Diffusion and Use of Genomic Innovations in Health and Medicine: Workshop Summary

Lyla M. Hernandez, Rapporteur, Roundtable on Translating Genomic-Based Research for Health

ISBN: 0-309-11677-5, 116 pages, 6 x 9, (2008)

DIFFUSION AND USE OF **GENOMIC INNOVATIONS** IN HEALTH AND MEDICINE

WORKSHOP SUMMARY

Lyla M. Hernandez, *Rapporteur*

Roundtable on Translating Genomic-Based Research for Health

Board on Health Sciences Policy

INSTITUTE OF MEDICINE
OF THE NATIONAL ACADEMIES

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

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

NOTICE: The project that is the subject of this report was approved by the Governing Board of the National Research Council, whose members are drawn from the councils of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine. The members of the committee responsible for the report were chosen for their special competences and with regard for appropriate balance.

This project was supported by contracts between the National Academy of Sciences and American College of Medical Genetics (Unnumbered contract); American College of Physicians (Unnumbered contract); American Medical Association (Unnumbered contract); AstraZeneca Pharmaceuticals, Inc. (Unnumbered contract); Blue Cross/Shield Association (Unnumbered contract); Centers for Disease Control and Prevention (CDC) (Contract No. 200-2005-13434); College of American Pathologists (Unnumbered contract); Department of Veterans Affairs (VA) (Contract No. V101(93) P-2238); Eli Lilly and Company (Contract No. LRL-0028-07); Food & Drug Administration (Contract No. 223012460); Genetic Alliance (Unnumbered contract); Genomics Health, Inc. (Unnumbered contract); GlaxoSmithKline, Inc. (Unnumbered contract); Health Systems Research, Inc. (Contract No. 07-H0116); National Human Genome Research Institute (Contract No. N01-OD-4-2139, TO#189); National Institute of Child Health and Human Development (Contract No. N01-OD-4-2139, TO#189); National Society of Genetic Counselors (Unnumbered contract); Secretary's Advisory Committee on Genetics, Health and Society (Contract No. N01-OD-4-2139, TO#189); and United Health Care (Unnumbered contract). Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the views of the organizations or agencies that provided support for the project.

International Standard Book Number-13: 978-0-309-11676-3
International Standard Book Number-10: 0-309-11676-7

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

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

Printed in the United States of America

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

Suggested citation: IOM (Institute of Medicine). 2008. *Diffusion and use of genomic innovations in health and medicine: Workshop summary.* Washington, DC: The National Academies Press.

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

INSTITUTE OF MEDICINE
OF THE NATIONAL ACADEMIES

Advising the Nation. Improving Health.

THE NATIONAL ACADEMIES

Advisers to the Nation on Science, Engineering, and Medicine

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

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

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

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

www.national-academies.org

PLANNING COMMITTEE ON DIFFUSION AND USE OF GENOMIC INNOVATIONS IN HEALTH AND MEDICINE*

WYLIE BURKE, M.D., Ph.D. (*Chair*), Professor and Chair, Department of Medical History and Ethics, University of Washington, Seattle

NAOMI ARONSON, Ph.D., Executive Director, Technology Evaluation Center, BlueCross/BlueShield Association, Chicago, Illinois

MOHAMED KHAN, M.D., Ph.D., Associate Director of Translational Research, Department of Radiation Medicine, Roswell Park Cancer Institute, Buffalo, New York

STEPHEN G. RYAN, M.D., Executive Director, Discovery Medicine and Epidemiology, AstraZeneca Pharmaceuticals, Wilmington, Delaware

KEVIN SCHULMAN, M.D., 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, North Carolina

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

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

ROUNDTABLE ON TRANSLATING GENOMIC-BASED RESEARCH FOR HEALTH*

WYLIE BURKE, M.D., Ph.D. (*Chair*), Professor and Chair, Department of Medical History and Ethics, University of Washington, Seattle
STEPHEN ECK, M.D., Ph.D., Vice President, Translational Medicine & Pharmacogenomics, Eli Lilly and Company, Indianapolis, Indiana
FAITH T. FITZGERALD, M.D., Professor of Medicine, Assistant Dean of Humanities and Bioethics, University of California, Davis Health System, Sacramento
GEOFFREY GINSBURG, M.D., Ph.D., Director, Center for Genomic Medicine, Institute for Genomic Sciences & Policy, Duke University, Durham, North Carolina
ALAN E. GUTTMACHER, M.D., Deputy Director, National Human Genome Research Institute, National Institues of Health, Bethesda, Maryland
R. RODNEY HOWELL, M.D., Special Assistant to the Director, National Institute of Child Health and Human Development, Bethesda, Maryland
KATHY HUDSON, Ph.D., Director, Genetics and Public Policy Center, Berman Bioethics Institute, Johns Hopkins University, Washington, District of Columbia
SHARON KARDIA, Ph.D., Director, Public Health Genetic Programs, Associate Professor, Department of Epidemiology, University of Michigan, School of Public Health, Ann Arbor
MOHAMED KHAN, M.D., Ph.D., Associate Director of Translational Research, Department of Radiation Medicine, Roswell Park Cancer Institute, Buffalo, New York
MUIN KHOURY, M.D., Ph.D., Director, National Office of Public Health Genomics, Centers for Disease Control and Prevention, Atlanta, Georgia
ALLAN KORN, M.D., Chief Medical Officer, Senior Vice President Clinical Affairs, BlueCross/BlueShield Association, Chicago, Illinois
DEBRA LEONARD, M.D., Ph.D., Professor and Vice Chair for Laboratory Medicine, Director of the Clinical Laboratories for New York-Presbyterian Hospital, Weill Cornell Medical Center of Cornell University, New York

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

MICHELE LLOYD-PURYEAR, M.D., Ph.D., Chief, Genetic Services Branch, Health Resources and Services Administration, Rockville, Maryland

ROBERT L. NUSSBAUM, M.D., Chief, Division of Medical Genetics, University of California, San Francisco, School of Medicine

TIMOTHY O'LEARY, M.D., Ph.D., Director, Biomedical Laboratory Research and Development Service, Director, Clinical Science Research and Development Service, Department of Veterans Affairs, Washington, District of Columbia

AMELIE G. RAMIREZ, Dr.P.H., Dielmann Chair, Health Disparities and Community Outreach Research, Director, Institute for Health Promotion Research, University of Texas Health Science Center at San Antonio

ALLEN D. ROSES, Ph.D., 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, North Carolina

STEPHEN G. RYAN, M.D., Executive Director, Discovery Medicine and Epidemiology, AstraZeneca Pharmaceuticals, Wilmington, Delaware

KEVIN SCHULMAN, M.D., 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, North Carolina

PATRICK TERRY, Director, Consumer Advocacy and Government Affairs, Genomic Health, Inc., Washington, District of Columbia

SHARON TERRY, President and CEO, Genetic Alliance, Washington, District of Columbia

STEVEN TEUTSCH, M.D., Ph.D., Executive Director, U.S. Outcomes Research, Merck & Co., Inc., West Point, Pennsylvania

MICHAEL S. WATSON, Ph.D., Executive Director, American College of Medical Genetics, Bethesda, Maryland

CATHERINE A. WICKLUND, M.S., CGC, President, National Society of Genetic Counselors, Associate Director, Graduate Program in Genetic Counseling, Assistant Professor, Department of Obstetrics and Gynecology, Northwestern University, Chicago, Illinois

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

IOM Staff

LYLA M. HERNANDEZ, M.P.H., Project Director
ERIN HAMMERS, M.P.H., Research Associate
ALEX REPACE, B.S., Senior Project Assistant

IOM Anniversary Fellow

LISA BARCELLOS, Ph.D., Assistant Professor, University of California, Berkeley

BOARD ON HEALTH SCIENCES POLICY*

*IOM boards do not review or approve individual workshops and are not asked to endorse conclusions and recommendations. The responsibility for the content of the summary rests with the workshop rapporteur and the institution.

E. ALBERT REECE, M.D., Ph.D., Vice President for Medical Affairs, Dean, School of Medicine, University of Maryland School of Medicine, Baltimore
LINDA ROSENSTOCK, M.D., M.P.H., Dean, School of Public Health, University of California, Los Angeles
MICHAEL J. WELCH, Ph.D., Professor of Radiology, Co-Director, Division of Radiological Sciences, Department of Radiology, Washington University School of Medicine, St. Louis, Missouri
OWEN N. WITTE, M.D., Investigator, HHMI, President's Chair in Developmental Immunology, David Geffen School of Medicine, University of California, Los Angeles

IOM Staff

ANDREW M. POPE, Ph.D., Director
AMY HAAS, Board Assistant
DONNA RANDALL, Financial Officer

Reviewers

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

Geoffrey Ginsburg, M.D., Ph.D., Director, Institute for Genomic Sciences & Policy, Duke University, Durham, North Carolina
David R. Nerenz, Ph.D., Director, Center for Health Services Research, Henry Ford Health System, Detroit, Michigan
Bernard M. Rosof, M.D., Senior Vice President for Health Affairs, North Shore-Long Island Jewish Health, Great Neck, New York
Sandra Suther, Ph.D., Assistant Professor, Behavioral Science & Health Education, Institute of Public Health, Florida A&M University, Tallahassee, Florida

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

Bradford H. Gray, Ph.D., Editor, *The Milbank Quarterly,* Principal Research Associate, The Urban Institute. Appointed by the Institute of Medicine, he was responsible for making certain that an independent examination of this report was carried out in accordance with institutional procedures and that all review comments were carefully considered. Responsibility for the final content of this report rests entirely with the author and the institution.

Contents

FIGURES

TABLES

1

Introduction[1]

The sequencing of the human genome has generated excitement about the potential of genomic innovations to improve medical care, preventive and community health services, and public health. Until fairly recently, genetic information was used primarily in the diagnosis of relatively rare genetic diseases, such as cystic fibrosis and Huntington's Disease, but a transformation in the use of genetic and genomic information is under way.

Genetic markers of increased risk for such chronic diseases as diabetes and coronary artery disease have been identified. Research on how genes influence the effects of drugs holds promise for helping physicians individualize drug therapy. Tests designed to help providers make treatment decisions based on variations in a patient's genome are being developed. The Department of Health and Human Services has launched a Personalized Health Care Initiative, one goal of which is to "link clinical and genomic information to support personalized health care"[2] (DHHS, 2007). It is anticipated that "genetic prediction of individual risks of disease and responsiveness to drugs will reach the medical mainstream in the next decade or so" (Collins and McKusick, 2001). To date, however, few of these promising discoveries have resulted in actual applications in medicine and health (Burke et al., 2006).

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

[2]Personalized health care, as defined by the Department of Health and Human Services, refers to medical practices that are targeted at individuals based on their specific genetic code in order to provide a tailored approach (www.hhs.gov/myhealthcare/glossary/glossary.html).

In 2007 the Institute of Medicine established the Roundtable on Translating Genomic-Based Research for Health. The purpose of the Roundtable is to foster dialogue and discussion that will advance the field of genomics and improve the translation of research findings to health care, public health, and health policy. As a first step in examining issues of translation of genomic innovations, the Roundtable decided to hold a workshop to gather information on three questions below. Information obtained from the workshop was then used to further discussion and exploration of the answers to these questions:

1. Are there different pathways by which new scientific findings move from the research setting into health care?
2. If so, what are the implications of those different pathways for genomics?
3. What can we learn from the translation of other new technologies as we seek to understand the translation of genome science into health care?

The December 4, 2007, workshop was moderated by Wylie Burke, chair of the Roundtable, and consisted of panel presentations in four areas: the process of translation of innovations, practical incentives and barriers to translation, translation of genomic technology at the clinical level, and opportunities and constraints for translation both within the United States and globally. A discussion period followed each panel. At the conclusion of the meeting Burke offered a summary of the day's presentations. While various types of genomic innovation were discussed, a number of presentations focused primarily on genomic testing technologies. The complete agenda can be found in Appendix A, and biographical sketches of the speakers are in Appendix B.

The following report summarizes speaker presentations and discussions. Any conclusions reported should not be construed as reflecting a group consensus, rather they are the statements and opinions of presenters and participants.

2

Translation of Innovations

A BROAD PERSPECTIVE

Robert M. Califf, M.D., MACC
Duke Translational Medicine Institute and
Duke University Medical Center

Biomedical science is advancing at an amazing rate, yet the translation of that science into better health outcomes has not kept pace. Much of this lag is due to non-technological reasons, including financing, regulation, and cultural issues. Another factor is that the rewards for researchers who promote innovation are increasingly disconnected with the healthcare needs of society at large.

Translation is a fragmented and uncoordinated process that, with few exceptions, takes 25 to 30 years from initial scientific discovery to the delivery of a therapy to the people who benefit most (Figure 2-1). While basic discoveries occur predominantly in academic medical centers funded by the National Institutes of Health (NIH), the process of translating these discoveries almost always begins in the medical products industry, where a basic discovery is followed up with a period of specifically directed preclinical activity intended to test whether the putative therapeutic target is indeed viable. The next step is determined by a decision-making process that comprises multiple steps and includes assessments that link financial support with the probability of success; if the decision is to move forward, then the next stage of development is undertaken by clinical research orga-

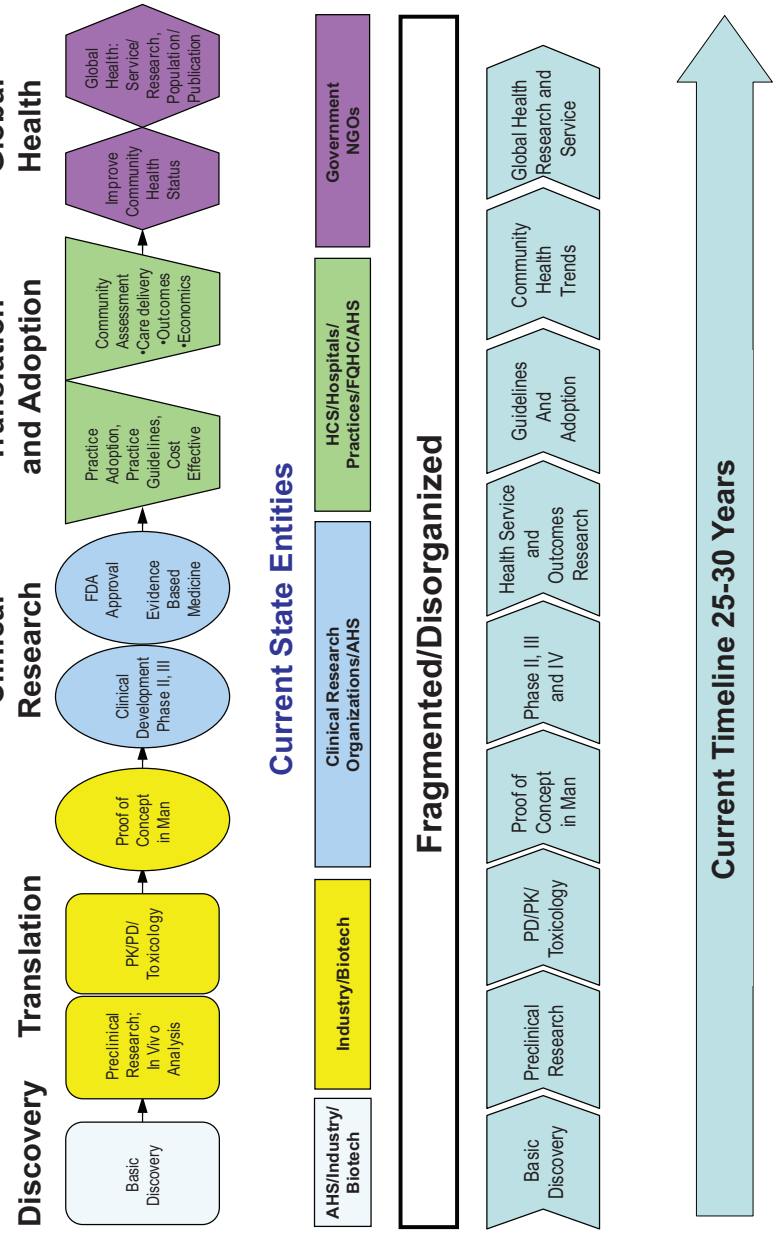


FIGURE 2-1 Translation of innovations.
SOURCE: Califf, 2007.

nizations from the medical products industry, contract research organizations, or academia.

The early period of human subjects research, commonly called "proof of concept" or phase I/IIa, is characterized by the introduction of novel therapies into either healthy volunteers or a carefully selected group of patients; if there are no red flags, this work is followed by a comprehensive set of clinical studies, known as phase III trials. Data from these phase III trials are then used by the U.S. Food and Drug Administration (FDA) and other national and international regulatory bodies to make decisions—based on criteria that vary depending on which division of the FDA is involved or which country is doing the evaluation—about whether the therapy is ready to be introduced into clinical use. After a therapy is approved, it is supposed to reach the appropriate people in the approved manner through a competitive system that includes health systems, hospitals, clinical practices, purchasers, and sales representatives for the product or technology. Ultimately, when the therapy's patent protection expires, its price will diminish, and the health of the entire community will benefit from the wider access thus afforded.

This system has generally worked well up to now, as evidenced by the steady decline in mortality in the United States since 1900, a decline only briefly interrupted by the 1918 flu pandemic. And while much of the decline during the first half of the 20th century was due to clean water, sewers, antibiotics, and better nutrition leading to a reduction in mortality from infectious diseases, a significant proportion of the decline since then has been attributable to advances in treatment, with the prevention of infant mortality and the treatment and prevention of cardiovascular disease playing the largest role.

Despite these achievements, however, key issues must still be addressed concerning the translation of scientific innovations into effective therapeutics. We now have information systems capable of providing detailed data on leading causes of death and disability, for example, and these data show that the benefits of technological advances have not been evenly distributed (Figure 2-2). Such information can be helpful in identifying new directions in which to focus the efforts of the translational enterprise.

Challenges Facing Translational Medicine

Our current general scheme of focusing on discovery science in academic centers and trusting for-profit industry to handle the diffusion of technology continues to be the most sensible path to follow. But along that path are major hurdles that must be cleared, particularly at the translational interfaces between discovery and commercialization and between commercialization and public health. In the arena of drugs and biologics, for

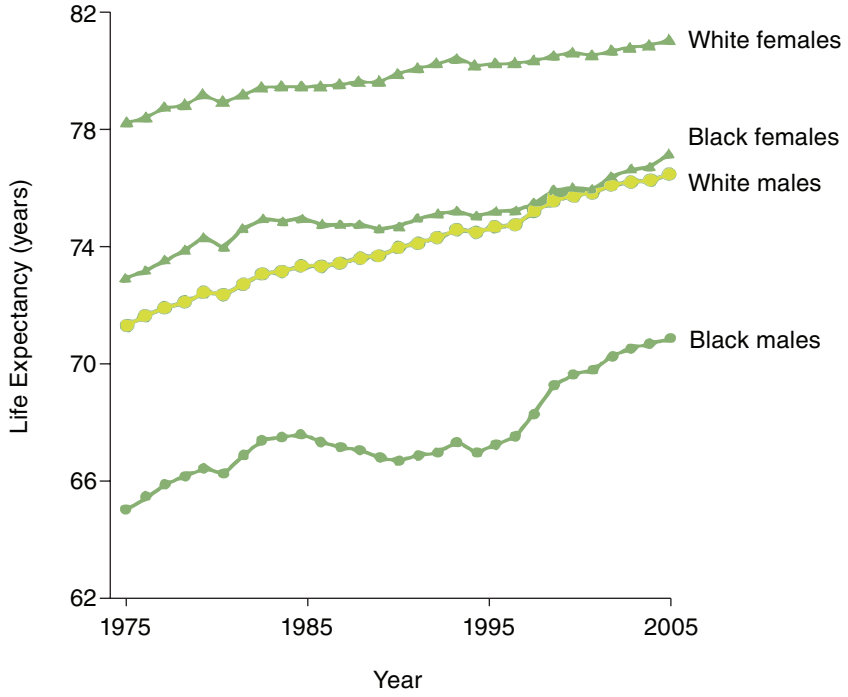


FIGURE 2-2 Life expectancy at birth.
SOURCE: Adapted from Harper et al., 2007.

instance, although novel targets afford bigger potential returns on investment, investors often shy away from them because of the risks entailed. Pursuing an already-proven target gives a much higher probability of success, which causes "follow-ons" to be seen as a better bet on average and leads investors to often—and understandably—choose the safer option. The net effect of these considerations is a risk-averse industry that pursues fewer novel, innovative pathways.

In the arena of genomics-based diagnostic testing and therapeutic decision making, for instance, the intersection of diagnostic testing and therapeutics is plagued with regulatory ambiguity, and the prospects for reimbursement are unsure. Such uncertainty directly affects willingness to invest. In terms of health services, enormous investment will be required to change current practices. Forces that encourage change in health care services (i.e., the Internet, consumerism, information technology, -omics,[1]

[1]-omics refers to a biological field of study that ends in the suffix *omics*, for example, genomics, proteomics, metabolomics.

medical technology, and Congress) are offset by countervailing pressures (i.e., regulation, financing, a fragmented marketplace, professional autonomy, and, once again, Congress). Many observers believe that these forces have created an equilibrium that discourages innovation, but there is no consensus about how that equilibrium can be changed while still maintaining the fundamental safety net created by the regulation of technologies through objective, empirical assessment of the balance of risk and benefit.

The high cost of developing a new product is one example of the difficulties facing innovation. A study conducted in 2003 by DiMasi and colleagues found that research and development costs for a new drug in the United States averaged a total of $800 million in 2000 dollars, up sharply from the estimated $231 million that such research and development cost in 1987 (in 1987 dollars) (DiMasi et al., 2003). The most recent published data provide an astonishing estimate of $1.4 billion per successfully developed drug. An important component of this figure is the cost of capital during the protracted period of drug development.

Unfortunately, the U.S. clinical research system is increasingly recognized as a bottleneck in the process of therapeutic development, as clinical research takes longer and is measurably more expensive to accomplish in the United States than in other countries, while the quality of the research itself may be inferior to that conducted in other parts of the globe. Furthermore, the application of therapies in the United States is measurably inefficient—not only are the costs of the therapies much higher here than in other countries, but the therapies have inferior results in terms of longevity and functionality of the population.

Another potential deterrent to innovation exists at the level of practice. The movement toward evidence-based medicine has pushed practitioners to have evidence for what they are doing. On balance this is clearly a favorable development. It gives patients and consumers much more confidence that the treatments they receive are appropriate to their needs and that they are administered correctly. The demand for evidence, however, can have a stultifying effect on innovation if it is employed ineffectively and without the application of modern methods and scientific insight.

Incentives should be developed to foster innovation. The current U.S. health care system has many incentives to seek efficiency in the delivery of technologically sophisticated, expensive approaches for those who can afford them. There is a great disincentive, however, to providing low-cost, efficient health care to the people who are experiencing most of the death and disability in the United States. Despite astounding advances in biology, ensuring that innovations reach those members of society who stand to benefit most from them—and thus that these innovations will have the largest possible effect on the rates of death and disability—is proving especially difficult.

Overcoming Translational Blocks

Along the translational pathway there are several blocks that slow progress from the identification of a potential biological system that could be attacked as a target to the translation of that concept into the first human studies. First, the high levels of risk involved in the process limit investment interest. Second, there is a large gap between scientific advances and the regulatory science needed to predict and evaluate product performance. Third, decision making is dominated by anecdote and intuition. In order to make a prediction about the success of a possible therapy, one must know what has succeeded and what has failed in the past and then use that information to understand the probabilities of success or failure in general. If only successful efforts are made public, however, there is little basis for understanding and determining which general approaches lead to greater success and thus for figuring out where to invest efforts and funding.

The Critical Path Initiative

This lack of data about the factors that underlie the success or failure of development efforts is a major motivating factor for the FDA's Critical Path Initiative,[2] which aims to create a "safe haven" for sharing knowledge that can accelerate translation while at the same time doing nothing to impair the drive for competitive advantage that stimulates creativity in our system.

The concepts of *pre-competitive* and *pro-competitive* spaces are key to understanding the strategy underlying the Critical Path Initiative. Generally speaking, pre-competitive knowledge advances a field as a whole before the point at which competition based on proprietary knowledge comes into play. An example of pre-competitive knowledge would be general knowledge about the operating characteristics of standard tests for pre-clinical toxicity required by the FDA. Currently, little is known about the true predictive value of these tests because abandoned projects are rarely discussed and almost never published, leaving an incomplete database of test results that renders any calculations about the value of the tests meaningless.

The pro-competitive space is characterized by mutual efforts toward development of new knowledge that in the past would have been proprietary but that, through collaboration, confers an equal advantage to all interests. An example would be a generally known biomarker that everyone

[2] "The Critical Path Initiative is FDA's effort to stimulate and facilitate a national effort to modernize the scientific process through which a potential human drug, biological product, or medical device is transformed from a discovery or 'proof of concept' into a medical product" (FDA, 2006).

can use. Individual companies usually do not have enough biological and clinical data to validate a biomarker, but a consortium of companies and academic institutions may be able to do so. Companies that make best use of publicly available information about the biomarker in developing therapeutics would be the ones to receive an advantage.

Continuing on the translation pathway illustrated in Figure 2-1, the next step is early-phase human studies. Many discoveries fail at this stage because of unanticipated off-target effects that are only detected in much later phase testing. A major recent example was the case of torcetrapib, a drug developed to treat abnormally low HDL cholesterol and prevent cardiovascular disease (Nissen et al., 2007). Torcetrapib failed in phase III trials, perhaps because of an unrecognized and completely unanticipated aldosterone-producing effect.

To identify these types of off-target effects before they cause harm to participants in large-scale clinical trials, it will be necessary to study human systems biology in greater detail. The traditional approach to early-phase human subjects research used in the pharmaceutical industry today (measuring pharmacokinetics, pharmacodynamics, and adverse events) does not address this problem, and a new approach that uses experimental medicine units capable of highly detailed systems measurement in human subjects is needed. Researchers will need to use modern technologies, such as gene expression analysis, proteomic and metabolomic profiling, and functional imaging, to study integrated physiology more effectively.

Once early-phase human studies have been conducted, research efforts move to the larger clinical trials. There seems to be a general assumption that we know how to conduct these clinical trials effectively. To the contrary, clinical trials are too expensive, too slow, and too often of doubtful quality. In fact, there are no standard definitions of quality for different types of trials (Baigent et al., 2008). Five years ago, a typical phase III trial in cardiovascular disease cost about $80 million to $140 million (Eisenstein et al., 2005, 2008). Currently many trials cost $300 million to $400 million, or even more. Such exorbitant costs become an inhibiting factor for therapeutic areas that require definitive data as a precondition to marketing.

The FDA Critical Path Initiative is seeking to transform the clinical research enterprise through the Clinical Trials Transformation Initiative. The goals of this project are to enhance knowledge and standards that improve the quality of clinical trials while eliminating practices that increase costs but provide no value in return (CTTI, 2007). Key players in these efforts include the FDA, industry, academia, patient advocates, and non-academic clinical research professionals.

Post-Marketing Research

Once a product has been approved for marketing and is released into the marketplace, it is still necessary to generate substantial additional evidence about the balance of safety and effectiveness in the post-marketing phase. Unfortunately, there is almost no money to support such research, which has the primary goal of improving the public health. Most funding for post-marketing studies comes from the company that markets the product, and most such trials are designed to expand the market for the product and thus to bolster its expected financial value to the company. Indeed, the decision about which studies to conduct is usually based on net present-value calculations, and a trial's sponsor will approve funding only if there is a high pre-test probability that the trial will lead to a desirable result. While these studies may give honest answers to the questions asked, the questions about translation that get asked under the current system are not the ones that would be asked if the welfare of the general public were the major concern. The Reagan-Udall Foundation, which was recently created as part of FDA renewal legislation, offers a public-private partnership to provide a venue in which such public-focused studies can be designed, but political maneuvering has so far blocked funding for this effort.

The endpoint of the translation pathway illustrated in Figure 2-1 is public and global health. There is a growing convergence between national healthcare issues and global ones. As is the case in the United States, financial incentives in many other countries emphasize practices that focus on expensive technology that benefits "paying customers," while incentives to provide basic health services receive less emphasis even as the understanding of ways to meet those basic health needs improves.

In Durham, North Carolina, with funding from the NIH's Clinical and Translational Science Awards, a study is underway whose goal is to develop a deeper understanding of the issues surrounding the delivery of basic health needs. Under the current reimbursement system, there is a tension between financial considerations and public health consideration in decisions about where to locate health clinics. In particular, the sites that are likely to result in profitable practices are not where the clinics would be located if the goal was to improve the overall health of Durham County, given that the greatest burden of death and disability is located in neighborhoods in which reimbursement is most adverse for provision of health-related services. Plans are now underway to harness geospatial-temporal mapping (Miranda et al., 2005), embedded personal health records, disease registries, and wireless monitoring capabilities in order to deploy low-cost technologies capable of delivering better, more affordable health care to the people who need it most. Providing incentives to develop technology aimed at serving people and neighborhoods with the greatest burden of premature

death and disability would result in an enormous redirection of innovative efforts. Indeed, the *New York Times* recently reported that the disparities in health outcomes as a function of education and income are widening rather than narrowing in the United States (Pear, 2008).

Positive Change: The Pediatric Exclusivity Program

Change, however, is possible, and the Pediatric Exclusivity Program provides a heartening example of how incentives for change can be created (FDA, 2005). In the 1990s the pediatric community became increasingly aware that many therapies used in children had no empirical data establishing their safety and effectiveness. The problem had its roots in a general sense that clinical trials in children were too risky; this community view in turn reinforced the reluctance of drug and device companies to engage in such trials. However, a determined coalition worked together to create legislation granting patent extensions to companies that agreed to evaluate their technologies by performing appropriate trials in children.

Since this program began in 1997, there has been a substantial increase in drug research for pediatric indications in addition to 138 labeling changes. Li and colleagues performed a meta-analytical study aimed at quantifying the economic return to industry for 6 months of pediatric exclusivity (Li et al., 2007). Nine drugs were studied, and net economic return and net–return-to-cost ratios were calculated. The median cost per written request[3] was $12.34 million. Net economic returns (minus $8.9 million to $507.9 million) and net minus return-to-cost ratios (minus 0.68 to 73.63) were highly variable, but, on balance, the net economic return to industry was favorable.

Benjamin and colleagues performed a meta-analysis of clinical trials completed for pediatric exclusivity in order to quantify the dissemination of study results (Benjamin et al., 2006). They evaluated 253 studies submitted to the FDA from 1998 to 2004. Of these, only 113 were published, and efficacy studies and trials that resulted in desirable labeling changes were most likely to be published. Unfortunately, a number of the negative findings received little or no attention in the pediatric community. Nonetheless, these studies represent a positive development. Prior to this program, many in the research community asserted that clinical trials in children were not practicable. Once the incentive was put in place, however, trials were indeed undertaken.

[3]The FDA issues the written request to the company. The written request describes in detail the studies needed to be eligible for pediatric exclusivity and the time frame for completion of those studies. A written request contains the indication, number of studies, sample sizes, and trials design required for eligibility.

Califf concluded by asking, if these incentives work, how can we deploy them in order to achieve the goals most crucial to the broad and equitable diffusion of biomedical innovations in society?

UNDERSTANDING TYPES OF INNOVATION AND IMPLICATIONS FOR POLICY

Kevin Schulman, M.D.
Duke University

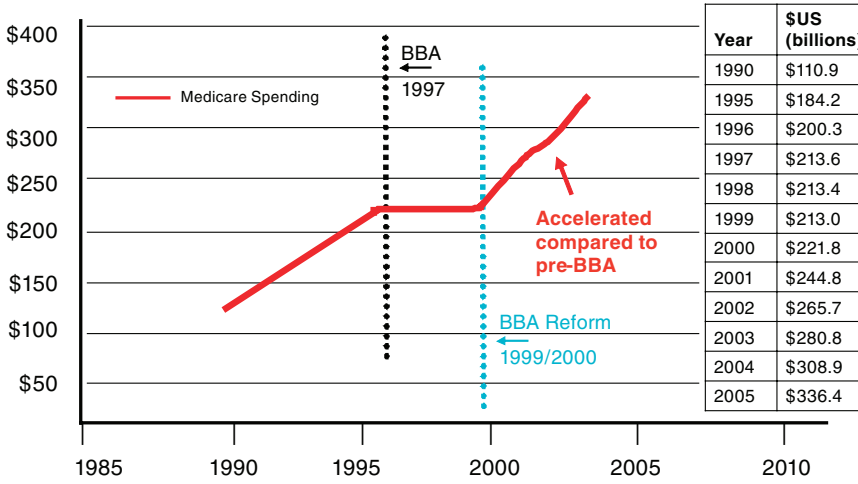
According to current estimates, by 2030 about 52 percent of the entire federal budget will be required to fund the Social Security and Medicare programs. Given the annual cash flow deficits in both Social Security and Medicare, these programs will be underfunded by $2 trillion by 2030. By mid-century underfunding will reach $7 trillion (Rettenmaier and Saving, 2004). The public policy debate has not yet faced the fact that there is not enough money to support these programs; yet this is a situation that needs to be discussed and debated soon.

How should one approach the issues of deficits and underinvestment in these programs? In the mid-1990s there was a crisis concerning the escalating health costs for Medicare. It was believed that the easiest way to fix the problem was to freeze Medicare spending, so the Balanced Budget Act was implemented. This was effective through the three years of the act, but, as shown in Figure 2-3, once the restrictions were removed, spending continued to increase, and the slope of that increase was steeper than before the act had been implemented (CMS, 2007). The market response to the policy of freezing expenditures was an unexpected acceleration in the costs of the program once controls were removed.

If controls did not work in the late 1990s, they most likely will not solve today's cost issues, will not solve quality issues, and will likely make things worse. So what course can be pursued?

Clay Christensen examined the role of innovation in the computer disk-drive industry and put forth some ideas that have relevance to a discussion on innovation in health care. He describes the process of innovation as resulting from entry of new firms and new business models in the marketplace. In his analysis, entry results from opportunities created when products outstrip the needs of the majority of the marketplace.

There is a distribution of demand by consumers for any new technology or innovation, and this distribution can be thought of as following a normal distribution. At the leading edge are early adopters, most of the popula-



Year	$US (billions)
1990	$110.9
1995	$184.2
1996	$200.3
1997	$213.6
1998	$213.4
1999	$213.0
2000	$221.8
2001	$244.8
2002	$265.7
2003	$280.8
2004	$308.9
2005	$336.4

FIGURE 2-3 Policy response: A budget freeze.
SOURCE: Adapted from CMS, 2007.

tion is in the middle, and there are some who are late adopters. The early adopters are of great interest to large firms since they have high demand for technology and innovation and are thought to be relatively price insensitive, Schulman said. Furthermore, the capacity of the majority of the market to use new technology increases at a much slower rate than the capacity of the early adopters.

For a technology company the high demand group is a great part of the market to satisfy because, if the company can develop products or services to meet the demands of this group, their products or services can be very profitable. As firms evolve their products to meet the demands of this specific subset of the population, however, an interesting phenomenon occurs. As a result of meeting the needs of the early adopters, the technology develops in such a way as to outperform the requirements of the majority of the market. In this situation, there is a gap between the performance of the existing technology and the needs of the majority of the market. This gap creates the opportunity for new firms to bring new products to the market that might be more limited in scope than the existing technology but might be a better match on price and quality for an important part of the market. Over time these new companies actually begin to meet the needs of the general population.

These two types of firms move ahead through two different types of advances. The first type of firm, the original innovator firm, moves forward

through sustained technology improvement. This type of innovation is called sustaining innovation. The second type of firm, which creates a new product and enters the market established by the first type, is called the "disruptive innovator." Joseph Schumpeter wrote about this phenomenon in 1911, saying, ". . . as a rule the new does not grow out of the old but appears alongside of it and eliminates it competitively . . ." (Schumpeter, 1911). The net result of this process of innovation is the creation of higher quality, lower cost products over time. While this is generally accomplished through the entry to the market of new firms with new business models, originator firms can respond to these threats. Technology innovation is a fundamental part of the market. In fact, in most markets, technology and organizational innovation drive cost and quality improvement.

How can these concepts be applied to health care? One of the things that distinguishes health care from, for example, the disk-drive industry that Christensen studied, is that the health care industry is regulated, with different aspects of it regulated to different degrees. A sustaining innovator in health care is above the regulatory barrier, that is, it has met the regulatory requirements. By contrast, the disruptive innovator that would like to enter the marketplace is below the regulatory barrier. Therefore, while the space for the disruptive innovator to enter the market is theoretically available, the regulations can deter entry.

Imagine, for example, a new player wanting to enter the highly regulated hospital market and compete with Duke Hospital, which is very profitable and also one of the most expensive hospitals in the country. A new competitor to the field would have to have billions of dollars to become an innovative competitor to Duke Hospital. In reality, therefore, the space for a disruptive innovator does not exist. The administrative barriers as well as the regulatory barriers effectively bar disruptive innovation.

Not all types of innovation are of equal interest from a policy perspective. From that perspective, there is a strong desire for innovation, but there is a willingness to pay a premium only for those innovations with the potential to be disruptive innovation. Since the policy goal is to improve quality and reduce costs, an implicit policy goal should be to encourage disruptive innovation and market entry to achieve this goal.

In practice, however, current medical reimbursement strategies reward the sustaining innovators with premiums, making it potentially very difficult for disruptive innovators to enter the health care market. Of course, one difficulty in encouraging disruptive innovation is that it is hard to determine in advance which technologies have the potential to become disruptive innovators.

There is an urgent need to better understand the relationship between incentives and market entry in order to foster technology innovation. To determine where to place incentives, one must first decide what innovations

are. Is molecular structure an innovation? Is the mechanism of action an innovation? What about the mode of delivery of an innovation? Is the fact that -omics is involved in some way an innovation? Is therapy that alters a treatment plan an innovation? It is important to answer these questions since the answers will shape the types of technology that are brought to the marketplace.

Nexium, a product used to treat heartburn and acid reflux disease, is one of the largest-selling drugs in the United States, with sales of over $5 billion in 2006 (Astrazeneca, 2006). It is an isomer of a previous product. It is an example of what Califf referred to earlier as a follow-on product. Is it a disruptive or sustaining innovation? Should this determination enter into price negotiations?

Another aspect of innovation that needs to be explored is the relationship between organizational and service innovation. What is service innovation, and can technology be a platform for that?

There are several types of organizational innovation. These are generally firm responses to competitive threats from market entry and this process is how originator firms respond to disruptive innovation. Firms can respond at several levels to new product or service creation, and some can adapt their business model to a new market environment over time. Many types of organizational innovation involve the development or acquisition of new business models. A good example of innovation in an internal exploratory environment is the Lockheed Skunk Works[4]—the place where many of the firm's new innovations and plans come from.

Corporate venture-capital companies make investments in small firms to acquire insights into new business models. They can also make acquisitions, especially exploratory acquisitions, to acquire new products or services. This process also involves divesting older models and older technologies. Interestingly, in terms of regulatory barriers, health information technology is one of the few areas in health care where there are not yet any regulatory barriers to disruptive innovation.

Schulman concluded by saying that the cost and quality pressures in health care are enormous and increasing. The easiest response is to freeze the system and lock in the status quo; the result of any such action is likely to be disappointing since this prevents organizational innovation in the

[4] "Skunk Works refers to both the division of the same name within the Lockheed Martin corporation and the organizational model popularized by that division's success at managing time-sensitive, complex projects. The latter sense is used in engineering and technical fields to describe a group within an organization given a high degree of autonomy and unhampered by bureaucracy, tasked with working on advanced or secret projects. The term is also used analogously in other fields, especially business, to describe any self-contained, semi-autonomous work-group or committee that directly manages its own projects" (http://en.wikipedia.org/wiki/Skunk_works, accessed January 18, 2008).

market. What is needed is a better understanding of the role of technology and of organizational innovation in the broader economy and especially in health care. If certain types of innovation can provide a solution to problems of cost and quality, then they should be part of the policy debate. If certain types of innovation can provide a solution, especially -omics, these efforts must be supported with strong market and policy messages. As stated earlier, there is an urgent need to better understand the relationship between incentives and market entry in order to foster technology innovation.

LESSONS FOR GENOMICS FROM OTHER TECHNOLOGIES

Annetine Gelijns, Ph.D.[5]
Columbia University

Advances in genetics have led to a remarkably improved understanding of the genetic and molecular basis of disease, and these advances are increasingly leading to the development of new interventions in such areas as genetic testing, gene-based therapy, and pharmacogenomics. These advances permeate life and even art (e.g., the catalogue of the Museum of Modern Art offers a framed print of one's own DNA). Advances in genetics also highlight the importance of the diffusion of innovations as well as the issue of how best to manage the challenges inherent in adopting and using genomic interventions.

Research into technological diffusion finds that diffusion typically follows an S-shaped course, with adoption proceeding slowly at first, then accelerating, and then slowing down as the saturation point is reached (Griliches, 1957). There are several factors that affect the speed at which diffusion occurs. The first of these factors is the characteristics of the technology itself. These characteristics include such things as available alternatives, the marginal benefits that the new technology offers, the severity and prevalence of the target illness, and the costs and complexity of adopting the technology. This last characteristic is a particularly important consideration for genomics technologies.

A second factor affecting the speed of diffusion is the collection of regulatory agencies and payers; these have become increasingly important

[5]This presentation was developed collaboratively by Annetine Gelijns, Ph.D., Alejandra Guerchicoff, Ph.D., Deborah D. Ascheim, M.D., Lawrence D. Brown, Ph.D., and Alan J. Moskowitz, M.D.

gatekeepers of the diffusion process in health care. A third factor is the characteristics and interests of potential adopters. For some health care technologies, physicians are the sole decision makers regarding adoption. For other technologies—liver transplant programs, for example—hospital administrators and boards of trustees become involved in the decision-making process. Finally, various economic, sociocultural, and ethical factors powerfully shape the diffusion process as well.

In the diffusion literature a technology is generally perceived as being static or constant; however, innovations continue to evolve as they enter clinical practice. As a result, decisions about adopting a technology are made in the face of considerable uncertainty about indications, populations, risks, and effectiveness. In recent years, the stakeholders—for examples, the FDA, payers, physicians, or patients—have sought more rigorous evidence to help guide adoption decisions. Each stakeholder brings its own distinct perspective to decisions that have major implications for quality, cost, and fairness, a fact that highlights the importance of understanding the preferences and the values of stakeholders.

It is only after a new technology is put into actual use in clinical practice that there can be significant downstream learning and innovation. Such learning and innovation falls into three broad categories.

First, after a new technology is put into practice, the medical profession typically refines the patient selection criteria within a given disease category. Coronary artery bypass graft surgery (CABG) is a case in point. Only four percent of patients treated with CABG a decade after its introduction would have met the eligibility criteria of the trials that established its initial value. These initial trials excluded the elderly, women, and patients with a range of comorbidities, all of whom are recipients of CABG today.

Second, the process of post-marketing innovation also includes the discovery of totally new and often unexpected indications for use. The history of pharmaceutical innovation is replete with such discoveries, such as happened, for instance, with alpha blockers. These were first introduced for hypertension, but 20 years later they are an important agent for the treatment of benign prostatic hypertrophy. The discovery of such new indications of use is an important public health and economic phenomenon and accounts, for example, for nearly half of the overall market for blockbuster drugs. Gelijns commented that it will be interesting to see how the introduction of pharmacogenomics might affect this dynamic.

The third type of downstream learning is the way that physicians gain knowledge about integrating a technology into the overall management of their patients. For example, the left ventricular assist device (LVAD) was approved by the FDA and the Centers for Medicare and Medicaid Services (CMS) in 2003 for end-stage heart-failure patients who were ineligible for

cardiac transplantation. After the device was approved, clinicians modified the operating technique in various ways. They discovered new ways to prevent infections, and they made changes in anticoagulation regimens. These changes led to a reduction in the adverse-event profile and a 25 percent reduction in the length of hospital stay.

These various types of post-marketing learning and innovation take place not only with therapeutic technologies but also with diagnostic technologies. Diagnostic technologies can be used to identify abnormalities, but uncertainty remains concerning how much they can be used to infer prognoses or the need for treatment. Several controversial examples include mammography and ductal carcinoma in situ (DCIS), prostate specific antigen (PSA) testing, and magnetic resonance imaging (MRI) evidence about unbled brain aneurisms. The uncertainties have resulted in significant variations in rates of further diagnostic testing and in treatment patterns, both nationally and internationally.

Genomic interventions may produce diagnostic technologies that enhance prognostic abilities. In the case of breast cancer, for example, many women receive adjuvant chemotherapy to prevent recurrence. Gene chips may identify women who have a high likelihood of developing such a recurrence and thereby allow targeting of such therapy more judiciously. But these technologies will also bring with them their own uncertainties—a positive test will not always indicate the development of disease, for instance, because a number of factors can also play a role, including variable expressivity, and environmental factors.

When new technologies are introduced into health care they may be relatively primitive, which accounts for some of the slowness of their diffusion. Actual use, however, produces downstream learning, which may lead to modifications in the technology itself or refinements of its application. One such refinement, for example, is better prognostic understanding of a genomic test as a better understanding develops about the interactions among genes and between genes and environmental factors (Burke and Psaty, 2007). Additionally, physicians become more knowledgeable about how to integrate these technologies with appropriate surveillance and treatment regimens for the whole spectrum of at-risk patients (Burke and Zimmern, 2004). The clinical utility[6] of such tests, however, will need to be confirmed in pragmatic clinical trials involving large, well-defined populations.

Evidence is a critical factor in the diffusion of technology. The FDA plays a key role in shaping the evaluation and adoption of technology in other fields, and the agency has taken a proactive role in the area of genom-

[6]Clinical utility is the degree to which a test alters medical management in a way that results in a net health benefit to the patients (IOM, 2005).

ics. Traditionally the FDA has regulated only those diagnostics that were marketed as kits, and CMS has had oversight over those diagnostics marketed to laboratories, but with DNA chip technology, such as the amplichip CYP 450, the FDA decided that genetic tests required a higher level of review. The value of diagnostic tests is harder to measure than that of therapeutic interventions, however. Premarketing trials are typically aimed at determining accuracy, and insights about clinical utility often emerge only in the post-marketing setting.

Uncertainty about new test interpretations may affect the adoption decisions of health care providers. For example, the AlloMap molecular expression test was developed to detect acute cellular rejection in heart-transplant patients. Premarketing studies suggested that while the AlloMap might have somewhat lower positive predictive value than biopsies, the fact that it is non-invasive gave it an advantage. After it was introduced, however, uncertainty about its clinical utility led many centers to use the test as an add-on and not a substitution—a common phenomenon with new diagnostic technologies. The reluctance to adopt the AlloMap as a substitute for biopsies was also influenced by the fact that cardiologists needed to become more comfortable interpreting the genomic information.

In the area of pharmacogenetics, the integration of diagnostic tests and drugs poses special challenges because it will require that historically separate regulatory pathways be brought together. One successful example of such integration is HER-2 testing and Herceptin, where both products were approved through the fast-track process in the same week, with coordinated labeling. This case may have been relatively straightforward in that there was a clear relationship between the biomarker and drug response and the drug resulted in survival benefit for a life-threatening condition. With tests that have more ambiguity about the ultimate value of the information, rigorously conducted studies in the pre- and post-marketing stages will be even more important.

Payers, who struggle with tradeoffs between costs and benefits, are exercising an increasingly important gatekeeper function through their coverage and reimbursement decisions. Although cost-effectiveness is not formally a coverage criterion for Medicare, many payers have adopted it as part of their decision-making process. Yet cost-effectiveness analyses of emerging novel technologies are challenging, partly because substantial innovation can be expected to take place after the technology goes to market. A strict adherence to a cost-effectiveness value such as $100,000 per life-year saved might eliminate some potentially valuable technologies before they have had the chance to prove their worth.

In the case of genomic technologies, cost-effectiveness analyses need to incorporate post-marketing innovation and learning-by-using sensitivity analyses in a more systematic manner. At the same time, payer decision mak-

ing may need to become flexible enough to allow for short-term inefficiencies in order to understand and benefit from long-term value. Still, optimal learning takes time and experience, and payers may be understandably uncomfortable in underwriting such learning. This raises the important questions: What models can be used? And are there public–private partnerships that can be used to capture post-introduction learning more efficiently?

Finally, the diffusion of genomic interventions is likely to be powerfully shaped by sociocultural factors. Even if genomic interventions are covered by insurers, patients may decide to pay out of pocket because of concerns about confidentiality and the potential for discrimination by employers and insurance companies. This, in turn, raises concerns about equity—for example, about lack of access to these technologies for those who do not have the means to pay.

Another issue concerns the diffusion of tests that would identify predispositions to future disease for which there are no cures, only treatments with limited effectiveness and major side effects. Patients may vary greatly in their decisions about whether to have the genetic test for Huntington's Disease, for example. Diffusion processes are fundamentally affected by patient preferences and by the public's perception of the value of health-risk information.

Gelijns concluded by saying that diffusion is a critical process by which the health, social, and economic rewards of an invention are ultimately reaped. Even more than that, however, diffusion is an integral part of the innovation process. It can be characterized as a learning process, and a fundamental aspect of learning is the reduction of uncertainty. Downstream learning can lead to changes in a technology or to refinements in its use. At the same time it poses new questions for basic and translational research and thereby enriches the ultimate payoff.

The determinants of diffusion in genomics are probably very similar to those for other medical technologies. Diffusion depends not only on the benefits that a new intervention provides, but also, importantly, on the institutional environment in which a technology is imbedded. Patients, consumers, and physicians need to understand what to do with new probabilistic risk information; the FDA must decide how to deal with genetic diagnostic tests and how best to regulate diagnostic drug combinations; insurers need to gain comfort with the interpretation of cost-effectiveness analyses of emerging novel genomic technologies; and, finally, the larger policy world will need to deal with privacy and confidentiality issues and the potential for discrimination.

DISCUSSION

Wylie Burke, M.D., Ph.D.
Moderator

A member of the audience commented that she believed Gelijns overestimated the use and the effect of cost-effectiveness analysis (CEA) in insurance decisions. The United Kingdom, she said, has adopted CEA and uses it for health-policy decisions, including those made by the National Health Service, but there is no mechanism in the United States for such analysis, and there is no systematic application of a quality threshold. While the Blue Cross and Blue Shield Association conducted a cost-effectiveness analysis regarding LVAD, it was done for educational purposes in order to understand the methodology.

Califf commented that the lack of use of cost-effectiveness analysis illustrates the chaos that exists in decision making about innovations. Some innovations are blocked, while others go forward despite extraordinary costs, but it is very difficult to understand the basis upon which the decisions are made.

Another audience member observed that it appears that as far as innovation is concerned, the health care system may reward small and relatively inconsequential changes and may sometimes create a disproportionate and negative response to rare events. Furthermore, the system prevents open exchange of information and creates many barriers to communication among affected parties—for example, between FDA and vendors or between those manufacturing or creating new devices or drugs and those who will be using them. This is the sort of situation described by the mathematician John Nash nearly 60 years ago—that is, that the optimal good is almost never achieved by the individual players optimizing their own individual results without being able to fully discuss how to jointly optimize the system (Nash, 1950). Could it be that the current health care system is so inhibited that the need to optimize individual results actually makes it impossible to introduce disruptive new technologies?

Schulman responded by saying that markets evolve through a private process, and that private process is being choked. Porter wrote a book describing the different things that each of the actors in society can do to improve things in the health field—how hospitals could serve the needs of the public better, how physicians and insurers could do better, and so on (Porter, 2006). But there is nothing in the book, Schulman said, that explains why any of these actors would actually move from their current position. The critical issue, he continued, is the marketplace. What are the

levers? How can opportunities for leveraging and enforcing innovation be created in this very public market?

Califf suggested that there are two important factors involved. First, there is an assumption that everyone agrees on what the ultimate state is or should be, which is not actually the case and which should be discussed further. If the goal were to optimize the longevity and quality of life of the American people as a population, that is not what we are now doing. But that may not be the goal everyone has in mind. Second, one must ask whether or not there actually is a need for disruption, as Schulman posits. If one makes a poor-quality transistor radio, the consequences are not great—the radio breaks, and it can easily be replaced. But if there is a poor-quality test that results in someone dying earlier than they otherwise might have, that should not be allowed to happen—a new test should not be allowed to enter the marketplace until studies have been conducted to show that it is worthwhile. Because the measurement of health status and health outcomes is so much more detailed and reliable than it once was, it is possible to measure what is being done and introduced.

One audience member asked the speakers to go into greater detail about why they said the prospects for introducing genomic innovations into the marketplace are poor. Califf responded that part of the problem is that the regulatory pathway for introduction is unclear, and that lack of clarity discourages investors. Furthermore, genomic innovations target relatively small groups of people who can really benefit from the innovation. If the innovation concerns a disease such as cancer, investors will invest because the potential payoff is so large. For anything else, finding investors is a problem because the market is smaller.

Gelijns said that one of the important issues is that the premarketing trials often focus on the accuracy of a test and that the ultimate clinical utility of these tests frequently emerges in the post-marketing setting. At that point one must deal with the issue of how best and most efficiently to obtain information about health and economic outcomes. Because individual stakeholders might not have enough incentive or means to conduct post-market testing, it is important to start thinking about new models of cooperation, such as public–private partnerships, that will pull together the various parts of the system to generate needed information and to improve the innovation process.

One questioner said that because he comes from a public health background he would like to see discussion about how the translation and diffusion of innovations take into account the end of the pathway—that is, improving the public's health. In many areas of medicine and public health it is known that if a particular action is taken, thousands of deaths in the population can be prevented, yet those things are not being done. It takes years to implement and diffuse proven innovations into practice.

On the other hand, there are new technologies, such as the genome-based technologies, that have uncertain added value compared with what is currently being done. These technologies are intended to replace or be added to the things that we know should be done but that are not actually being translated into practice.

So the question is, is there a process or an organizing principle that helps sort out or distinguish what is ready to be introduced from what is not? Furthermore, given the complicated schema of translation and the multiple factors and players that are involved in this process, where is the role of the evidence-based guideline?

Schulman responded that, intellectually, the use of clinical-practice guidelines for genomic innovations is an exciting area. There is still relatively little clinical information available, however, so how these guidelines will fit into the marketplace is uncertain.

Califf agreed with the questioner that there are things we know now that could be used for better health or treatment. For example, a person admitted to a hospital with an acute coronary event has a 33 percent likelihood of getting the wrong dose of any thrombotic drugs that are prescribed. A simple serum creatinine in body mass index will give information about the correct dose, but it is often not used. Yet there is discussion of diffusing even more sophisticated genomic technology into outpatient settings or unsupervised settings. These new tests should not be unleashed upon the public without evidence that they will help rather than harm. On the other hand, when there is an important disruptive technology that can make a big difference, there should be some special approach that allows people to develop the evidence with some protection while a determination is made about whether the innovation is valuable.

Gelijns said that a major question is how to create incentives for gathering information as the technology keeps changing in the post-marketing setting.

Another questioner, stuck by the idea of post-market innovation and how much is learned when something is put into practice or use, asked if expanding regulations and the increasing drive for evidence-based practice of medicine are going to squelch innovation.

Schulman said that one argument is that we need at least some level of evidence on the technology innovation side that there is some effectiveness. Then there is also the need for service innovation. Today's system is costly and less effective than it should be, therefore service innovation is necessary.

Califf stated that creativity is needed both in customizing (some say personalizing) medical care for patients and in delivering services. People running health systems find it impossible, given today's finances, to actually do what people need to have done in order to make them healthier, he

said. Furthermore, companies, in order to make money, study the wrong things because of the way incentives are structured. If one could create a more streamlined system and actually define the value of the products as opposed to developing marketing schemes that are tangential to the value, one could focus on defining what actually works.

Gelijns mentioned that the history of pharmaceutical innovation is full of new and often unexpected indications of use that emerge in the post-marketing setting. What is needed may be a more streamlined process for looking at evidence, not only of unexpected side effects, but also of unexpected benefits. One important issue is how much room to leave for experimentation.

One audience member said that he thought that a major player in innovation and translation—the pharmaceutical companies—was being ignored. In 1998 a drug that had a 4 percent complication rate of hypersensitivity syndrome (abacavir) was released with an accelerated approval that required risk-management studies to be conducted. As a result, some early-stage pharmacogenetics were conducted. At the time, it was the only product in its class. Other products in the class have now entered the marketplace, and each of them has a different type of adverse event.

From the company's point of view, if there was a highly accurate test that demonstrated that the adverse event for its drug could be avoided, the company would have a competitive advantage to getting that put on the label. That is, in fact, what happened. There is now a test that can identify, with greater than 99 percent specificity, the people who will suffer hypersensitivity syndrome if they take abacavir. For the first time there is a diagnostic test for a drug allergy.

Month by month the sales of this test, which did not come out simultaneously with the drug, have quadrupled, according to the audience member. Yet this information has not been published in any of the scientific journals. This example illustrates the fact that if there is a competitive situation and market share is at stake, a huge incentive exists to make the investment needed to implement the test, even if there is no reimbursement for the test. It might even be considered unethical to give the drug without testing.

It is important, the audience member concluded, not to restrict thinking about incentives to academia and the government, but rather to expand incentives to include those who can change the system, such as pharmaceutical companies.

3

Practical Incentives and Barriers to Translation

TRANSLATING MEDICAL INNOVATIONS WITH APPROPRIATE EVIDENCE

Sean Tunis, M.D., M.Sc.
Center for Medical Technology Policy

There is an important tension between innovation and the opportunity for post-market learning or evidence, and the risks and benefits of this tension need to be better understood. Post-market learning can occur only in an environment where payers have started to pay for something with presumably less evidence than they might have wanted. The idea that we can encourage post-market learning in an environment where the evidence requirements are becoming more rigorous is likely to change dramatically.

To be covered by Medicare, an item or service must be determined to be reasonable and necessary. The working definition for reasonable and necessary is that there is adequate evidence to conclude that the item or service improves net health outcomes, is generalizable to the Medicare population, and is as good as or better than currently covered alternatives. The key question here is, what constitutes adequate evidence? Unfortunately, the evidentiary bar is not well defined, which is part of the reason for the tension between innovation and evidence. From the point of view of a company developing products that it wishes to bring to market and for which it hopes to be reimbursed, the evidentiary target is fuzzy.

The approach that Medicare takes to determine whether adequate

evidence exists for the reasonableness and necessity of diagnostic tests has two main components. First, the evidence must be adequate to determine whether the test provides more accurate diagnostic information than existing tests. Second, if the test provides more accuracy, the evidence must be adequate to determine how the changed accuracy affects health outcomes. For example, does it change patient management, and do those changes in patient management actually improve outcomes?

At the request of the Centers for Medicare and Medicaid Services (CMS), experts at Duke University conducted an evidence review on the use of positron emission tomography (PET) in scanning for Alzheimer's disease. The review concluded that there was adequate evidence to conclude that PET scanning has better sensitivity and specificity than clinical evaluation by an expert neurologist. The experts also constructed a decision model which determined that because available treatments had very limited efficacy and were relatively safe (i.e., treatments for dementia are basically nontoxic and not very effective), the new diagnostic information available from the PET scans had essentially no effect on patient management—that is, it would not change patient outcomes. Furthermore, the small false-negative rate of PET scans might lead to withholding treatment and might lead to worse outcomes than empirically treating anyone with a clinical diagnosis of dementia.

In light of this, Medicare policy is to not cover PET scans for Alzheimer's disease except in the context of a prospective clinical trial that would evaluate whether the information from PET scans changed patient management in other important ways. Medicare agreed to cover tests in the context of such a study, and a proposal for such a study was developed by scientists at the University of California at Los Angeles and submitted to the National Institutes of Health, but it was never funded. So the practical effect of this policy is that Medicare does not pay for any PET scanning for Alzheimer's disease.

What do other payers need in order to determine whether to pay for a diagnostic test? One major private payer's policy on the clinical utility of ambulatory electocardiograms (ECG) is that ambulatory ECG is considered experimental and investigational because of the lack of peer-reviewed published reports of prospective clinical trials on the effectiveness of the distinct features of ECG in improving clinical outcomes over standard cardiac event-monitoring services. What this payer is saying, in other words, is that in order to qualify the service for reimbursement one would need to conduct a prospective study of ambulatory ECG versus Holter monitoring, and the results would need to demonstrate that some important clinical outcome was changed as a result of the use of ambulatory ECG.

Such a study has never been done, Tunis said, and it is unlikely that any company manufacturing ambulatory ECGs will ever conduct such a study.

Thus, while such evidence requirements may be desirable from an evidence-based medicine perspective, it may not be a feasible evidence threshold to use as a condition for reimbursement.

Another example of evidentiary review is a retrospective gene-expression profiling for breast cancer that was conducted in October 2006 by the California Technology Assessment Foundation. That assessment found that the predictive accuracy of Onco*type* Dx is high for recurrence (although the test was never compared to standard risk-assessment tools) and that the National Surgical Adjuvant Breast and Bowel Project (NSABP) Protocol B-14 showed that low-risk patients randomized to chemotherapy and followed for 10 years did no better than did those who did not undergo chemotherapy. In this case, there were 10 years of frozen specimens that could be used in the study. The TAILORx[1] and MINDACT trials[2] (10,000 and 6,000 patients, respectively) are now under way.

For Onco*type* Dx, the California contractor for Medicare initially issued a draft decision for non-coverage which was reversed because of strong feedback from clinical oncologists disagreeing with the draft decision. Furthermore, administrative law judges were reversing denials of payment when the issue was brought before them. The important point is that there is no national policy from Medicare about how much evidence CMS will consider adequate to conclude that there is clinical utility on this test or any molecular diagnostic.

Different payers have different evidence requirements for what is sufficient to determine the clinical utility of a diagnostic test, which makes it very difficult for a company developing a product to know how to design its clinical research portfolio. Payers, physicians, and patients are demanding more evidence on comparative effectiveness and value. Yet the evidence requirements for coverage are poorly defined, inconsistent, and, in some cases, not feasible.

Furthermore, there is a major problem created by the fact that reimbursement and regulatory evidence requirements are not aligned with one another. There is frequently a mismatch between what the payers would like to know and what the regulators would like to know. This means that even

[1] "The Trial Assigning IndividuaLized Options for Treatment (Rx), or TAILORx, will examine whether genes that are frequently associated with risk of recurrence for women with early-stage breast cancer can be used to assign patients to the most appropriate and effective treatment" (National Cancer Institute. http://www.cancer.gov/clinicaltrials/digestpage/TAILORx, accessed January 22, 2008).

[2] "Microarray for Node-Negative Disease may Avoid Chemotherapy (MINDACT) was originally designed to compare the ability of a 70-gene prognostic profile versus clinical and pathological criteria to identify women with node-negative breast cancer who are unlikely to benefit from adjuvant chemotherapy" (Tuma, R. S. 2005. Trial and error: Prognostic gene signature study design altered. *Journal of the National Cancer Institute* 97(5):331-333).

with regulatory approval there is no certainty that reimbursement approval will be forthcoming. The next major area of contention may well be the evidentiary framework of the payers concerning molecular diagnostic tests.

The Center for Medical Technology Policy (CMTP), a private, nonprofit corporation, has begun work on issues relevant to the discussion of creating high-quality evidence of clinical effectiveness[3] and clinical utility from the perspective of decision makers, that is, from the perspective of payers, clinicians, and patients. For the past two years funding for CMTP has come primarily from foundations, but it is now changing to a membership-funded model with health plan and life science company memberships. The primary mission of CMTP is to support collaborative activities among stakeholders that will improve the quality and efficiency of prospective studies of new medical technologies.

One of CMPT's projects is to create coverage-guidance documents that will provide a clear, well-defined, and consistent target for what evidence is necessary to demonstrate clinical effectiveness and clinical utility across a broad range of technologies. The primary audience for these documents is product developers. The documents are analogous to FDA guidance documents, but FDA guidance documents articulate evidence requirements for regulatory approval. The idea behind the CMPT documents is that they should serve as companion documents to these FDA documents and articulate the specific evidence requirements for reimbursement or for coverage. The purpose of the documents is to reduce uncertainty, increase consistency, and incorporate a notion of feasibility.

To develop these coverage-guidance documents CMPT will work with multiple stakeholder workgroups—for example: payers, product developers, clinical organizations, and patient groups—to determine what the evidence requirements should be. Then there will be a web-based, iterative, public-comment process on the draft documents. The first effort being undertaken is to develop an evidence-guidance document on gene-expression profiling for breast cancer. The next topic will be wound-healing interventions.

Tunis concluded by saying that the definition of evidence requirements for clinical utility, clinical effectiveness, and comparative effectiveness[4] should be well-defined—and not just defined by payers or by evidence-based medicine experts. They need to be defined in a collaborative way that

[3]Clinical effectiveness is defined as "the extent to which specific clinical interventions when deployed in the field for a particular patient or population do what they are intended to do, that is, maintain and improve health and secure the greatest possible health gain from the available resources" (NHSE, 1996).

[4]Comparative effectiveness is the "comparison of one diagnostic or treatment option to one or more others. . . . Primary comparative effectiveness research involves the direct generation of clinical information on the relative merits or outcomes of one intervention in comparison to one or more others" (Buckley, 2007).

involves the perspectives of people who understand what it takes to develop these products as well as the perspectives of the patient and the clinician.

ASSESSING TECHNOLOGY FOR USE IN HEALTH AND MEDICINE

Naomi Aronson, Ph.D.
Blue Cross and Blue Shield Association

In examining evidence it is important to distinguish between three kinds of policy, Aronson said: medical policy, coverage policy, and payment policy. Medical policy is based on scientific evidence and does not consider costs or coverage issues. Technology assessment is used in the support of medical policy. Medical policy essentially operationalizes two health plan contract provisions, those describing what an investigational service is and those describing what a medically necessary service is.

Coverage policy, by contrast, is determined through a contract with purchasers of health plan policies; these purchasers are largely employers. In designing the benefits, their cost-effectiveness may be considered. The clearest example of this is with drug benefits, where decisions include such factors as cost-equivalent substitutability.

Finally, there is payment policy, which is the contract between health plans and medical professionals and providers.

The 39 independent Blue Cross and Blue Shield (BC/BS) plans around the country make their own coverage decisions, but they do look to the Technology Evaluation Center (TEC) at the Blue Cross and Blue Shield Association (BC/BSA) for evidence-based analysis because that is the basis for their coverage decisions. With 100 million total members in the 39 independent plans, BC/BS covers one in three Americans.

The purpose of the Technology Evaluation Center (www.bcbs.com/tech) is to provide rigorous assessment of clinical evidence. The TEC is staffed by physicians, epidemiologists, research scientists, medical librarians, and pharmacists, who are employees of the BC/BSA. Nothing is released as a TEC assessment until it has been approved by an independent expert medical advisory panel under whose authority the TEC operates. The advisory panel is composed of academic clinical researchers and specialty society appointees, including an appointee from the American College of Medical Genetics, an association that was recently added because of the complexity and importance of the area of medical genetics.

Only 4 of the 17 votes on the panel are allotted to plan clinicians, which is an important point and emphasizes the independence and scientific

rigor of the process. The staff presents its analysis for the panel to decide whether the technology under consideration improves health outcomes. Does it improve length of life, quality of life, or the ability to function? If the panel judges the evidence as supportive of improvement, the report is approved.

During the past three years, the TEC has conducted more than 300 technology assessments, all of which can be viewed at www.bcbs.com/tech. The TEC takes the position that the process of technology assessment should be transparent so that stakeholders can understand the level of evidence used. While the TEC cannot consult with companies, its staff will hold teleconferences with any company that wishes to understand better how the TEC might approach the evidence concerning the company's technology. The TEC is also an evidence-based practice center for the Agency for Healthcare Research and Quality.

Interest in genomics at the TEC is not new, Aronson said. Ten years ago the TEC assessed BRCA-1 and BRCA-2, found that they met the TEC criteria, and recommended that they should be offered and accompanied by genetic counseling. During the past year the TEC has placed a strong focus on genomics because it understands that genomics is an area that is both rapidly evolving and that can be somewhat confusing and intimidating to the average clinician.

There are two roles that the TEC can play in genomics. The first is to assess specific currently emerging technologies. The second is "horizon scanning"—looking ahead to see what important technologies are approaching. Assessments of specific emerging technologies have included gene-expression profiling of breast cancer and genetic testing for long-QT Syndrome (LQTS). Horizon scanning has examined cardiovascular pharmacogenomics, cancer pharmacogenomics, and genomics of neurologic disorders.

The focus of the TEC is on patient-outcome efficacy, that is, improved health. This is compatible with the ACCE[5] evaluation model and the framework developed by the Centers for Disease Control and Prevention.

If one understands clinical validity,[6] how can the case for clinical utility be made? When can it be made from inference? When does it need to be directly demonstrated? In an ideal world one would always have direct evidence for clinical utility. Randomized controlled trials (RCTs) are expensive, however, and are typically not the norm in the area of diagnostic

[5] "ACCE, which takes its name from the four components of evaluation—analytic validity, clinical validity, clinical utility and associated ethical, legal and social implications—is a model process for evaluating data on emerging genetic tests" (http://www.cdc.gov/genomics/gtesting/ACCE.htm, accessed January 24, 2008).

[6] "The clinical validity of a genetic test defines its ability to detect or predict the associated disorder (phenotype)" (http://www.cdc.gov/genomics/gtesting/ACCE.htm, accessed January 24, 2008).

testing. The reality is that the case for clinical utility generally relies heavily on indirect evidence, using a causal chain of logic, inference, and linkage of various bodies of literature, from the diagnostic performance of the test to the effect on patient management and, ultimately, to the effect on health outcomes. Bona fide health outcomes for diagnostic tests include avoidance of other tests and avoidance of an invasive procedure.

The following example illustrates how the TEC puts together an assessment. Computed tomographic angiography (CTA) has been proposed as a noninvasive alternative to invasive coronary angiography for the evaluation of coronary artery disease. When the findings on CTA are negative, invasive angiography is not necessary, but those results with significant stenosis (positive CTA findings) need to be confirmed by invasive angiography. Comparing the health outcomes of the two technologies involves considering such factors as the number of catheterizations avoided, the risks and the effects of a false negative CTA, the effects of added radiation (since CTA involves a substantially higher dose of radiation), and the effects of extra cardiac findings. The TEC assessment found that the evidence is insufficient to draw conclusions about the effect of CTA on health outcomes. Therefore, CTA does not meet the TEC criterion that requires being able to draw conclusions concerning the effect of the technology on health outcomes.

Another example involves the assessment of genetic testing versus the use of clinical criteria for identifying people with LQTS, a condition that marks individuals as being at risk for lethal arrhythmias. Such individuals are typically under age 40 and usually have a family history of the condition. The TEC assessment found that the genetic test is accurate in identifying the presence of a mutation but that the diagnostic accuracy for identifying LQTS is not clear because there is no true gold standard for clinical diagnosis of LQTS. The assessment did find, however, that genetic testing identifies more individuals that may have LQTS than are identified through clinical diagnosis alone.

The opinion of the medical advisory panel was that there is value in uncovering additional information because LQTS is an underdiagnosed condition. It is treatable with beta blockers, but if LQTS is not identified it can have catastrophic results. In this situation, it is not possible to conduct the kind of quantitative modeling of health outcomes that was done with CTA. However, a qualitative analysis showed that the genetic test had the potential to identify more patients with LQTS, who would then receive low-risk treatment with beta blockers, thereby forestalling the potential catastrophe of untreated disease.

There are many associations in genomics, and more information is rapidly becoming available. However, the relationship between evidence and clinical validity is not well defined. In the previous example one could infer that the genetic test for LQTS would improve health outcomes, but when is direct evidence needed?

One example of the need for direct evidence can be found in the area of lung cancer screening. Lung cancer screening is controversial because it is unclear whether there is any value to early detection. Improved accuracy in detection must be viewed cautiously because of the potential for lead-time bias,[7] length bias,[8] and overdiagnosis bias.[9] In light of these complexities, the National Cancer Institute is carrying out the National Lung Cancer Screening Trial in order to address the question of whether there is value in using spiral computed tomography (spiral CT) for the early detection of disease. Approximately 50,000 patients, both smokers and former smokers, are participating in the trial. They were randomized to receive either spiral CT or X-rays.

This situation can be compared to some of the controversies surrounding the use of genotyping in the decision whether to initiate warfarin dosing. It is not easy to compare genotyping to a reference standard. There are many intervening variables that contribute to warfarin dosing, and there is a narrow window between an effective therapeutic dose that prevents clotting and a too-high dose that leads to bleeding. For these reasons, one needs direct evidence, and that can only be accumulated with prospective trials of dosing algorithms to compare personalized warfarin starting doses with standard dosing in terms of bleeding outcomes. A small trial has been conducted, but the results were not encouraging. With a correct algorithm, however, there may be definitive results. Several trials are currently under way to find out.

While evidence of clinical effectiveness is the cornerstone of the BC/BS plans' medical and coverage policies, cost-effectiveness and affordability are also pressing issues. Every health plan in business is operating under state regulators. State regulators may have slight differences in their investigational and medical-necessity language, but those differences are not substantial. The key point is that there is no contract language that allows payers to use cost-effectiveness as a standard or a criterion for coverage. The medical-necessity language does, however, specify that more will not be paid in order to achieve the same results.

The difficulty is that there is no contract language that addresses those situations where there are only incremental benefits compared to costs, and there are many new technologies that have such incremental benefits compared to their costs. The TEC has produced some cost-effectiveness

[7]Lead-time bias means that there is a longer time between diagnosis and death, even though death is not delayed.

[8]Length time bias means that slower-growing tumors are more likely to be detected, which biases the resulting data to imply a better prognosis than the actual prognosis.

[9]Overdiagnosis bias is when screening detects cancer that would not, within the lifetime of the individual, have developed into disease.

analyses[10] of technologies that met the criteria of improved outcomes, but because there are no clear cost-effectiveness thresholds that are scientifically prescribed, cost-effectiveness analysis is not a solution to the affordability problem. Just because something is of value, does that mean it is affordable?

The affordability problem is real and there is a substantial gap between health care insurance premiums and workers' wages and inflation. Employers, who are the main source of insurance in this country, are dealing with this problem by shifting costs onto employees. Additionally, there are currently about 47 million uninsured individuals in this country, which is slightly more than the number of people insured by Medicare and somewhat more than half the number insured by BC/BS (Center on Budget and Policy Priorities, 2006). A sustainable health care system is one that is affordable. As products are designed and brought into the market, thinking about long-term sustainability is a key to long-term success.

Health plans want to make evidence-based decisions, but there are considerable challenges to obtaining good evidence on outcomes for both therapeutic interventions and diagnostic tests. For diagnostic tests, indirect evidence can be used if that evidence is based on performance, in which inferences can be made about clinical utility. But when there are complex associations and intervening variables, direct evidence is necessary. Ultimately, while the TEC process is not aimed at costs, cost-effectiveness and affordability are pressing concerns and will shape the success or failure of technologies, Aronson concluded.

INTEGRATING GENETIC TECHNOLOGY INTO A HEALTH CARE SYSTEM

Wylie Burke, M.D., Ph.D.
University of Washington

The movement of genetics into the health care system is marked by three major trends, Burke said. First, information that was previously handled by medical geneticists and a few specialists is now moving into more of a specialty–primary care mix. This requires addressing the barriers that exist as that transition is made. Second, although genetics historically has used information as an endpoint that did not improve health care outcome, such

[10] "Cost-effectiveness analysis is a measure or evaluation of the cost of an intervention relative to its impact, usually expressed in dollars per unit of effect" (Modeste, 1996).

improvement in outcome is now possible. This shift requires thinking about the use of genetics in a different way—a way that is more like how other health care information is used. Finally, there is movement from a limited amount of information to a great deal of—maybe too much—information. In the past the worry might have been about what was not known, whereas now the worry is about managing the information that is available.

There are three ways that genetic research can provide health benefits. The first is through the use of tests to identify genetic diagnosis or to identify genetic risk. The second is the application of gene-expression panels and other kinds of genetic technology that will enable improved disease classification. Third, some innovative therapies have been developed, and there is hope for more. The latter two are likely to have the biggest benefits over time.

In thinking about tests for genetic diagnosis and risk assessment, it is important to acknowledge the different kinds of tests that are currently available. Genetic tests differ in penetrance of the genotype from low to high. Historically, most applications of medical genetics have involved high-penetrance genotypes—that is, genotypes where, in those cases where genetic information is available, there is a great deal of certainty about what the clinical experience of a patient is going to be. The current wave of tests, however, is of a very different sort: they are much more probabilistic.

Another difference from other areas of health care is that in genetics it may be the case that the available information has no connection with measures to improve the disease course, whereas at other times the measures are available. Historically, medical genetics was defined as a practice that told people about very high risks for which there was not much that could be done and for which the main intervention would be making decisions about reproduction or selective abortion. This led to the development in 1975 of a medical-genetics standard that still applies in those cases where a geneticist is dealing with information regarding a highly penetrant genotype for which there is no treatment. This standard calls for non-directed counseling and is described as "An attempt to help the individual or family to . . . choose the course of action which seems to them appropriate in view of their risk, their family goals, and their ethical and religious standards, and to act in accordance with that decision" (Ad Hoc Committee on Genetic Counseling, 1975).

The first challenge, then, is that medical genetics has a standard that is very counseling-intensive and personnel-intensive, requiring time with individuals to provide them with information that is very charged. But genetics now has a growing presence in such routine medical procedures as the obstetrical screening for trisomy 21. There are also some carrier-screening tests (e.g., Tay-Sachs, hemoglobinopathies, and cystic fibrosis) that are now part of routine health care.

In the future, genomic technologies will likely make it possible to provide many more tests of the carrier-screening variety or of a prenatal-diagnostic variety. One concern is that these tests will provide people with more information about risks to the fetus, and without also giving people the opportunity for in-depth counseling, this information might move individuals along a pathway of selective termination, which is what frequently happens with an abnormal result.

There is, then, a barrier related to the quality of care and ethical practice. How, in often time-pressured practice settings, can one deliver the kind of counseling that medical genetics standards would suggest should be delivered?

There are also cases of high-penetrant[11] conditions where there is an opportunity for treatment benefit, such as newborn screening. Newborn screening is a very successful program that looks for individuals with specific genetic conditions for which there are definitive treatments to improve health outcomes and also where there is a time urgency. A question arises because there are many more conditions for which tests are available than was the case when newborn screening started in the 1960s. Thus it has become necessary to consider what sort of time urgency and what level of outcome improvement is sufficient to justify incorporating tests into this kind of mandated screening program.

There are other examples of conditions with high penetrance that have the opportunity for treatment, such as colorectal cancer. There are two relatively rare hereditary conditions that involve a very high lifetime risk. One is hereditary non-polyposis colon cancer (HNPCC), which has a prevalence of one person in 500 and is correlated with an 80 percent lifetime risk of colorectal cancer. HNPCC also increases the risks of endometrial cancer and ovarian cancer. Screening for HNPCC should start in the early twenties. A primary-care provider might see 10 or 12 individuals with this condition during his or her career.

The second hereditary condition that increases the risk of colorectal cancer is familial adenomatous polyposis (FAP). Its prevalence is 1 person in 8,000, or rarer. A practitioner might never see a single case, but picking up individuals with this condition is important. In people with this condition there is a 100 percent lifetime risk of colorectal cancer, and prophylactic subtotal colectomy is recommended.

If genetic testing for these conditions follows the medical genetic standard, there would be pre-test counseling, family assessment, determination

[11]Penetrance is the probability of developing a disease (or some other outcome of interest) given that an individual has a particular genotype. The penetrance of a genotype is often estimated by examining the proporation of people with a particular genotype who develop the disease or outcome of interest.

of who is at risk, explanation of the limitations of current testing technology, and a description of treatment options if the test is positive, all delivered in a very counseling-intensive way.

Typically, however, these families come to genetics only after a fairly dramatic occurrence of cancer in the family. There is a need to do better prospectively. Families like this should be identified in primary care and specialty practice. The challenge in doing that is that it will require physicians to become sophisticated about the continuum of family history, which exists in virtually all common diseases.

A geneticist would look at a configuration such as appears on the right side of Figure 3-1 and be able to determine that the individual is almost certainly either HNPCC or FAP. Testing would be done to determine which it is, and family members would begin the pathway for prevention.

The middle of Figure 3-1 illustrates a family that does not meet the criteria and is unlikely to be at high risk. There may be some risk if information about different members of the family is missing. Based on knowledge of colorectal cancer epidemiology, however, individuals in that middle family are likely to benefit from starting routine colon cancer screening at 40 instead of 50. On the left side of Figure 3-1 is a family history where a grandfather died with colon cancer at 80. Such a family history is not an indicator or a red flag.

Primary-care physicians as well as all the specialists who may come into contact with this kind of family history must become more sophisticated about making these kinds of distinctions and then referring their patients to the appropriate specialists. There is a great deal of evidence showing that physicians do not do a good job of making these distinctions and are not comfortable doing so. One major barrier to physicians doing a better job

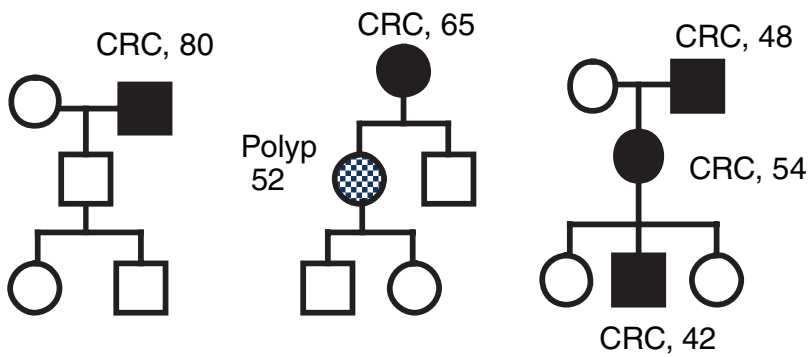


FIGURE 3-1 Continuum of family history of colorectal cancer.
SOURCE: Burke, 2007.

is that there is no current funding mechanism for adequately reimbursing physicians for spending the time that it takes to complete a careful family history assessment.

Apolipoprotein E (APOE) testing is an example of a test with a relatively low predictive value. An APOE4 genotype predicts increased risk for Alzheimer's disease. At present there are no treatments available that can reduce the risks of Alzheimer's disease in individuals with this genotype, so there is no way to use this genetic test to improve health outcomes.

The information from the test is potentially actionable in other ways, however. For example, in one very small study about 150 adults whose parents had Alzheimer's disease were offered APOE genotyping; the statistically significant finding of the study was that individuals who tested positive for APOE4 were more likely to buy long-term care insurance than individuals who either received a negative result or chose not to learn the result (Roberts et al., 2003). There is also a suggestion that APOE4-positive individuals are more likely to buy health insurance and less likely to buy life insurance.

A patient, then, might consider APOE4 testing to be of value and actionable even if health outcomes are not affected. One of the challenges for a health care system is to decide whether this and similar uses of information is something that should be considered part of health care.

Another example involves warfarin. As mentioned briefly above, variants in the genes encoding the enzymes VKORC1 and CYP2C9 can affect how the body responds to warfarin. Only about 35 percent or 40 percent of the variance in response is explained by these genetic variants, however, which means this is not the sole predictor of how individuals are going to respond to warfarin or what dose they require (Rieder et al., 2005). Physicians are fairly good at determining warfarin dosing and are good at monitoring reactions, so it is uncertain whether genetic test information about variance will help doctors better manage their patients. This is the kind of situation where additional data are needed.

Another issue in pharmacogenetic testing is ancillary risk information. Many pharmacogenetic variants provide valid information about risks for diseases other than the one for which the test was conducted, and some of those risks are entirely unrelated to the purpose for which the pharmacogenetic testing is done. In some tests there are two or more risks. Undoubtedly, more ancillary risk will be discovered. Is this a good or a bad thing? The answer may well depend upon the test.

A particular transcription factor (TCF7L2 variant DG10S468) has been identified as associated with a relative risk for Type 2 diabetes. The effect is statistically significant, but it is small. The major issue then is, what is the clinical utility? The action that would be taken after getting the results of this test is the kind of action that everyone should be taking, test or no

test. Therefore the test does not lead to any particularly useful advice to patients.

It will be increasingly feasible to examine variance in multiple genes that contribute to a common disease, Burke said. In such situations it will be possible to identify a very small proportion of the population that has high risk, where the positive predictive value is much higher than would normally be obtained with a single-gene variant. An example of this is age-related macular degeneration. Using variation in three genes, researchers were able to estimate that a small percentage of the population (about 1 percent) had a risk of greater than 50 percent for age-related macular degeneration (Maller et al., 2006). One would want to engage in careful monitoring with this group and, when definitive preventive therapies are developed, these individuals would be the first candidates to receive treatment.

What one also finds in the course of these tests is that most people have risks that are a little bit above or a little bit below the population average. How can one manage that information in order to extract from it what is clinically useful without getting distracted by a lot of information that is not clinically useful for most people?

There are early indicators that genetics has a powerful ability to characterize and classify disease. Both HER-2-neu amplification (which identifies candidates for herceptin therapy) and gene-expression profiling for breast cancer are good examples of this ability. These are harbingers of an important way in which genetics will provide tools to improve practice in the future. There are also a few novel therapies, such as Gleevec, where an RNA therapy is directed toward a messenger RNA of the virus.

Looking at the idea of different pathways from genetic research to clinical benefit, much of the discussion is focused on the current management of genetic information and what benefits might flow from it. Indeed, very substantial benefits are beginning to emerge. It appears that disease classification and innovative therapy will be two major contributions of genetic research and, perhaps, the major contributions from genetic research to health care. Because those contributions are predicated on the goal of improving health outcomes, the technology assessment issues are analogous to the technology assessment issues for other kinds of health care and pose the same kinds of challenges.

An interesting question is the extent to which tests for gene variance that are associated with increased risk will become an important modulator of either disease classification or innovative therapies. That is, will these tests help achieve greater benefit?

Part of the complexity, as with other innovations, is point-of-service information. That is, how do we integrate these innovations? Clinicians and health care systems are concerned about billable services that improve

health care. Clinicians clearly want evidence-based guidelines. Furthermore, there is a need to think about high-quality, cost-effective methods for performing the kind of education and counseling needed by the patient. How to accomplish that in a less resource-intensive way or in a more efficient way is a major challenge, Burke concluded.

VIEW FROM THE TRENCHES: CHALLENGES AND OPPORTUNITIES IN PERSONALIZED MEDICINE

Brad Gray
Genzyme Genetics

Diffusion of genomic innovations into the practice of health care through new product launches requires a balancing of economic risk and reward, Gray said.

The old paradigm in medicine is a series of actions: observation of a disease condition and action to treat it; an observation of response; and then a correction if the desired response is not achieved immediately. When this leads to innovation and improvement, it is deemed a success. In the long run this trial-and-error medicine can lead to great innovation, but in the short term, for the individual patient, it can provide a long, arduous path to identifying the correct treatment approach.

There is a new paradigm for personalized medicine, however, one in which complex testing (some of which is genomic, some of which is proteomic, and some of which is other technologies) plays a central role in linking observation to tests and therapy. In such a paradigm, observation is followed by a test that provides specific information for better decision making. This, in turn, is followed by the action, which would be the therapeutic choice or regimen that leads to a predictable response, thereby breaking the cycle of trial and error.

A series of technological innovations has made it possible to categorize diseases much more specifically, transforming what had been gross categorization into very narrow classifications based on genomics, and this, in turn, has improved care significantly. One hundred years ago, all blood cancers would have been classified as one disease, the disease of the blood. Over time, however, it was recognized that the cancers of the blood could be divided into leukemias and lymphomas, and later it was understood that there are actually several different types of leukemias and lymphomas. With our current ability to understand the specific protein expression, the

specific morphology, and the genetics of these diseases, they are being further disaggregated, and it now appears that there may be tens of different diseases in what was once categorized simply as blood disease. As a result, treatments can be tailored to the specific disease, providing a significant reduction in the risk of dying from the disease in the near term. This pattern is a striking example of what will happen in many other diseases, Gray said, with the first improvements probably occurring among the cancers and then later moving into the rest of the disease burden.

There are a growing number of drugs on the market that are tied to specific tests that help identify those patients who would benefit from them (see Figure 3-2). In general, the tests are used to answer one of the following four questions. Which drug should be used? How should the dose be tailored for a specific patient? How can it be confirmed that the drug is actually working for that patient? Is there a response observed? A variation of the third question is, is the response strong enough to say the disease has been cured?

The timeline of personalized medicine can be divided into three phases, Gray said. The first phase is fear, the second relates to value, and the final phase is acceptance. A few tests have gone all the way through these three phases, but the vast majority are still stuck somewhere in the middle of this timeline continuum because of the barriers encountered.

Which Drug Should I Use?		
Breast Cancer	Tamoxifen®	ER/PR
Breast Cancer	Herceptin®	HER2
Leukemia, Chronic Myelogenous	Gleevec®	BCR-ABL
Colorectal Cancer	Erbitux®	EGFR
Lung Cancer	Tarceva®	EGFR
Leukemia,MDS	Revlimid®	Deletion (5q)
How Much of the Drug Do I Need?		
Colorectal Cancer	Camptosar®	UGT1A1
Is the Drug Working?		
Leukemia, Chronic Myelogenous	Gleevec®	Quant BCR-ABL
Leukemia, Chronic Myelogenous	Gleevec®	BCR-ABL mutations
Is My Disease Gone?		
Leukemia, Chronic Lymphocytic	Campath®	Minimal Residual Disease

FIGURE 3-2 Personalized drugs available today.
SOURCE: Gray, 2007.

In the first phase (fear) there are a number of barriers. Pharmaceutical companies are concerned about their markets being constricted in size by the narrowing of the definition of the disease or its indication. Payers are understandably concerned about making sure that, as these additional tests are performed, there is actually a reduction in cost or an improvement in outcome that appropriately compensates for the additional expense. There are physicians who are concerned that testing will constrict the way they practice medicine. Patients may worry that if a test result comes back negative, they may actually be denied access to a treatment they see as important to their health or survival. Regulators are concerned about how to address the complexities of genomic innovations. Finally, the diagnostics industry, which sees genomic testing as an opportunity, also sees extreme risks and uncertainties concerning the clinical value of these tests as well as risks and uncertainties pertaining to regulation, market adoption, and reimbursement.

In 2005 Genzyme Genetics made a significant push in the field of personalized medicine, focusing explicitly on tests that could be directly tied to a therapeutic intervention. The company was aggressive in licensing technologies with early but promising clinical data that had been published in reputable journals. The company then worked quickly to get those technologies into the marketplace, believing that physicians would be convinced of their value as the data grew stronger and that a test that helped determine the dosing of a therapy would be a compelling value proposition.

Two tests that Genzyme Genetics brought to market offer enlightening examples and shaped the way in which the company currently thinks about new product development. First, the company aggressively brought to market the UGT1A1 test associated with irinotecan. In June 1996 the FDA had approved irinotecan for second-line treatment for colorectal cancer, and over the next several years a series of studies indicated a connection between polymorphisms in the UGT1A1 gene and toxicities stemming from irinotecan dosing. In June 2005, in response to those studies, the FDA approved the addition of information to the irinotecan label warning that a different starting dose should be considered for people who were homozygous carriers of a certain allele in UGT1A1. Very quickly thereafter, in August 2005, the FDA approved a device for detecting this allele that was manufactured by Third Wave.

Genzyme Genetics was very excited about this technology, believing that an FDA label that included a recommendation of the use of the test would be extremely compelling and that eliminating these toxicities would be universally desired, and so the company worked with Third Wave to bring the test to the U.S. market very quickly. In December 2005, Genzyme Genetics launched the UGT1A1 polymorphism testing service. There was strong clinical evidence for the usefulness of this testing—strong enough

for the FDA to change the label—and there was an FDA-approved device that met all the criteria for clinical validity. From the company's perspective this was a very promising situation, one that seemed as positive as a situation could be.

The experience, however, turned out to be quite different. Physicians said such things as, "I don't need a test because I can start patients on irinotecan, and when side effects occur, I lower the dose, stop a cycle, or stop treatment," or "I monitor bilirubin level, so do not need to test." Physicians who were willing to test asked what dose to use if the patient did have the polymorphism because the dosage and administration section of the drug label did not offer details about what to do if a polymorphism was found. Some physicians decided that the specific polymorphism was fairly rare so that it was not worth testing all patients.

From these experiences, the company learned that clinical utility data are not sufficient to change clinical practice. Physicians will use work-around solutions when they are modestly effective. Additionally, the inclusion of a test or genomic information in a drug-package insert does not necessarily lead to testing. Finally, package inserts must be clear on the implication of the testing results for dosing, or else physicians will struggle to interpret them. After an initial pulse of interest, physicians in the United States have largely disregarded the use of UGT1A1 testing when prescribing irinotecan.

Dosing with irinotecan without UGT1A1 testing results in about $1,000 in additional costs because of adverse events, Gray said. Theoretically, that additional cost could be eliminated by testing every patient and dosing accordingly. Therefore, $1,000 is the value Genzyme Genetics would assign for the value of the test. The company, however, is reimbursed based on the current procedural terminology (CPT) code, where the dollars associated with the activities that are used to perform this test are totaled, yielding about $310. That figure, then, is the reimbursement for the test. Therefore, Gray said, the test is delivering three times the value of its cost—a compelling argument from a health-economic perspective.

One might argue that, for innovations such as this to flourish in the future, a larger portion of the health-economic value delivered to the system should be captured by the company making the test. In the case of the UGT1A1 test, the company struggled to drive adoption but captured only a fraction of the value being delivered.

A second high-profile product that Genzyme Genetics became involved in is the use of epidermal growth factor receptor (EGFR) testing. Mutations in the tyrosine kinase domain of the EGFR govern response to tyrosine kinase inhibitors (TKIs) in non-small-cell lung cancer (NSCLC). The first TKI for non-small-cell lung cancer was gefitinib, which was approved by the FDA in May 2003 for third-line treatment of advanced or metastatic NSCLC. Very shortly afterward, some prominent publications appeared that

discussed the way that mutations in this protein would predict the response or non-response to gefitinib. Then, in November 2004, the FDA approved a second drug in the class for second-line treatment of advanced or metastatic NSCLC: Tarceva from Genentech. In response, Genzyme Genetics aggressively pursued worldwide exclusive licensing of EGFR mutation testing. The company paid more than it had ever paid for an intellectual-property license and quickly drove a test to market.

Soon afterward publications emerged that seemed to question the utility of EGFR mutation testing for driving dosing. Since that time there has been disagreement about which is the correct biomarker to predict response to this class of drugs. In July 2006 the C-Path Institute announced an effort to try to resolve the question of biomarkers in NSCLC cancer, but results are not yet available.

When this product was taken to market, only a small minority of NSCLC patients who received TKIs—probably less than 5 percent—actually received the test, Gray said. The penetration is highest in the leading academic centers, where there is a willingness and an ability to navigate the nuances of the emerging evidence. Community physicians, on the other hand, have generally been reluctant to adopt this approach. They are confused about the multiple-testing options, and they use what they consider clinical information (e.g., patient's race, smoking habits) as a proxy for the mutation status. Furthermore, because TKIs are most often used as the last line of treatment in these patients, there is a reluctance to do a test that would suggest that certain patients will not respond.

The company learned several things from this experience. First, the connection between genetics and treatment is not always clear. Community physicians need education and assistance in understanding conflicting evidence. Robust clinical-utility data will be required to drive adoption by community physicians, who will continue to substitute work-around solutions when they are modestly effective. Furthermore, community physicians are not inclined, in general, to deselect patients from treatment. A test that selects patients in is much easier to sell than one that selects out, especially when there are few alternatives for those patients, Gray said.

The adoption curve for EGFR testing is still heading upward. While the EGFR mutation test has not been adopted as rapidly as a new drug therapy typically would be, the indicators are moving in the right direction. The National Comprehensive Cancer Network (NCCN) guidelines for non-small-cell lung cancer include the test, a point which Genzyme Genetics believes will help community physicians gain comfort with the utility of the test.

Based on past experience, then, Genzyme Genetics has revised its criteria for bringing new personalized medicine tests to market. First, for the company to invest in a test, the test needs to represent the only reliable way

to obtain information. When there are low-cost work-around approaches (e.g., measuring a bilirubin count or assessing smoking status), there is too much commercial risk to proceed.

Second, clinical evidence is absolutely critical to driving adoption. There must be proof-of-concept data from inventors, or it must be feasible to run a decisive experiment at reasonable cost and in a reasonable amount of time if the company is going to pursue the innovation.

Third, because reimbursement in the testing sector of the health care system has traditionally not been based on value but on activity-based costing, the economics must support investment in clinical and market development. The reimbursement path must be attractive, either by virtue of its intrinsic coding or because there is the possibility of making a compelling case to be reimbursed on a different basis than activity-based costs. Furthermore, the company will look for places to invest where intellectual property and know-how is available on an exclusive basis. In situations where only a non-exclusive product is offered, the company will not be able to justify the investment required to perform clinical research or to navigate the regulatory system.

Concerning licensors and inventors, whether in academic medical centers or in small companies, Gray said that they must be educated about the experience of Genzyme Genetics in this area. The company is now looking for a partnership structure that will provide the needed return on investment, given the risk the company would be making. Genzyme Genetics is still committed to personalized medicine, but with a far more realistic and cautious approach.

To overcome the barriers and to help innovators bring genomics to the market quickly and effectively, several things are needed. The first is education. More information about the new tests must be given to physicians and to health care providers. Furthermore, organizations such as the NCCN need to develop and provide clinical-practice guidelines. Such guidelines will help interpret information about the test for the benefit of community physicians, who will be playing a more important role in testing than they have in the past. There is also a need to start education about diagnostics and genetics early in medical school.

A second requirement is better data. Industry-wide cooperation is needed to collect and analyze data in a timely manner on the best use and outcomes with diagnostics.

Finally, Gray concluded, the reimbursement system must compensate the innovators for their expense and risk if innovation is to continue. That means that there must be movement toward reimbursement based on value delivered by the test rather than according to activity-based costs. Furthermore, reimbursement must appropriately take into account the regulatory burden undertaken to deliver the test to market.

DISCUSSION

Wylie Burke, M.D., Ph.D.
Moderator

One audience member asked whether the issue of legal liability might drive the adoption of some testing, even in advance of clear clinical utility. Gray responded that where the utility of the test is clear and where there is a very clear way to use the information, physicians will likely see performing the test as reducing their liability. With most tests, however, utility is not always clear, and there might be disagreement about how to use the information. Therefore, while liability is a factor, it is difficult to generalize about whether it will promote or inhibit the adoption of new genomic technologies.

Another participant said that she was struck by the parallel between the current state of genetic medicine practice and HIV. When HIV was first recognized, there was fear and stigmatization associated with the diagnosis. As effective treatments were developed and understanding of the disease increased, being HIV positive changed from a death sentence to a chronic disease. Still, because of the earlier stigmatization, there are still many controls and consents that must be included in counseling and other efforts surrounding HIV.

As one moves forward with integrating genetics into routine medical practice, the audience member continued, it will be important to evaluate what is occurring and to not maintain all the stigmatization and consent requirements that surround genetics today and that contribute to the lack of use of this information. Burke responded that it will be important to stratify and recognize differences in genomic medicine. When the possible result of a test is a terminated pregnancy, it is likely that there will still be a need for counseling, but the level of counseling needed will be different, for instance, in the case of a pharmacogenetic test.

Another audience member noted that several speakers used the warfarin example in their presentations. Physicians have a 40- to 50-year history of giving warfarin or Coumadin. Physicians also know that the risk of using Coumadin is highest in the first month and trails off by the third month of use. By the time a patient has been on the drug for years, the dose is rock solid unless there are changes in other medications.

The target of interest should be the detection or prevention of an adverse event during the first three months. That is a much smaller market than the millions of people already on Coumadin. The calculations are very different if one is trying to detect a rare adverse event in a defined small population versus whether one is using disease-based diagnostics,

such as APOE. If there were an effective drug for Alzheimer's disease and the dosage depended on their APOE genotype, that becomes an entirely different matter.

One member of the audience noted that Gray, in his presentation, had described two incentives that diagnostic companies have for generating good-quality clinical data, had discussed the concept of value-based reimbursement, and had explored the ideas of gaining monopoly in a test through the use of intellectual property and a biomarker. What is unclear, the questioner said, is how, without a monopoly and a biomarker, one can squeeze value-based reimbursement out of the payers. If one does not have a monopoly, then the diagnostic companies are simply all going to compete with each other and drive down the price.

Gray responded that the questioner was correct and that the situation is borne out by the examples presented. Those examples illustrate that the innovations that have achieved value-based reimbursement are all cases in which a company owns intellectual property or know-how that cannot be replicated by competitors.

The final question for the panel involved the use of direct or indirect evidence in technology assessment. Direct evidence from clinical trials is preferable, the questioner said, but very few genomic innovations will proceed along that pathway. Indirect evidence, if one can construct the biological pathway, makes sense. The problem is, as Tunis described, that the evidence lines are not clear. What the FDA requires is different from what the third-party payers use. Industry wants the incentive to invest, that is, they want to recoup their investments.

The problem is how to proceed. CMS tried the concept of coverage with evidence development. Could something like that work for innovations that may be close to showing some clinical utility but that still need a clinical trial to demonstrate the additional benefit?

Tunis replied that, for certain clinical applications, obtaining definitive evidence of clinical utility is going to be extremely lengthy, burdensome, and costly. Part of the new paradigm may require that the payer become comfortable with reimbursement tied to indirect evidence or to some threshold of clear clinical validity plus promising evidence of clinical utility with the subsequent documentation or verification of clinical utility occurring in post-market.

This may be generally true for diagnostics, Tunis said. The evidentiary burden of demonstrating an effect of diagnostics on clinical outcomes through RCTs is heavy, whether it is for genetic testing or for CT angiography. Therefore, some kind of conditional reimbursement that presumes that some of the additional questions about clinical utility will eventually be answered—not before reimbursement but after—is going to have to be part of the new approach.

4

Translation of Genomic Technology at the Clinical Level

A PRIMARY-CARE PROVIDER VIEW OF TRANSLATING GENOMIC INNOVATION

Alfred O. Berg, M.D., M.P.H.
University of Washington

Primary care is of growing importance in the translation of genomic innovations, and genomic innovations will achieve a bigger foothold in this country only if they penetrate into primary care, Berg said. Most of the coming innovations in risk profiling for chronic diseases and many of the pharmacogenomic applications will be very important components of primary care.

Primary care is generally thought of as family medicine, general internal medicine, and general pediatrics, but obstetrics and gynecology can also be included. Collectively, these medical specialties account for more than half of all visits to physicians in the United States, who were estimated to number more than 500,000 in 2006.

Primary-care physicians serve as the personal medical home for most patients. They are the first medical contact for most patients and are turned to by most patients when a new complaint arises. Primary-care physicians attempt to be comprehensive. Patients can bring any complaint, interest, or concern to them, and the primary-care physician should be able to assist them, either directly by providing services, or indirectly, by making appropriate referrals.

An important component of primary care is that it is continuous, allowing the physician and the patient to develop a relationship over time. This makes it possible for primary-care physicians to accumulate information about personal family history that would not be available to other specialists.

Primary-care physicians try to be community- and population-focused. A doctor can practice high-quality primary care only if he or she knows the community, knows what is prevalent in the community, knows the risks in the community, and knows what the community's health concerns are.

An important characteristic of primary-care practice is that the physicians see common problems. They specialize in breadth of knowledge and expertise. At the same time, they need to recognize patterns that suggest the unusual. In order to practice in this way, primary-care physicians need information systems and decision support. Because they have a very high volume of practice, their support systems must work on time and all the time. Primary-care physicians cannot wait until the evening or the next day to come up with answers.

In primary care, medical tests and interventions must be appropriate for populations in which rare conditions are actually rare. Tests with even small errors can have magnified effects. A test that has a 99.9 percent specificity can still be a catastrophe in primary care if the condition is rare because positive tests will often be false positives, requiring a further cascade of medical testing and intervention. Rare conditions are rare in primary care, as they are in populations.

For primary-care physicians to incorporate a new test or innovation, several conditions must be met. First, a new test or innovation must be available, feasible, and acceptable to the patient. It has to do what it says it does. It has to be accurate and reproducible. It has to improve clinical outcomes that patients would notice and care about when compared to current practice. For example, changes in laboratory values are usually not enough in primary care because the patient expects to actually experience improvement with a new test or innovation. A new innovation should not increase adverse effects. Finally, it should be "worth it"—that is, the patient should think it is worth it either with insurance or with out-of-pocket payments. The calculation that goes into determining worth is more complex and nuanced than what typically goes into a cost-effectiveness analysis.

Primary-care clinicians need authoritative advice. No one can keep up with the staggering volume of medical information or make sense of all the volumes of literature. Authoritative advice can help clinicians deal with complex decisions by identifying the key factors important to decision making. Furthermore, authoritative advice has the potential to improve the quality of physician decision making. Such advice can provide justification to patients, payers, and the legal system by presenting the criteria used to make decisions.

Clinical guidelines are one useful form of authoritative advice. They help transmit medical knowledge, and they assist in making decisions. Clinical guidelines are a way to set clinical norms that can be used in quality improvement, in privileges and credentialing, and in payment and cost control and that can be useful in medico-legal evaluations as well.

Evidence-based guidelines have three major hallmarks. First, they need to be explicit, meaning that they state exactly how to proceed. They need to be transparent, in that all of the information has to be available to users and to consumers. Finally evidence-based guidelines must be publicly accountable.

A report of the Institute of Medicine concluded that, in order to be useful, a clinical guideline must specify the clinical condition, the health practice, the target population, the health care setting, the type of clinician, and the purpose of the guideline (IOM, 1992). Therefore, it is not enough to ask, "Is test X a good test?" One must ask, "Is test X a good test in this particular clinical scenario?"

The Agency for Healthcare Research and Quality (AHRQ) has further specified process characteristics for developing a clinical guideline. One of the characteristics addresses panel selection. Selection could be based on expertise, which might include individuals with conflict of interest, or selection could identify individuals with a general perspective on evidence so as to avoid conflicts of interest. A second characteristic is that the problem for which the guideline is being developed needs to be specified, as does the literature search strategy. There must also be explicit statements about how the literature is analyzed, the criteria used to judge the quality of the literature, how the evidence is summarized, and how one moves from the evidence to the rationale.

Every guideline process has some level of subjectivity. There is no generally agreed-upon algorithm for moving from evidence to recommendation. A linkage between evidence and the recommendation must always be made, however. Furthermore, the decision-making process needs to be specified as explicitly as possible. Another process characteristic is that the guideline needs to focus on clinical outcomes—that is, not simply on intermediate outcomes such as laboratory tests or knowledge, but on actual clinical outcomes that patients or families would notice and care about. Also, clinical guidelines must be sensitive to cost and practicality.

AHRQ has developed a list of desirable attributes for guidelines. Guidelines should be valid, that is, they should be based on criteria that are public and accountable so that validity can be assessed. Guidelines need to be reliable. This means when guidelines are used in similar circumstances, there should be a similar outcome each time. Guidelines must also be applicable, flexible, clear, multidisciplinary, and documented.

In the area of genetics and genomics, Berg described primary-care physicians as being skeptical of genetic exceptionalism, that is, of the

claims that genetic information is qualitatively different from other types of information. And, indeed, many non-genomic tests in current use produce exactly the same kind of information that is promised for genetic tests—they provide information about risk, prognosis, and response to drugs and other therapies. Furthermore, they too have ethical, legal, and social consequences.

The one area where genomic tests are unquestionably unique is in their ability to provide information about family members. In almost every other regard, there are many tests used every day by primary-care physicians that provide the same kinds of information that is promised by genetic and genomic testing.

There are thousands of genomic tests available, but there is little regulation of those tests. There is also direct-to-consumer and direct-to-physician marketing. The result is that clinicians and consumers are confused and need reliable advice. There are precedents in providing reliable advice, such as the United States Preventive Services Task Force, which evaluates preventive interventions. The Centers for Disease Control and Prevention (CDC) in partnership with AHRQ is sponsoring an initiative called Evaluation of Genomic Applications in Practice and Prevention (EGAPP). The EGAPP working group has no regulatory authority and is an independent, non-federal, multidisciplinary panel. Those selected to serve on the panel do not have extensive financial or other ties to various stakeholder groups that would have a stake in the recommendations of the group; they were selected in a manner aimed at minimizing conflicts of interest. As a result, there are a number of generalists on the panel who have a fairly broad view of such things as laboratory testing, primary care, and ethical, legal, and social issues.

The first EGAPP guideline, which was released in December, relates to the use of CYP-450 testing in decision making about the kind and dose of selective serotonin reuptake inhibitor (SSRI) that should be used in patients newly diagnosed with major depression. As noted earlier, it is very important to be explicit and clear about the clinical context in which the guideline is going to be used. This guideline is very specific. It is addressed to a primary-care clinician seeing a patient with a new diagnosis of major depression for whom an SSRI is being considered as treatment, and it may not be at all useful in a different clinical scenario.

There are a number of other EGAPP reviews underway concerning such things as testing for early detection of ovarian cancer, testing before placing a patient on an antidepressant drug, testing for family-related colon cancer, testing for response to treatment for colon cancer, genetic profiling for cardiac risk, and breast cancer gene-expression profiling.

The quantity and quality of evidence that supports testing in typical practice settings has been disappointing. Research designs in the published

literature are weak. Some potentially important data are proprietary and cannot be examined. Furthermore, there is very little evidence on potential benefits and harms and no head-to-head comparisons with current practice. Comparison with current practice is one of the things that primary-care physicians are looking for. As mentioned earlier, for example, there are 50 years of practice in using warfarin. Could a test actually improve on the clinical outcomes? The tests have not typically been evaluated in real-patient populations but rather only in research centers. There is little information about cost and cost-effectiveness compared with current practice and essentially no information about the ethical, legal, and social implications, particularly for family members.

EGAPP will be publishing a paper that documents the outcomes it considers to be of importance, one of which is family outcomes. Additionally, EGAPP has specified a number of things that it believes are potentially important when genomic technologies are evaluated.

Genomic innovations that can be used to assess risk or to guide therapy hold great promise, and primary-care physicians are very interested in them. A major issue, however, will be recognizing the importance of appropriateness in primary-care settings; many of the most exciting tests that are being discussed do have important implications for whole populations that are typical of primary care. Additionally, new testing technologies must improve on what is in current use. In the next three to five years there are likely to be few examples of genetic tests that will meet standards for common use in the typical primary-care practice.

There is an enormous need for more and better-quality research on the effects of testing on clinical outcomes, both good and bad, with publicly available results. Having high-quality information about actual outcomes of testing is critically important, Berg concluded.

INTRODUCING A GENOMIC INNOVATION TO CLINICAL PRACTICE

Steven Shak, M.D.
Genomic Health

There has been a great deal of discussion about recent genomic innovations and the question of whether there is adequate evidence to validate their clinical utility. One challenge for the companies engaged in developing such innovations is to actively collaborate and to fund the studies to obtain that evidence. Because patients urgently need genomics to be translated into

clinical practice, it is important for various stakeholders to work together to conduct the right studies that identify potential breakthroughs that work as well as those that do not work.

When Genomic Health was started in 2000, Shak said, little had been done to bring biomarkers into oncology practice. Despite thousands of papers on biomarkers, there were few with any actual diagnostic tests for use in oncology. Genomic Health developed Onco*type* DX, the first diagnostic, multi-gene expression test for breast cancer treatment planning which has been commercially available since 2004. There is clinical evidence from multiple independent studies demonstrating the test's ability to predict the likelihood of breast cancer recurrence as well as the magnitude of chemotherapy benefit. Thus the test is useful for individual patients in judging their own likely benefits from chemotherapy and how much those benefits will likely exceed the risks of treatment.

Use of Onco*type* DX has been growing over the years and more than 39,000 tests have been performed for more than 6,000 physician orders. Furthermore, the test is now reimbursed by Medicare and other major payers. And recently an American Society of Clinical Oncology (ASCO) clinical practice guideline recommended the use of Onco*type* DX for node-negative, estrogen-receptor-positive breast cancer. That is the group of women for which the test was developed, and it accounts for about half the women diagnosed with breast cancer in the United States.

There were a number of challenges to realizing the promise of Onco*type* DX. Bringing this test to clinical practice required that multiple, independent clinical studies be conducted that were rigorous in terms of design, performance, and analysis as well as comparing the test to standard measures. Assay precision, standardization, and control, well-known principles in the field of laboratory medicine, were of extreme importance. It was vital to show clinical utility in order to show that Onco*type* DX would meet the needs of patients, physicians, payers, regulators, and even investors. Finally, even after meeting those challenges, there is still a need for continuing research.

Over the years there had been great innovation in the field of cancer. In the past century cancer treatment was largely one-size-fits-all. Tumors were diagnosed based on their site of origin. Clinicians knew that there were marked individual differences in breast cancers, lung cancers, and other tumors, but they did not have the tools or the insight to analyze those differences and make practical use of them. During the 1990s, however, new technologies such as gene expression assays were developed that now allow careful measurement and quantification of individual genes in a tumor in order to better understand individual differences.

Genomic Health chose to optimize a new technology for quantitative analysis of gene expression for use with tumor blocks. Every time a cancer

is diagnosed and a tissue sample is taken by either biopsy or surgery, that sample is sent to the pathology laboratory; there the tissue is "pickled" in a fixative and placed into wax or paraffin, where it can be stored indefinitely. It is this fixed paraffin embedded tissue, the tumor block, that is sliced, sectioned, stained, and examined by the pathologist in order to make the diagnosis of cancer. In the past, this material was not considered useful for conducting molecular studies, but scientists at Genomic Health undertook an effort to look at RNA in tumor blocks in a precise and quantitative way and, after two years of effort, developed an assay system that enabled them to do so. This was important for two reasons. The first is that it allowed the test to be practically useful to examine tumors as routinely processed by hospitals and pathology laboratories at the time of diagnosis. The second, and more important, reason was that it enabled Genomic Health to analyze data on women who had been diagnosed with breast cancer in the past and whose outcomes were known, which, in turn, lead to the development and validations of a multi-gene test for breast cancer treatment planning.

Genomic Health's technology uses real-time quantitative reverse transcription polymerase chain reaction (RT-PCR) to quantify RNA. It is reliable, sensitive and specific, it has a wide dynamic range, and it is highly reproducible. A series of studies was performed to optimize the assay system so that it would work with the tumor blocks and define and minimize all sources of assay variability.

With this innovation, Genomic Health developed a strategy and a plan for developing a breast cancer genomic test that was focused, from the very beginning, on the challenges facing physicians and patients. Physicians and patient advocates said that what was needed for node-negative, estrogen-receptor patients at the time of diagnosis was an increased ability to pick out the truly low-risk patients and, most importantly, to determine who would benefit from the cytotoxic chemotherapy.

The critical need for more individual information can be appreciated by reviewing the clinical trial results which examined the benefit of chemotherapy. The landmark trial that changed the care of breast cancer was performed by the National Surgical Adjuvant Breast and Bowel Project (NSABP) from 1988 to 1997. This was a controlled trial that randomized patients to either Tamoxifen alone or to Tamoxifen plus chemotherapy. The results of this study demonstrated the benefit of chemotherapy (see Figure 4-1), and, based on these results, chemotherapy became the recommended treatment for the vast majority of patients.

The study did find, however, that more than 85 percent of women will survive without recurrence with just Tamoxifen and no chemotherapy. By definition, Shak said, the vast majority of women have been overtreated because it was not known which women were going to suffer recurrence or who would benefit by adding chemotherapy.

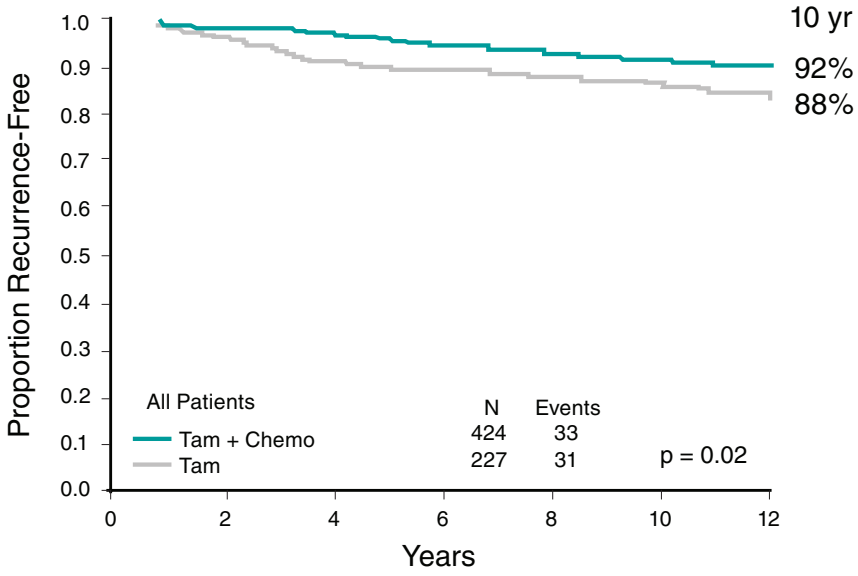


FIGURE 4-1 NSABP B-20 clinical trial (1988-1997). Tamoxifen vs. Tamoxifen + Chemotherapy—All 651 patients.
SOURCE: Shak, 2007.

Another illustration of the problem can be seen in a case presented in 2002 to an audience at an ASCO conference. The case was a 40-year-old woman with an invasive ductal carcinoma. She had a node-negative tumor of 1.1 centimeters. The estrogen receptor/progesterone receptor (ER/PR) was positive, and HER-2 was negative. The presenter asked the physicians in the audience what they would use to treat this patient. Fifty-four percent said they would use hormonal therapy, while 45 percent stated they would use hormonal therapy plus chemotherapy. There was no consensus.

The challenge, then, was to determine who would benefit by the addition of chemotherapy. It was to meet this challenge that Genomic Health applied the principles of drug development to the process of developing a diagnostic test. In particular, the company applied the principle of doing multiple studies with a logical sequence and rigor at each step, essentially analogous to the phase I, phase II, and phase III drug development trials.

The first series of studies was developmental and was designed to examine a set of genes to identify whether genes matter and, if so, which genes. Two hundred fifty candidate genes were analyzed in a total of 447 patients

from three separate studies,[1] which eventually led to a 21-gene profile and an algorithm for calculating a Recurrence Score[2] (Cobleigh et al., 2003; Esteban et al., 2003; Paik et al., 2003).

After defining a specific assay (in this case an assay of 21 genes), Genomic Health conducted two clinical-validation studies to test that particular assay independently in a prospective way on archival tissue from well-defined cohorts. The first study was performed in collaboration with NSABP to examine the Recurrence Score in the landmark NSABP B-14 clinical trial. The second study was performed in collaboration with the Division of Research at Kaiser Permanente to examine the Recurrence Score in a large, community-based observational study.

Finally, treatment-benefit studies of Onco*type* DX were undertaken. A study of the NSABP B-20 patients was made to determine the magnitude of the chemotherapy benefit as a function of the 21-gene Recurrence Score assay, with a Recurrence Score provided for each individual tumor. Patients who were randomized in NSABP B-20 to tamoxifen or to tamoxifen plus either CMF or MF chemotherapy were eligible. The primary analysis was prespecified to examine the tamoxifen-treated patients compared with those patients treated with tamoxifen plus either CMF or MF chemotherapy.

The hypothesis was that patients with a Recurrence Score of less than 18 would be at lower risk. In fact, the study by Paik and colleagues indicated that the risk of recurrence in patients with a Recurrence Score less than 18 was very low, and there was little evidence that chemotherapy was of benefit (Paik et al., 2004). For women with a Recurrence Score that was intermediate—that is, a score of 18 to 30—there was an increased risk of recurrence, although the benefit of chemotherapy was uncertain for this group. For women with a Recurrence Score of greater than or equal to 31 there was a clear, large benefit from chemotherapy. These individual differences make a compelling argument for the use of the Recurrence Score to individualize treatment.

The regulations and principles of the Clinical Laboratory Improvement

[1]The first study, the National Surgical Adjuvant Breast and Bowel Project (NSABP) B-20 study, was a multicenter study in which tissue was analyzed from 233 patients in a homogeneous patient cohort characterized by having histologically negative nodes, estrogen-receptor-positive tumors, and treatment with tamoxifen and no other systemic therapy. The second study was a single-site study at Rush Presbyterian-St. Luke's Hospital in which tissue was analyzed from 78 patients, all characterized by having more than 10 positive nodes and treatment predominantly with chemotherapy or Tamoxifen, or both. The final study, at Providence St. Joseph's Hospital, was a single-site study in which tissue was analyzed from 136 patients who were either node positive or negative, ER positive or negative, and treated with tamoxifen or chemotherapy.

[2]"The Recurrence Score is a number between 0 and 100 that corresponds to a specific likelihood of breast cancer recurrence within 10 years of your initial diagnosis" (http://www.genomichealth.com/oncotype/about/patresults.aspx, accessed January 15, 2008).

Amendments (CLIA) require that all assay methods and procedures be defined prior to clinical validation studies. These methods and procedures include specimen eligibility, reagent qualification, instrument validation, controls and calibrators, and linearity, precision, and reproducibility. In the case of Genomic Health's clinical validation study, it took 6 to 9 months to finalize the assay format and show its analytical performance.

The Onco*type* DX process is complicated. It involves multiple steps and requires more than 150 standard operating procedures and more than 90 forms to run the test, but the attention to detail assures that the Recurrence Score obtained on an individual tumor a year ago is going to be the same as the score today, and next year, and the year after that. Traditionally, women have been described as having tumors that are simply ER positive or ER negative. With the new molecular, quantitative assay, however, one can see that it is not the case that there are just two types of breast cancer. There is continuous biology and considerable heterogeneity in estrogen-receptor-positive tumors.

Prior to the use of Onco*type* DX, if one used conventional measures to define patients for treatment, only 7 percent of all women who were node negative and ER positive would be found to be low risk using National Comprehensive Care Network (NCCN) guidelines; the vast majority would be designated high risk, and chemotherapy would be recommended for them. With Onco*type* DX, however, many patients are classified differently. For those classified as low risk by conventional measures, about a third may be undertreated because their Recurrence Score is intermediate or high. Conversely, there are many more women (almost half) who are judged by the conventional measures to be at high risk when, in fact, their tumors have a low Recurrence Score (RS < 18) and they might obtain only a minimal or no benefit from chemotherapy.

There are currently four separate studies examining the use of Onco*type* DX as a guide to treatment planning in clinical practice. One study by Oratz and colleagues showed that 25 percent of treatment recommendations changed with the use of Onco*type* DX (Oratz et al., 2007).

Several important factors facilitated the introduction of Onco*type* DX. First, the suite of clinical utility studies was designed to meet the needs of patients, physicians, payers, and regulators. The acceptance of Onco*type* DX by physicians and payers is driven by a number of factors, the most prominent being published evidence. Publications do matter and are a method for community education and understanding. Before the publication of clinical validation studies and the presentation of chemotherapy benefit data, there was very little use of the test, but following publication use increased. The investment made in helping physicians, health care providers, and payers understand clinical research and how to interpret data was key to getting them to accept the test.

Another important factor was federal funding. Genomic Health could turn to the NSABP to collaborate on these studies. The NSABP had the foresight—and the federal government supplied the funding—to collect tumor blocks 15 years ago and save them, not knowing that anyone could develop an assay in the future that might be used on them.

One challenge was technology assessment criteria. The various groups involved in assessment do not use uniform criteria or evidence, even if they are using the same data. Each group assesses a new technology from its own perspective and context, which may or may not be the most useful or relevant. In defining the conceptual framework for evaluation of clinical effectiveness, it is important that the focus be on transparency and putting breast cancer patients first. There is a need for studies that establish clinical utility by showing directly, or by inference, that use of the test will improve outcomes and spare toxicity and health care resources. The studies must also rigorously compare the new test with traditional measures for decision making.

In the first quarter of 2006, Medicare provided coverage of Onco*type* DX. Reimbursement had a dramatic effect on the use of the test, and, as major payers began to reimburse for the test, use increased.

Another factor that mattered in the translation of Onco*type* DX was the conduct of treatment-decision studies. Although the validation studies were conducted on archival tissue from women treated in the past, it was important to conduct studies that examined the experience of physicians actually using the test in practice. These studies showed that treatment decisions changed. It was also important to conduct health economic studies.

There are two threats or challenges for Onco*type* DX. Traditionally, the reimbursement system has generously reimbursed for drugs and therapeutics and has reimbursed diagnostics poorly. This has clear consequences and is a clear threat to continued innovation by the diagnostics industry. The second threat is regulatory uncertainty. If the path is clear, one can plan how to proceed. But it is very difficult to make adequate plans in an uncertain regulatory environment.

Continued research is needed. Reimbursement for the test is now at $3,650. A significant portion of that money is going into new studies that will examine additional questions in breast cancer and that will begin to look at other tumors such as colon cancer, prostate cancer, as well as treatment selection for targeted drugs.

The road to realizing the promise of genomics is difficult. It takes innovation. It takes multiple, well-designed clinical studies. One must pay incredible attention to assay precision and standardization. One must focus on clinical utility and reimbursement from the beginning all the way through to the end of the translational path. Finally, one must understand the importance of a team working together. None of this would have been possible without an

incredible group of people in industry, academia, the regulatory environment, the National Cancer Institute and the Cooperative Groups, the breast cancer advocates, and patients and their families.

The women who participated in the landmark clinical trial 20 years ago, donating their tissue at that time, would likely feel very good that their participation in that trial is now helping women today, perhaps their daughters or granddaughters, to obtain more informed breast cancer treatment, Shak concluded.

DISCUSSION

Wylie Burke, M.D., Ph.D.
Moderator

One questioner asked Dr. Shak to comment on infrastructure development within National Institutes of Health (NIH)-funded research, given that Genomic Health was able to take advantage of data that had been collected from randomized controlled trials (RCTs) and that it appears that such opportunities will be limited in the future. Shak responded that funding to collect and save such data is a critical issue. Cooperative groups are still conducting multiple trials. Providing funding to collect data and save tumor blocks would enable others to use those data in the future in order to learn about optimizing treatment. Observational studies might be used to answer some questions, Shak said, but one must be very cautious in using such studies to look at treatment questions because often the treatments instituted are not done without bias.

One participant stated that RCTs have been identified as the gold standard for research but are rarely used for the study of genomic innovations. Given this, should there be major effort undertaken to determine how to make the best use of or optimize observational data? Is the knowledge for how to do that already available, or is generating that knowledge a task that also must be done?

Berg said that the evidence on the Onco*type* DX appears to still be in the indirect category. There are trials underway to answer the question directly about whether use of a test actually changes not only the clinician's recommendation for treatment but also the patient's choice and her ultimate clinical outcome. What kind of resources would it take to actually answer that question? Will a lot of very good-quality, indirect evidence be enough or, for such an important disease with such an important outcome as breast cancer, should one insist on a properly conducted randomized

trial or trials in order to find out whether the test has the clinical outcomes that are promised?

Shak responded that the current evidence is clearly indirect. The most direct and logical approach would have been to do an RCT, but there was no support in the community for that approach, nor were the resources available to conduct such a trial. The TAILORx trial that is being conducted assumes that the Onco*type* DX test does what it says. The patients who have a low score are all given hormonal therapy. Those with a high score are all given chemotherapy. The critical question that is being asked is, for the group of women for whom there is uncertainty about the benefit of chemotherapy (those in the mid range), what is the actual effect of chemotherapy?

When this study is completed it will be possible to use the new technologies developed between now and the end of that study to examine the tumor blocks and, it is hoped, have a definitive test. One can return to these cohorts and conduct multiple studies in a rigorous manner in order to build confidence in the test.

An audience member asked where biobanking might fit in the discussion of evidence and data. Some countries have national biobank systems, and there are smaller biobank systems in the United States. What, she asked, do the presenters think about biobanking, either as a national effort or as a more comprehensive local effort, perhaps a Framingham-for-biobanks approach?

Burke answered that in biobanking large amounts of genomic information and clinical information are combined. There is a great deal of empirical research, currently in the early stages, that is looking at participant attitudes, researchers' concerns, and IRB personnel concerns around issues of biobanking. The issues are complex, with much concern about harm and participant safety.

The example of using stored tumor blocks to ask a question about a therapeutic intervention that was not anticipated illustrates the value of biobanking. For that particular example, the concerns were lessened because the later research addressed precisely the issue that the original samples were collected to address—that is, improved outcomes for breast cancer patients. The concern with biobanks is that samples collected for one purpose might be used for other purposes. This is likely an area around which continued discussion will occur.

One questioner asked if prospective trials will be used to examine whether Onco*type* DX is having an effect on the clinical outcomes of women with breast cancer, rather than just on the change in the decision about whether to add chemotherapy to the treatment.

Shak said that the National Cancer Institute is talking about methodology for conducting prospective trials on archival tissue. This is a well-

accepted approach in many circles, especially in situations where multiple studies have been conducted and the results have been the same. What is important is rigor. There are good prospective studies and bad ones. In terms of conducting studies on archival tissue, there are good studies and, probably, bad ones. The goal is to become well educated about being able to assess genomic tests, their advantages, and their disadvantages.

Another comment was that Onco*type* DX is a diagnostic, developed as a business model to be performed as a laboratory test and, as such, is not subject to the same regulatory environment that in vitro diagnostic (IVD) test kits have to deal with. Given this, what will be the effect of the new IVD multivariate index assays (MIA) draft guidance from the FDA? How will that affect Genomic Health's business model? And how will all the different pathways of bringing diagnostic tests to market compare with one another?

Shak responded that there will be a number of gene-expression tests that will be done by different groups in different ways, some by reference laboratories, some by kits. That is where transparency and rigorous assessment of tests will be important.

One audience member thanked Genomic Health for pioneering a great approach based on archival tissues and for a product development that set a high bar for the rest of the industry to follow. He then asked Shak to describe the magnitude of the effort involved, including the money invested in the development of Onco*type* DX before it was launched and the amount of ongoing investment in clinical research. He also asked Shak to comment on whether such an investment was going to be typical of this type of diagnostic development.

Shak responded that the company probably spent somewhere between $50 million and $100 million over about 7 years to arrive at its current situation. Genomic Health is now working on colon cancer and has been able to take advantage of some of the lessons learned with Onco*type* DX, so some things are less costly. One of the great things in the field of genomics, Shak stated, is that creative people at the bench are developing better ways of solving problems, so some costs should decrease.

It is important to emphasize that much of the cost relates to the quality control required to obtain the same result again and again. All of the reagents used are quality controlled. They need to meet particular specifications, as do all the machines. The laboratory personnel are all highly trained and licensed. It is an investment that is often underestimated.

An advantage of using archival studies is that the costs of enrollment and long-term follow up of patients in clinical trials are avoided. Incurring such costs would have made it prohibitive for Genomic Health to undertake the studies. In a way, then, Genomic Health has been able to leverage the investment of the community's resources for the benefit of patients.

It will still be a challenge to proceed in other areas, such as colon cancer, particularly if coverage by payers becomes a major issue.

One audience member noted that Berg had talked about the importance of clinical guidelines. It is important to think about how to increase the number of clinical guidelines and, more important, how to accelerate the adoption of those guidelines. The data on adoption indicate that it is slow and painful. It is probable that EGAPP will help accelerate the generation of professional guidelines on genetics, but how can adoption of those guidelines by primary-care and specialty providers be accelerated?

Berg responded that one must be careful which guidelines are implemented and how they are implemented. Keeping that in mind, however, the way to accelerate adoption of innovation is to tie it to reimbursement. When there is system support (e.g., through development of guidelines) and reimbursement, adoption is often rapid. A member of the audience pointed out that just because something is tied to reimbursement does not mean it is a good thing. Burke responded that it is important to properly align reimbursement incentives.

One questioner noted that Shak described paying close attention to what the providers and patients viewed as clinical utility, which in this situation was avoiding the unnecessary use of chemotherapy. From that perspective, Berg was asked to comment on where he believes the gains in genomics might be. For example, will gains be in the area of pharmacogenetics, or are they more likely to be in some category like disease classification? What are the pressure points in primary-care practice where genomics might play a role?

Berg responded by pointing out that about 500 women need to be screened to prevent 1 breast-cancer-related death in 5 years in a 50-year-old. That means that for a woman walking into a primary-care practice who says, "I'm 60 and I think I need a mammogram," the response would be, "Great. You can have your mammogram."

In the example described above of the early trials with adjuvant chemotherapy, the difference between the two groups—those who had the chemotherapy and those who didn't—was 4 percent. One woman out of 25 benefited from the chemotherapy. If a woman is willing to take a 1-in-500 chance of benefit with a mammogram, is there any feasible scenario where she is going to turn down a 1-in-25 chance of benefiting from chemotherapy with or without the test? This is why it is so important to test these things in practice, to find out what recommendations are made to the patients, what decisions are actually made, and whether the test makes a difference.

A strong argument can be made for prospective clinical trials with Onco*type* DX, even at this point, because it is not yet known how patients and clinicians will actually use the test, whether they will respond to the

risk information in a novel way or whether they will still opt for treatment with chemotherapy. In one of the trials there was a 1-in-50 chance of benefiting. Is that still enough for a woman to say she wants the chemotherapy? At this point, no one knows.

Burke said that this illustrates the issues of how many of these kinds of questions must be answered pre-market, how many of them can appropriately be answered post-market, and how the infrastructure necessary for conducting the studies can be created.

One audience member noted that a point had been made earlier about educating physicians. From a primary-care perspective, how can one make physicians more receptive to new innovations? Berg responded that the American Academy of Family Physicians had an annual clinical focus on genetics and genetic testing. It was disappointing in that clinicians said that the kinds of educational offerings were not things that their patients were asking for and they were not things that seemed to be common enough in their practice for them to make the investment in learning. Although acknowledging that he is not an expert in the area, Berg said that he believed that primary-care physicians need to see how the innovation is going to directly benefit their patients and make their practice life better. If one can make those connections, the educational program is much more likely to work.

One audience member commented that the compelling question for payers is whether the outcome would be better if Onco*type* DX were to be used. There are common tools that are used to make decisions about chemotherapy. While there is a substantial benefit from using chemotherapy, there are also risks. What is impressive about the reclassification studies is that one could compare the incremental benefit due to use of Onco*type* DX with the benefits from a conventional tool and link that to an outcome known from a randomized controlled trial, albeit one that had been conducted in the past. Although this study approach does not fit easily into the clear conceptual boxes of indirect and direct evidence, the evidence appears more direct, the speaker said.

Women in the archival data were women who were on tamoxifen, which is the less prevalent treatment today. Today the more common treatment is aromatase inhibitors. To make a further inference to aromatase inhibitors would be an example of indirect evidence, where one needs to link bodies of literature and make assumptions.

Berg said that ultimately there will always be some subjectivity in moving from evidence to a recommendation. One could probably determine a way of objectively characterizing the evidence, as Shak and Aronson have done. But the question is, for a given clinical situation, for whom is the evidence enough? Is it enough for the patient? Is it enough for the clinician? Is it enough for the payer? The EGAPP project is trying to answer those

questions for primary-care clinicians. The answer that EGAPP arrives at might be quite different than the answer that a group of oncologists would settle on. The context—that is, the clinical setting—matters enormously in making that subjective assessment about evidence.

Even though the evidence for Onco*type* DX can be characterized as very good, high-quality, almost direct evidence, it is not quite there yet, Berg said. Does it matter? The question is, for whom might it be enough? Maybe it is enough for current investigators. Maybe it is enough for payers. Is it enough for all clinicians or for all patients? That question cannot be answered.

Aronson responded that the issue is quite complex. But what is persuasive is that for women making a decision about whether or not to have chemotherapy, Onco*type* DX is a better predictive tool.

Another audience member drew attention to the fact that it took between $50 million and $100 million to get the test out, which did not include the cost of the original study from which the blocks were obtained. If that cost were included, the figure would be much higher. And this is just the first generation of the test. Will a prospective randomized trial be done for every generation, and, if not, how will the information be obtained? If it has to be done post-marketing, then there is a problem because the payers say, "There must be a decent RCT before this will be paid for." The data cannot be obtained, however, unless the patients get the test, and the test will not be given unless it is paid for. This is a quandary.

Berg responded that this is a common problem. In the case of treatments for post-traumatic stress disorder (PTSD), for example, out of the 13 treatments available there was adequate evidence to support only 1 of them. And this is a condition that affects millions of people, not just veterans. PTSD has been known for 30 years. There has been enormous public investment, but the studies are difficult and costly. The question becomes, how much is society willing to invest in answering questions about treatment of PTSD? And how does answering those questions compete with answering other important questions such as have been discussed today?

How much money would it take to actually answer the question about Onco*type* DX and improved outcomes if we wanted to do an RCT at public expense? What would the trade-offs be with respect to other priorities for research? These are much larger questions than can be answered in this workshop. They are extremely important, Berg said, but it should not be assumed that the area of genomics will be treated any differently than any other area. There are many areas in clinical practice that desperately need better quality research.

Shak commented that one thing that might help is if one moves on to other innovations when the early data on a particular innovation do not look compelling. There are two things needed for success: resources must

be focused on solving compelling problems, and, if the solutions developed are not important, if they do not make a large difference, then maybe the effort to continue is not justified.

Another participant expressed concern about the discussion of evidence-based medicine (EBM) because most of what is done in medicine today is not completely evidence-based, and it most certainly is not supported at the level of RCTs. Physicians incorporate a great deal of retrospective data when making their judgments. Is it reasonable to say that in the future everything must be prospective, randomized trial data? Presumably the answer is no.

When discussing EBM, one must be careful not to equate EBM with prospective, multi-center RCTs as the only way that physicians should make decisions. Not everything will be tested. Whether a drug works better in a 20-year-old versus a 21-year-old is not something worth testing. Judgment must be used.

It is not reasonable for insurers to require multi-center randomized controlled prospective trials if they are not paying for them, the speaker continued. Who will pay for these trials? If a high level of evidence is required, someone must pay to obtain that evidence. Physicians use evidence to improve their practice, but if it is incorrectly used or unreasonably required, evidence can be a bar to innovation.

Berg agreed but said that in the absence of evidence, no one is suggesting that nothing should happen. Clinicians and patients are accustomed to dealing with uncertainty; that is what much of the practice of medicine is about. Nonetheless, it is extremely important that someone draw a scientific line in the sand and say what it is we know and what we do not know, confronting and looking at the evidence objectively in order to determine whether it is worth the investment to fix.

There may be many questions that are not worth investing in, but there are many questions where we are currently saying, "Well, is it okay to substitute our judgment?" when in reality an investment should be made. The investment is long overdue, because there are many, many important clinical questions that could be answered but that are not being answered. Finding the answers is a public good that all should support, Berg said.

5

Opportunities and Constraints for Translation of Genomic Innovations

THE GLOBAL PERSPECTIVE

Stuart Hogarth
University of Nottingham

Finding something valuable can be difficult, Hogarth said. Innovations in genomics have been much more difficult and taken much longer to develop than many initially hoped.

Innovation is important, but most innovations fail, in many cases simply because they are not very good. Despite this, it is important to support innovation even while acknowledging that many innovations are not successful and never can be.

Some innovations are radical, but most are incremental. In thinking about innovation policy, one must think about the importance not just of major breakthroughs, such as finding a new biomarker and discovering its association with a disease or response to a drug, but also of incremental innovation. In the case of cystic fibrosis, for example, the development of robust, reliable test kits was just as important as the initial identification of the mutations in the cystic fibrosis transmembrane conductance regulator gene.

One must also think about the importance of the diffusion and use of genomic innovations. Indeed, diffusion and use may be more important than innovation in many ways. Science and technology innovation policy generally focuses too much on innovation and not enough on the diffusion

and use of existing technologies. It might be more important, for instance, to ensure that everyone is using two or three really good new tests than wondering how to encourage the use of another 100 that do not offer a significant advantage over existing technology.

Hogarth has been involved in a project to examine policy issues surrounding the evaluation and regulation of genetic tests. As part of the project, interviews and workshops were conducted with over 80 individuals from key stakeholder groups (industry, clinicians, patient groups, regulators, and policy makers) in Europe, Canada, the United States, and Australia.

The project classified policy issues into three areas: incentives and infrastructure for generating a robust evidence base for new innovations; regulatory mechanisms for the independent evaluation of evidence; and systems for ensuring that doctors, patients, health care policy makers, and payers have access to accurate and comprehensive information presented in a way that can be easily understood.

Other work in the area includes a project on information policy for pharmacogenetics and two reports for the Canadian government, one on regulating pharmacogenomics and another on the clinical application of molecular diagnostic technologies.

Genomic innovation transcends national boundaries. Multinational companies are involved, and there are global markets for the products. International research is being done by such organizations as the Human Genome Organisation (HUGO) and the Human Proteome Organisation (HUPO). There is transnational regulation and standard setting being carried out by such groups as the International Conference on Harmonization (ICH) and the International Organization for Standardization (ISO).

Innovations in genomics are affected by nongovernmental organizations, such as the Organisation for Economic Co-operation and Development and the World Health Organization, as well as by research funders with a global reach, such as the Bill and Melinda Gates Foundation. Innovation is also affected by transnational agreements such as the General Agreement on Tariff and Trade and the Agreement on Trade-Related Aspects of Intellectual Property Rights. The European Union crosses national borders and heavily influences innovation within Europe.

Genomics varies around the world in terms of the organization of health care delivery systems, the regulatory frameworks for innovation, and the economic incentives and infrastructure. On the other hand, there are a number of policy reports from across the world that express, in different ways, shared policy concerns.

The first such shared concern is that, in some cases, genomic innovations such as genetic tests have been moving into routine clinical practice too quickly and without enough independent evaluation.

The second concern relates to capacity building. Our health care systems need their capacity built through education and expansion of the workforce. There is a further need to enhance capacity in the specialty of clinical genetics and also to diffuse capacity more broadly across the health care system.

The third concern is the opposite of the first one: Some observers worry that, rather than moving too quickly, innovation is moving too slowly because of regulation and gate-keeping. The activities of regulators need to be understood in the context of changing policy priorities. Limiting the inappropriate use of new technology and controlling health care expenditures continue to be major concerns surrounding the health care system; but in the last decade or so there has been a marked shift in emphasis, and now an imperative to support the health care innovation process is emerging as a significant policy concern. Licensing agencies, such as the Food and Drug Administration (FDA) and the European Medicines Evaluation Agency (EMEA), and technology assessment bodies, such as the United Kingdom's National Institute for Clinical Excellence (NICE), are beginning to reconceptualize their roles in the innovation process. In particular, they are beginning to move from a strictly gate-keeping role, in which they evaluate evidence for the safety and effectiveness of new technologies, to a more collaborative or facilitative role.

This new policy orientation is taking concrete shape in programs such as the FDA's Critical Path Initiative and the Innovative Medicines Initiative in Europe, which is linked to EMEA's Road Map strategy. In the United Kingdom there is also the Clinical Research Collaboration, which is attempting to bring together key groups such as NICE, the National Health Service regulatory bodies, medical researchers, industry, and patients in order to create a new system of health care innovation.

Some of these initiatives involve new models of evaluation, while others involve new strategies for assisting the development of the evidence base for a new technology by providing either incentives (for instance, through conditional reimbursement) or the infrastructure for data collection. The new initiatives are often focused primarily on therapeutics, but they also have implications (and potential) for diagnostics innovation (not least because many are designed to support pharmacogenetic testing with new drugs).

The translation of pharmacogenomics into clinical practice has generally been slow. One factor that may be delaying the development of new pharmacogenomic products is a lack of clarity in the regulatory response to pharmacogenomic data. Other factors are the complexity of the science and various structural issues in the pharmaceutical industry. The result of these issues is what Hogarth referred to as a pipeline problem.

The first pipeline problem can be found in drug discovery and development. Biomarkers are frequently seen as the solution to this prob-

lem, but there are also problems in the discovery and development of biomarkers.

Regulatory agencies are uniquely positioned, given their responsibility for the development and enforcement of standards for drugs and devices, to shift the focus of the pharmaceutical industry from its preferred blockbuster drug model, which is aimed at broad populations, to a model that is more targeted. The regulatory agencies are also well positioned to encourage the participation of diagnostics companies in working toward this goal.

Pharmacogenomics, although providing an example of a novel approach to drug development, is but one aspect of a more general trend. The FDA's Critical Path Initiative and EMEA's Roadmap both see pharmacogenomics at the heart of a broader agenda for the enhanced use of novel biomarkers in drug development, diagnosis, and screening and the review of existing clinical trial design and statistical tools for drug evaluation. This agenda represents a shift in the role of regulatory agencies from guardians of public safety to a wider public health mission as supporters of translational medicine.

In general, regulatory authorities are moving cautiously, seeking to ensure that they do not act prematurely in a fast-developing area of science. Still, a number of general trends can be identified. One of these trends is the establishment of new mechanisms for voluntary sharing of genomic data, which is being done outside the formal approval process in FDA and is also being carried out in EMEA's pharmacogenomics briefing meetings and within a similar process in Japan. A second trend is the development of guidance on regulatory processes and types of data needed. A third is organizational restructuring in regulatory agencies. A fourth is the approval of new products and the relabeling of existing ones. And a fifth is a broad-based move toward international cooperation and harmonization.

There can be no doubt that the FDA is leading the way, in part because it has prominent champions of pharmacogenomics among its leadership and in part because it has far greater resources to bring to bear on this field than any other organization. A comparison of FDA and EMEA, for instance, shows that the FDA has 20 full-time staff in its interdisciplinary pharmacogenomics review group, while EMEA has none in its equivalent pharmacogenetics working group. However, the EMEA is also very active, albeit at a slower speed and smaller scale, reflecting both the resources available and the complex political relationship between EMEA and European member states. Regulatory agencies in individual European member states have little or no interest in pharmacogenomics.

While there are shared concerns, there are also some major differences between the United States and Europe. For example, the FDA has devoted considerable resources toward and places great importance on the relabeling of existing drugs as a strategic plan for promoting the use

of pharmacogenomics. Thus far, labeling updates have been advisory or cautionary rather than mandatory.

EMEA has been far more reluctant to relabel than the FDA. EMEA's authority in this area is limited since it appears that in those cases where drug approval was given on a state-by-state basis, then updating the drug label is the responsibility of the individual member states. Relabeling to include pharmacogenomic data does not seem to be a priority issue for the member states' regulatory agencies.

Just as is the case with the FDA, the EMEA has approved drugs co-developed with tests (e.g., Herceptin). Unlike the FDA, however, the EMEA does not have a diagnostics division and has no legal authority over the regulation of diagnostic tests. Authority for the regulation of medical devices under the European In Vitro Diagnostic (IVD) Directive resides at the member state level. Therefore, while the EMEA can evaluate the performance of a test codeveloped with a drug and can include strong recommendations for the use of testing as part of the drug label, it cannot mandate the use of a particular test kit. Furthermore, this regulatory gap means that the EMEA does not feel empowered to issue guidance on codevelopment.

No action has been taken at the European level by the expert groups that guide device regulation, and while the IVD Directive permits individual member states to take action when they deem it necessary, none has done so in relation to pharmacogenomics. EMEA officials, who are committed to the ideal of harmonization through the ICH process, would prefer to avoid a situation where individual member states take action.

This raises the issue of the need for a coherent and consistent regulatory framework for genetic tests. This has not happened on an international basis because of a series of regulatory gaps—different regulatory gaps in different countries. In the United States, for example, the primary regulatory gap is that, historically, the FDA has not regulated laboratory-developed tests as medical devices. By contrast, in Europe and Australia laboratory-developed tests are regulated as medical devices.

There have been some interesting developments over the past few years. Perhaps the most important one in Europe is that the IVD Directive will be revised and the risk classification system is probably going to change. It is likely that genetic tests will be classified as moderate risk rather than receiving the low-risk classification that they have in today's system. In Australia there has been a complete revision of the IVD regulations, primarily to address the issue of laboratory-developed tests and genetic tests. Australia has issued some guidance concerning nutrigenetic tests. Elsewhere, Canada has provided some guidance on pharmacogenetic tests.

Industry has emphasized the importance of clarity in regulatory guidance and the need to strike a balance between enhancing regulations and the creation of a clear pathway to market. One problem in the European

system now is that no standards or guidance for genomic tests are being generated.

Another issue of importance that crosses national boundaries is the issue of sustainable business models. The traditional IVD innovation model is an incremental process involving multiple parties. One starts with laboratory-developed tests and gradually works toward test kits at higher levels of automation. In keeping with this innovation model, the traditional IVD business model is based on intellectual property (IP) in test platforms rather than in biomarkers. Essentially, this business model leads to intense competition between companies, which offer different ways of testing for the same biomarkers. But with little protection on investment, relatively low margins, and little experience or infrastructure for clinical evaluation, the traditional sector is ill-equipped to undertake large-scale clinical studies. Furthermore, there is no economic incentive to invest in the kind of clinical studies discussed in this workshop. The use of a model with weak intellectual property rights in biomarkers has led to a situation where no one party is responsible for developing the data on the clinical validity of a new test. Academic studies and professional advocates have filled the gap, often promoting tests on the back of ad hoc clinical experience.

A lack of biomarker IP has created a disincentive for generating clinical data. Any one manufacturer who undertakes such clinical studies will be developing the market not simply for itself but also for all the other manufacturers, who will bear none of the risks but will share in the benefits. Indeed, the structure of the market is deliberately exploited by some IVD companies that specialize in being "fast followers," the first on the market with a "me-too" test. The problem is summed up by the industry maxim, "It's hard to be first."

There are a number of disruptive new business models appearing among companies that develop and market medical tests, and there is some evidence that the emerging field of molecular diagnostics has disrupted the traditional model in a number of ways. A number of companies have appeared that are developing genetic tests based on patent protection of the gene and its association with disease. **The emerging market for gene expression** and proteomic tests is based on similar strong intellectual property rights being claimed by companies like Genomic Health, Agendia, Avaria Dx, Correlogic, and Exact Sciences.

Strong intellectual property rights for biomarkers allow companies to charge higher prices for their tests for a longer period of time before the arrival on the market of competing products. **Higher reimbursement rates** are being seen for some new tests, including Genomic Health's Onco*type* Dx test, which costs $3,460, and Agendia's MammaPrint test, which costs $3,000. When companies have greater certainty of a return on their investment, they are more likely to invest in substantial clinical studies to generate

a proper evidence base for their tests. This anticipated return also gives small companies the leverage to access the money needed for clinical studies; they can raise money from venture capitalists or find a bigger partner, either a major diagnostics manufacturer, or a major reference laboratory. So IP has become an important incentive for funding clinical studies for new molecular diagnostics, and this new model can help to address oversight concerns about the lack of clinical data to support novel tests by offering clear incentives to generate that data.

There are concerns about this business model, however. The issue of pricing leads one to consider the particular regulatory challenges presented by monopolies. Market failures are a major justification for regulatory action and it is a well-established tenet of regulatory practice that the existence of a monopoly is in itself a market failure which provides strong justification for regulatory action. In particular, regulators will try to protect against abuses of the monopoly situation by making sure that consumers have access to goods and services of a decent quality and at a reasonable price.

IP in biomarkers can lead to monopolistic provision of tests, and the homebrew loophole has made it even more attractive for companies to develop their tests as in-house tests which are carried out on a monopolistic basis by the test developer or by two or three exclusive licensees. Many clinicians and laboratory directors have opposed this, arguing that the monopolistic provision circumvents the traditional (informal) methods of test evaluation, with in-house tests being subject to peer review in the field. They are concerned that it creates a situation where the only people who can perform a new test are those with a vested interest in its promotion, which in turn could lead to a situation where companies, in order to recoup their research and development investment, may make strong clinical claims for their tests at a stage when the evidence base is still developing. In recent years there has been repeated controversy over emergent IP-protected tests, with little agreement about when tests are ready for routine clinical use. The novelty and complexity of many of the tests involved only heightens concerns.

Another new business model is the rise of consumer genetics. In this model companies offer their tests directly to consumers. Some have suggested that this business model is a way to overcome some of the hurdles of translation. By taking the test directly to consumers, for example, one does not have to address the issue of physician reluctance to adopt. Consumer genetics is a disruptive business model, Hogarth said, because it marks the first time that new tests go directly from research to a consumer offer. There is significant national and regional variation in regulatory attitudes to direct-to-consumer testing which may affect this business model.

Business issues faced by IVD companies have regional variations. For example, venture-capital funding is far more available in the United States

than it is in Europe. Market size is also important; this can be seen, for instance, in the way that Canadian biotech companies that develop new tests will launch them first in the United States, next in Europe, and then, finally, in Canada. Of the 13 companies engaged in the gene-expression market, only 4 are located outside the United States, which illustrates the degree to which innovation is heavily focused on the United States.

In terms of the IVD industry and business models, then, there are a number of policy options to consider. One option is to support a radical restructuring of the traditional industry so as to move toward supporting the new model of biomarkers and monopolistic provision of tests. Another option is to focus on developing mechanisms for addressing market failures of the traditional model. Neither of these options will work on its own, however.

New business models are largely unproven and therefore cannot be relied upon. Intellectual property may turn out to be a poorly structured incentive, or it may be unavailable in many cases. What is needed is to take a case-by-case approach, supporting multiple innovation pathways. Such an approach is a much greater challenge for policy makers.

Another major issue is third-party reimbursement for genomic innovations. Companies are greatly concerned about this issue, not just in the United States but also in Europe. Reimbursement is a very powerful gatekeeper and has been the de facto regulator of genetic tests since payers frequently set stricter evidence standards than those established by licensing authorities. The Roche Amplichip is a good example. In 2004 it became the first pharmacogenetic microarray to gain FDA approval, but since that time the test has been rejected in a number of negative health technology assessment reports in the United States, Canada, and Europe.

Clearly reimbursement decisions can have a profound effect on clinical uptake of new tests. Yet if payers are informal regulators, then they face the same challenges as licensing authorities: how to wield that power responsibly and how to balance thorough evaluation with the encouragement of innovation. One option is conditional reimbursement—that is, paying for new tests but only on the basis that there is systematic data collection post market. Conditional reimbursement is one way of dealing with decision making under uncertainty and is also a way in which health care systems and payers can facilitate the process of evidence development. This model has been adopted by CMS in its Coverage with Evidence Development program and it is being used in the Netherlands, Germany, and Australia.

As can be seen, there are shared problems and policy concerns that cross national borders. There are also some interesting examples of international cooperation and harmonization. Inevitably, however, there is international competition. Each country, even within Europe, wants to promote its own biotechnology, pharmaceutical, and diagnostic sectors. There is

also variation in the capacity for action, based on many different kinds of structural issues.

The best innovators, Hogarth concluded, may ultimately not benefit the most from their innovations because they may not be the ones that are best at diffusion.

FINDING VALUE IN TRANSLATION OF GENOMIC-BASED RESEARCH

Deborah Marshall, Ph.D.[1]
McMaster University

Value in pharmacogenomics has recently taken on new importance, Marshall said. There are a number of reasons for this. For example, there is broader availability of pharmacogenomic testing for some commonly used drugs. The FDA has issued guidance about maximizing translation of pharmacogenomics from the bench to the bedside, including requirements to submit pharmacogenomic data alone and in combination with tests and treatments.[2] The Critical Path Initiative, which is intended to address the pipeline problem of getting pharmacogenomics to the bedside, is playing a role as well, and there are concerns about adverse drug reactions of these new technologies. Finally, there is concern about increasing prescription drug costs.

The new buzzword is value. Dr. Harold Varmus, former director of the NIH, has asked, "How much will the expanded use of genetic information further escalate the cost of healthcare, and who will pay for it?" (Varmus, 2002). These questions are not surprising given that there has been an 80 percent growth in the number of new drugs that are being prescribed, a 100 percent growth in new device patents, and a 1,500 percent growth in diseases with identified gene tests (Ferrusi, 2007).

What is "value" in genomic-based translational research? The Secre-

[1]This presentation was developed collaboratively by Deborah Marshall, Ph.D., of McMaster University and Kathryn Phillips, Ph.D., of the University of California at San Francisco.

[2]Three guidances are relevant. They are: Pharmacogenomic Data, March 2005 Procedural. http://www.fda.gov/cder/guidance/6400fnl.pdf (accessed June 2, 2008); Guidance for Industry. Pharmacogenomic Data Submissions—Companion Guidance, August 2007 Procedural. http://www.fda.gov/cder/guidance/7735dft.pdf (accessed June 2, 2008); and Realizing the Promise of Pharmacogenomics: Opportunities and Challenges, *Draft Report of the Secretary's Advisory Committee on Genetics, Health, and Society*. http://www4.od.nih.gov/oba/sacghs/SACGHS_PGx_PCdraft.pdf (accessed June 2, 2008).

tary's Advisory Committee on Genetics, Health, and Society (SACGHS) has suggested that for successful adoption into clinical practice, a pharmacogenomic test has to have analytic validity (that is, be an accurate test for the genotype), it must be clinically valid (the test has accuracy for the clinical outcome), and it must have clinical utility (that is, it has the ability to inform clinical decision making, prevent adverse outcomes, or predict outcomes).

There must also be economic value. Measuring economic value in pharmacogenomics involves three different elements: an evaluation of the cost of illness, criteria for cost-effectiveness, and criteria for economic viability. In examining the cost of illness, one examines the size of the problem in monetary terms: What is the relevant population, and what is the cost of disease burden? To determine cost-effectiveness one examines efficiency measured as marginal cost per unit of effectiveness of the new innovation versus the standard care. Finally, in considering economic viability, one takes the perspective of societal net benefit. To what extent is value-based pricing possible, as opposed to cost-based pricing? What is a fully informed patient willing to pay for the innovation?

HER-2 neu and trastuzumab provide good examples to illustrate each of these elements. The cost-of-illness framework (see Table 5-1) has five components: prevalence of the condition for drug treatment, mutation prevalence, utilization, drug expenditures, and condition expenditures. For HER-2 the population would be those patients with metastatic breast cancer or, in its new indication, early breast cancer. One also needs to know

TABLE 5-1 Data for Cost-of-Illness of Pharmacogenomics

Relevant Data	Description	Example HER-2 and Trastuzumab
Prevalence of condition for drug treatment	Size of the population for testing	Prevalence of patients with metastatic BC
Mutation prevalence	Size of the population in which testing could impact outcome	20–30% of BC patients overexpress HER-2
Utilization	Extent to which testing will be undertaken	Test costs $100 to $400
Drug expenditures	Testing could change drug utilization	Annual cost of treatment ~$30 to $80K
Condition expenditures	Measure clinical outcomes of testing on condition	25% increase in median survival

SOURCE: Adapted from Phillips and Van Bebber, 2005.

what the mutation prevalence is, that is, the size of the population in which testing could affect outcome. In the situation with HER-2, about 20 to 30 percent of breast cancer patients would over-express the HER-2 protein.

Other data require focus on utilization, drug expenditures, and condition expenditures. In terms of drug expenditures, testing will affect how the drug is used, so one needs to think about the annual cost of the treatment. In the case of HER-2 and trastuzumab, the cost might fall somewhere between $40,000 and $80,000 per year per patient. Finally, data related to clinical outcomes are necessary. For the example in Table 5-1, there is a 25 percent increase in median survival.

Moving to cost-effectiveness, one examines the difference in costs divided by the difference of effects between the two different paradigms. The mathematical expression is

$$\frac{\text{Cost (A)} - \text{Cost (B)}}{\text{Effect (A)} - \text{Effect (B)}}.$$

The new paradigm uses pharmacogenomics; the old paradigm is the standard care delivered. Some of the key factors and test characteristics for which pharmacogenomic testing would likely be cost-effective are shown in Table 5-2. The higher the prevalence of the mutation—that is, the more frequently it appears in the population—the more likely it is that there will be a favorable cost-effectiveness ratio. A favorable cost-effectiveness ratio is also more likely if there is a very strong association between the

TABLE 5-2 Criteria for Cost-Effectiveness of Pharmacogenomics

Factors	Characteristics Favoring Cost-Effectiveness
Prevalence of mutation	Variant allele frequency is relatively high
Severity of disease and outcomes avoided	Severe outcome, high mortality, significant impact on quality of life, or expensive medical care costs
Drug monitoring	Monitoring of drug response currently not practiced or difficult
Gene and outcome association	Strong association between gene variant and clinically relevant outcomes
Test performance and cost	A rapid and relatively inexpensive, but accurate test is available

SOURCE: Veenstra et al., 2000, and Phillips et al., 2004.

gene variant and clinically relevant outcomes. Finally, when one is looking at pharmacogenomic testing and treatment combinations, the best cost-effectiveness ratios arise from rapid, accurate, and relatively inexpensive tests (Veenstra et al., 2000; Phillips et al., 2004).

The third approach for examining economic value is concerned with market economics and value-based pricing. To support pharmacogenomic innovation, at least initially, the health system marketplace must provide attractive economics and a sustainable franchise to both the diagnostics and the treatment manufacturers. There needs to be a place where the product can be introduced in a viable way.

In examining market economics, one must determine to what extent value-based pricing is possible. The first criterion is that the test must be able to identify an appropriate patient population or subpopulation and to demonstrate the improved response. Second, value-based, flexible pricing for both the test and drug will provide stronger incentives for innovation. Third, there needs to be some kind of intellectual property protection—which is not as common in the diagnostic industry as in pharmaceuticals—in order to encourage and facilitate the innovation. Finally, there must be some kind of additional regulatory market protection aimed at facilitating innovation in this context (Garrison and Austin, 2007; Trusheim et al., 2007).

The era of blockbuster drugs is past, but there are opportunities for sufficient financial return through charging a premium price for the higher efficacy of a pharmacogenomic innovation, even in a smaller target population. An extreme example is provided by the situation with orphan drugs, but a more pertinent example is Gleevec. In this case the company was able to generate revenue of $2.5 billion even though only about 55,000 patients were eligible for this treatment. The drug generated an average revenue per patient per year of about $44,000 (Trusheim et al., 2007). The question is, how sustainable will this be in the long run, given the likelihood of disruptive competition that could improve performance and decrease costs?

There are many challenges in assessing value and these have implications for the translation of pharmacogenomic technologies to benefit patient outcomes. In order to be of value, pharmacogenomics must fill a knowledge gap that is clinically important to the diagnosis, prognosis, and treatment of patients. However, as discussed earlier, data and evidence of effectiveness are lacking. There is an ongoing debate about whether observational data can provide sufficient evidence of clinical utility, but not all genetic tests can be put through randomized controlled trials. When direct evidence is not available, one must consider methods for obtaining indirect evidence, including modeling approaches.

In the HER-2 example, no secondary data set was available to find real-world utilization of the test, so a chart review was conducted. This review

found wide variation in the types of testing performed. Most people received immunohistochemistry, fewer received the fluorescent in-situ hybridization (FISH) test, and some received both. There was variation in trastuzumab use by HER-2/neu status. Importantly, only 56 percent of the patients had documented evidence of actually having a clearly positive test in order to obtain treatment. This raises questions about whether the testing is being done appropriately and whether testing is a requirement for treatment.

A second challenge to assessing value is that there are very few economic models for pharmacogenomics. It is important to conduct economic modeling in order to understand the downstream consequences of the pharmacogenomic testing-treatment paradigm. In the long run, one must demonstrate value for adoption and reimbursement purposes. While the hurdles traditionally have been lower for diagnostics, the situation is changing. It may well be that future requirements for diagnostics will be relatively similar to those for pharmaceuticals.

Historically, diagnostics have been less studied than drugs. Up-front testing costs are perceived to be higher than downstream savings. Most products are not evaluated early enough. Analyses are usually conducted after the intervention has been adopted, yet these are not as useful. Again, HER-2 is a great example. A systematic review by Phillips and Van Bebber found only 11 cost-effectiveness studies, only 1 of which looked at HER-2/neu, even though it had been approved in 1998 (Phillips and Van Bebber, 2004). However, an update in 2007 (Ferrusi et al., 2007) found that there are now 15 cost-effectiveness studies for HER-2, and 7 are for early-stage breast cancer. The reason that few cost-effectiveness analyses are conducted may well be because most payers in the United States do not require cost-effectiveness analyses.

There is a need to model very complex clinical pathways, particularly for test-treatment combinations. Yet most modeling efforts have not adequately considered testing variability, that is, sensitivity, specificity, sequencing, and timing of the tests. For HER-2, most of the models have assumed perfect testing conditions. Those that examined testing accuracy did not include any consideration of the sequence in which tests were administered or of the fact that there were alternative tests available with very different performance characteristics. Nor did the models look at utilization of the test in terms of how often it was actually applied in a particular population.

The one model that did examine testing as an issue found that there was a huge difference in the incremental cost-effectiveness ratio depending on which test was used and in which sequence it was used. The cost-effectiveness ratio was either a few thousand dollars per quality-adjusted-life-year or, when a different sequence was used, it was more than $150,000 per quality-adjusted-life-year (Elkin et al., 2004). This demonstrates that the testing

sequence and how it is modeled makes a huge difference in what the cost-effectiveness of that test/treatment process would be.

Another issue in developing an evidence base is the lack of information about performance of the test in a real-world context. None of the models, for instance, looked at real-world utilization, by examining claims data to understand how frequently the test is applied, followed by what treatment decisions the clinicians make based on the test information. Another issue is the need to consider multiple populations. For example, about 80 percent of people with Lynch Syndrome have an increased risk of colorectal cancer. Certainly these patients should be tested, but relatives should also be tested.

A final challenge is the need to build an evidence base for pharmacogenomics that can be used in cost-effectiveness models. There is a lack of evidence about applications. There are numerous studies about genetic associations, but less information is available about what one should do with that information.

It is important to reiterate that evidence is needed concerning many things—analytic validity, clinical validity, clinical utility, availability and utilization, and the effect on economic outcomes and on the entire population health burden. One approach to building the evidence base has been provided by the Evaluation of Genomic Applications in Practice and Prevention Working Group, which was described earlier. Another project that has been proposed to the National Institutes of Health concerns cancer and personalized medicine. That project, which is called the Cancer and Personalized Medicine Research Study, is aimed at building an evidence base from an economic perspective.

One element necessary in any effort to build an evidence base is an examination of utilization. Utilization research needs to explore who has access and who uses the available technologies. Real-world data will be needed for this, perhaps claims data or chart review. It is also very important to understand patient and provider preferences, since these preferences will influence the adoption of new technologies. One approach is to use stated-preference methods, which not only give quantitative estimates of individuals' preferences but also allow one to calculate willingness to pay for the technologies. Finally, there is the economic element, the "What is value?" question. One needs to understand the downstream consequences of these technology test interventions with respect to their cost-effectiveness.

Pharmacogenomics is an inevitable trend for the future, Marshall said. There are many promising new technologies, but a key aspect of success in the long run will be the ability to demonstrate value to payers, providers, and patients. There are multiple challenges, but building the evidence base that captures the health burden, utilization, clinical utility, and cost-effectiveness of pharmacogenomics will be critical, Marshall concluded.

DISCUSSION

Wylie Burke, M.D., Ph.D.
Moderator

One audience member noted that Hogarth said in his presentation that innovation sometimes happens too quickly and that there are those who believe this has been the case with genomics. What, the questioner asked, do the trends for genomics look like from the two presenters' perspectives? Hogarth responded that there is no single answer to that question, but that it depends on the technologies. For consumer genetics (or direct-to-consumer genetics), which has increased significantly, clinical geneticists and research scientists in the United Kingdom think that translation into practice is premature. Some of the most significant and potentially fruitful innovation appears to be occurring in the gene-expression market, particularly in oncology.

Marshall responded that she believes the translation process is working at about the correct pace. There are rapid adopters and slow adopters, and one needs both when there is something new. One also needs regulation, but not too much because fast adopters and fast innovators must be allowed to get ahead of the curve, thereby enabling the remainder to catch up. The bottom line, however, is that it takes a great deal of time and energy to collect all the data needed—10 to 15 years for randomized controlled trials. On the one hand, one wants to be sure the new technologies do not cause harm, but on the other hand, too much restriction could inhibit innovation.

Another audience member said that the discussion appeared to be primarily from the perspective of those who are involved with diagnostics and those involved with reimbursement. There are a number of other genomic innovations that have proven uses. For example, one presenter said that a drug company might not want to go into genomics because it will decrease market share. But decreasing market share should not be a barrier since a tremendous amount of money can be made on a small market, as shown by the example of Gleevec. One never gets the whole market.

It seems reasonable for a company to look to pharmacogenomics as a way to get to "proof of concept," a critical stage in drug development. Gleevec is a great example of how a drug with a presumed niche indication and that is tied to biomarker, can become a blockbuster. Pharmacogenetics—biomarkers, in the context of drug development—is an important issue that needs greater exploration.

6

Concluding Remarks

GENERAL OBSERVATIONS

Wylie Burke, M.D., Ph.D.
Moderator

These comments are oriented around some general themes that seemed to emerge from the presentations and workshop discussion, Burke said. One theme is that innovation is important but that most innovations fail, perhaps because they are not good innovations.

Both technological and organizational innovations are important drivers of cost and quality improvement. Therefore innovation should be encouraged, even recognizing that many innovations will fail.

The environment in which an innovation occurs must be taken into account, as should the barriers that innovation confronts.

Many innovations are incremental, but the willingness to pay a premium price tends to occur only for disruptive innovations. Genomics may provide disruptive innovations, but it will also provide incremental ones. Disruptive business models are in play as well.

The context for genomic innovation is a complex, fragmented health care system that incorporates a fair amount of uncertainty about the regulatory environment. This uncertainty makes it difficult for innovators to plan. If the goal is efficient health care for the improvement of the health of the population, then incentives are not aligned well to encourage innovations that will achieve that goal.

Not only is the U.S. health care system today fragmented and complex, but it is a system that delivers very poor value for the money as compared to health care systems in other developed countries.

Innovation is a learning process. Most new technologies are relatively primitive when they get to market, and benefits and risks often are refined over time. Sometimes, unanticipated new uses of technology occur as people use the technology. There is a question about how the efficiency of that process can be increased since it frequently takes a long time to obtain a good understanding of the uses and value of a technology. One approach might be to think explicitly about capacity development—that is, not just encouraging innovation but encouraging the capacity to incorporate new innovations into the health care system.

The theme heard over and over in the workshop was evidence. Everyone agrees, although from a variety of perspectives, that evidence is needed to support innovation in health care. It was suggested that evidence is needed to show that an innovation will make a difference in outcomes that patients will notice. Evidence that shows an innovation will help clinicians do better at something they are already doing or do something helpful and beneficial that they previously were unable to do is also important.

From a health technology assessment viewpoint, randomized clinical trials (RCTs) are a gold standard, but they are not the usual way in which technology is evaluated because there are not very many RCTs. Health technology assessment usually needs to use a combination of indirect evidence and causal inference, and that leads to a number of serious questions. How does one make optimal use of indirect evidence? When should RCTs be insisted upon?

What is the role of the post-market evaluation process; that is, for that early diffusion process where there is learning about the technology? What is the role for evaluation at that stage, either RCTs or observational data? What kinds of infrastructures would promote the best use of post-market evaluation?

Throughout the discussion of evidence there was a recurrent idea that further work on evidence standards is needed. One important point made is that test developers face different needs from regulators than they face from payers. There is a need to think about the hurdles for regulation and payment being the same.

When is a leap of inference justified?

What is the role of cost-effectiveness or other economic indicators in determining whether something comes to market?

Is there any way to enable health technology assessment to occur earlier in the process, beginning even pre-market?

Another theme in the discussion of evidence was that methods for evaluation should be transparent. Some models were presented for how one might go about this kind of evaluation.

Discussion highlighted a need to ensure that clinicians and health care systems have authoritative sources of information, that evidence evaluation occurs in a transparent way using established standards and methodologies, and that evaluations for clinicians and the health care system are completed by a disinterested source.

Ultimately, it is clear that providers and patients want data on outcomes. Even if an innovation enters clinical use without outcome data, there is a need to acquire those data.

There are tensions in all of these issues, and those tensions need to be acknowledged. The level of regulation that is appropriate for genomic innovation is an area of tension. The level of evidence needed before proceeding to market is a tension, as is the role of conditional coverage as a mechanism to ensure that better data are obtained over time and after something comes to market. Then there is the question of what happens if the data say the innovation does not work.

Finally, there was a fair amount of discussion about encouraging innovation, but with the caveat that innovation should not cause harm. What harm is and how it can be avoided are issues requiring discussion.

Those were some of the big themes in today's workshop, Burke concluded.

References

Ad Hoc Committee on Genetic Counseling, American Society of Human Genetics. 1975. Genetic counseling. *American Journal of Human Genetics* 27:240-242.

Astrazeneca. 2006. *Annual results presentation*. Jon Symonds. http://www.astrazeneca.com/sites/7/imagebank/typearticleparam511711/astrazeneca-fourth-quarter-2006-jrs-script.pdf.pdf (accessed January 18, 2008).

Baigent, C., F. E. Harrell, M. Buyse, J. R. Emberson, and D. G. Altman. 2008. Ensuring trial validity by data quality assurance and diversification of monitoring methods. *Clinical Trials* 5(1):49-55.

Benjamin, D. K., P. B. Smith, M. D. Murphy, R. Roberts, L. Mathis, D. Avant, M. Califf, and J. S. Li. 2006. Peer-reviewed publication of clinical trials completed for pediatric exclusivity. *JAMA* 296:1266-1273.

Buckley, T. 2007. *The complexities of comparative effectiveness*. Biotechnology Industry Organization. http://bio.org/healthcare/compeffective/20071025.pdf (accessed April 29, 2008).

Burke, W. 2007. *Integrating genetic technology into a health care system*. Presentation at the Institute of Medicine workshop on diffusion of genomic innovations in health and medicine. December 4. Washington, DC.

Burke, W., and B. M. Psaty. 2007. Personalized medicine in the era of genomics. *JAMA* 298(14):1682-1684.

Burke, W., and R. L. Zimmern. 2004. Ensuring the appropriate use of genetic tests. *Nature Reviews Genetics* 5(12):955-959.

Burke, W., M. J. Khoury, A. Stewart, and R. Zimmern. 2006. The path from genome-based research to population health: Development of an international public health genomics network. *Genetics in Medicine* 8:451-458.

Califf, R. 2007. *Translation of innovations: A broad perspective*. Presentation at the Institute of Medicine workshop on diffusion of genomic innovations in health and medicine. December 4. Washington, DC.

Center on Budget and Policy Priorities. 2006. *The number of uninsured Americans is at an all-time high*. http://www.cbpp.org/8-29-06health.htm (accessed January 16, 2008).

CMS (Center for Medicare and Medicaid Services). 2007. *Board of Trustees 2007 report*. http://www.cms.hhs.gov/ReportsTrustFunds/downloads/tr2007.pdf (accessed February 12, 2008).

Cobleigh, M. A., P. Bitterman, J. Baker, M. Cronin, M.-L. Liu, R. Borchik, B. Tabesh, J.-M. Mosquera, M. G. Walker, and S. Shak. 2003. Tumor gene expression predicts distant disease-free survival (DDFS) in breast cancer patients with 10 or more positive nodes: High throughput RT-PCR assay of paraffin-embedded tumor tissues. *Proceedings of the American Society of Clinical Oncology* 22:Abstract #3415.

Collins, F. S., and V. A. McKusick. 2001. Implications of the human genome project for medical science. *JAMA* 285:540-544.

CTTI (Clinical Trials Transformation Initiative). 2007. *Welcome to CTTI*. https://www.trialstransformation.org/ (accessed March 11, 2008).

DHHS (Department of Health and Human Services). 2007. *Personalized health care*. http://www.hhs.gov/myhealthcare/ (accessed January 14, 2008).

DiMasi, J. A., Hansen, R. W., and H. G. Grabowski. 2003. The price of innovation: New estimates of drug development costs. *Journal of Health Economics* 22(2):151-185.

Eisenstein, E. L., P. W. Lemons 2nd, B. E. Tardiff, K. A. Schulman, M. K. Jolly, and R. M. Califf. 2005. Reducing the costs of phase III cardiovascular clinical trials. *American Heart Journal* 149(3):482-488.

Eisenstein, E. L., R. Collins, B. S. Cracknell, O. Podesta, E. D. Reid, P. Sandercock, Y. Shakhov, M. L. Terrin, M. A. Sellers, R. M. Califf, C. B. Granger, and R. Diaz. 2008. Sensible approaches for reducing clinical trial costs. *Clinical Trials* 5(1):75-84.

Elkin, E. B., M. C. Weinstein, E. P. Winer, K. M. Kuntz, S. J. Schnitt, and J. C. Weeks. 2004. HER-2 testing and trastuzumab therapy for metastatic breast cancer: A cost-effectiveness analysis. *Journal of Clinical Oncology* 22(5):854-863.

Esteban, J., J. Baker, M. Cronin, et al. 2003. Tumor gene expression and prognosis in breast cancer: Multi-gene RT-PCR assay of paraffin-embedded tissue. *Proceedings of the American Society of Clinical Oncology* 22:Abstract #3416.

FDA (U.S. Food and Drug Administration). 2005. *Pediatric exclusivity*. http://www.fda.gov/cder/ogd/pedtrac.htm (accessed March 11, 2008).

FDA. 2006. *Critical path opportunities report*. http://www.fda.gov/oc/initiatives/criticalpath/reports/opp_report.pdf (accessed February 6, 2008).

Ferrusi, I. L., D. A. Marshall, N. A. Kulin, and K. A. Phillips. 2007. *A systematic review of economic analyses of HER2 testing and trastuzumab therapy*. Unpublished.

Garrison, L., and M. J. F. Austin. 2007. The economics of personalized medicine: A model of incentives for value creation and capture. *Drug Information Journal* 41:501-509.

Gray, L. 2007. *View from the trenches: Challenges and opportunities in personalized medicine*. Presentation at the Institute of Medicine workshop on diffusion of genomic innovations in health and medicine. December 4. Washington, DC.

Griliches, Z. 1957. Hybrid corn: An exploration in the economics of technological change. *Econometrica* 25(4):501-522.

Harper, S., J. Lynch, S. Burris, and G. D. Smith. 2007. Trends in the black-white life expectancy gap in the United States, 1983-2003. *JAMA* 21(297):1224-1232.

IOM (Institute of Medicine). 1992. *Guidelines for clinical practice: From development to use*. Washington, DC: National Academy Press.

IOM. 2005. *Implications of genomics for public health: Workshop summary*. Washington, DC: The National Academies Press.

Li, J. S., E. L. Eisenstein, H. G. Grabowski, E. D. Reid, B. Mangum, K. A. Schulman, J. F. Goldsmith, M. D. Murphy, R. M. Califf, and D. K. Benjamin. 2007. Economic return of clinical trials performed under the pediatric exclusivity program. *JAMA* 297:480-488.

Maller, J., S. George, S. Purcell, J. Fagerness, D. Altshuler, M. J. Daly, and J. M. Seddon. 2006. Common variation in three genes, including a noncoding variant in CFH, strongly influences risk of age-related macular degeneration. *Nature Genetics* 38:1055-1059.

Miranda, M. L., J. M. Silva, M. A. Overstreet Galeano, J. P. Brown, D. S. Campbell, E. Coley, C. S. Cowan, D. Harvell, J. Lassiter, J. L. Parks, and W. Sandele. 2005. Building geographic information system capacity in local health departments: Lessons from a North Carolina project. *American Journal of Public Health* 95(12):2180-2185.

Modeste, N. N. 1996. *Dictionary of public health promotion and education: Terms and concepts*. Thousand Oaks, CA: SAGE Publications.

Nash, J. F. 1950. Equilibrium points in N-person games. *Proceedings of the National Academy of Sciences* 36:48-49.

NHSE (National Health Service Executive). 1996. *Promoting clinical effectiveness: A framework for action in and through the NHS*. Leeds, UK: National Health Service Executive.

Nissen, S. E., J. C. Tardif, S. J. Nicholls, J. H. Revkin, C. L. Shear, W. T. Duggan, W. Ruzyllo, W. B. Bachinsky, G. P. Lasala, and E. M. Tazcu. 2007. Effect of torcetrapib on the progression of coronary atherosclerosis. *New England Journal of Medicine* 356(13):1304-1316.

Oratz, R., D. Paul, A. L. Cohn, and S. M. Sedlacek. 2007. Impact of a commercial reference laboratory test recurrence score on decision making in early-stage breast cancer. *Journal of Oncology Practice* 3(4):182-186.

Paik, S., S. Shak, G. Tang, C. Kim, J. Baker, M. Cronin, R. Baehner, M. Walker, D. Watson, T. Park, J. Bryant, and N. Wolmark. 2003. *Multi-gene RT-PCR assay for predicting recurrence in node negative breast cancer patients—NSABP studies, B-20 and B-14*. San Antonio Breast Cancer Symposium 2003. Abstract #16.

Paik, S., S. Shak, G. Tang, C. Kim, H. Joo, J. Baker, M. Cronin, D. Watson, J. Bryant, J. Costantino, and N. Wolmark. 2004. *Expression of the 21 genes in the Recurrence Score assay and prediction of clinical benefit from tamoxifen in NSABP study B-14 and chemotherapy in NSABP study B-20*. San Antonio Breast Cancer Symposium 2004. Abstract #24.

Pear, R. 2008 (March 23). Gap in life expectancy widens for the nation. *New York Times*. http://www.nytimes.com/2008/03/23/us/23health.html (accessed March 25, 2008).

Phillips, K. A., and S. L. Van Bebber. 2004. A systematic review of cost-effectiveness analyses of pharmacogenomic interventions. *Future Medicine* 5(8):1139-1149.

Phillips, K. A., and S. L. Van Bebber. 2005. Measuring the value of pharmacogenomics. *Nature Reviews Drug Discovery* 4(6):500-509.

Phillips, K. A., D. L. Veenstra, S. Ramsey, S. L. Van Bebber, and J. Sakowski. 2004. Genetic testing and pharmacogenomics: Issues for determining the impact to health care delivery and costs. *American Journal of Managed Care* 10(7):425-432.

Porter, M. E. 2006. *Redefining health care: Creating value-based competition on results*. Cambridge, MA: Harvard Business School Press.

Rettenmaier, A. J., and T. R. Saving. 2004. *The 2004 Medicare and Social Security trustees reports*. National Center for Policy Analysis, Study No. 266. http://www.ncpa.org/pub/st/st266/ (accessed February 12, 2008).

Rieder, M. J., A. P. Reiner, B. F. Gage, D. A. Nickerson, C. S. Eby, H. L. McLeod, D. K. Blough, K. E. Thummel, D. L. Veenstra, and A. E. Rettie. 2005. Effect of VKORC1 haplotypes on transcriptional regulation and warfarin dose. *New England Journal of Medicine* 352:2285-2293.

Roberts, J. S., S. A. LaRusse, H. Katzen, P. J. Whitehouse, M. Barber, S. G. Post, N. Relkin, K. Quaid, R. H. Pietrzak, A. Cupples, L. A. Farrer, T. Brown, and R. C. Green. 2003. Reasons for seeking genetic susceptibility testing among first-degree relatives of people with Alzheimer disease. *Alzheimer Disease & Associated Disorders* 17(2):86-93.

Schumpeter, J. A. 1911. *The theory of economic development: An inquiry into profits, capital, credit, interest, and the business cycle*. Cambridge, MA: Harvard University Press.

Shak, S. 2007. *Introducing a genomic innovation to clinical practice*. Presentation at the Institute of Medicine workshop on diffusion of genomic innovations in health and medicine. December 4. Washington, DC.

Trusheim, M. R., E. R. Berndt, and F. L. Douglas. 2007. Stratified medicine: Strategic and economic implications of combining drugs and clinical biomarkers. *Nature Reviews Drug Discovery* 6(4):287-293.

Varmus, N. 2002. Getting ready for gene-based medicine. *New England Journal of Medicine* 347(19):1526-1527.

Veenstra, D. L., M. K. Higashi, and K. A. Phillips. 2000. Assessing the cost-effectiveness of pharmacogenomics. *AAPS PharmSci* 2(3):Article 29. http://www.aapsj.org/view.asp?art=ps020329 (accessed February 13, 2008).

Appendix A

Workshop Agenda

Roundtable on Translating Genomic-Based Research for Health
Board on Health Sciences Policy
Tuesday, December 4, 2007

Agenda

Workshop on Diffusion and Use of Genomic Innovations in Health and Medicine

Auditorium
National Academies Building
2101 Constitution Avenue, NW
Washington, DC 20001

8:30-8:40 Welcome and Overview of Workshop

WYLIE BURKE
Roundtable Chair
Professor and Chair
Department of Medical History and Ethics
University of Washington School of Medicine

8:40-10:00 ***Panel on Translation of Innovations***

8:40-9:00 Translation of Innovations: A Broad Perspective

ROBERT CALIFF
Vice Chancellor for Clinical Research
Professor of Medicine
Duke University

9:00-9:20 Innovation: Understanding Types of Innovation and Implications for Policy

KEVIN SCHULMAN
Professor, Medicine and Business Administration
Director, Center for Clinical and Genetic Economics
Associate Director, Duke Clinical Research Institute
Duke University

9:20-9:40 Technologic Diffusion: Lessons for Genomics from Other Technologies

ANNETINE GELIJNS
Professor, Health Policy, Management, and Surgical Science
Co-Director, International Center for Health Outcomes and Innovation Research
Columbia University

9:40-10:00 Discussion

10:00-10:20 **BREAK**

10:20-12:00 ***Panel on Practical Incentives and Barriers to Translation***

10:20-10:40 Translating Medical Innovations with Appropriate Evidence

SEAN TUNIS
Founder and Director
Center for Medical Technology Policy

10:40-11:00 Assessing Technology for Use in Health and Medicine

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

11:00-11:20 Integrating Genetic Technology into a Health Care System

WYLIE BURKE
Roundtable Chair
Professor and Chair
Department of Medical History and Ethics
University of Washington School of Medicine

11:20-11:40 View from the Trenches: Challenges and Opportunities in Personalized Medicine

BRAD GRAY
Vice President of Business and Strategic Development
Genzyme Genetics

11:40-12:00 Discussion

12:00-1:00 **LUNCH**

1:00-2:30 ***Panel on Translation of Genomic Technology at the Clinical Level***

1:00-1:30 A Primary Care Provider View of Translating Genomic Innovation

ALFRED BERG
Professor
University of Washington

1:30-2:00 Introducing a Genomic Innovation to Clinical Practice

STEVEN SHAK
Chief Medical Officer
Genomic Health

2:00-2:30 Discussion

2:30-2:50 **BREAK**

2:50-4:00 ***Panel on Opportunities and Constraints for Translation of Genomic Innovations***

2:50-3:10 The Global Perspective

STUART HOGARTH
Research Associate
Department of Public Health and Primary Care
University of Cambridge

3:10-3:30 The U.S. Perspective

DEBORAH MARSHALL
Professor, McMaster University
Vice President, Global Health Economics and Outcomes
i3 Innovus

3:30-4:00 Discussion

4:00-4:30 Summing Up

WYLIE BURKE
Roundtable Chair
Professor and Chair
Department of Medical History and Ethics
University of Washington School of Medicine

4:30 **ADJOURN WORKSHOP**

Appendix B

Biographical Sketches of the Workshop Speakers

Naomi Aronson, Ph.D., is the executive director of the Blue Cross and Blue Shield Association Technology Evaluation Center (TEC). She has overseen TEC's development as a nationally recognized technology assessment program and an Evidence-based Practice Center (EPC) of the Agency for Healthcare Research and Quality (AHRQ). Dr. Aronson has directed over 300 technology assessments and 10 evidence reports for AHRQ. She has published articles in *Annals of Internal Medicine*, *Journal of the National Cancer Institute*, *Cancer*, *Journal of the American College of Surgeons*, *Academic Radiology*, *Journal of Family Practice*, and *Gastrointestinal Endoscopy*. She represented the private sector on a U.S. Agency for International Development team providing technical assistance to the Hungarian government on building evidence-based medicine capacity in the national health insurance system. She is a member of the Institute of Medicine Forum on Drug Discovery Translation and Development, and the Steering Committee of the Chicago-Area DEcIDE Research Center and the National Business Group on Health Committee on Evidence-Based Benefit Design. Previously, Dr. Aronson was a member of Northwestern University faculty, specializing in sociology of science and medicine. She was also a post-doctoral fellow in the Science, Technology and Society Program at the Massachusetts Institute of Technology and received research awards from the National Science Foundation and the American Council of Learned Societies. Dr. Aronson's academic research focused on how the organization of scientific specialties in biomedical and clinical research affects the process of scientific discovery.

Alfred O. Berg, M.D., M.P.H., received his professional education at Washington University, St. Louis, the University of Missouri, Columbia, and the University of Washington, Seattle. He is board certified in Family Medicine and in General Preventive Medicine and Public Health, and was elected to the Institute of Medicine of the National Academy of Sciences in 1996. In 2004 he received the Thomas W. Johnson Award for career contributions to family medicine education from the American Academy of Family Physicians. Dr. Berg's research has focused on clinical epidemiology in primary care settings. He has served on many expert panels using evidence-based methods to develop clinical guidelines, including chairmanship of the United States Preventive Services Task Force, co-chair of the otitis media panel convened by the Agency for Health Care Policy and Research, chair of the CDC STD Treatment Guidelines panel, member of the AMA/CDC panel producing Guidelines for Adolescent Preventive Services, member of the Institute of Medicine's Immunization Safety Review Committee, and chair of the Institute of Medicine's Committee on the Treatment of Post-traumatic Stress Disorder. He currently chairs the CDC's panel on Evaluation of Genomic Applications in Practice and Prevention.

Wylie Burke, M.D., Ph.D., is professor and chair of the Department of Medical History and Ethics at the University of Washington. She received a Ph.D. in Genetics and an M.D. from the University of Washington and completed a residency in Internal Medicine at the University of Washington. She was a medical genetics fellow at the University of Washington from 1981 to 1982. Dr. Burke was a member of the Department of Medicine at the University of Washington from 1983 to 2000, where she served as associate director of the Internal Medicine Residency Program from 1988 to 1994 and as founding director of the University of Washington's Women's Health Care Center from 1994 to 1999. She was appointed chair of the Department of Medical History in October 2000. She is also an adjunct professor of Medicine and Epidemiology and an associate member of the Fred Hutchinson Cancer Research Center. She was a visiting scientist at the Centers for Disease Control and Prevention in 1998 and is a fellow of the American College of Physicians. She has served on the NIH National Advisory Council for Human Genome Research and the Secretary's Advisory Committee on Genetic Testing. Dr. Burke's research addresses the social, ethical and policy implications of genetic information, including genetic test evaluation, the development of practice standards for genetically based services and genetics education for health professionals. She is also the director of the University of Washington Center for Genomics and Healthcare Equality, a Center of Excellence in Ethical, Legal, and Social Implications (ELSI) research funded by the National Human Genome Research Institute. Dr. Burke is a member of the Institute of Medicine.

Robert M. Califf, M.D., is vice chancellor for clinical research and professor of medicine in the Division of Cardiology at Duke. Former director of the Duke Clinical Research Institute, he became head of the Duke Translational Medicine Institute in 2006. A native of South Carolina, Califf graduated from Duke University, summa cum laude and Phi Beta Kappa, in 1973 and from the Duke University School of Medicine in 1978, where he was selected for Alpha Omega Alpha. He completed his internship and residency at the University of California at San Francisco and his fellowship in cardiology at Duke University. He is board-certified in internal medicine (1984) and cardiology (1986) and is a fellow of the American College of Cardiology (1988). Califf has served as an editor for the first and second editions of the landmark textbook, Acute Coronary Care, published by Mosby Inc., and is the editor in chief of Mosby's American Heart Journal. He is a section editor for the Textbook of Cardiovascular Medicine and has been an author or coauthor of more than 600 peer-reviewed journal articles. He is a contributing editor for theheart.org, an online information resource for academic and practicing cardiologists. Dr. Califf's role as Director of the Duke Translational Medicine Institute, which is funded in part by an NIH Clinical and Translational Science Award (CTSA), includes service as co-chairman of the Principal Investigators Steering Committee of the CTSA.

Annetine Gelijns, Ph.D., is co-director (with Alan Moskowitz) of the International Center for Health Outcomes and Innovation Research (InCHOIR), and an associate professor of Surgical Sciences in the Department of Surgery, College of Physicians and Surgeons, and the Division of Health Policy and Management of the Mailman School of Public Health, Columbia University, New York City. She is also a division chief in the Department of Surgery. Her current research focuses on measurement of the long-term clinical outcomes and economic impact of clinical interventions, patient safety research, and the factors driving the development and diffusion of medical technology. She has special expertise in cardiovascular disease, particularly in the design, coordination, and analysis of multi-center left ventricular assist devices (LVAD) trials. She has been the director of the Data Coordinating Center for the National Institutes of Health (NIH)-sponsored REMATCH trial, and is the principal investigator (PI) or co-PI of several newer generations of LVAD trials. She also will direct the Data Coordinating Center for the SCCOR grant on the biology of long-term LVAD implantation, for which NIH funding is pending. Before coming to Columbia in 1993, she directed the Program on Technological Innovation in Medicine at the Institute of Medicine, National Academy of Sciences. Dr. Gelijns has been a consultant to various national and international organizations, including the World Health Organization (WHO) and the

Organization for Economic Cooperation and Development (OECD), Paris, France. She holds a Ph.D. from the medical faculty and the department of science policy, University of Amsterdam, and a master's degree in law from the University of Leyden, the Netherlands.

Brad Gray joined Genzyme Genetics, a division of Genzyme Corporation, in September 2006 as vice president of Business and Strategic Development. In this role, Brad leads the efforts to grow Genzyme's laboratory diagnostic testing services through licensing and acquisition. In addition, he works closely with other members of the management team to strengthen the existing business and strategically invest in building Genzyme's premier reference lab. Previously, he held several positions within Genzyme's Corporate Development group, including leading Genzyme Ventures, a corporate venture capital fund focused on furthering the business development goals of Genzyme Corporation. Prior to joining Genzyme in November 2004, he was an engagement manager in the Boston office of McKinsey & Company, a management consulting firm. During his four years at McKinsey, he worked with senior healthcare executives in the United States and Europe on a broad range of issues including pharmaceutical and diagnostic product strategy, post-merger integration, organization design, and operational turnarounds. He holds a B.S. in Chemical Engineering from the Massachusetts Institute of Technology and a B.A. in Economics & Management from Oxford University, where he studied as a Marshall Scholar.

Stuart Hogarth is a visiting research fellow at the Institute for Science and Society at the University of Nottingham. His research interests include the innovation processes in the drugs and diagnostics industries and the regulatory issues emerging from novel healthcare technologies. He is working on a Wellcome Trust funded project examining policy issues in the evaluation of clinical genetic tests for common complex conditions. The project's forthcoming report examines the policy options for improving both the regulatory landscape for genetic tests and the incentives needed to encourage test developers to generate high-quality clinical data. For this work he has received an FDA Leveraging/Collaboration Award. He was lead author of a recent report for Health Canada on international developments in the regulation of pharmacogenomics and a briefing for the Human Genetics Commission on the regulation of commercial genetic testing services in the United Kingdom. He participated in the drafting of the OECD's guidelines on quality assurance for molecular genetic testing.

Deborah Marshall, Ph.D., M.H.S.A., is an associate professor in the Department of Epidemiology and Biostatistics at the Centre for Evaluation of Medicine at McMaster University in Canada. She has managed numerous

health economics studies including prospective evaluations and decision analysis models for a variety of health conditions. Dr. Marshall is also the vice president of Global Health Economics and Outcomes Research at i3 Innovus, where she is responsible for scientific leadership in health economics and outcomes research. She previously worked in government, academic and industry research settings commissioning, conducting and authoring evaluations of various health technologies in national Health Technology Assessment agencies and for Bayer Diagnostics in California. Dr. Marshall's history of peer-reviewed research grant funding includes a number of cancer screening studies evaluating cost-effectiveness and patient preferences. Dr Marshall is a co-investigator on a recently submitted NCI Program Project Grant proposal on Personalized Medicine for Colorectal and Breast Cancer, a grant from the Blue Shield of California Foundation on the policy challenges for personalized medicine, and a grant from Genome Canada on Genome-Specific Approaches to Therapy in Childhood (GATC). Dr. Marshall has held leadership positions for health care professional organizations including the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) and the International Society for Health Technology Assessment. Dr. Marshall has authored or co-authored more than 40 papers published in peer-reviewed journals as well as a number of book chapters, and other published technical reports. She has served on the Editorial Board for *International Journal for Technology Assessment in Health Care* since 1998, and is a referee for various health economic and health policy journals.

Kevin Schulman, M.D., is a professor of medicine in the Duke University School of Medicine, where he also serves as the director of the Center for Clinical and Genetic Economics and as an associate director of the Duke Clinical Research Institute. He holds a joint appointment as a professor of business administration in Duke University's Fuqua School of Business, where he is the director of the Health Sector Management Program. Dr. Schulman also holds appointments in the Center for Health Services Research in Primary Care in the Durham VA Medical Center, the Duke Center for Clinical Health Policy Research, and the Trent Center for Bioethics, Humanities and History of Medicine. His research interests include economic evaluation in clinical research; health services research and policy, including access to care and the impact of reimbursement and regulatory policies on clinical practice; and medical decision making, especially in patients with life-threatening conditions.

Steven Shak, M.D., is chief medical officer and co-founder of Genomic Health, Inc., a company pioneering the practical application of genomics to clinical practice. Genomic Health has used new molecular diagnostic

technologies and rigorous clinical studies to develop the Oncotype DX™ breast cancer assay. From July 1996 to October 2000, he served in various roles in Medical Affairs at Genentech, Inc., most recently as senior director and staff clinical scientist. From November 1989 to July 1996, Dr. Shak served as a director of Discovery Research at Genentech, where he was responsible for Pulmonary Research, Immunology, and Pathology. He led the clinical team that gained approval for Herceptin®, a targeted biologic treatment for metastatic breast cancer. He also initiated the cancer clinical trials of the anti-angiogenesis agent, anti-VEGF (Avastin™). In addition, Dr. Shak discovered Pulmozyme®, a mucus-dissolving enzyme that is approved worldwide for the treatment of the genetic disease, Cystic Fibrosis. Prior to joining Genentech, he was an assistant professor of Medicine and Pharmacology at New York University School of Medicine. He holds a B.A. in Chemistry from Amherst College and an M.D. from New York University School of Medicine, and completed his post-doctoral training at University of California at San Francisco.

Sean Tunis, M.D., M.Sc., is the founder and director of the Center for Medical Technology Policy in San Francisco, where he is working with health care decision makers and stakeholders to support the rapid evaluation and effective use of new medical technologies. He is also a principal at Rubix Health, which consults with early-stage life sciences companies on reimbursement strategy designed around developing reliable evidence of product value. Through September of 2005, Dr. Tunis was the director of the Office of Clinical Standards and Quality and chief medical officer at the Centers for Medicare and Medicaid Services (CMS). In this role, he had lead responsibility for clinical policy and quality for the Medicare and Medicaid programs, which provide health coverage to over 100 million U.S. citizens. Dr. Tunis supervised the development of national coverage policies, quality standards for Medicare and Medicaid providers; quality measurement and public reporting initiatives, and the Quality Improvement Organization program. As chief medical officer, Dr. Tunis served as the senior advisor to the CMS Administrator on clinical and scientific policy. He also co-chaired the CMS Council on Technology and Innovation. Dr. Tunis joined CMS in 2000 as the director of the Coverage and Analysis Group. Before joining CMS, Dr. Tunis was a senior research scientist with the Lewin Group, where his focus was on the design and implementation of prospective comparative effectiveness trials and clinical registries. Dr. Tunis also 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 Committee on Labor and Human Resources, where he participated in policy development regarding pharmaceutical and device regulation. He received a B.S. degree in History of Science from Cornell University, and a

medical degree and masters in Health Services Research from the Stanford University School of Medicine. Dr. Tunis did his residency training at UCLA and the University of Maryland in Emergency Medicine and Internal Medicine. He is board certified in Internal Medicine and holds adjunct faculty positions at Johns Hopkins and Stanford University School of Medicine.