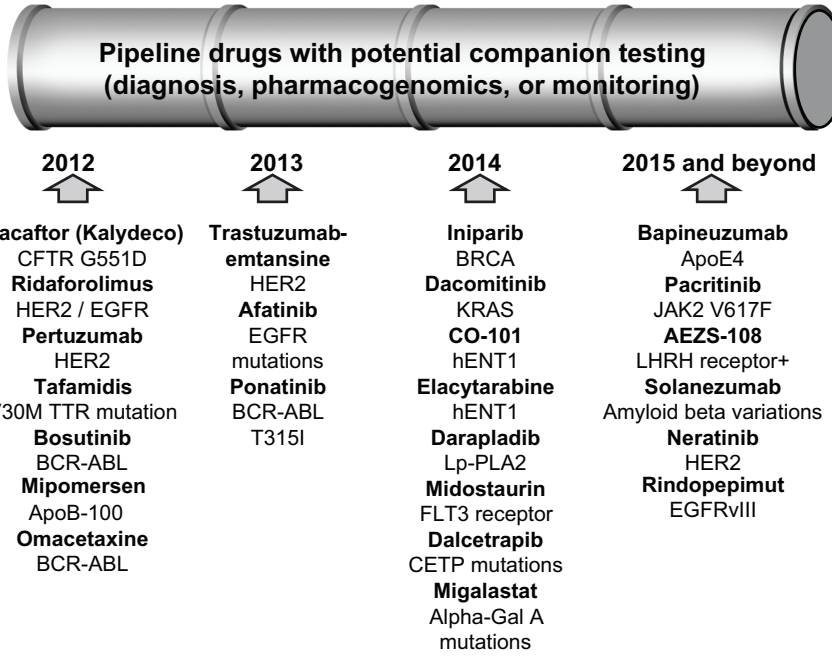

FIGURE 5-3 Many drugs with potential companion tests are in the development pipeline.
NOTE: APO, apolipoprotein; BCR-ABL, breakpoint cluster region-abelson; BRCA, breast cancer susceptibility gene; CETP, cholesteryl ester transfer protein; CFTR, cystic fibrosis transmembrane conductance regulator; EGFR, epidermal growth factor receptor; FLT3, fms-related tyrosine kinase 3; Gal, galactosidase; hENT1, human solute carrier family 29 (nucleoside transporters), member 1; HER2, human epidermal growth factor receptor 2; JAK2, Janus kinase 2; KRAS, v-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog; LHRH, luteinizing hormone-releasing hormone; Lp-PLA2, lipoprotein-associated phospholipid A2; TTR, transthyretin.
SOURCE: Gabriela Lavezzari, workshop presentation, May 24, 2012.

struggle with payers' questions about whether a diagnostic test is clinically useful and cost-effective. Payers ask: How well does the test perform? Do the test results change subsequent care? Does the change in care lead to better outcomes? What is the impact on overall cost?

Lavezzari has been working with diagnostics companies to help them understand and answer each of these questions. In the process, she has helped create different business models to help diagnostics companies advance their products. The offerings are integrated, extending from pre- to postproduct launch. For example, Lavezzari helps diagnostics companies

educate physicians and patients about new tests, create biobanks, and prove clinical utility.

Stakeholders look at approvals in different ways, Lavezzari noted. Off-label use of drugs is a good example of this dichotomy. Pharmacy benefit management companies follow FDA approvals for drug prescriptions, whereas guidelines groups can vary from this approved use on the basis of their reviews of agent-cancer combinations. Off-label prescriptions can be blocked from being dispensed because of this difference, leaving the patient stuck in the middle between the medical benefit and the pharmacy benefit. Pharmaceutical companies also need to be a part of the discussion on this issue, she said.

Lavezzari said that the most important step to improve the generation of evidence for molecular diagnostics in oncology would be to overcome the segmentation within the community, particularly between the medical and pharmacy benefit. "We cannot work on our own. The payer, the physician, and the patient that ultimately has to take the medication all have to work together."

ASSESSING CLINICAL UTILITY WITH REAL-WORLD EVIDENCE

Major changes in the development of drugs and companion diagnostics have been forcing pharmaceutical companies to adopt new models, said Greg Rossi, vice president, Payer and Real World Evidence, at AstraZeneca UK. Information about disease is rapidly increasing, providing more potential therapeutic targets and identification of biomarkers. At the same time, however, the costs of development are growing rapidly, as are the evidentiary hurdles to be overcome. These barriers are causing many oncology drugs to be looked at as marginal candidates for development because the population sizes tend to be low, the risks associated with development high, and the evidentiary standards demanding. Inevitably, companies wonder in these circumstances about the security of their returns on investment, Rossi said. "We absolutely believe . . . in evidence-based medicine," said Rossi. "But there is an opportunity cost associated with that evidence development."

Today, about 15 companion diagnostics have been approved for 7 drugs. Since 2000, the total value of the market for such drugs has risen from about $3 billion to around $18 billion, Rossi said. This market has become highly valuable, but it is also a difficult one in which to operate. As diseases are divided into subgroups, smaller and smaller populations fall into those categories, and the costs of evidence generation go up, resulting in price increases for many of the agents. "The affordability of health care is really important as we think about the cost of development and the return on investment," said Rossi.

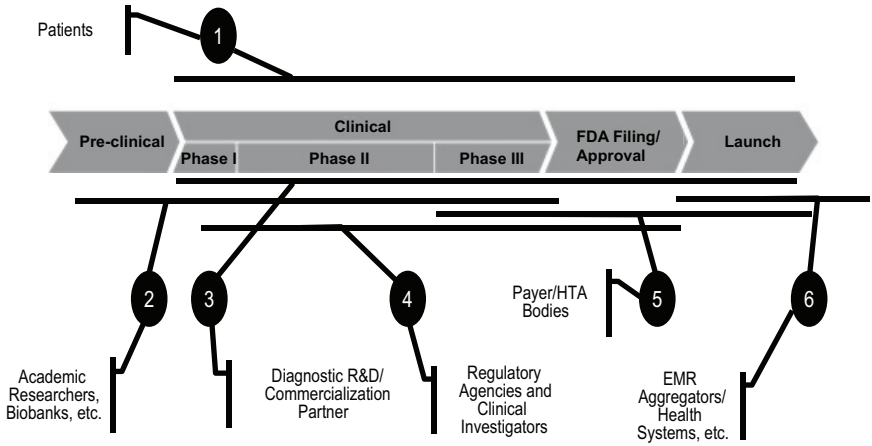

FIGURE 5-4 Partnerships are essential in developing companion diagnostics.
NOTE: EMR, electronic medical record; FDA, U.S. Food and Drug Administration; HTA, health technology assessment; R&D, research and development.
SOURCE: Greg Rossi, workshop presentation, May 24, 2012.

Operating in this market requires multiple partnerships throughout the drug development process (see Figure 5-4). Patients are at the forefront, Rossi said. There is a contract with patients to make sure that the appropriately rigorous methodology and analysis are being performed in clinical trials to ensure that benefits are being maximized. There is also a need to team up at various stages with academic researchers, companies, regulatory agencies, payers, and health care systems in order to make sure that the right type of evidence is being generated. "We can't do this without partnership," he said.

Categorization of Companion Diagnostics

Woodcock (2010) divided personalized health care options into three categories. In true drug-diagnostic codevelopment, the clinical validity and utility have been demonstrated at the time of launch. With rescue diagnostics, retrospective or prospective analysis can be done to develop evidence to facilitate a label change. And with retrofit diagnostics, new information allows a drug already on the market to be used with a new or refined set of patients who can benefit from that drug.

This is a useful categorization, said Rossi, because it serves as a reminder that drug-diagnostic codevelopment is rarely a sequential and seamless process. For example, crizotinib started as a MET inhibitor before

the development program was refocused on patients with ALK-positive non-small-cell lung cancer.[2]

He also briefly described differences among countries in assessments of clinical utility. In Italy, for example, all uses of high-cost oncology drugs go into a national registry so that practice patterns, clinical outcomes, and value can be assessed (Russo et al., 2010). "As we start thinking about real-world evidence, there are examples, in some countries and in many centers in this country, of innovative ways to think about how you collect data . . . to start answering the questions that we know are necessary and important," he added.

Real-World Evidence

Real-world evidence is a somewhat amorphous term, Rossi admitted. He focused mostly on observational studies, splitting them into prospective observational studies and retrospective analyses.

The ACCE model process has defined clinical utility as "the balance of benefits and harms associated with the use of a test in practice, including improvement in relevant outcomes and the usefulness of added value in decision making compared with not using the test."[3] Observational studies should suffice to achieve this outcome, Rossi said, but in cancer, separating the clinical difference from the noise of the assay remains very difficult. Studies of electronic medical records and claims databases could be useful in determinations of clinical utility, but these are currently hampered by such challenges as a lack of integrated data on important patient clinical characteristics, a lack of pathologic and diagnostic data, and difficulty collating all associated direct and consequential costs.

Innovative assessments of real-world evidence have the potential to monitor practice patterns before and after the introduction of a technology, assess adherence to treatment guidelines, and monitor the total impacts on costs. They also can assess generalizability through comparisons of high-level clinical outcomes with evidence from prior intervention trials. They can generate hypotheses about putative benefits and risks of competing strategies, and they can inform prospective registry designs and collaborations around those designs. "But we are going to struggle right now to get into some of the clinical data," said Rossi, "and we are not going to be

[2] For a full discussion of the development of crizotinib and its companion Vysis ALK Break Apart FISH Probe test A, see IOM, 2012c.

[3] The ACCE name is derived from the components of the model—analytic validity, clinical validity, clinical utility, and ethical, legal, and social implications. More information about the model is available at http://www.cdc.gov/genomics/gtesting/ACCE/FBR/index.htm (accessed August 13, 2012).

able to have comparative effectiveness types of questions [answered using] electronic medical records [EMRs] today."

Rossi said that the most important step to improve the generation of evidence for molecular diagnostics in oncology would be to standardize clinical information, including data in EMRs, to provide information for the broader community. In this way, it would be possible to access much larger sample sizes, which will be particularly important with smaller subgroups of patients. Computerized natural language recognition could help extract valuable information from EMRs.

Rossi also said that informed consents for the tissues that are collected can produce limited degrees of freedom for how those samples are used. That is a key issue for the use of samples from biobanks, especially globally, since some standardization initiatives are under way in the United States.

The overarching challenge, said Rossi, is to use the totality of evidence from interventional studies, observational studies, registries, and other sources, rather than competing evidentiary approaches, to produce better outcomes for patients.

CONCLUDING REMARKS

In concluding the meeting, Robert McCormack stated that the meeting helped clarify the hurdles that are preventing genomic tests from reaching patients. Stakeholders need to focus on defining when it is absolutely necessary to conduct RCTs and when other forms of evidence may be acceptable, he said. All the stakeholders have recognized their roles in the process and are trying to clarify what their needs are for demonstrating clinical utility. Speakers made a number of recommendations (see Box 5-1) that, if acted on, could spark much more progress. "It is going to take time and much more dialogue and partnership, but we have moved the ball down the field today," McCormack said.

BOX 5-1
Proposals Made by Individual Speakers

This box compiles the suggestions made by individual speakers at the workshop to advance the development of measures of clinical utility for molecular diagnostics in oncology. These suggested actions should not be seen as recommendations of the workshop, but they are promising ideas for further discussion and possible implementation.

Definitions and Standards

- To demonstrate that clinical utility has been achieved, the concept of clinical utility needs to be better defined. (McCormack)
- To determine whether test results lead to changes in practice that can be linked to improved health outcomes, clear, predictable, and consistent standards of evidence need to be developed that can be used to judge diagnostic technologies. (Tunis)
- An evidentiary framework needs to articulate the minimal evidence necessary before the clinical application of a genomic technology is warranted. (Freedman)
- Well-defined measures of clinical utility need to be accepted, and then the magnitude of the impact on those measures needs to be set to justify the adoption of a test into guidelines or for regulatory approval. (Lyman)
- Focusing more on value than on cost-effectiveness in assessments of molecular diagnostics will enable analyses to be descriptive in addition to prescriptive and will allow consideration of the full context of care. (Phillips)

Evidence Generation

- A provisional period of several years during which payers cover part of the costs of a molecular diagnostic's use would allow additional evidence to be gathered, after which the test could be accepted only if it produces substantial improvements in health outcomes. (Newcomer)
- For a biomarker to progress to a clear clinical test, it should have significant and independent value, be validated by clinical testing, be feasible and reproducible, and be widely available with quality control. (Benson)
- Test development needs to be rigorous, using meaningful and well-designed studies, proper statistical analysis, independent external validation, and interdisciplinary expertise. (McShane)
- Appropriate control groups are necessary to determine whether a marker is predictive and distinguishes a group that benefits from a treatment. (McShane)
- Studies of clinical utility should be conducted in settings that are relevant to more real-world clinical decisions. (Freedman)
- Clear priorities for comparative-effectiveness research could ensure that limited resources are used to resolve the most compelling questions. (Freedman)

- Unmet medical needs that require prospective randomized trials to develop their evidence bases need to be identified; this will also allow limited resources to be applied appropriately. (Schilsky)
- A clearinghouse of data on clinical utility from various sources could be used both in guidelines development and in deciding whether to cover or not cover the clinical use of a molecular diagnostic. (Leonard)
- A biomarker study registry could aid in identifying relevant biomarker studies for overviews and meta-analyses, make study protocols available, and help reduce nonpublication bias and selective reporting. (McShane)
- Creating an accelerated review and publication format specifically for personalized medicine assays could overcome the extended and biased review cycles in traditional publications. (Doheny)
- New strategies involving transformation of the research infrastructure to "learning systems" could allow continual additions to the knowledge base. (Freedman)
- An incentive structure for providers to put patients on clinical trials needs to be enabled. (Schilsky)

Sample and Data Collection

- Expanded access to well-annotated specimens, including alternative sources of specimens, would be especially useful for the development of molecular diagnostics. (McShane)
- The collection of blood and tissue from every patient with cancer, including patients who die, could greatly advance research. (Doheny)
- "Cancer information donors" could volunteer to provide information for cancer research even without participating in a clinical trial. (Schilsky)
- Standardization of clinical information, including data in electronic medical records, would be a valuable source of evidence for molecular diagnostics in oncology. (Rossi)
- A registry of patients with apparent false positives is needed in the development of measures of clinical utility. (Doheny)
- An integrated, outcomes-based database could be used to better understand external validity, inform unmet need assessments/trial designs, and identify variation in practice/hypotheses for detailed interventional studies. (Rossi)

Application in the Clinic

- The evidence generated and analyzed to demonstrate clinical utility needs to be adapted to the clinical setting. (McCormack)
- Clinical utility needs to receive earlier and more intense focus, with more education about how to interpret the results of tests. (McShane)
- FDA guidance should be applied to laboratory-developed tests. (Bast)
- Stakeholders need to be willing to accept that less-than-gold-standard randomized controlled trials may need to serve as sufficient evidence to make regulatory, payment, and clinical decisions. (Schilsky)

continued

BOX 5-1 Continued

- Generation of evidence for molecular diagnostics in oncology would help overcome segmentation within the provider community, including the divide between the medical benefit and the pharmacy benefit. (Lavezzari)
- The totality of evidence from interventional studies, observational studies, registries, and other sources needs to be combined to produce better outcomes for patients. (Rossi)
- Observational/database analyses could be used to augment interventional study data to assist managed entry for new technologies. (Rossi)

The Patient's Perspective

- The term "personalized medicine" should not be used because treatments can be targeted but not yet personalized to the individual level. (Collyar)
- A more relevant term than "clinical utility" for most patients is "personal utility" or "personal guidance." (Collyar)
- Patients need to get test results quickly and in clear language, and test results need to be updated as the test or a person's condition changes. (Collyar)

Partnerships

- Collaborations among cancer centers are essential, particularly to investigate rare cancers. (Freedman)
- A multitude of stakeholders having a role in evidence generation could lead to better studies. (Freedman)
- A major commitment of patients, insurers, government agencies, private institutions, and clinicians will be needed to foster partnerships aimed at innovation and technology development. (Benson)
- Modeling could be useful in determining evidence gaps and prioritizing efforts but will require consensus across stakeholders on what are considered reasonable assumptions. (Lyman)
- The sharing of emerging biomarker data can enrich research databases thereby informing the understanding of practice patterns and clinical outcomes in the real-world setting. (Rossi)

References

ACS (American Cancer Society). 2011. *Colorectal cancer facts and figures, 2011–2013*. Atlanta, GA: American Cancer Society.

ACS. 2012. *Cancer facts and figures, 2012*. Atlanta, GA: American Cancer Society.

Altman, D. G., I. Simera, J. Hoey, D. Moher, and K. Schulz. 2008. EQUATOR: Reporting guidelines for health research. *Lancet* 371(9619):1149–1150.

Altman, D. G., L. M. McShane, W. Sauerbrei, and S. E. Taube. 2012. Reporting recommendations for tumor marker prognostic studies (REMARK): Explanation and elaboration. *PLoS Medicine* 9(5):e1001216.

Andre, F., L. M. McShane, S. Michiels, D. F. Ransohoff, D. G. Altman, J. S. Reis-Filho, D. F. Hayes, and L. Pusztai. 2011. Biomarker studies: A call for a comprehensive biomarker study registry. *Nature Reviews Clinical Oncology* 8(3):171–176.

Baggerly, K. A., and K. R. Coombes. 2009. Deriving chemosensitivity from cell lines: Forensic bioinformatics and reproducible research in high-throughput biology. *Annals of Applied Statistics* 3(4):1309–1334.

Bristow, R. E., R. S. Tomacruz, D. K. Armstrong, E. L. Trimble, and F. J. Montz. 2002. Survival effect of maximal cytoreductive surgery for advanced ovarian carcinoma during the platinum era: A meta-analysis. *Journal of Clinical Oncology* 20(5):1248–1259.

CDC (Centers for Disease Control and Prevention). 2011. Vital signs: Colorectal cancer screening, incidence, and mortality—United States, 2002–2010. *Morbidity and Mortality Weekly Report* 60(26):884–889.

CMS (Centers for Medicare & Medicaid Services). 2005. *Demonstration of improved quality of care for cancer patients undergoing chemotherapy*. Baltimore, MD: Centers for Medicare & Medicaid Services.

Cochrane, A. L. 1972. *Effectiveness and efficiency: Random reflections on health services*. London: Nuffield Provincial Hospitals Trust.

Edelman, M. J., D. Watson, X. Wang, C. Morrison, R. A. Kratzke, S. Jewell, L. Hodgson, A. M. Mauer, A. Gajra, G. A. Masters, M. Bedor, E. E. Vokes, and M. J. Green. 2008. Eicosanoid modulation in advanced lung cancer: Cyclooxygenase-2 expression is a positive predictive factor for celecoxib + chemotherapy—Cancer and Leukemia Group B trial 30203. *Journal of Clinical Oncology* 26(6):848–855.

Elkin, E. B., D. A. Marshall, N. A. Kulin, I. L. Ferrusi, M. J. Hassett, U. Ladabaum, and K. A. Phillips. 2011. Economic evaluation of targeted cancer interventions: Critical review and recommendations. *Genetics in Medicine* 13(10):853–860.

Ferreira-Gonzalez, A., S. Teutsch, M. S. Williams, S. M. Au, K. T. Fitzgerald, P. S. Miller, and C. Fomous. 2008. U.S. system of oversight for genetic testing: A report from the Secretary's Advisory Committee on Genetics, Health and Society. *Personalized Medicine* 5(5):521–528.

Freidlin, B., L. M. McShane, and E. L. Korn. 2010. Randomized clinical trials with biomarkers: Design issues. *Journal of the National Cancer Institute* 102(3):152–160.

Goff, B. A., J. W. Miller, B. Matthews, K. F. Trivers, C. H. Andrilla, D. M. Lishner, and L. M. Baldwin. 2011. Involvement of gynecologic oncologists in the treatment of patients with a suspicious ovarian mass. *Obstetrics and Gynecology* 118(4):854–862.

Harris, L., H. Fritsche, R. Mennel, L. Norton, P. Ravdin, S. Taube, M. R. Somerfield, D. F. Hayes, J. R. C. Bast, and American Society of Clinical Oncology. 2007. American Society of Clinical Oncology 2007 update of recommendations for the use of tumor markers in breast cancer. *Journal of Clinical Oncology* 25(33):5287–5312.

Hayes, D. F., A. D. Thor, L. G. Dressler, D. Weaver, S. Edgerton, D. Cowan, G. Broadwater, L. J. Goldstein, S. Martino, J. N. Ingle, I. C. Henderson, L. Norton, E. P. Winer, C. A. Hudis, M. J. Ellis, D. A. Berry, and Cancer and Leukemia Group B Investigators. 2007. HER2 and response to paclitaxel in node-positive breast cancer. *New England Journal of Medicine* 357(15):1496–1506.

Henderson, I. C., D. A. Berry, G. D. Demetri, C. T. Cirrincione, L. J. Goldstein, S. Martino, J. N. Ingle, M. R. Cooper, D. F. Hayes, K. H. Tkaczuk, G. Fleming, J. F. Holland, D. B. Duggan, J. T. Carpenter, E. Frei, 3rd, R. L. Schilsky, W. C. Wood, H. B. Muss, and L. Norton. 2003. Improved outcomes from adding sequential paclitaxel but not from escalating doxorubicin dose in an adjuvant chemotherapy regimen for patients with node-positive primary breast cancer. *Journal of Clinical Oncology* 21(6):976–983.

Howlader, N., A. M. Noone, M. Krapcho, N. Neyman, R. Aminou, S. F. Altekruse, C. L. Kosary, J. Ruhl, Z. Tatalovich, H. Cho, A. Mariotto, M. P. Eisner, D. R. Lewis, H. S. Chen, E. J. Feuer, and K. A. Cronin. 2012. *SEER cancer statistics review, 1975–2009 (vintage 2009 populations)*. Bethesda, MD: National Cancer Institute.

Howlader, N., A. M. Noone, M. Krapcho, N. Neyman, S. F. Altekruse, C. L. Kosary, M. Yu, J. Ruhl, Z. Tatalovich, H. Cho, A. Mariotto, D. R. Lewis, H. S. Chen, E. J. Feuer, and K. A. Cronin (Eds.). 2013. *SEER cancer statistics review, 1975–2010*. Bethesda, MD: National Cancer Institute.

IOM (Institute of Medicine). 2009. *Initial national priorities for comparative effectiveness research*. Washington, DC: The National Academies Press.

IOM. 2011a. *Generating evidence for genomic diagnostic test development: Workshop summary*. Washington, DC: The National Academies Press.

IOM. 2011b. *Clinical practice guidelines we can trust*. Washington, DC: The National Academies Press.

IOM. 2011c. *Finding what works in health care: Standards for systematic reviews*. Washington, DC: The National Academies Press.

IOM. 2012a. *Genome-based diagnostics: Clarifying pathways to clinical use: Workshop summary*. Washington, DC: The National Academies Press.

IOM. 2012b. *Evolution of translational omics: Lessons learned and the path forward*. Washington, DC: The National Academies Press.

IOM. 2012c. *Genome-based therapeutics: Targeted drug discovery and development: Workshop summary*. Washington, DC: The National Academies Press.

Jacobs, I., D. Oram, J. Fairbanks, J. Turner, C. Frost, and J. G. Grudzinskas. 1990. A risk of malignancy index incorporating CA 125, ultrasound and menopausal status for the accurate preoperative diagnosis of ovarian cancer. *British Journal of Obstetrics and Gynaecology* 97(10):922–929.

Ladabaum, U., G. Wang, J. Terdiman, A. Blanco, M. Kuppermann, C. R. Boland, J. Ford, E. Elkin, and K. A. Phillips. 2011. Strategies to identify the Lynch syndrome among patients with colorectal cancer: A cost-effectiveness analysis. *Annals of Internal Medicine* 155(2):69–79.

McShane, L. M., D. G. Altman, W. Sauerbrei, S. E. Taube, M. Gion, G. M. Clark. 2005. Reporting recommendations for tumor marker prognostic studies (REMARK). *Nature Clinical Practice Urology* 2(8):416–422.

Menon, U., A. Gentry-Maharaj, R. Hallett, A. Ryan, M. Burnell, A. Sharma, S. Lewis, S. Davies, S. Philpott, A. Lopes, K. Godfrey, D. Oram, J. Herod, K. Williamson, M. W. Seif, I. Scott, T. Mould, R. Woolas, J. Murdoch, S. Dobbs, N. N. Amso, S. Leeson, D. Cruickshank, A. McGuire, S. Campbell, L. Fallowfield, N. Singh, A. Dawnay, S. J. Skates, M. Parmar, and I. Jacobs. 2009. Sensitivity and specificity of multimodal and ultrasound screening for ovarian cancer, and stage distribution of detected cancers: Results of the prevalence screen of the UK collaborative trial of ovarian cancer screening (UKCTOCS). *Lancet Oncology* 10(4):327–340.

Meyer, R. 2011. *Contrasting explanatory and pragmatic randomized controlled trials in oncology*. http://meetinglibrary.asco.org/sites/meetinglibrary.asco.org/files/Educational%20Book/PDF%20Files/2011/zds00111000072.pdf (accessed October 8, 2013).

Molinaro, A. M., R. Simon, and R. M. Pfeiffer. 2005. Prediction error estimation: A comparison of resampling methods. *Bioinformatics* 21(15):3301–3307.

Moore, R. G., S. MacLaughlan, and J. R. C. Bast. 2010. Current state of biomarker development for clinical application in epithelial ovarian cancer. *Gynecologic Oncology* 116(2):240–245.

Moore, R. G., M. C. Miller, P. Disilvestro, L. M. Landrum, W. Gajewski, J. J. Ball, and S. J. Skates. 2011a. Evaluation of the diagnostic accuracy of the risk of ovarian malignancy algorithm in women with a pelvic mass. *Obstetrics and Gynecology* 118(2 Pt 1):280–288.

Moore, H. M., A. B. Kelly, S. D. Jewell, L. M. McShane, D. P. Clark, R. Greenspan, D. F. Hayes, P. Hainaut, P. Kim, E. A. Mansfield, O. Potapova, P. Riegman, Y. Rubinstein, E. Seijo, S. Somiari, P. Watson, H. U. Weier, C. Zhu, and J. Vaught. 2011b. Biospecimen reporting for improved study quality (BRISQ). *Cancer Cytopathology* 119(2):92–101.

Muggia, F. M., P. S. Braly, M. F. Brady, G. Sutton, T. H. Niemann, S. L. Lentz, R. D. Alvarez, P. R. Kucera, and J. M. Small. 2000. Phase III randomized study of cisplatin versus paclitaxel versus cisplatin and paclitaxel in patients with suboptimal stage III or IV ovarian cancer: A gynecologic oncology group study. *Journal of Clinical Oncology* 18(1):106–115.

Neuss, M. N., C. E. Desch, K. K. McNiff, P. D. Eisenberg, D. H. Gesme, J. O. Jacobson, M. Jahanzeb, J. J. Padberg, J. M. Rainey, J. J. Guo, and J. V. Simone. 2005. A process for measuring the quality of cancer care: The quality oncology practice initiative. *Journal of Clinical Oncology* 23(25):6233–6239.

O'Connell, M. J., I. Lavery, G. Yothers, S. Paik, K. M. Clark-Langone, M. Lopatin, D. Watson, F. L. Baehner, S. Shak, J. Baker, J. W. Cowens, and N. Wolmark. 2010. Relationship between tumor gene expression and recurrence in four independent studies of patients with stage II/III colon cancer treated with surgery alone or surgery plus adjuvant fluorouracil plus leucovorin. *Journal of Clinical Oncology* 28(25):3937–3944.

Phillips, K. A. 2008. Closing the evidence gap in the use of emerging testing technologies in clinical practice. *Journal of the American Medical Association* 300(21):2542–2544.

PricewaterhouseCoopers. 2009. *The new science of personalized medicine: Translating the promise into practice*. http://www.pwc.com/us/en/healthcare/publications/personalized-medicine.jhtml (accessed August 31, 2012).

Ramsey, S. D., D. Veenstra, S. R. Tunis, L. Garrison, J. J. Crowley, and L. H. Baker. 2011. How comparative effectiveness research can help advance "personalized medicine" in cancer treatment. *Health Affairs* 30(12):2259–2268.

Ramsey, S. D., W. E. Barlow, A. M. Gonzalez-Angulo, S. Tunis, L. Baker, J. Crowley, P. Deverka, D. Veenstra, and G. N. Hortobagyi. 2013. Integrating comparative effectiveness design elements and endpoints into a Phase III, randomized clinical trial (SWOG S1007) evaluating Oncotype DX-guided management for women with breast cancer involving lymph nodes. *Contemporary Clinical Trials* 34(1):1–9.

Russo, P., F. S. Mennini, P. D. Siviero, and G. Rasi. 2010. Time to market and patient access to new oncology products in Italy: A multistep pathway from European context to regional health care providers. *Annals of Oncology* 21(10):2081–2087.

Saltz, L. B., D. Niedzwiecki, D. Hollis, R. M. Goldberg, A. Hantel, J. P. Thomas, A. L. Fields, and R. J. Mayer. 2007. Rinotecan fluorouracil plus leucovorin is not superior to fluorouracil plus leucovorin alone as adjuvant treatment for stage III colon cancer: Results of CALGB 89803. *Journal of Clinical Oncology* 25(23):3456–3461.

Scher, H. I., S. Halabi, I. Tannock, M. Morris, C. N. Sternberg, M. A. Carducci, M. A. Eisenberger, C. Higano, G. J. Bubley, R. Dreicer, D. Petrylak, P. Kantoff, E. Basch, W. K. Kelly, W. D. Figg, E. J. Small, T. M. Beer, G. Wilding, A. Martin, M. Hussain, and Prostate Cancer Clinical Trials Working Group. 2008. Design and end points of clinical trials for patients with progressive prostate cancer and castrate levels of testosterone: Recommendations of the Prostate Cancer Clinical Trials Working Group. *Journal of Clinical Oncology* 26(7):1148–1159.

Simon, R. M., S. Paik, and D. F. Hayes. 2009. Use of archived specimens in evaluation of prognostic and predictive biomarkers. *Journal of the National Cancer Institute* 101(21): 1446–1452.

Skates, S. J. 2012. Ovarian cancer screening: Development of the risk of ovarian cancer algorithm (ROCA) and ROCA screening trials. *International Journal of Gynecological Cancer* 22(Suppl 1):S24–S26.

Smith, I. E. 2001. Efficacy and safety of herceptin in women with metastatic breast cancer: Results from pivotal clinical studies. *Anticancer Drugs* 12(Suppl 4):S3–S10.

Sparano, J. A., and L. J. Solin. 2010. Defining the clinical utility of gene expression assays in breast cancer: The intersection of science and art in clinical decision making. *Journal of Clinical Oncology* 28(10):1625–1627.

Subramanian, J., and R. Simon. 2010. Gene expression-based prognostic signatures in lung cancer: Ready for clinical use? *Journal of the National Cancer Institute* 102(7):464–474.

Teutsch, S. M., L. A. Bradley, G. E. Palomaki, J. E. Haddow, M. Piper, N. Calonge, W. D. Dotson, M. P. Douglas, A. O. Berg, and Evaluation of Genomic Applications in Practice and Prevention Working Group. 2009. The Evaluation of Genomic Applications in Practice and Prevention (EGAPP) initiative: Methods of the EGAPP working group. *Genetics in Medicine* 11(1):3–14.

Ueland, F. R., C. P. Desimone, L. G. Seamon, R. A. Miller, S. Goodrich, I. Podzielinski, L. Sokoll, A. Smith, J. R. van Nagell, Jr., and Z. Zhang. 2011. Effectiveness of a multivariate index assay in the preoperative assessment of ovarian tumors. *Obstetrics and Gynecology* 117(6):1289–1297.

USPSTF (U.S. Preventive Services Task Force). 2012. Screening for Prostate Cancer: U.S. Preventive Services Task Force Recommendation Statement. *Annals of Internal Medicine* 157(2):120–134.

Vogel, V. G., J. P. Costantino, D. L. Wickerham, W. M. Cronin, R. S. Cecchini, J. N. Atkins, T. B. Bevers, L. Fehrenbacher, E. R. Pajon, Jr., J. L. Wade, 3rd, A. Robidoux, R. G. Margolese, J. James, S. M. Lippman, C. D. Runowicz, P. A. Ganz, S. E. Reis, W. McCaskill-Stevens, L. G. Ford, V. C. Jordan, N. Wolmark, and National Surgical Adjuvant Breast and Bowel Project. 2006. Effects of tamoxifen vs raloxifene on the risk of developing invasive breast cancer and other disease outcomes: The NSABP study of tamoxifen and raloxifene (STAR) P-2 trial. *Journal of the American Medical Association* 295(23):2727–2741.

Wong, W. B., S. D. Ramsey, W. E. Barlow, L. P. Garrison, Jr., and D. L. Veenstra. 2012. The value of comparative effectiveness research: Projected return on investment of the RxPONDER trial (SWOG S1007). *Contemporary Clinical Trials* 33(6):1117–1123.

Woodcock, J. 2010. Assessing the clinical utility of diagnostics used in drug therapy. *Clinical Pharmacology and Therapeutics* 88(6):765–773.

Appendix A

Workshop Agenda

Evidence for Clinical Utility of Molecular Diagnostics in Oncology: A Workshop

May 24, 2012

PEW DC Conference Center
901 E Street, NW
Washington, DC 20004

Workshop Objectives:

- To assess the evidentiary requirements for clinical validity and clinical utility of molecular diagnostics that are used to guide treatment decisions for patients with cancer.
- To discuss methodologies, including innovative models related to demonstrating the evidentiary requirements that meet the needs of all stakeholders.
- To consider innovative, sustainable research collaborations for generating evidence of clinical utility that involve multiple stakeholders.

8:30 A.M. **Welcoming Remarks and Charge to Workshop Speakers and Participants**

Robert McCormack, *Workshop Co-Chair*
Head of Technology Innovation and Strategy
Veridex, LLC

8:50 A.M. **Stakeholder-Informed Methods for Evaluating Clinical Utility**

Sean Tunis
Director
Center for Medical Technology Policy

9:05–11:50 A.M. **EVIDENCE UTILIZATION**

Session Moderator: Elizabeth Mansfield, U.S. Food and Drug Administration

9:05–9:35 A.M. **Guideline Development**

Gary H. Lyman
Professor of Medicine; Director, Comparative Effectiveness and Outcomes Research–Oncology
Duke University School of Medicine and the Duke Cancer Institute

Al B. Benson III
Professor of Medicine
Associate Director for Clinical Investigations
Robert H. Lurie Comprehensive Cancer Center
Northwestern University

9:35–10:05 A.M. **Payer Perspectives**

Elaine Jeter
Medical Director
Palmetto GBA

Lee Newcomer
Senior Vice President, Oncology
United HealthCare Corporation

10:05–10:20 A.M. **BREAK**

10:20–10:35 A.M. **Provider Perspective**

Lloyd Everson
Vice Chairman and Founder
The U.S. Oncology Network

10:35–10:50 A.M. **Academic Health System Perspective**

Robert Bast
Vice President for Translational Research; Internist and Professor of Medicine, Department of Experimental Therapeutics, Division of Cancer Medicine; Harry Carothers Wiess Distinguished University Chair for Cancer Research, University of Texas MD Anderson Cancer Center

10:50–11:05 A.M. **Patient Perspective**

Deborah E. Collyar
President
Patient Advocates in Research

11:05–11:50 A.M. **Discussion with Speakers and Participants**

11:50 A.M.–12:35 P.M. **LUNCH**

12:35–3:00 P.M. **STUDY DESIGN AND ANALYSIS**

Session Moderator: Patricia Deverka, Center for Medical Technology Policy

12:35–12:55 P.M. **Evolution of Translational Omics: Lessons Learned and the Path Forward**

Debra Leonard
Professor and Vice Chair, Department of Pathology and Laboratory Medicine; Director of the Clinical Laboratories
Weill Cornell Medical Center

12:55–1:15 P.M. **Comparative-Effectiveness Research Methodologies for Generating and Synthesizing Evidence for Cancer Genomics**

Andrew N. Freedman
Chief, Clinical and Translational Epidemiology Branch
Epidemiology and Genomics Research Program
Division of Cancer Control and Population Sciences
National Cancer Institute

1:15–1:35 P.M. **Designing Studies to Evaluate Biomarkers for Clinical Applications**

Lisa M. McShane
Senior Mathematical Statistician
Biometric Research Branch
Division of Cancer Treatment and Diagnosis
National Cancer Institute

1:35–1:55 P.M. **Assessing Cost-Effectiveness for Oncology-Based Molecular Diagnostics**

Kathryn Phillips
Professor of Health Economics and Health Services Research
Director and Founder, University of California, San Francisco (UCSF), Center for Translational and Policy Research on Personalized Medicine
Department of Clinical Pharmacy/School of Pharmacy, UCSF Institute for Health Policy Studies, and UCSF Comprehensive Cancer Center
UCSF

1:55–2:15 P.M. **Advancing the Utility of Oncology Diagnostics**

Noel Doheny
Chief Executive Officer
Epigenomics, Inc.

2:15–3:00 P.M. **Discussion with Speakers and Participants**

3:00–3:15 P.M. **BREAK**

3:15–5:15 P.M **ADVANCING MOLECULAR DIAGNOSTICS FOR ONCOLOGY**

Session Moderator: Margaret Piper, Blue Cross and Blue Shield Association

3:15–3:30 P.M. **Biomarker Studies in Multi-Center Cancer Clinical Trials: The Role of Cooperative Groups**

Richard Schilsky
Professor of Medicine
Chief, Section of Hematology-Oncology
Deputy Director, Comprehensive Cancer Center
University of Chicago

3:30–3:45 P.M. **Partnering for the Cure: An Innovative Role for Academia in Oncology Drug and Diagnostic Development**

Howard I. Scher
D. Wayne Calloway Chair in Urologic Oncology
Sidney Kimmel Center for Prostate and Urologic Cancers
Chief, Genitourinary Oncology Service
Memorial Sloan-Kettering Cancer Center

3:45–4:00 P.M. **Patient Approaches to Generating Evidence**

Deborah E. Collyar
President
Patient Advocates in Research

4:00–4:15 P.M. **Novel Partnership Strategies for Using Outcomes Data to Develop Clinical Utility Evidence**

Gabriela Lavezzari
Director of Development, Diagnostics
Express Scripts

4:15–4:30 P.M.	**Assessing Clinical Utility with Real-World Evidence**
	Greg Rossi Vice President, Payer and Real-World Evidence AstraZeneca UK
4:30–5:15 P.M.	**Discussion with Speakers and Participants on Paths Forward**
5:15–5:30 P.M.	**FINAL REMARKS**
5:15–5:30 P.M.	**Concluding Remarks**
	Robert McCormack, *Workshop Co-Chair* Head of Technology Innovation and Strategy Veridex, LLC
5:30 P.M.	**ADJOURN**

Appendix B

Speaker Biographical Sketches

Robert C. Bast, Jr., M.D., is vice president for translational research at the University of Texas MD Anderson Cancer Center. His office facilitates translation of new strategies, drugs, and devices from the laboratory to the clinic, as well as the movement of human material and data from the clinic to the laboratory. Dr. Bast's group coordinates programs to train physician-scientists and clinician-investigators, facilitate development of multi-investigator grants, provide instrumental shared resources, to develop faculty inventions, and enhance collaborations with pharmaceutical companies.

Dr. Bast received his B.A. *cum laude* from Wesleyan University and his M.D. *magna cum laude* from Harvard Medical School. After completing a medical internship at Johns Hopkins Hospital, he served as a research associate at the National Cancer Institute (NCI). Returning to Boston, Dr. Bast completed a medical residency at the Peter Bent Brigham Hospital and a fellowship in medical oncology at the Dana-Farber Cancer Institute. He joined the faculty at Harvard as an assistant professor and was subsequently appointed associate professor at the Dana-Farber Cancer Institute. Dr. Bast was recruited to the Duke University Medical Center in 1984 as professor of medicine, microbiology, and immunology to codirect the Division of Hematology-Oncology and to serve as clinical director of the cancer center. In 1987, he became the director of the Duke Comprehensive Cancer Center, and, in 1992, he was named Wellcome Clinical Professor of Medicine in honor of R. Wayne Rundles. In July 1994, Dr. Bast was recruited to head the Division of Medicine at MD Anderson Cancer Center and to fill the Harry Carothers Wiess Chair for Cancer Research. In 2000,

Dr. Bast was appointed vice president for translational research. In 2004, he became the Harry Carothers Wiess Distinguished University Professor for Cancer Research. Overall, Dr. Bast has published more than 600 articles and chapters and has edited the textbook *Cancer Medicine*. He has been recognized by the Institute for Scientific Information as one of the most frequently cited scientists in his field (top 0.5 percent). He continues to care for patients with breast and ovarian cancer and has been listed in the *Best Doctors of America* and in *America's Top Physicians*.

Al B. Benson III, M.D., FACP, FASCO, is a professor of medicine in the Division of Hematology/Oncology at Northwestern University's Feinberg School of Medicine in Chicago, Illinois. He is also the associate director for clinical investigations at the Robert H. Lurie Comprehensive Cancer Center, an NCI-designated Comprehensive Cancer Center, at Northwestern University. In addition, he is an attending physician at Northwestern Memorial Hospital, a staff physician at Jesse Brown Veterans Affairs Medical Center, and a consultant to the Rehabilitation Institute of Chicago.

Dr. Benson earned his medical degree at the State University of New York at Buffalo after which he completed an internal medicine residency at the University of Wisconsin Hospitals in Madison, Wisconsin. He was an assistant professor of medicine at the University of Illinois and co-medical director for the National Public Health Service in Champaign, Illinois. He then served as a clinical oncology and research fellow at the University of Wisconsin Clinical Cancer Center, where he received an American Cancer Society Fellowship Award.

Dr. Benson is active on numerous professional committees, often serving as an officer. He is a recipient of the American Society of Clinical Oncology (ASCO) Statesman Award (Fellow of ASCO) and has served on a number of committees. He has been a member of the Task Force on Quality of Cancer Care, the co-chair of ASCO's Colorectal Cancer Guidelines Subcommittee, the Stage II Colon Cancer Guidelines Panel, and the Guidelines Panel for Use of Radiofrequency Ablation for Colorectal Cancer Hepatic Metastases. He also is the chair of both the Eastern Cooperative Oncology Group Gastrointestinal and Data Monitoring Committees and a member of the National Cancer Institute Gastrointestinal Steering Committee. In addition, he is a past president of the Illinois Medical Oncology Society, past president of the Association of Community Cancer Centers (ACCC), and an executive committee member and immediate past chair of the board of directors of the National Comprehensive Cancer Network (NCCN). He is the chair of both the NCCN Hepatobiliary Guidelines Panel and a member of the Colon, Rectal and Anal Panel and a member of the Pancreatic Cancer and Neuroendocrine Panels. He is a past president of the International Society of GI Oncology. Dr. Benson is a member of the Scientific Board of

Directors of the Patient Advocate Foundation, the National Patient Advocate Foundation, and Friends of Cancer Research.

Dr. Benson is on the editorial board of the *ASCO Connection*, *American Health and Drug Benefits*, *Personalized Medicine in Oncology*, *Journal of Comparative Effectiveness Research*, and *Gastrointestinal Cancer Research*, among other publications. His most recent national awards include the NCCN Rodger J. Winn Award and the ACCC Outstanding Achievement in Clinical Research Award.

Deborah E. Collyar has been a leader in cancer patient advocacy since 1991, utilizing her extensive business expertise to bridge gaps between science, medical providers, and patients. Her advocacy spans many diseases and in-depth programs with many academic and private institutions, nonprofits, government agencies, companies, and patients. Ms. Collyar founded the Patient Advocates in Research international network after successfully developing research patient advocacy throughout cancer centers and cooperative groups. Ms. Collyar has chaired the Cancer and Leukemia Group B Committee on Advocacy, Research Communication, Ethics, and Disparities and served on many NCI, American Association for Cancer Research, and ASCO committees. Currently, she serves on the NCI's Investigational Drug Steering Committee, Experimental Therapeutics program, and the Cancer Immunotherapy Network. She also works with the Clinical and Translational Science Awards Sentinel Network to help researchers learn about community needs.

Patricia Deverka, M.D., M.S., M.B.E., is a senior research director at the Center for Medical Technology Policy (CMTP). Her current research responsibilities, including the Center for Comparative Effectiveness Research in Cancer Genomics and the CER Institute, reflect her broad expertise in comparative-effectiveness research and genomics and personalized medicine. Before joining CMTP, Dr. Deverka was on the faculty at the Institute for Pharmacogenomics and Individualized Therapy at the University of North Carolina at Chapel Hill, where her research focused on issues such as coverage and reimbursement and improving the evidence base for pharmacogenomic tests. Earlier in her career, she worked as an outcomes researcher in the pharmaceutical industry and was vice president of scientific affairs at Medco Health Solutions, where she led several pharmacogenomic-related initiatives. Dr. Deverka has a medical degree from the University of Pittsburgh and is board certified in general preventive medicine and public health. She also obtained a master's degree in bioethics from the University of Pennsylvania and completed a fellowship at Duke University's Institute for Genome Sciences and Policy.

Noel Doheny joined Epigenomics in May 2011 as chief executive officer (CEO) of the U.S. subsidiary Epigenomics, Inc. He has more than 30 years of experience in the field of diagnostics, with more than 20 years in senior management. Prior to joining Epigenomics, Mr. Doheny held positions as CEO of OpGen; senior vice president for the Molecular Diagnostics Division of Affymetrix, Inc.; vice president of Pre-Analytical Solutions and as a member of the executive committee at QIAGEN; and president and CEO of BioStar, Inc. He has built several operating teams from the ground up, including the commercial teams to launch novel products at companies such as Ciba Corning, Biostar, and OpGen. Mr. Doheny obtained degrees in biology and chemistry from West Virginia University and attended Georgetown University for postgraduate studies in biochemistry prior to moving into industry.

Lloyd K. Everson, M.D., completed his medical education at Harvard University in 1969. He has recently served as vice-chairman and member of the board of directors of U.S. Oncology, Inc. In 1993, Dr. Everson became a member of the board of directors, president, and chief operating officer of American Oncology Resources, Inc., which was the predecessor organization of U.S. Oncology, Inc. In 2009, Dr. Everson formed the Life Beyond Cancer Foundation, for which he serves as chairman of the board.

Dr. Everson has served in numerous leadership roles in medical education through his career, including the practice of oncology and hematology, cancer research, and academic and administrative medicine. His postgraduate education included clinical and research positions at Cornell Medical Center and New York Hospital, Memorial Sloan-Kettering Cancer Center, and the National Cancer Institute. Dr. Everson is a member of the board of directors of the Intercultural Cancer Council and a member of the board of directors of C-Change. He also serves as a member of the Sigma Alpha Epsilon Foundation board of directors and is a member of the board of directors of Interfaith of the Woodlands, Texas. He has served in leadership positions in several national organizations, including the NCI Board of Scientific Counselors and the NCI National Cancer Advisory Board. Dr. Everson has also served as president of the Association of Community Cancer Centers, medical director of the Indiana Regional Cancer Center, associate chairman of the Eastern Cooperative Oncology Group, and member of the CEO Roundtable on Cancer.

Dr. Everson is the recipient of many honors and awards, including the University of North Dakota Sioux Award for Outstanding Alumni and the Houston Lifetime Outstanding Achievement Award.

Andrew N. Freedman, Ph.D., is the chief of the Epidemiology and Genomics Research Program's (EGRP's) Clinical and Translational Epidemiology

Branch. He oversees EGRP's research portfolio and initiatives that focus on factors that influence cancer progression, recurrence, new primary cancers, survival and other treatment outcomes, and factors associated with cancer development among individuals with underlying diseases and conditions.

In 1997, Dr. Freedman joined the National Cancer Institute's Division of Cancer Control and Population Sciences (DCCPS) as a molecular epidemiologist in the Applied Research Program's Risk Factor Monitoring and Methods Branch. He developed and supported a program of research in cancer risk prediction, genetic susceptibility testing, pharmacoepidemiology, and pharmacogenomics and managed research contracts, interagency and cooperative agreements, and a grant portfolio pertaining to these research areas. Dr. Freedman also directed multidisciplinary molecular, clinical, and translational epidemiology studies within the HMO Cancer Research Network; Department of Veterans Affairs medical system; NCI's Surveillance, Epidemiology, and End Results Program; and the National Institutes of Health (NIH)-AARP Diet and Health Study. He is internationally recognized for his work in molecular cancer epidemiology and cancer risk prediction.

In the areas of pharmacoepidemiology and pharmacogenomics, Dr. Freedman has developed research collaborations with several NIH institutes and centers and other agencies within the U.S. Department of Health and Human Services (HHS). He is chair of the Trans-NCI Pharmacoepidemiology and Pharmacogenomics Working Group and represents the NCI in the Trans-NIH Pharmacogenomics Working Group and the Institute of Medicine (IOM) Roundtable on Translating Genomic-Based Research for Health. Before joining the DCCPS, Dr. Freedman was a postdoctoral research fellow in the Genetic Epidemiology Branch of NCI's Division of Cancer Epidemiology and Genetics.

Elaine Jeter, M.D., is the Palmetto GBA Part B medical director for jurisdiction 1 (J1) where she is currently working to implement the Molecular Diagnostic Services Program. She is a graduate of the Medical University of South Carolina (MUSC) and is board certified in clinical and anatomic pathology and in blood banking/transfusion medicine. Dr. Jeter received her undergraduate degree from the State University of New York and her master's from the University of South Carolina. She was an academic physician at MUSC for 10 years and in private pathology practice in Columbia, South Carolina, for a number of years. Dr. Jeter joined Palmetto GBA in 2005.

Gabriela Lavezzari, Ph.D., M.B.A., is director of development at Express Scripts–Medco, where she works toward proving the clinical utility and cost-effectiveness of a broad range of health care innovation products.

Before joining Medco, Dr. Lavezzari was the laboratory director for

Theranostics Health, a biotech company in Rockville, Maryland, focused on the discovery of new cancer theranostics biomarkers. There she managed the laboratory services and led the research and development efforts for biomarker discovery. Prior to that position, Dr. Lavezzari was a senior laboratory specialist with the AIDS Clinical Trials Group (ACTG) at Social and Scientific Systems. There she facilitated technical management and operations of ACTG lab structure domestically and internationally.

Dr. Lavezzari received her Ph.D. in biology at the University of Milano, Italy, followed by her postdoctoral work at the National Institute of Neurological Disorders and Stroke. She also received her M.B.A. at the New York Institute of Technology.

Debra Leonard, M.D., Ph.D., received her M.D. and Ph.D. from the New York University School of Medicine and is currently professor and vice-chair for laboratory medicine in the Department of Pathology and Laboratory Medicine as well as director of the clinical laboratories for New York–Presbyterian Hospital's Cornell campus (NYPH–WCMC). She is also director of the Pathology Residency Training Program at NYPH–WCMC. Dr. Leonard was previously director of molecular pathology at the University of Pennsylvania School of Medicine and is a nationally recognized expert in molecular pathology. She has served on several national committees that develop policy for the use of genetic and genomic technologies and information, including most recently the Secretary's Advisory Committee on Genetics, Health and Society that advises the Secretary of HHS. Dr. Leonard is editor of two molecular pathology textbooks and has spoken widely on various molecular pathology test services, the future of molecular pathology, and the impact of gene patents on molecular pathology practice. Dr. Leonard is interested in the use of genomic technologies in the practice of medicine to improve patient outcomes.

Gary H. Lyman, M.D., M.P.H., FASCO, FRCP (Edinburgh), is professor of medicine and director of comparative effectiveness and outcomes research–oncology at Duke University and the Duke Cancer Institute. Dr. Lyman is also a senior fellow at the Duke Center for Clinical Health Policy Research. Dr. Lyman received his undergraduate and medical degrees from the State University of New York in Buffalo and completed internal medicine residency at the University of North Carolina at Chapel Hill. He subsequently completed a clinical hematology/oncology fellowship at the Roswell Park Memorial Institute and a postdoctoral fellowship in biostatistics at the Harvard School of Public Health and the Dana Farber Cancer Center. Dr. Lyman previously served as professor of medicine, director of medical oncology and chief of medicine at the H. Lee Moffitt Cancer Center and Research Institute. During this period he served as professor of medicine

and professor of epidemiology and biostatistics as well as associate chairman of the Department of Medicine at the University of South Florida. Prior to relocating to Duke University in 2007, Dr. Lyman held positions as Thomas Ordway Professor of Medicine and cancer center director at the Albany Medical College and subsequently professor of medicine, associate cancer center director, and director of health services and outcomes research at the University of Rochester and Strong Memorial Hospital.

Dr. Lyman is active in ASCO, serving as chair-elect of the ASCO Clinical Practice Guideline Committee after having chaired the Methodology Subcommittee for several years. Dr. Lyman also chairs several ASCO guideline panels and is a member of the ASCO Biomarkers Guideline Working Group, the Comparative Effectiveness Research Task Force, and the Cost of Care Task Force. In 2010, Dr. Lyman received the prestigious ASCO Statesman Award and was recently elected to the ASCO board of directors. Dr. Lyman is an advisor to the U.S. Food and Drug Administration (FDA) and the Oncology Drug Advisory Committee. He is editor-in-chief of *Cancer Investigation* and the peer review editor for ASCO's *Journal of Oncology Practice* and serves on the editorial board of the *Journal of Clinical Oncology* and several other subspecialty journals. In addition to serving as a fellow of ASCO, Dr. Lyman is a fellow of the Royal College of Physicians (Edinburgh), the American College of Physicians, the American College of Preventive Medicine, and the American College of Clinical Pharmacology. His research interests include personalized cancer supportive care, comparative effectiveness studies of targeted therapies and biomarkers, mathematical and statistical prognostic and predictive models, advanced methods of evidence synthesis in support of clinical practice guidelines and population studies of patterns of cancer treatment, and the impact of health disparities on the quality of cancer care.

Elizabeth Mansfield, Ph.D., is the director of the personalized medicine staff in the Office of In Vitro Diagnostic and Radiological Health (OIR) in the Center for Devices and Radiological Health, FDA, where she is developing a program to address companion and novel diagnostic devices. She was previously a senior policy analyst in OIR, managing policy and scientific issues. From 2004 to 2006, Dr. Mansfield served as the director of regulatory affairs at Affymetrix, Inc. She previously served in other positions at FDA, including scientific reviewer and genetics expert. Dr. Mansfield received her Ph.D. from Johns Hopkins University and completed further postdoctoral training at NCI and the National Institute for Arthritis, Musculoskeletal, and Skin Diseases.

Robert McCormack, Ph.D., is currently head of technology innovation and strategy for Veridex, LLC. He was formerly the director of technol-

ogy assessment for Ortho-Clinical Diagnostics, focusing on novel cellular and molecular cancer technology. In 2005, he assumed the role of vice president of scientific and medical affairs at Veridex, LLC, a Johnson & Johnson start-up dedicated to the development and commercialization of novel cancer diagnostic tests. His group successfully conducted clinical trials to launch the first molecular test for assessing axillary nodal status in women diagnosed with breast cancer. Prior to this position, in 2001 he was appointed general manager of the cellular diagnostics group at Veridex. The cellular diagnostics group successfully launched their first product in 2004 for the detection and enumeration of circulating tumor cells in patients with metastatic breast cancer. He joined Johnson & Johnson in 1998 as vice president of clinical affairs for Ortho-Clinical Diagnostics. Under his direction, Ortho-Clinical Diagnostics became the first diagnostics company to gain FDA approval for hepatitis assay testing on random access automation for clinical laboratories. In 1995, he joined Sanofi Diagnostics Pasteur as director of clinical and regulatory affairs and worldwide group leader for cancer diagnostics. Dr. McCormack spent his early career in genetic, molecular, and cellular research at the University of Minnesota, 3M, and Hybritech. He transitioned to clinical and regulatory affairs at Hybritech and was part of the team that successfully gained FDA approval for prostate-specific antigen in the early detection of prostate cancer. Dr. McCormack received his B.S. degree in medical technology from the University of Wisconsin, River Falls, and his M.S. and Ph.D. degrees from the University of Minnesota in hematology and immunology, respectively.

Lisa M. McShane, Ph.D., is a senior mathematical statistician in the Biometric Research Branch in the Division of Cancer Treatment and Diagnosis at the NCI. She earned her Ph.D. in statistics from Cornell University. Since 1996, Dr. McShane has worked closely with the NCI Cancer Diagnosis Program and Cancer Therapy Evaluation Program on statistical matters relating to the development and use of tumor markers for prognosis, prediction, and disease monitoring. She is a member of the NCI Program for the Assessment of Clinical Cancer Tests Strategy Group.

Dr. McShane's statistical interests and publications have covered a diverse set of topics, including statistical methods for the analysis of high-dimensional genomic data, multiple comparisons methods, surrogate end points, measurement error adjustment methods, laboratory quality control and assay reproducibility assessment, and spatial statistics. She has also been statistical coauthor on many biomedical papers covering such topics as genomic studies in breast, colon, and lung cancer; colorectal epithelial cell proliferation; serum markers in prostate cancer; molecular characterization of ovarian tumors; Parkinson's disease; motor control disorders; stroke;

and Creutzfeldt-Jakob disease. She is a co-author of the book *Design and Analysis of DNA Microarray Investigations*.

Dr. McShane is a frequent speaker at national and international statistics meetings and oncology meetings. She has presented numerous statistical lectures, didactic lectures, and discussions on the design and analysis of biomarker studies, including gene expression microarray studies. In 2008, Dr. McShane was awarded a prestigious NIH Director's Award in recognition of her work on trial designs to assess predictive biomarkers for their utility in therapeutic decision making for patients with cancer.

Lee N. Newcomer, M.D., M.H.A., is senior vice president at United Healthcare (UHC), with strategic responsibility for oncology, genetics, and women's health. Prior to rejoining UHC in 2006, Dr. Newcomer was a founding executive of Vivius, a consumer-directed venture that allowed consumers to create their own personalized health plans. From 1991 to 2000, Dr. Newcomer held a number of positions at UHC, including chief medical officer. His work there emphasized the development of performance measures and incentives to improve clinical care. Prior to initially joining UHC, he was medical director for CIGNA Health Care of Kansas City.

Dr. Newcomer is a board certified medical oncologist; he practiced medical oncology for 9 years in Tulsa, Oklahoma, and Minneapolis, Minnesota (Park Nicollet Clinic). He is the former chairman of Park Nicollet Health Services, an integrated system of more than 700 physicians and a 400-bed hospital. The group is nationally recognized for its leadership in quality, safety, and lean processes.

Dr. Newcomer earned a B.A. degree from Nebraska Wesleyan University, an M.D. degree from the University of Nebraska College of Medicine, and an M.H.A. from the University of Wisconsin–Madison. He completed his internship and residency in internal medicine at the University of Nebraska Medical Center and a fellowship in medical oncology at the Yale University School of Medicine.

Kathryn A. Phillips, Ph.D., is professor of health economics and health services research at the University of California, San Francisco (UCSF), and founder/director of the UCSF Center for Translational and Policy Research on Personalized Medicine. She has appointments in the UCSF Department of Clinical Pharmacy, the UCSF Institute for Health Policy Studies, and the UCSF Comprehensive Cancer Center. Dr. Phillips holds degrees from the University of California, Berkeley, Harvard University, and the University of Texas at Austin and previously spent 8 years working for the federal government.

Dr. Phillips's research focuses on using quantitative tools to examine policy issues on how health care is organized, delivered, and financed in

the United States. Her emphasis is on translation of new technologies and personalized medicine and its impact on clinical care, health economics, and health policy. She has published more than 100 peer-reviewed articles in such leading journals as *Journal of the American Medical Association*, *New England Journal of Medicine*, and *Health Affairs*. She also serves on the editorial board for 6 journals and has led or participated in approximately 50 funded research grants, with continuous funding from NIH as a principal investigator since 1993.

Dr. Phillips has served as an adviser to many organizations, including the IOM, the President's Council of Advisors on Science and Technology, Genome Canada, FDA, the Australia Pharmaceutical Benefits Committee, and the Medicare Services Advisory Committee, and regularly speaks to national and international groups. Dr. Phillips also consults with a number of biotech start-ups and venture capital firms and serves on the Novartis Molecular Diagnostics Advisory Board.

Margaret Piper, Ph.D., M.P.H., is the director of genomics resources at the Blue Cross Blue Shield Association Technology Evaluation Center (TEC), an Agency for Healthcare Research and Quality (AHRQ)-funded Evidence-based Practice Center (EPC). She has been with TEC since 1994, joining the staff full-time in 1999. Her experience at TEC has focused on systematic reviews of medical technology, including topics in autoimmunity and transplantation, oncology, laboratory medicine, and genomics/genetic testing. Dr. Piper has written more than 30 TEC systematic reviews and reports and has co-written 4 AHRQ-EPC reports. Among other outreach activities, Dr. Piper has served on CMS's Medicare Evidence Development and Coverage Advisory Committee and on a work group for the Institute for Quality in Laboratory Medicine, and she currently serves on the working group for the Centers for Disease Control and Prevention (CDC)-funded Evaluation of Genomic Applications in Practice and Prevention (EGAPP) project. Her roles with the EGAPP working group include establishing methods and processes for evidence-based evaluation of genetic tests; prioritizing and selecting topics for review; participating in technical expert panels for commissioned evidence reviews; and developing conclusions or recommendations on the basis of the evidence. In addition to these activities, Dr. Piper has given presentations on evidence-based evaluation of genetic tests at meetings organized by the IOM, AHRQ, and the NCI. Prior experience includes more than 13 years of managing a variety of clinical diagnostic laboratory departments in both academic hospital and commercial clinical laboratory settings, designing and evaluating new laboratory diagnostics for the biomedical industry, consulting with physicians, publishing, and volunteer teaching for professional organizations in laboratory medicine. In 2000, Dr. Piper received a Distinguished Service Award from the Ameri-

can Society of Clinical Pathologists Commission on Continuing Education. Following a mid-career NCI fellowship in cancer prevention and control, which included obtaining an M.P.H. in epidemiology, Dr. Piper gained experience in cancer epidemiology at the NCI and subsequently at the CDC, with a focus on cancer genetics. Dr. Piper has a B.S. in molecular biology (University of Wisconsin–Madison), a Ph.D. in immunology (Duke University), and an M.P.H. in epidemiology (Emory University).

Greg Rossi, Ph.D., is the vice president of payer and real-world evidence for AstraZeneca UK. His responsibilities include overseeing coverage and reimbursement submissions and real-world evidence (outcomes) studies and ensuring that development programs address evidence requirements of global payers. He received his doctorate in molecular biology and protein chemistry from University College London in 1993. Since then he has worked in the biotechnology/pharmaceutical industry.

Between 1996 and 2007, Dr. Rossi was at Amgen, Inc., where he held a number of roles in regulatory affairs, clinical development, and global product strategy. He was involved in the clinical development and registration of a number of products, with responsibilities including negotiations with global reimbursement authorities and U.S. payers. Between 2005 and 2007, Dr. Rossi was responsible for coleading the development and commercialization strategies of Amgen's therapeutic oncology pipeline in Phase II and III clinical development.

In 2007, Dr. Rossi joined Genentech Inc./Roche Pharmaceuticals, where he was responsible for outcomes research, health economics, and global pricing. These responsibilities included providing input into the design of Phase II, III, and IV studies, conducting observational research studies, developing global pricing policies, and generating health technology/reimbursement dossiers for coverage and reimbursement submissions. Dr. Rossi is the author of more than 25 clinical research manuscripts and numerous abstracts.

Howard I. Scher, M.D., is the first incumbent of the D. Wayne Calloway Chair in urologic oncology and chief of the Genitourinary Oncology Service at the Sidney Kimmel Center for Urologic and Prostate Cancers at Memorial Sloan-Kettering Cancer Center (MSKCC). One of the foremost genitourinary medical oncologists in the country, he also serves as principal investigator of the NIH Specialized Program of Research Excellence in prostate cancer at MSKCC, the Department of Defense Prostate Cancer Research Program (PCRP), and the Prostate Cancer Foundation–funded Prostate Cancer Clinical Trials Consortium, a 13-center research collaborative funded by the PCRP and the Prostate Cancer Foundation. Dr. Scher's work is focused on three critical areas: developing treatments that target

specific signaling pathways that contribute to prostate cancer growth, developing noninvasive methods to determine whether these agents are working, and improving the way drugs and biomarkers are evaluated in the clinic.

In addition, as professor of medicine at the Joan and Sanford Weill Medical College of Cornell University, Dr. Scher has mentored a new generation of physician-scientists who are already making important contributions to the field of prostate cancer research. He is the recipient of the Donald S. Coffey–Prostate Cancer Foundation Physician-Scientist Award and the Distinguished Alumnus Award. Dr. Scher also serves on numerous national scientific advisory boards and is a reviewer for many journals, including *New England Journal of Medicine*, *Lancet*, *Lancet Oncology*, *Cancer Cell*, *Nature*, and *Journal of Clinical Oncology*. He has written extensively and has published more than 400 peer-reviewed articles in scientific journals.

Richard L. Schilsky, M.D., is professor of medicine and chief, Section of Hematology/Oncology in the Department of Medicine, Biological Sciences Division, at the University of Chicago. He earned his M.D. at the University of Chicago Pritzker School of Medicine in 1975. Following a residency in internal medicine at the University of Texas Southwestern Medical Center and Parkland Memorial Hospital, he received training in medical oncology and clinical pharmacology at the National Cancer Institute from 1977 to 1981. He then served as assistant professor of medicine at the University of Missouri–Columbia School of Medicine from 1981 to 1984 and then returned to the University of Chicago. Dr. Schilsky previously served as director of the University of Chicago Cancer Research Center (1991–1999) and as associate dean for clinical research (1999–2007).

From 1995 to 2010, Dr. Schilsky served as chairman of the Cancer and Leukemia Group B, an NCI-sponsored national cancer clinical trials group. He is presently a member of the board of directors of the Alliance for Clinical Trials in Oncology and of CTNet, the statewide clinical trials network of Texas.

An international expert in gastrointestinal malignancies and cancer pharmacology, Dr. Schilsky has served on a number of peer review and advisory committees for the NCI and previously served as chair of the Oncologic Drugs Advisory Committee for FDA. Dr. Schilsky is past chair of the NCI Board of Scientific Advisors and has served as a member of the Clinical and Translational Research Advisory Committee. Dr. Schilsky also served as a member of the board of directors of the American Society of Clinical Oncology and as ASCO president from 2008 to 2009. He is presently a member of the board of directors of the Conquer Cancer Foundation.

He is also a member of the external advisory committees of several comprehensive cancer centers, including the Roswell Park Cancer Center, the MD Anderson Cancer Center, the Huntsman Cancer Institute, the Simmons Cancer Center at the University of Texas–Southwestern, the Tisch Cancer Institute at Mount Sinai Hospital, and the UCSF Helen Diller Family Comprehensive Cancer Center.

Dr. Schilsky is an associate editor of *Clinical Cancer Research*, *Seminars in Oncology*, and the *Journal of the National Cancer Institute*; senior associate editor of *Molecular Oncology*; and a member of the editorial boards of *Cancer Prevention Research*, *Cancer Investigation*, *Annals of Surgical Oncology*, *Journal of Cancer Research and Clinical Oncology*, and several other journals. He has published more than 285 articles and book chapters in the medical literature and is the editor of 4 books.

Sean Tunis, M.D., M.Sc., is the president and chief executive officer of the Center for Medical Technology Policy (CMTP) in Baltimore, Maryland. CMTP's main objective is to improve the quality, relevance, and efficiency of clinical research by providing a neutral forum for collaboration among experts, stakeholders, and decision makers. Dr. Tunis was a member of the IOM Committee on Initial National Priorities for Comparative Effectiveness Research. He advises a wide range of domestic and international public and private health care organizations on issues of comparative effectiveness, evidence-based medicine, clinical research, reimbursement, and health technology policy.

Through September of 2005, Dr. Tunis was the chief medical officer at CMS, where he had lead responsibility for clinical policy for the Medicare and Medicaid programs. Previously he served as the director of the Health Program at the Congressional Office of Technology Assessment and as a health policy advisor to the U.S. Senate, where he worked on pharmaceutical and device policy issues.

Dr. Tunis trained at the University of California, Los Angeles, and the University of Maryland in internal medicine and emergency medicine. He holds adjunct faculty positions at the Center for Health Policy at Stanford University, the Department of Internal Medicine at the Johns Hopkins University School of Medicine, and the Department of Surgery at UCSF.

Appendix C

Statement of Task

An ad hoc planning committee will plan and conduct a public workshop to discuss research methodologies related to demonstrating and assessing the clinical validity and clinical utility of molecular diagnostics that are used to guide treatment and management of oncology patients. The goal of the workshop will be to advance discussions among a broad array of stakeholders, including patients, providers, policy makers, payers, test developers, guideline developers, and others, around evidence needs for informed clinical and health policy decision making. The planning committee will develop the workshop agenda, select and invite speakers and discussants, and moderate the discussions. An individually authored summary of the workshop will be prepared by a designated rapporteur in accordance with institutional policy and procedures. This workshop will be held jointly with the Center for Medical Technology Policy.

Appendix D

Registered Attendees

Scott Allocco
BioMarker Strategies

Eric Assaraf
WRG

Erin Balogh
Institute of Medicine

Robert Bast
University of Texas MD Anderson Cancer Center

Jessica Bayh
Applied Policy

Al Benson III
Robert H. Lurie Comprehensive Cancer Center, Northwestern University

Marian Birkeland
National Comprehensive Cancer Network

Mary Bordoni
Personalized Medicine Coalition

Jennifer Bowman
American Clinical Laboratory Association

Pamela Bradley
American Association for Cancer Research

Joel Brill
Predictive Health, LLC

Jeffrey Bush
BD

Khatereh Calleja
AdvaMed

Allen Egon Cholakian
IRDF Project Harvard/Columbia

Deborah Collyar
Patient Advocates in Research

Jeff Cossman
USDS, Inc.

Robin Harper Cowie
Biodesix, Inc.

Claude Desjardins
Johns Hopkins Bloomberg School of Public Health

Patricia Deverka
Center for Medical Technology Policy

Vamil Divan
Credit Suisse

Noel Doheny
Epigenomics, Inc.

Sarah Donovan
Avalere

Georgia Dunston
Howard University

Tania Dutta
Center for Medical Technology Policy

Christopher Earl
Innotrove, LLC

Raith Erickson
Complete Genomics

Lloyd Everson
The U.S. Oncology Network

Kelly Filipski
National Cancer Institute

Andrew Freedman
National Cancer Institute

Gwen Gordon
Personalized Medicine Coalition

Daniel Grosu
Illumina, Inc.

Valerie Hutchins
HillCo HEALTH

Usman Iqbal
Sanofi

Ram Iyer
Inova Health Systems

Leila Jamal
Kennedy Krieger Institute

Elaine Jeter
Palmetto GBA

Brett Johnson
International Cancer Advocacy Network

Roger Klein
University of South Florida School of Medicine

Hon-Sum Ko
U.S. Food and Drug Administration

Nicole Kuderer
Duke University

Hanns Kuttner
Hudson Institute

Gabriela Lavezzari
Express Scripts

Debra Leonard
Weill Cornell Medical Center

Rachel Lindor
Arizona State University College of Law

Tracy Lively
National Cancer Institute

Gary Lyman
Duke University School of Medicine; Duke Cancer Institute

Julie Lynch
U.S. Department of Veterans Affairs

Elizabeth Mansfield
U.S. Food and Drug Administration

Stephen Marmaras
Biotechnology Industry Organization

Robert McCormack
Veridex, LLC

Robert McDonough
Aetna

Mary McGrane
U.S. Department of Agriculture

Lisa McShane
National Cancer Institute

Kala Menon
Booz Allen Hamilton

Donna Messner
Center for Medical Technology Policy

Douglas Moeller
McKesson Health Solutions

Padmaja Mummaneni
U.S. Food and Drug Administration

Bill Murray
Cepheid

David Nelson
EPIC Sciences

Lee Newcomer
United HealthCare Corporation

R. Hannes Niedner
5AM Solutions, Inc.

Danielle Pambianco
HillCo HEALTH

David Parkinson
Nodality

Scott Patterson
Amgen, Inc.

Edmund Pezalla
Aetna

Kathryn Phillips
University of California, San Francisco, Center for Translational and Policy Research on Personalized Medicine

Steve Phurrough
Center for Medical Technology Policy

Margaret Piper
Technology Evaluation Center, Blue Cross Blue Shield Association

Bruce Quinn
Foley Hoag, LLP

Samantha Roberts
Friends of Cancer Research

Jeffrey Roche
Centers for Medicare & Medicaid Services

Greg Rossi
AstraZeneca UK

Drew Saelens
Applied Policy

Howard Scher
Memorial Sloan-Kettering Cancer Center

Richard Schilsky
Comprehensive Cancer Center, University of Chicago

Michele Schoonmaker
Cepheid

Beverlyn Settles-Reaves
Howard University

Paul Sheives
Biotechnology Industry Organization

Naoko Simonds
National Cancer Institute

Mary Lou Smith
Research Advocacy Network

Stuart Spector
Center for Medical Technology Policy

Merianne Spencer
Center for Medical Technology Policy

Sylvia Trujillo
American Medical Association

Sean Tunis
Center for Medical Technology Policy

Jessica Walrath
Friends of Cancer Research

Corinne Warren
Center for Medical Technology Policy

Richard Wenstrup
Myriad Genetic Laboratories

Elizabeth West
FasterCures

Wendy Wifler
Agendia, Inc.

Dara Wright
Affymetrix

Marc Williams
Geisinger Health System

Amy Zettle
Potomac Research Group