Learning What Works: Infrastructure Required for Comparative Effectiveness Research: Workshop Summary

ISBN
978-0-309-12068-5

588 pages
6 x 9
PAPERBACK (2011)

LeighAnne Olsen, Claudia Grossman, and J. Michael McGinnis; The Learning Health System Series; Institute of Medicine

THE LEARNING HEALTH SYSTEM SERIES

ROUNDTABLE ON VALUE & SCIENCE-DRIVEN HEALTH CARE

LEARNING WHAT WORKS

Infrastructure Required for Comparative Effectiveness Research

Workshop Summary

LeighAnne Olsen, Claudia Grossmann, and J. Michael McGinnis

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

This project was supported by the Agency for Healthcare Research and Quality, America's Health Insurance Plans, AstraZeneca, Blue Shield of California Foundation, Burroughs Wellcome Fund, California Health Care Foundation, Centers for Medicare & Medicaid Services, Charina Endowment Fund, Department of Veterans Affairs, Food and Drug Administration, Johnson & Johnson, sanofi-aventis, and Stryker. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the view of the organizations or agencies that provided support for this project.

International Standard Book Number-13: 978-0-309-12068-5
International Standard Book Number-10: 0-309-12068-3

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). 2011. *Learning What Works: Infrastructure Required for Comparative Effectiveness Research: 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

ROUNDTABLE ON VALUE & SCIENCE-DRIVEN HEALTH CARE*

Denis A. Cortese (*Chair*), Emeritus President and Chief Executive Officer, Mayo Clinic; Foundation Professor, ASU
Donald Berwick, Administrator, Centers for Medicare & Medicaid Services (*ex officio*)
David Blumenthal, National Coordinator, Office of the National Coordinator for Health IT (*ex officio*)
Bruce G. Bodaken, Chairman, President, and Chief Executive Officer, Blue Shield of California
David R. Brennan, Chief Executive Officer, AstraZeneca PLC
Paul Chew, Chief Science Officer and CMO, sanofi-aventis U.S., Inc.
Carolyn M. Clancy, Director, Agency for Healthcare Research and Quality (*ex officio*)
Michael J. Critelli, Former Executive Chairman, Pitney Bowes, Inc.
Helen Darling, President, National Business Group on Health
Thomas R. Frieden, Director, Centers for Disease Control and Prevention (*designee*: **Chesley Richards**) (*ex officio*)
Gary L. Gottlieb, President and CEO, Partners HealthCare System
James A. Guest, President, Consumers Union
George C. Halvorson, Chairman and Chief Executive Officer, Kaiser Permanente
Margaret A. Hamburg, Commissioner, Food and Drug Administration (*ex officio*)
Carmen Hooker Odom, President, Milbank Memorial Fund Board
Ardis Hoven, Board Chair, American Medical Association
Brent James, Chief Quality Officer and Executive Director, Institute for Health Care Delivery Research, Intermountain Healthcare
Michael M. E. Johns, Chancellor, Emory University
Craig Jones, Director, Vermont Blueprint for Health
Cato T. Laurencin, Vice President for Health Affairs, Dean of the School of Medicine, University of Connecticut
Stephen P. MacMillan, President and Chief Executive Officer, Stryker
Mark B. McClellan, Director, Engelberg Center for Healthcare Reform, The Brookings Institution
Sheri S. McCoy, Worldwide Chairman, Johnson & Johnson Pharmaceuticals Group
Elizabeth G. Nabel, President, Brigham and Women's Hospital

*Formerly the Roundtable on Evidence-Based Medicine. Institute of Medicine forums and roundtables do not issue, review, or approve individual documents. The responsibility for the published workshop summary rests with the workshop rapporteurs and the institution.

Mary D. Naylor, Professor and Director of Center for Transitions in Health, University of Pennsylvania
Peter Neupert, Corporate Vice President, Health Solutions Group, Microsoft Corporation
William D. Novelli, Former CEO, AARP; Professor, Georgetown University
Jonathan B. Perlin, Chief Medical Officer and President, Clinical Services, HCA, Inc.
Robert A. Petzel, Under Secretary, Veterans Health Administration (*ex officio*)
Richard Platt, Professor and Chair, Harvard Medical School and Harvard Pilgrim Health Care
John C. Rother, Group Executive Officer, AARP
John W. Rowe, Professor, Mailman School of Public Health, Columbia University
Susan Shurin, Acting Director, National Heart, Lung, and Blood Institute (*ex officio*)
Mark D. Smith, President and CEO, California HealthCare Foundation
George P. Taylor, Assistant Secretary for Health Affairs (Acting), Department of Defense (*designee*: **Michael Dinneen**) (*ex officio*)
Reed D. Tuckson, Executive VP and Chief of Medical Affairs, UnitedHealth Group
Frances M. Visco, President, National Breast Cancer Coalition

Workshop Planning Committee

John W. Rowe (*Chair*), Mailman School of Public Health, Columbia University
Stuart H. Altman, Brandeis University
Kathleen A. Buto, Johnson & Johnson
Carolyn M. Clancy, Agency for Healthcare Research and Quality
W. David Helms, AcademyHealth
Mark B. McClellan, Brookings Institution
Peter R. Orszag, Congressional Budget Office

Roundtable Staff

Christie Bell, Financial Associate
Patrick Burke, Financial Associate (until December 2009)
Andrea Cohen, Financial Associate (until December 2008)
Claudia Grossmann, Program Officer
Kiran Gupta, Research Assistant (until May 2009)
Katie Jakubs, National Institutes of Health detailee
J. Michael McGinnis, Senior Scholar and Executive Director

LeighAnne Olsen, Program Officer
Daniel O'Neill, Research Associate (until December 2008)
Stephen Pelletier, Consultant
Brian Powers, Senior Program Assistant
Valerie Rohrbach, Program Assistant
Julia Sanders, Program Assistant
Robert Saunders, Program Officer
Ruth Strommen, Intern
Leigh Stuckhardt, Program Associate
Kate Vasconi, Senior Program Assistant (until January 2011)
Pierre Yong, Program Officer (until May 2010)
Catherine Zweig, Senior Program Assistant (until June 2010)

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 process. We wish to thank the following individuals for their review of this report:

Barbara Alving, National Institutes of Health
Meryl Bloomrosen, American Medical Informatics Association
Jean Paul Gagnon, sanofi-aventis
Larry A. Green, University of Colorado at Denver

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 **Samuel Nussbaum,** Wellpoint, Inc. Appointed by the National Research Council and 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 authoring committee and the institution.

Institute of Medicine
Roundtable on Value & Science-Driven Health Care[1]
Charter and Vision Statement

The Institute of Medicine's Roundtable on Value & Science-Driven Health Care has been convened to help transform the way evidence on clinical effectiveness is generated and used to improve health and health care. Participants have set a goal that, by the year 2020, 90 percent of clinical decisions will be supported by accurate, timely, and up-to-date clinical information, and will reflect the best available evidence. Roundtable members will work with their colleagues to identify the issues not being adequately addressed, the nature of the barriers and possible solutions, and the priorities for action, and will marshal the resources of the sectors represented on the Roundtable to work for sustained public–private cooperation for change.

****

The Institute of Medicine's Roundtable on Value & Science-Driven Health Care has been convened to help transform the way evidence on clinical effectiveness is generated and used to improve health and health care. We seek the development of a ***learning health system*** that is designed to generate and apply the best evidence for the collaborative healthcare choices of each patient and provider; to drive the process of discovery as a natural outgrowth of patient care, and to ensure innovation, quality, safety, and value in health care.

Vision: Our vision is for a healthcare system that draws on the best evidence to provide the care most appropriate to each patient, emphasizes prevention and health promotion, delivers the most value, adds to learning throughout the delivery of care, and leads to improvements in the nation's health.

Goal: By the year 2020, 90 percent of clinical decisions will be supported by accurate, timely, and up-to-date clinical information, and will reflect the best available evidence. We feel that this presents a tangible focus for progress toward our vision, that Americans ought to expect at least this level of performance, that it should be feasible with existing resources and emerging tools, and that measures can be developed to track and stimulate progress.

Context: As unprecedented developments in the diagnosis, treatment, and long-term management of disease bring Americans closer than ever to the promise of personalized health care, we are faced with similarly unprecedented challenges to identify and deliver the care most appropriate for individual needs and conditions. Care that is important is often not delivered. Care that is delivered is often not important. In part, this is due to our failure to apply the evidence we have about the medical care that is most effective—a failure related to shortfalls in provider knowledge and accountability, inadequate care coordination and support, lack of insurance, poorly aligned payment incen-

[1] Formerly the Roundtable on Evidence-Based Medicine.

tives, and misplaced patient expectations. Increasingly, it is also a result of our limited capacity for timely generation of evidence on the relative effectiveness, efficiency, and safety of available and emerging interventions. Improving the value of the return on our healthcare investment is a vital imperative that will require much greater capacity to evaluate high priority clinical interventions, stronger links between clinical research and practice, and reorientation of the incentives to apply new insights. We must quicken our efforts to position evidence development and application as natural outgrowths of clinical care—to foster health care that learns.

Approach: The IOM Roundtable on Value & Science-Driven Health Care serves as a forum to facilitate the collaborative assessment and action around issues central to achieving the vision and goal stated. The challenges are myriad and include issues that must be addressed to improve evidence development, evidence application, and the capacity to advance progress on both dimensions. To address these challenges, as leaders in their fields, Roundtable members will work with their colleagues to identify the issues not being adequately addressed, the nature of the barriers and possible solutions, and the priorities for action, and will marshal the resources of the sectors represented on the Roundtable to work for sustained public–private cooperation for change.

Activities include collaborative exploration of new and expedited approaches to assessing the effectiveness of diagnostic and treatment interventions, better use of the patient care experience to generate evidence on effectiveness, identification of assessment priorities, and communication strategies to enhance provider and patient understanding and support for interventions proven to work best and deliver value in health care.

Core concepts and principles: For the purpose of the Roundtable activities, we define science-driven health care broadly to mean that, ***to the greatest extent possible, the decisions that shape the health and health care of Americans—by patients, providers, payers, and policymakers alike—will be grounded on a reliable evidence base, will account appropriately for individual variation in patient needs, and will support the generation of new insights on clinical effectiveness.*** Evidence is generally considered to be information from clinical experience that has met some established test of validity, and the appropriate standard is determined according to the requirements of the intervention and clinical circumstance. Processes that involve the development and use of evidence should be accessible and transparent to all stakeholders.

A common commitment to certain principles and priorities guides the activities of the Roundtable and its members, including the commitment to: the right health care for each person; putting the best evidence into practice; establishing the effectiveness, efficiency, and safety of medical care delivered; building constant measurement into our healthcare investments; the establishment of healthcare data as a public good; shared responsibility distributed equitably across stakeholders, both public and private; collaborative stakeholder involvement in priority setting; transparency in the execution of activities and reporting of results; and subjugation of individual political or stakeholder perspectives in favor of the common good.

Foreword

In its role as adviser to the nation to improve health, the Institute of Medicine (IOM) focuses on ensuring that Americans receive care based on the best scientific knowledge—from identifying key opportunities to better support innovation and discovery in basic research to accelerating the translation of biomedical discoveries into clinical practice and providing guidance on developing the systems and workforce to ensure delivery of the care most appropriate to each patient. In several seminal reports, the IOM has highlighted key shortfalls in the safety, quality, and effectiveness of health care delivered—shortfalls compounded by the unsustainable growth of national healthcare expenditures and the recent strains on the national and global economy. Through convening activities and consensus reports, the IOM has also sought solutions to these issues by gathering stakeholder perspectives on key barriers and opportunities for improvement and by developing recommendations for moving toward a healthcare system driven by evidence and focused on delivering care of greater value to patients.

Convened in 2006, the IOM Roundtable on Value & Science-Driven Health Care has contributed to these discussions and articulated a vision for a learning health system, in which evidence is both applied and generated as a natural course of care. The Roundtable has explored key components of such a system though its Learning Health System series of workshops and publications. A central dimension of the Roundtable's work toward this long-term goal is expanding the capacity to meet the acute, near-term need for evidence of comparative effectiveness information. Such was the focus of the seventh workshop in the Roundtable's Learning Health System series, Learning What Works: Infrastructure Required for Comparative Effective-

ness Research. Held on July 30–31, 2008, this workshop gathered experts from across disciplines and sectors to explore the nature of the work required to develop insights on the comparative effectiveness of clinical interventions and care processes and to assess the current and needed capacity to expand and improve this work. This publication presents a summary of the workshop presentations and discussions and provides the beginnings of a roadmap for moving forward, especially important in the context of the passage of the American Recovery and Reinvestment Act (ARRA) of 2009. Included in ARRA is $1.1 billion of federal funds to increase national capacity for clinical effectiveness research, underscoring the pressing nature of the demand and the importance of developing coordinated and efficient capacity for the work needed.

I thank the members of the Roundtable staff for facilitation of Roundtable activities, as well as the sponsors, who make this work possible: the Agency for Healthcare Research and Quality, America's Health Insurance Plans, AstraZeneca, Blue Shield of California Foundation, Burroughs Wellcome Fund, California Health Care Foundation, Centers for Medicare & Medicaid Services, Charina Endowment Fund, Commonwealth Fund, Department of Veterans Affairs, Food and Drug Administration, Johnson & Johnson, sanofi-aventis, and Stryker. I would also like to offer my personal thanks to the Roundtable members themselves for their dedication.

Harvey V. Fineberg, M.D., Ph.D.
President, Institute of Medicine

Preface

The Institute of Medicine's Roundtable on Value & Science-Driven Health Care provides a trusted venue for sustained discussion and collaboration between national experts and health system stakeholders on issues important to improving the generation and application of evidence for healthcare decisions. The Roundtable has set the goal that by 2020, 90 percent of clinical decisions will reflect and be supported by accurate, timely, and up-to-date evidence. Anchoring their work toward this goal is a focus on three dimensions of the challenge:

1. Accelerating progress toward the long-term vision of a **learning health system,** in which evidence is both generated and applied as a natural product of the care process.
2. Expanding the capacity to meet the acute, near-term need for evidence of **comparative effectiveness** to support medical care that is most effective and produces greatest value.
3. Improving **public understanding** of the nature of evidence, the dynamic character of evidence development, and the importance of insisting on medical care that reflects the best evidence.

As illustrated by previous workshops and publications in the Roundtable's Learning Health System series, the nation's capacity to develop information on the comparative effectiveness of clinical interventions and care processes falls far short of the need. Evidence development often fails to meet the needs of healthcare decision makers—lacking applicability to the broad patient populations encountered in clinical practice or consid-

eration of individual patient variation and preference. Capacity to capture and analyze clinical data relevant to point-of-care decisions is growing but remains fragmented and inefficient, often due to proprietary or privacy concerns. And systems for synthesizing, translating, and applying evidence to clinical practice remain limited.

To explore and assess the infrastructure needed (e.g., skills, workforce, methods, coordination, information networks) to expand the nation's capacity to develop and apply comparative effectiveness information, the Roundtable convened the workshop, Learning What Works: Infrastructure Required for Comparative Effectiveness Research. Emerging from discussion at the 2-day workshop was the need for coordinated and tailored infrastructure development efforts that engage key healthcare system stakeholders—patients, providers, policy leaders, information technology experts, health services researchers, health economists, and educators, among others. The discussion highlighted key gaps in current and past work, including the fragmentation and redundancy of clinical effectiveness activities; the need to take better advantage of opportunities provided by emerging health information technologies and of established local and international capacities for evidence development, synthesis, and translation; the need to develop study designs and methods that can keep pace with the development of treatments and diagnostics and support dynamic, real-time approaches to learning; and a growing appreciation for comparative effectiveness research as a discipline, with unique workforce training and skills development needs. Also noted was the counter-productivity of the current balkanization of clinical effectiveness research investments and opportunities provided by public–private collaboration. This publication summarizes the presentations and key discussion points and is rich with insights that are important to inform needed infrastructure development.

We would like to acknowledge those individuals and organizations that donated valuable time toward the development of this workshop summary, including all present at the workshop and the workshop presenters, who not only offered valuable comments but also further developed their presentations into the manuscripts contained within this summary. We would also like to thank those who provided their counsel by serving on the planning committee for this workshop, including Stuart Altman, Ph.D. (Brandeis University); Kathy Buto, M.P.A. (Johnson & Johnson); Carolyn Clancy, M.D. (Agency for Healthcare Research and Quality); David Helms, Ph.D. (AcademyHealth); Mark McClellan, M.D., Ph.D. (Brookings Institution); Peter Orszag, Ph.D. (at the time of the workshop, Congressional Budget Office; currently, Office of Management and Budget); and John Rowe, M.D.

(*Chair*) (Mailman School of Public Health, Columbia University).[1] Roundtable staff, including Katharine Bothner, Lori Burns, Alex Goolsby, Kiran Gupta, Katie Jakubs, LeighAnne Olsen, Daniel O'Neill, Kate Sharaf, Ruth Strommen, and Catherine Zweig, were instrumental in coordinating the 2-day workshop in July 2008 and translating the workshop proceedings and discussion into this workshop summary. We would also like to thank Greta Gorman, Michele de la Menardiere, Abbey Meltzer, Robert Pool, Bronwyn Schrecker, Vilija Teel, Jackie Turner, and Jordan Wyndelts for helping to coordinate the various aspects of review, production, and publication.

The need to improve the effectiveness and value of health care is all the more pressing as economic disparities increase within the United States and globally. The development of methods to reduce costly system inefficiencies and waste and improve outcomes for each individual patient is integral to the sustainability of our healthcare system, and the implementation of infrastructure to support comparative clinical effectiveness research will provide a solid foundation.

Denis A. Cortese, M.D.
Chair, Roundtable on Value & Science-Driven Health Care

John W. Rowe, M.D.
Member, Roundtable on Value & Science-Driven Health Care, and *Chair*, Workshop Planning Committee

J. Michael McGinnis, M.D., M.P.P.
Executive Director, Roundtable on Value & Science-Driven Health Care

[1] Institute of Medicine planning committees are solely responsible for organizing the workshop, identifying topics, and choosing speakers. The responsibility for the published workshop summary rests with the workshop rapporteurs and the institution.

Contents

*Appendixes A-E are not printed in this book. They are available online at http://www.nap.edu/catalog.php?record_id=12214.

Abbreviations and Acronyms

AAAS	American Association for the Advancement of Science
AAP	American Academy of Pediatrics
AASHTO	American Association of State Highway and Transportation Officials
ACA	Affordable Care Act (2010)
ACC	American College of Cardiology
ACOG	American Congress of Obstetricians and Gynecologists
ADHD	attention deficit hyperactivity disorder
ADNI	Alzheimer's Disease Neuroimaging Initiative
AF	atrial fibrillation
AHA	American Heart Association or American Hospital Association
AHC	academic health center
AHIP	America's Health Insurance Plans
AHRQ	Agency for Healthcare Research and Quality
AIDS	acquired immunodeficiency syndrome
ALS	amyotrophic lateral sclerosis
AMA	American Medical Association
AMCP	Academy of Managed Care Pharmacy
AMD	age-related macular degeneration
AMI	acute myocardial infarction
AQA	Ambulatory Care Quality Alliance
ARRA	American Recovery and Reinvestment Act
ASC	active surveillance culturing

BC	Biomarkers Consortium
BMI	biomedical informatics
BMS	bare-metal stent
CABG	coronary artery bypass graft
CADTH	Canadian Agency for Drugs and Technologies in Health
CARE ICDJ	Querying Author about Definition
CATIE	Clinical Antipsychotic Trials in Intervention Effectiveness
CCN	Cardiac Care Network (Ontario)
CCR	Center for Cancer Research
CDC	Centers for Disease Control and Prevention
CDMS	chronic disease management system
CDR	Common Drug Review (Canada)
CDRH	Center for Devices and Radiological Health
CDSR	Cochrane Database of Systematic Reviews
CE	comparative effectiveness
CEAP	Committee for Assessment of Diagnostic and Therapeutic Procedures (France)
CED	coverage with evidence development
CENTRAL	Cochrane Central Register of Controlled Trials
CEPP	Committee for Assessment of Devices and Health Technologies (France)
CER	comparative effectiveness research
CEVG	Cochrane Eyes and Vision Group
CHD	coronary heart disease
CI	confidence interval
CIHR	Canadian Institutes of Health Research
CIS	clinical information system
CMS	Centers for Medicare & Medicaid Services
CMTP	Center for Medical Technology Policy
CNS	central nervous system
COE	Center for Outcomes and Evidence
COMPUS	Canadian Optimal Medication Prescribing and Utilization Service
COPD	chronic obstructive pulmonary disease
COPR	Council of Public Representatives
COX	cyclo oxygenase
CPI	Critical Path Initiative
CPOE	computerized physician order entry
CPR	computer-based patient record
CPSC	Consumer Product Safety Commission
CRG	Cochrane HIV/AIDS review group
CRN	Cancer Research Network

CS	computer science
CT	computed tomography
CTSA	Clinical and Translational Science Awards
CTSC	Clinical and Translational Science Center
DAG	directed acyclic graph
DBS	deep brain stimulation
DCRI	Duke Clinical Research Institute
DEcIDE	Developing Evidence to Inform Decisions about Effectiveness
DERP	Drug Effectiveness Review Project
DES	drug-eluting stent
DHS	Department of Homeland Security
DNA	deoxyribonucleic acid
DOD	Department of Defense
DOE	Department of Energy
DSM	Diagnostic and Statistical Manual of Mental Disorders
EBM	evidence-based medicine
EC	executive committee
ECG	electrocardiogram
EHR	electronic health record
EMEA	European Medicines Agency
EMR	electronic medical record
EPC	evidence-based practice center
EUnetHTA	European Network for Health Technology Assessment
FDA	Food and Drug Administration
FDAAA	FDA Amendments Act of 2007
FFRDC	federally funded research and development center
FNIH	Foundation for the National Institutes of Health
FOMC	Federal Open Market Committee
FTC	Federal Trade Commission
FTE	full-time equivalent
FY	fiscal year
G-BA	Federal Joint Committee (Germany)
GBS	Guillain-Barré Syndrome
GCGH	Grand Challenges in Global Health
GERD	gastroesophageal reflux disease
GIN	Guidelines International Network
GKS	gamma knife surgery
GPRD	General Practices Research Data (United Kingdom)

GRADE	Grading of Recommendations Assessment, Development, and Evaluation Working Group
HAI	hospital acquired infection
HAS	Haute Autorité de Santé (France)
HCDS	health care delivery system
HECS	highly effective clinical services
HEI	Health Effects Institute
HER2	human epidermal growth factor receptor 2
HHS	Department of Health and Human Services
HIMSS	Health Information Management Systems Society
HIPAA	Health Insurance Portability and Accountability Act
HIT	health information technology
HIV	human immunodeficiency virus
HMO	health maintenance organization
HMORN	HMO Research Network
HQA	Hospital Quality Alliance
HR	hazard ratio
HRT	hormone-replacement therapy
HSI	Homeland Security Institute
HSR	health services research
HSRProj	Health Services Research Projects in Progress
HTA	health technology assessment
HVI	Heart and Vascular Institute
IC	institute and center
ICD	implantable cardiac defibrillator
ICER	Institute for Clinical and Economic Review
ICES	Institute for Clinical Evaluative Sciences
IDA	Institute for Defense Analyses
IH	international health
ILI	influenza-like-illness
INHATA	International Network of Agencies for Health Technology Assessment
IOM	Institute of Medicine
IQWiG	Institute for Quality and Efficiency (Germany)
IRB	institutional review board
ISO	independent scientific organization
IT	information technology
JHU	Johns Hopkins University

LLNL	Lawrence Livermore National Laboratory
LVEF	left ventricular ejection fraction
MAS	Medical Advisory Secretariat
MCV	meningococcal conjugate vaccine
MEDCAC	Medicare Evidence Development and Coverage Advisory Committee
MedPAC	Medicare Payment Advisory Commission
MI	myocardial infarction
MIS	management information systems
MIT	Massachusetts Institute of Technology
MMA	Medicare Prescription Drug, Improvement, and Modernization Act
MOHLTC	Ministry of Health and Long-Term Care (Ontario)
MRI	magnetic resonance imaging
MRSA	methicillin-resistant *Staphylococcus aureus*
MSAC	Medical Services Advisory Committee (Australia)
NAE	National Academy of Engineering
NANSAID	non-aspirin, non-steroidal anti-inflammatory drug
NAS	National Academy of Sciences
NC2D	National Consortium of Clinical Databases
NCCN	National Cancer Care Network
NCHRP	National Cooperative Highway Research Program
NCHS	National Center for Health Statistics
NCI	National Cancer Institute
NDAI	National Data Aggregation Initiative
NGO	nongovernmental organization
NHLBI	National Heart, Lung, and Blood Institute
NHS	National Health Service (United Kingdom)
NICE	National Institute for Health and Clinical Excellence (United Kingdom)
NIH	National Institutes of Health
NMCES	National Medical Care Expenditure Survey
NNSA	National Nuclear Security Administration
NRC	National Research Council
NRT	nicotine replacement therapy
NSAID	non-steroidal anti-inflammatory drug
NSF	National Science Foundation
NYHA	New York Heart Association
NYU	New York University

OAI	Osteoarthritis Initiative
ODE-I	Office of Drug Evaluation I
OFPP	Office of Federal Procurement Policy
OHSU	Oregon Health and Science University
OHTAC	Ontario Health Technology Assessment Committee
OS	Office of the Secretary
OSD	Office of the Secretary of Defense
OSTP	Office of Science and Technology Policy
OTA	Office of Technology Assessment
OTC	over-the-counter
PACS	Picture Archiving and Communication System
PATH	Program for the Assessment of Technology in Health
PBAC	Pharmaceutical Benefits Advisory Committee (Australia)
PCI	percutaneous coronary intervention
PCORI	Patient-Centered Outcomes Research Institute
PCR	polymerase chain reaction
PDUFA	Prescription Drug Use Fee Act
PET	positron emission tomography
POC	point of care
PPP	public–private partnership
PSA	prostate-specific antigen
PSAC	Priority Setting Advisory Committee
PTSD	post-traumatic stress disorder
QI	quality improvement
QIO	quality improvement organization
R&D	research and development
RCT	randomized controlled trial or randomized clinical trial
RNI	radionuclide imaging
RWJF	Robert Wood Johnson Foundation
S&T	science and technology
SBU	Swedish Council on Technology Assessment in Health Care
SEER	Surveillance, Epidemiology and End Results
SES	socioeconomic status
SHRPII	Strategic Highway Research Programs II
SNOMED	Systematized Nomenclature of Medicine
SPARQL	SPARQL Protocol and RDF Query Language
SSNDI	Social Security National Death Index
SSRI	selective serotonin reuptake inhibitor
STEMI	ST-segment elevation myocardial infarction

STPI	Science and Technology Policy Institute
STS	Society of Thoracic Surgeons
SUNY	State University of New York
TAB	Technology Assessment Board
TAPP	transabdominal pre-peritoneal
TCA	tricyclic antidepressant
TEC	Technology Evaluation Center
TEP	totally extraperitoneal
TGA	Therapeutic Goods Administration (Australia)
THETA	Toronto Health Economics and Technology Assessment Collaboration
tPA	tissue plasminogen activator
TRB	Transportation Research Board
UAB	University of Alabama at Birmingham
UC	University of California
UMLS	unified medical language system
UNC	University of North Carolina
URI	upper respiratory infection
USPSTF	U.S. Preventive Services Task Force
USQA	U.S. Quality Algorithms
USRDS	U.S. Renal Data System
VA	Department of Veterans Affairs
VATAP	VA Technology Assessment Program
VBID	value-based insurance design
VHA	Veterans Health Administration
VSD	Vaccine Safety Datalink
WHO	World Health Organization

Summary

INTRODUCTION AND OVERVIEW

The demand for better evidence to guide healthcare decision making is increasing rapidly for a variety of reasons, including the adverse consequences of care administered without adequate evidence, emerging insights into the proportion of healthcare interventions that are unnecessary, recognition of the frequency of medical errors, heightened public awareness and concern about the very high costs of medical care, the burden on employers and employees, and the growing proportion of health costs coming from out of pocket (Fisher and Wennberg, 2003; Fisher et al., 2003a, 2003b; IOM, 2000, 2001, 2008a; McGlynn et al., 2003; Wennberg et al., 2002). Although nearly $2.5 trillion was spent in 2009 on health and medical care in the United States, only a very small portion of that amount—perhaps less than one tenth of 1 percent—was devoted to learning what works best in health care, for whom, and under what circumstances.

To improve the effectiveness and value of the care delivered, the nation needs to build its capacity for ongoing study and monitoring of the relative effectiveness of clinical interventions and care processes through expanded trials and studies, systematic reviews, innovative research strategies, and clinical registries, as well as improving its ability to apply what is learned from such study through the translation and provision of information and decision support. Several recent initiatives have proposed the development of an entity to support expanded study of the comparative effectiveness of interventions. To inform policy discussions on how to meet the demand for more comparative effectiveness research (CER) as a means of improving

the effectiveness and value of health care, the Institute of Medicine (IOM) Roundtable on Value & Science-Driven Health Care convened a workshop on July 30–31, 2008, titled Learning What Works: Infrastructure Required for Comparative Effectiveness Research. Box S-1 describes the issues that motivated the meeting's discussions: the substantial and growing interest in activities and approaches related to CER; the lack of coordination of key activities, such as the selection and design of studies, synthesis of existing evidence, methods innovation, and translation and dissemination of CER information; shortfalls and widening gaps in the workforce needed in all areas of CER; the opportunities presented by the recent calls for expanded resources for work on the comparative effectiveness of clinical interventions; the growing appreciation of the infrastructure needed to support this work; and the need for a trusted, common venue to identify and characterize the need categories, begin to estimate the shortfalls, consider approaches to addressing the shortfalls, and identify priority next steps.

BOX S-1
Issues Motivating the Discussion

1. Substantial demand for greater insights into the comparative clinical effectiveness of clinical interventions and care processes to improve the effectiveness and value of health care.
2. Expanded interest and activity in the work needed—e.g., comparative effectiveness research, systematic reviews, innovative research strategies, clinical registries, coverage with evidence development.
3. Currently fragmented and largely uncoordinated selection of studies, study design and conduct, evidence synthesis, methods validation and improvement, and development and dissemination of guidelines.
4. Expanding gap in workforce with skills to develop data sources and systems, design and conduct innovative studies, translate results, and guide application.
5. Opportunities presented by the attention of recent initiatives and the increasing possibility of developing an entity and resources for expanded work on the comparative effectiveness of clinical interventions.
6. Growing appreciation of the importance of assessing the infrastructure needed for this work—e.g., workforce needs, data linkage and improvement, new methodologies, research networks, technical assistance.
7. Desirability of a trusted, common venue to identify and characterize the need categories, begin to estimate the shortfalls, consider approaches to addressing the shortfalls, and identify priority next steps.

The goal of the workshop was to clarify the elements and nature of the needed capacity, solicit quantitative and qualitative assessments of the needs, and characterize them in a fashion that will facilitate engagement of the issues by policy makers. Two assumptions guided the discussions but were not explored as part of the workshop: resources will be available to expand work on the comparative effectiveness of medical interventions, and, given recent public discourse on the need for a stronger focus on the work, a designated entity would be developed with a formal charge to coordinate the expanded work.

The workshop gathered leading practitioners in health policy, technology assessment, health services research, health economics, information technology (IT), and health professions education and training to explore, through invited presentations, the current and future capacity needed to generate new knowledge and evidence about what works best, including skills and workforce, data linkage and improvement, study coordination and result dissemination, and research methods innovation. Participants explored, in both qualitative and quantitative terms, the nature of the work required, the IT and integrative vehicles required, the skills and training programs required, the priorities to be considered, the role of public–private partnerships, and the strategies for immediate attention while considering the long-term needs and opportunities. Through the course of the workshop, a number of common themes and implications emerged. These are indicated below, along with a number of possible follow-up actions identified for Roundtable consideration.

Since the meeting, three events have occurred with significant implications for the infrastructure necessary for comparative effectiveness research: (1) the American Recovery and Reinvestment Act of 2009 (ARRA) included $1.1 billion for the conduct of CER; (2) formal assessments by the IOM and the federal government have recommended priorities for such research; and (3) the Accountable Care Act of 2010 (ACA) established an independent Patient-Centered Outcomes Research Institute (PCORI). See Appendixes C, D, and, E for additional background. Accordingly some of the information has been updated as appropriate to bring the text current with 2011 circumstances.

Comparative Effectiveness Research and the Roundtable on Value & Science-Driven Health Care

The IOM's Roundtable on Value & Science-Driven Health Care provides a trusted venue for key stakeholders to work cooperatively on innovative approaches to the generation and application of evidence that will drive improvements in the effectiveness and efficiency of medical care in the United States. Participants seek the development of a *learning health system*

that enhances the availability and use of the best evidence for the collaborative healthcare choices of each consumer and healthcare professional, that drives the process of discovery as a natural outgrowth of patient care, and that ensures innovation, quality, safety, and value in health care. As leaders in their fields, Roundtable members work with their colleagues to identify issues not being adequately addressed, determine the nature of the barriers and possible solutions, and set priorities for action. They marshal the energy and resources of the sectors represented on the Roundtable to work for sustained public–private cooperation for change.

This work is focused on the three major dimensions of the challenge:

1. accelerating progress toward the long-term vision of a *learning health system,* in which evidence is both applied and developed as a natural product of the care process,
2. expanding the capacity to meet the acute, near-term need for evidence of *comparative effectiveness* to support medical care that is maximally effective and produces the greatest value,
3. improving *public understanding* of the nature of evidence, the dynamic character of evidence development, and the importance of insisting on medical care that reflects the best evidence.

Roundtable members have set a goal that *by the year 2020, 90 percent of clinical decisions will be supported by accurate, timely, and up-to-date clinical information and will reflect the best available evidence.* To achieve this goal, Roundtable members and their colleagues work to identify priorities for action on those key issues in health care where progress requires cooperative stakeholder engagement. Central to these efforts is the *Learning Health System* series of workshops and publications that collectively characterize the key elements of a healthcare system that is designed to generate and apply the best evidence about the healthcare choices of patients and providers as well as identify barriers to the development of such a system and opportunities for progress.

Each meeting is summarized in a publication available through the National Academies Press. Workshops in this series include the following:

- The Learning Healthcare System (July 20–21, 2006)
- Judging the Evidence: Standards for Determining Clinical Effectiveness (February 5, 2007)
- Leadership Commitments to Improve Value in Healthcare: Toward Common Ground (July 23–24, 2007)
- Redesigning the Clinical Effectiveness Research Paradigm: Innovation and Practice-Based Approaches (December 12–13, 2007)

- Clinical Data as the Basic Staple of Health Learning: Creating and Protecting a Public Good (February 28–29, 2008)
- Engineering a Learning Healthcare System: A Look to the Future (April 28–29, 2008)
- Learning What Works: Infrastructure Required for Learning Which Care is Best (July 30–31, 2008)
- Value in Health Care: Accounting for Cost, Quality, Safety, Outcomes, and Innovation (November 17–18, 2008)
- The Healthcare Imperative: Lowering Costs and Improving Outcomes (May, July, September, December, 2009)
- Digital Infrastructure for the Learning Health System: The Foundation for Continuous Improvement in Health and Health Care (July, September, October, 2010)

This publication summarizes the proceedings of the seventh workshop in the Learning Health System series, which focused on the infrastructure needs—e.g., methods, coordination capacities, data resources and linkages, workforce—for developing an expanded and efficient national capacity for CER. A synopsis of the key points from each of the sessions is included in this chapter, with more detailed information on session presentations and discussions found in the chapters that follow. Sections of the workshop summary not specifically attributed to an individual are based on the presentations, background papers, and discussions associated with the workshop, and reflect the views of this publication's rapporteurs, not those of the IOM Roundtable on Value & Science-Driven Health Care.

Day 1 featured two keynote speakers who provided a vision for developing an infrastructure that can contribute to an evidence base of what works best for whom, as well a sense of some of the potential returns from health care driven by evidence (Chapter 1), and presentations by speakers asked to characterize the nature of the work (Chapter 2), the information networks (Chapter 3), and the talent (Chapter 4) needed to carry out that vision. Day 2 featured discussions focused on identifying priority items for implementation to meet current shortfalls and opportunities to build upon existing public–private partnership efforts (Chapter 5). Chapter 6 provides a summary of the final session's discussion to outline key elements of a roadmap for progress, suggest some "quick hits" for immediate implementation, and opportunities to build needed support; this chapter also highlights common themes from the meeting's discussions and suggestions on opportunities for follow-up actions by the Roundtable. An overview of the topics discussed in specific manuscripts is provided in Table S-1.

A white paper, authored by staff in 2007 and titled *Learning What Works Best: The Nation's Need for Evidence on Comparative Effectiveness*

TABLE S-1 Overview of the Specific Aspects of Comparative Effectiveness Research (CER) Infrastructure Addressed in This Publication's Manuscript

Chapter	Manuscript and Author(s)	CER Research Methods and Settings	Clinical Data Development and Use	Health Information Technology	Evidence Review and Synthesis	Coordination and Dissemination	Workforce Education and Training	International CER Efforts
1	The Nation's Need for Evidence on Comparative Effectiveness in Health Care: Learning What Works Best *J. Michael McGinnis et al.*	■	■	■	■	■	■	■
	A Vision for the Capacity to Learn What Care Works Best *Mark B. McClellan*	■	■	■			■	
	The Potential Returns from Evidence-Driven Health Care *Gail R. Wilensky*		■			■		
2	The Cost and Volume of Comparative Effectiveness Research *Erin Holve and Patricia Pitman*	■	■				■	
	Intervention Studies That Need to Be Conducted *Douglas B. Kamerow*		■			■		
	Clinical Data Sets That Need to Be Mined *Jesse A. Berlin and Paul E. Stang*		■					
	Knowledge Synthesis and Translation That Need to Be Applied *Richard A. Justman*				■	■		

3	Methods That Need to Be Developed *Eugene H. Blackstone et al.*	■	■	■				
	Coordination and Technical Assistance That Need to Be Supported *Jean R. Slutsky*				■			
	Electronic Health Records: Needs, Status, and Costs for U.S. Healthcare Delivery Organizations *Robert H. Miller*			■				
	Data and Information Hub Requirements *Carol C. Diamond*		■	■				
4	Integrative Vehicles Required for Evidence Review and Dissemination *Lorne A. Becker*				■	■		■
	Comparative Effectiveness Workforce—Framework and Assessment *William R. Hersh et al.*						■	
	Toward an Integrated Enterprise—The Ontario, Canada, Case *Sean R. Tunis et al.*					■	■	■
5	Information Technology Platform Requirements *Mark E. Frisse*			■				

continued

TABLE S-1 Continued

Chapter	Manuscript and Author(s)	CER Research Methods and Settings	Clinical Data Development and Use	Health Information Technology	Evidence Review and Synthesis	Coordination and Dissemination	Workforce Education and Training	International CER Efforts
5	Data Resource Development and Analysis Improvement *T. Bruce Ferguson, Jr., and Ansar Hassan*		■					
	Practical Challenges and Infrastructure Priorities for Comparative Effectiveness Research *Daniel E. Ford*	■				■	■	
	Transforming Health Professions Education *Benjamin K. Chu*			■			■	
	Building the Training Capacity for a Health Research Workforce of the Future *Steven A. Wartman and Claire Pomeroy*	■					■	
	Public–Private Partnerships *Carmella A. Bocchino et al.*	■	■	■				
6	The Roadmap—Policies, Priorities, Strategies, and Sequencing *Stuart Guterman et al.*	■	■	■	■	■	■	■

in Health Care, provided important context for the workshop discussions. The executive summary of that white paper and the full manuscript are included in Chapter 1 and Appendix A, respectively. Appendix B includes evidence summaries of research questions identified and other materials relevant to discussion in a paper in Chapter 2. Appendixes C and D present the recommendations of two groups for priority studies in CER: *Initial National Priorities for Comparative Effectiveness Research*, an Institute of Medicine report; and the Federal Coordinating Council for Comparative Effectiveness Research *Report to the President and Congress*. Appendix E contains the portions of the ACA relevant to the structure, funding, and charge of PCORI. The workshop agenda, biographical sketches of the workshop participants, and a list of workshop attendees can be found in Appendixes F, G, and H, respectively.

COMMON THEMES

Common themes that emerged from the 2 days of discussion are summarized in Box S-2 and elaborated in the text that follows:

- ***Care that is effective and efficient stems from the integrity of the infrastructure for learning.*** The number of medical diagnostics and treatments available to patients and caregivers is increasing, but the knowledge about their effectiveness—in particular, their comparative effectiveness—is not keeping pace. This is in part a function of the rate of change, but it is also a product of capacity that is both underdeveloped and, as several participants noted, substantially fragmented, which leads to gaps, inefficiencies, and inconsistencies in the work. The accelerating rate of change in the interventions requiring effectiveness assessment compels a substantial shoring up in the level of effort, the nature of the effort, and the coordination of the effort in order to produce the necessary insights into the right care for different people under different circumstances.
- ***Coordinating work and ensuring standards are key components of the evidence infrastructure.*** Several presentations highlighted the point that substantial activity is currently under way in effectiveness research, including work on comparative effectiveness, but the activities are fragmented and often redundant in both structure and function. The fact that the application of evidence lags behind its production is in part a function of the disparate and "siloed" approaches between and within organizations seeking and developing information. The notions of infrastructure for evidence development therefore also include the capacity for greater coordination in the setting of study priorities; the development of systematic

BOX S-2
Infrastructure Required for Comparative Effectiveness Research: *Common Themes*

- Care that is effective and efficient stems from the integrity of the infrastructure for learning.
- Coordinating work and ensuring standards are key components of the evidence infrastructure.
- Learning about effectiveness must continue beyond the transition from testing to practice.
- Timely and dynamic evidence of clinical effectiveness requires bridging research and practice.
- Current infrastructure planning must build to future needs and opportunities.
- Keeping pace with technological innovation compels more than a head-to-head and time-to-time focus.
- Real-time learning depends on health information technology investment.
- Developing and applying tools that foster real-time data analysis is an important element.
- A trained workforce is a vital link in the chain of evidence stewardship.
- Approaches are needed that draw effectively on both public and private capacities.
- Efficiency and effectiveness compel globalizing evidence and localizing decisions.

decisions for the conduct of CER, systematic reviews, and guideline development; and the need to ensure the consistent translation of developed information. The identification of priority conditions, evaluation, and evidence gaps is needed in order to target limited resources, especially for high-cost or high-volume procedures and interventions.

- ***Learning about effectiveness must continue beyond the transition from testing to practice.*** "The learning process cannot stop when the label is approved," one meeting participant pointed out. Premarket testing for the safety and effectiveness of various interventions cannot assess the results for all populations or the circumstances of use and differences in practice patterns, so gathering information as interventions are applied in practice settings should represent a key focus in designing the infrastructure to learn which care is best. Local coverage decisions and private insurer use

of coverage with evidence development approaches were cited as opportunities to learn as a part of the care process.

- ***Timely and dynamic evidence of clinical effectiveness requires bridging research and practice.*** Although historical insulation of clinical research from the regular delivery of healthcare services evolved to facilitate data capture and control for confounding factors, it may not adequately inform the real-world setting of clinical practice. With the prospect of enhanced electronic data capture at the point of care on real-world patient populations, and statistical approaches to improve analysis, as well as increasing demand to keep pace with technologic innovation, this divide increasingly limits the utility of research results. Efforts under way to better engage health delivery organization, practitioners, patients, and the community in research prioritization, conduct, and results dissemination should be supported and expanded.
- ***Current infrastructure planning must build to future needs and opportunities.*** Research is often driven more by the methods than the questions. In fact, both are important, and infrastructure planning must account for both the key emerging healthcare questions and the key emerging CER opportunities. Emerging questions include those related to the management of multiple co-occurring chronic diseases of increasing prevalence in an aging population, the improved insights into individual variation relevant to both treatments and diagnostics, and the impact of innovation in shortening the lifecycle of any particular intervention. Emerging tools include innovations in trial design, the development of new statistical approaches to data analysis, and the development of electronic medical and personal health records.
- ***Keeping pace with technological innovation compels more than a head-to-head and time-to-time focus.*** Much of the current discussion about CER has emphasized the need for more clinical trials and more head-to-head studies. Although there are numerous examples of diagnostic and treatment interventions for which such studies are needed, the notion of a research process that essentially offers periodic and static determinations is inherently limited. Especially with the rapid pace of change in the nature of interventions and the difficulty, expense, and time required to develop studies—and the challenges of ensuring the generalizability of results in the face of limitations of the traditional approach to randomized controlled trials (RCTs)—a first-order priority for effectiveness research is the establishment of infrastructure for a more dynamic, real-time approach to learning. Leveraging new tools, such as health information technology (IT) should allow for

a more networked and distributed approach to information sharing and evidence creation.

- ***Real-time learning depends on HIT investment.*** It was noted that collecting data is the most time-intensive part of trials and studies, and IT is critical to streamlining this work. Moreover, it is the key to accelerated learning from broader-based clinical experience. We heard that "[t]he type of learning needed cannot be paper based." The increasing complexity of the factors involved in understanding the effectiveness of clinical options under different circumstances requires a blend of database access and computing power that can only be provided from broadly applied HIT. Although not in itself sufficient to produce the information required for better medical care management, it is a necessity for the continuous improvement expected of a learning health system. A policy framework for privacy and security will be necessary to build and maintain public trust that information will be protected as it is shared.
- ***Developing and applying tools that foster real-time data analysis is an important element.*** The scope and scale of evidence needs suggests that innovation is needed across the range of research methods, from making clinical trials faster and less expensive to moving beyond randomized trials to better address practical circumstances, using registries, observational databases, and other emerging data resources. If full advantage is to be taken of HIT, statistical tools and analytic algorithms that can be embedded in databases to allow real-time insights will be important. Similarly, tools are needed that will allow findings to be drawn from databases built on different vendor platforms, using semantic technology to integrate currently disparate medical data, in order to develop the next generation of statistical tools for the analysis of clinical data, including the building of models that allow insights to be generated by virtual studies.
- ***A trained workforce is a vital link in the chain of evidence stewardship.*** As in many other domains, progress in CER will relate to the capacity to develop and maintain the broad and diversely skilled workforce needed. Mention was often made of that factor as a determining element as well for progress in development of the learning health system. Given the pace of change in the number and variety of clinical interventions as well as in the tools and approaches to assessing them, there is a need to ensure that these developing opportunities are matched by the skills of the workforce. This includes training and education in the methodologies of research design, the translation of research, guideline development, and the maintenance and mining of clinical records. But it also

includes giving attention to reorienting the education of frontline caregivers around their emerging responsibilities for access, interpretation, and discussion with patients of a dynamic evidence base, as well as helping to ensure the availability and integrity of the clinical data that shape conclusions on evidence.

- ***Approaches are needed that draw effectively on both public and private capacities.*** Several times in the course of the meeting it was pointed out that although the total investment in CER in the United States is substantial, it is inefficient because of the absence of a vehicle for common priority setting and coordination of efforts and because the work on effectiveness done by private companies in product development and testing is usually not accessible to the broader community. In aggregate, private investment often far exceeds public investment in assessing a given intervention, but even when available, studies associated with an enterprise with a commercial stake may be viewed suspiciously. Several models are in development to establish public–private collaborative efforts to improve the efficiency and effectiveness of the work.
- ***Efficiency and effectiveness compel globalizing evidence and localizing decisions.*** Two presentations featured international work, including the Cochrane Collaboration on evidence synthesis, and efforts in Ontario, Canada, to develop and apply insights about the comparative effectiveness of clinical interventions. Reference was made throughout the meeting to work going on elsewhere in the world and, in particular, to work at the National Institute for Health and Clinical Excellence in the United Kingdom. This brought clearly into play the need to ensure that, where possible, common work to assess an intervention's clinical effectiveness—or collective work to assess the body of evidence—be collaborative and well coordinated across boundaries, while also being mindful that different cultural and policy environments may lead to different decisions at the local level.

Key Factors and Needs

Workshop speakers described a number of implications of the current state of play for the development of an infrastructure for CER (Box S-3). These included the following:

- ***Several elements are involved in infrastructure development.*** Developing the infrastructure for CER has at least five dimensions: (1) putting in place the physical capacity, i.e., the hardware; (2) developing the analytic tools and methods; (3) training the work-

BOX S-3
Key Factors and Needs for Expanded Comparative Effectiveness Research Capacity

- Several elements are involved in infrastructure development:
  - putting in place the physical capacity, i.e., the hardware;
  - developing the analytic tools and methods;
  - training the workforce needed;
  - establishing processes for efficient and effective operation; and
  - shaping the strategy for attention and phasing.
- Strategies and priorities for infrastructure application include the following:
  - conduct of systematic reviews,
  - conduct of primary research,
  - clinical registry resources,
  - introduction of health information technology throughout practice,
  - fostering public and private collaboration, and
  - keeping focus on the utility and impact of a networked world.

force needed; (4) establishing processes for efficient and effective operation; and (5) shaping the strategy for attention and phasing. Presentations at the meeting described and discussed in qualitative terms the needs and challenges in each of these dimensions and offered "opening bid" quantitative estimates on the needs for the IT infrastructure, as well as for investments in human capital. Refinements of these first approximations will be needed, as will additional clarity on the analytic tools, processes, and strategies for a stronger infrastructure for research into effective health care.

- ***Strategies and priorities for infrastructure application.*** The dimensions noted above represent in certain ways the functional dimensions of relevance to the infrastructure that is needed for effectiveness research. There are phasing considerations as well, in part driven by the ability and need to take actions even without additional resources and in part driven by the time required to set in motion the necessary activities. Suggestions for key strategies and priorities for progress included the following:
  - ***Conduct of systematic reviews.*** There is an immediate need to improve the conduct, coordination, and consistency of systematic reviews—a point that, in effect, echoed the recommendations of the 2008 IOM report *Knowing What Works in Health Care: A Roadmap for the Nation.*

- ***Conduct of primary research.*** Similarly, the approach to primary research on effectiveness needs a more systematic means of determining priorities, better tools and more streamlined designs, and a deeper bench workforce to do the work.
- ***Clinical registry resources.*** In making the transition to a pattern of real-time, continuous learning in health care—in effect, creating a beta approach to clinical data systems that generate learning—the technologies for clinical registries and in the field of registry development, maintenance, and improvement will need to be strengthened.
- ***Introduction of HIT throughout practice.*** In the area of IT development, the issues include the need to install appropriate hardware in virtually every clinical setting, the incorporation into operating software of design elements that are pegged to research activities and embedded analytic tools, the incorporation of design elements used in decision assistance, and training of the required workforce to work with this technology.
- ***Fostering public and private collaboration.*** The longer-term development needed to sustain the growth and improvement of the infrastructure will include the design of approaches that foster meaningful public and private collaboration in support of the research activities.
- ***Keeping focus on the utility and impact of a networked world.*** Also important to guide strategy development in the long term are approaches designed to take advantage of the resources emerging in our increasingly networked world—the opportunities for which hints are provided by recent developments, such as the Patients Like Me Web site, the HMO Research Network, the registries of the Society of Thoracic Surgeons, and even information made available by such resources as Google and Wikipedia.

CONTEXT, PRESENTATION, AND DISCUSSION SUMMARIES

Background for workshop discussions was provided by an IOM staff-authored background brief that illustrates the case for expanded CER, provides an overview of current CER activities and needs, and briefly discusses relevant issues not under consideration at the workshop (e.g., financing and structure of a new entity to coordinate CER work). Workshop presentations focused on key infrastructure needs and identified components of and existing capacity for CER, provided qualitative and quantitative assessments of what is needed to meet the demand, and suggested options for strengthening and building upon existing infrastructure. Much of the discussion focused

on the prioritization of these needs and how to develop the beginnings of a roadmap of specific immediate steps and priority actions needed to move from where we are to where we need to be. The background brief, workshop presentations, and meeting discussions are summarized below—with expanded discussion included in the main body of the text.

The Need and Potential Returns for Comparative Effectiveness Research

Enhancing the capacity for CER is not an end in itself but is rather a means to begin guiding the development of a healthcare system in which care is evidence driven and focused on providing care of value to individual patients. The staff-authored issue brief, provided as background for meeting discussions, and two presentations provided an important starting point for workshop discussions by summarizing current CER capacity, outlining a vision for—and suggesting the potential returns of—an evidence- and value-driven healthcare system.

The Nation's Need for Comparative Effectiveness Research[1]

The nation's capacity has fallen far short of its need for producing reliable and practical information about the care that works best. Medical-care decision making is now strained, at both the level of the individual patient and the level of the population as a whole, by the growing number of diagnostic and therapeutic options for which there is insufficient evidence to make a clear choice (IOM, 2008a). As reviewed in the background paper provided to workshop participants, these developments have fundamental implications for health prospects, and to capture and use them effectively and efficiently will require a proportionate commitment to understand their advantages and appropriate applications. It is a problem in both capacity investment and resource allocation. If only 1 percent of the nation's healthcare bill were devoted to understanding the effectiveness of the care purchased, the total investment would be more than $20 billion annually. In contrast, even accounting for the support from all private and public sources, the aggregate national commitment to assessing clinical interventions is still likely well under 1 percent.[2]

[1] At the request of the Roundtable's sustainable capacity working group, Roundtable staff developed an Issue Overview on national capacity for CER (IOM, 2007). This paper was provided as part of the meeting briefing materials to inform workshop discussion and is summarized briefly here (S-16–S-19). The complete white paper can be found in Appendix A of this publication.

[2] The American Recovery and Reinvestment Act of 2009 provided $1.1 billion of funds to the National Institutes of Health, Agency for Healthcare Research and Quality, and the Secretary of Health and Human Services for activities related to CER.

Activities currently under way to assess the effectiveness of healthcare interventions are broad but underresourced and fall far short of the need (IOM, 2007). In addition to the contributions of industry through phase 3 and 4 trials, several government agencies support CER, including the Agency for Healthcare Research and Quality (AHRQ), which has a specific mandate and a small appropriation for CER. The total of all appropriations to all federal agencies—the National Institutes of Health (NIH), the Veterans Health Administration, the Department of Defense, the Centers for Medicare & Medicaid Services (CMS), the Food and Drug Administration (FDA), AHRQ, and the Centers for Disease Control and Prevention—for all health services research amounts to about $1.5 billion, only a modest portion of which is devoted to CER and which is far below the industry level (AcademyHealth, 2005). Additional work, also modest, is undertaken by certain of the larger healthcare delivery organizations. Evidence synthesis activity is supported by the insurance industry, professional societies, healthcare organizations, and government. AHRQ has established a network of 13 AHRQ-sponsored evidence-based practice centers that review literature and produce evidence reports including comparative effectiveness reviews. Organizations interested in evidence reviews will often draw upon syntheses performed by several well-established technology assessment entities (IOM, 2008a).

The most pressing needs of clinicians and their patients center on the development of reliable studies on which to base their decisions. These needs have been characterized in various ways, and they can be grouped into the key areas indicated in Box S-4 (Buto and Juhn, 2006; Clancy, 2006; Health Industry Forum, 2006; Hopayian, 2001; Kupersmith et al., 2005; Rowe et al., 2006). The related key challenges are summarized in Table S-2.

To narrow the rapidly growing gap between the available evidence on clinical effectiveness and the evidence necessary for sound clinical decision making, various organizations and recent public articles have called for the creation of a new entity and a quantum increase—several billion dollars—for CER (IOM, 2008a; Kupersmith et al., 2005; Wilensky, 2005). The various approaches suggested for building the required capacity can be grouped into four categories according to the funding patterns for their support (Box S-5). Each of the approaches is based on an existing or recent model. Although presented as discrete models for discussion purposes, they are not mutually exclusive.

Other implementation considerations include those related to funding and program management. Suggestions for the funding mechanism range from a direct annual federal appropriation or a small set-aside from the Medicare Trust Fund to the structuring of proportionately matching contributions including set-asides from Medicare fund expenditures, from private health insurance premiums, or from manufacturer research and

BOX S-4
Prominent Comparative Effectiveness Research Activities and Needs

1. Studies of comparative effectiveness ("head to head")
2. Systematic reviews of comparative effectiveness
3. Assessment of comparative value/cost effectiveness
4. Coordinated priority setting and execution
5. Improved study designs and research methods
6. Better linkage of studies of efficacy, safety, and effectiveness
7. Appropriate evidence standards consistently applied
8. Consistent recommendations for clinical practice
9. Guidance for coverage and funding
10. Dissemination, application, and public communication

SOURCE: IOM, 2007.

TABLE S-2 Prominent Comparative Effectiveness Research Activities and Needs—Key Challenges

Issue	Key Challenges
Head-to-head studies	Scant resources; rapidly increasing need; comparison choice
Systematic reviews	Few primary studies; inconsistent methods; uncoordinated
Comparative value insights	Little agreement on metrics or role of costs; cost fluctuation
Priority setting	Fragmentation; inefficiency; no mechanism for coordination
Study designs and tools	Clinical trial time/cost/limits; large dataset mining methods
Research life-cycle links	Efficacy–effectiveness disjuncture; postapproval surveillance
Evidence standards	Standards not adapted to needs; inconsistency in application
Practice guidance	Disparate approaches; conflicting recommendations
Coverage guidance	Narrow evidence base; limited means for provisional coverage
Application tools	Public misperceptions; incentive structures; decision support

SOURCE: IOM, 2007.

BOX S-5
Models for Enhancing Capacity

Incremental funding augmentations
- Incremental model

Publicly funded entity
- Executive branch agency model
- Independent government commission model
- Legislative branch office model

Privately funded entity
- Operating foundation model
- Investment tax credit cooperative model

Public–private funded entity
- User fee public model
- Federally funded research and development center public model
- Independent cooperative model
- Independent quasi-governmental authority model

SOURCE: IOM, 2007.

development expenditures (Hopayian, 2001; Health Industry Forum, 2006; Kupersmith et al., 2005; Wilensky, 2005). There can be many variations on these themes, but the key concept is related less to the source of the funds invested than to the value of the return for the outcomes and efficiency of the nation's health care.

Because of the challenges to increasing CER through a simple appropriation to an existing agency—the difficulty of marshaling an appropriation at a sufficient level, the agency's lack of political independence, the limited ability to draw on other agencies—much of the recent discussion has focused on three of the independent models, often with blended public and private funding (Buto and Juhn, 2006; Kupersmith et al., 2005; Wilenksy, 2005). As independent entities, each of these approaches assumes the establishment of a governing board composed of stakeholders and charged with priority setting, broad budget allocation, and fiduciary responsibility for execution of the program of activities. These approaches differ in the degree of insulation between the stakeholder priority setting and the conduct of the scientific studies as well as in the ways the studies would be managed, the involvement of existing agencies, and the reporting of results.

To this end, the ACA (2010) established the Patient-Centered Outcomes Research Institute (PCORI) as an independent non-profit organization to

assist in informing the health decisions of "patients, clinicians, purchasers, [and] policy-makers." The ACA appropriated to the PCORI Trust Fund $10 million, $50 million, and $150 million for fiscal year 2010-2012. Additionally, $150 million plus $1 per Medicare part A and B enrollee has been appropriated for 2013 and $150 million plus $2 for each A/B enrollee, each year from 2014-2019. As outlined in the Act, PCORI will set a national agenda for research priorities, fund entities that conduct priority research, improve clinical effectiveness research methods, and ensure transparency and broad dissemination of its findings. It will be overseen by a Governing Board, comprised of 19 members appointed by head of the Government Accountability Office, as well as 2 ex officio representatives from the Agency for Healthcare Research and Quality and the National Institutes of Health. For more information on PCORI, see Appendix E.

A Vision for the Capacity to Learn Which Care Is Best

The growing support for CER represents an important first step in transforming health care. Mark B. McClellan, director, Engelberg Center for Health Care Reform, from the Brookings Institution emphasized that as the infrastructure required to expand the nation's capacity for CER is identified and prioritized, policy makers will need to consider how these elements can serve the longer-term goal of developing a learning health system. Efforts to improve the key infrastructure elements and data networks, methods, and workforce should also consider how to best build upon current health system capacities. Key advances needed for these elements include supporting a virtual approach to linking databases through the development of needed standards and incentives; advancing innovative approaches to clinical trials to facilitate their conduct in real-world settings, as well as improved statistical and epidemiologic methods; and a focus on developing a broad, cross-disciplinary workforce with capabilities in biostatistics, epidemiology, decision analysis, health economics, health services research, and program evaluation. To take best advantage of the many efforts already under way, infrastructure is also needed to promote the sharing and learning from the diverse experiences of all stakeholders. Public–private partnerships are one possible approach to helping organizations share information and learn more quickly about what works best.

McClellan also suggested that for infrastructure development, form should follow function, and he identified four critical evidence gaps that need to be addressed by CER and a learning health system. First, to move beyond evaluating the average impact of a treatment in a population and toward targeted medicine, researchers need a better understanding of current care and how this might vary from patient to patient. Large epidemiologic datasets will be useful to develop the disease models or natural

histories that provide such baselines for future evaluations. The development and use of these data resources have several implications for infrastructure, including the need to develop and implement complete standards for data collection, clinical trials, and electronic records.

Establishing the means to monitor the safety of medical therapies and products is another key evidence gap. Developing data networks and requisite methods of analysis will help to support the creation of a national, virtual infrastructure—as endorsed in the FDA Amendments Act of 2007—for monitoring product use, including adverse reactions. Such infrastructure may eventually serve as an important component of infrastructure for evidence generation—by supporting studies that compare the safety and effectiveness of treatments in different subgroups.

A third and related need is developing a reliable and relevant evidence base on the comparative effectiveness of treatment options to help physicians and patients make the best possible healthcare decisions. At present, conducting carefully randomized studies in real-world situations on practical treatment questions can be difficult as well as costly and time consuming. Moreover, by the time a large randomized trial is completed, the information may be outdated. The key challenge is to move beyond approaches that generate evidence about the overall average effect—in one population versus another—to the efficient development of information relevant to particular types of patients. In this respect, work is needed to determine limitations, methodological challenges, and needed improvements in data collection methods, as well as to develop agreement on the amount and type of evidence needed for decision-making purposes.

Finally, infrastructure development must aim to address the evidence gap related to understanding effective treatment strategies and policies. Current capacity cannot contend with the impending exponential growth in the complexity of medical decision making. Subtle differences in the management of chronic diseases and practice patterns affecting chronic disease management often result in broad variations in care delivered. In the absence of information on these technologies and strategies, care provided can be only marginally beneficial or even harmful. McClellan also noted that some suggest it should be possible to reduce costs in Medicare by 20 percent without consequences for patient outcomes if these variations are addressed. However, these practices often cannot be assessed in a simple RCT, but require study in real-world settings. Such studies could be very useful in closing the gap between what we know works and what is delivered in medical practice, as well as in understanding underlying issues related to the coordination and integration of care that constitute a major problem in the current healthcare system. Infrastructure needed to address this challenge should involve broad collaboration among stakeholders, as well as developing consensus on the best methodological approaches. A

particular focus is needed on methods development for studies in real-world settings, including observational approaches.

The Potential Returns from Evidence-Driven Health Care

Interest in the potential of comparative clinical effectiveness information to help Americans learn to "spend smarter" is part of a drive towards the increased availability and use of evidence to guide medical practice. According to Gail R. Wilensky, senior fellow at Project HOPE (Health Opportunities for People Everywhere), the potential ability for better information to improve health outcomes and also to help moderate spending increases is enormous. To capture some of the potential savings that CER could bring, consideration needs to be given to which approaches, data resources, and analyses will be most useful in producing the information needed. Data will be available from many sources, and it will be important to find ways to identify and address issues related to study design limitations and biases, as well as to reduce the costs and time required for the collection of new prospective data for comparative effectiveness trials and studies. Anticipated need for expanded CER might begin with an initial investment of several hundred million dollars and then ramp up to $4–$6 billion a year. Several steps are critical to ensuring a return on this investment. First, a center should be established and charged with creating better information on comparative clinical effectiveness. Second, priority setting for comparative effectiveness analyses should focus on medical conditions in high-cost, high-volume areas as well as areas that are subject to substantial practice variation. It will also be important to consider issues of clinical relevance, disease burden, and the various subgroups that are particularly affected. Third, it is essential to recognize that all stakeholders need to be a part of the decision-making process, as a slowdown in spending rates will have broad effects across the healthcare system.

Wilensky suggested that an immediate infrastructure priority should be to develop a common national data source that captures what is known about the likely clinical outcomes of various treatments for relevant population subgroups. Moreover, although cost- and clinical-effectiveness information should be considered in reimbursement and even clinical decisions, Wilensky underscored the importance—for technical and political reasons—of keeping these analyses and the places where they are conducted separate. Finally, as important as it is to have information available on clinical and cost effectiveness, the potential gains will not be achieved unless the reimbursement system is changed to make better use of information to reward health care of value rather than just paying more for doing more. In this respect, she emphasized the importance of expecting and allowing for different players to use this information differently, the need to make CER

information available to help guide reimbursement policies that reward good clinical outcomes, and the importance of taking advantage of the diversity in healthcare delivery by tying local coverage decisions to evidence development. Finally, it will be particularly important to provide legislative authority to CMS to introduce what is known about clinical and cost effectiveness into its reimbursement decisions.

The Work Required

CER aims to determine what works best for whom under which circumstances to inform the healthcare decisions of patients, physicians, and policy makers. Developing information that is relevant and understandable to these end users requires the efficient conduct of a range of activities, including primary research, synthesis, and translation. Efficient use of resources depends on prioritization of research questions, coordination of disparate but related efforts, and advancing research methods and data resources. The presentations described in Chapter 2 discussed the nature of the specific types of the work required, clarified what is known about the current capacity, illustrated the opportunities to improve care presented by expanded investment, and offered initial suggestions about policies or activities for progress.

Cost and Volume of Current Comparative Effectiveness Research

As policy discussions about CER gather momentum, there continues to be a lack of awareness of the current scale of CER. Erin Holve, senior manager at AcademyHealth, presented the results of a survey of the costs and volume of current CER. The study was based on stakeholder interviews with research funders and researchers as well as a review of databases that track health research. In this work, CER was defined as an examination of the effectiveness of the risk and benefits of two or more healthcare services or treatments used to treat a specific disease (e.g., pharmaceuticals, medical devices, medical procedures, other treatment modalities in appropriate real-world settings). Results from both sources suggest there are currently more than 600 studies under way in the area. Three primary research categories were considered: head-to-head trials (including pragmatic trials), observational studies (including registry studies, prospective cohort studies, and database studies), and syntheses and modeling (including systematic reviews). Data from the interviews demonstrate a wide range in the cost of conducting CER, both across and within study designs, although costs clustered by study design (Table S-3). Interviewees emphasized a need for multi-disciplinary training to expose researchers to a variety of methods in trials, observational studies, and syntheses. Finally, the study revealed an

TABLE S-3 Costs of Various Comparative Effectiveness Studies

Type of Study		Cost
Head to head	Randomized controlled trials:	
	Smaller	$2.5m–$5m
	Larger	$15m–$20m
Observational	Registry studies	$2m–$4m
	Large prospective cohort studies	$800k–$6m
	Small retrospective database studies	$100k–$250k
Synthesis	Simulation/modeling studies	$100k–$200k
	Systematic reviews	$200k–$350k

absence of clear definitions regarding the scope of comparative effectiveness as well as a limited understanding of the appropriate methods for conducting CER. These definitional and organizational issues may be an impediment to coordinating future CER activities.

Intervention Studies That Need to Be Conducted

To illustrate the multifaceted need for comparative effectiveness information on procedures, devices, pharmaceuticals, diagnostics, and health systems and to highlight some of high-priority studies that might be needed, Douglas B. Kamerow, chief scientist at RTI International, presented results from a stakeholder work group convened to pilot a process for identifying candidate comparative effectiveness studies. The process resulted in the adoption of selection criteria—including the importance of the conditions being treated or prevented, the current availability of effective treatments or preventive interventions, lack of definitive knowledge about the relative effectiveness of available treatments, research plausibility, and study type heterogeneity. These criteria were used to select among candidates nominated in the following comparative effectiveness categories: diagnostic studies drug–drug comparisons, health services systems studies, preventive interventions, surgical studies, and treatment studies across modalities. Potential comparative effectiveness studies identified using this process are

listed in Table S-4, and brief evidence reviews developed by Kamerow for each item are provided in Appendix B. Importantly, this exercise illuminated a number of challenges facing those who seek to prioritize the work needed. Key lessons learned included the existence of many opportunities for research that can make a difference in costs and outcomes; the importance of establishing an explicit and transparent process and of considering stakeholder perspectives and input during the process of nomination, review, and selection; and the need for carefully defined research questions and the utilization of the full range of study designs and methods. Many of these issues and perspectives were similarly reflected in the report on CER priorities mandated under ARRA and conducted by the IOM, the summary of which is found at Appendix C.

Clinical Data Sets That Need to Be Mined

Reflecting on the opportunities to better use routinely captured electronic data to generate insights on benefit and safety, Jesse A. Berlin, vice president of pharmacoepidemiology at Johnson & Johnson, noted that the appropriate use of existing data, as well as creative new uses of existing data collection mechanisms, will be crucial to improving healthcare decision making. Reviewing the major strengths and limitations of currently available administrative data in addressing questions of comparative effectiveness and safety, he noted that most existing databases currently used in observational studies of pharmaceuticals were created for a purpose other than research. These purposes include allowing payers to track expenditures in order to manage costs (in insurance claims databases, the largest and most prevalent type), manage patients (electronic medical records), and manage purchasing and capacity (facility-specific databases). These databases can provide relatively inexpensive and rapid access to clinical data and analyses of pharmaceutical exposures within a quantifiable source population, and these data reflect healthcare decisions and outcomes as they were actually made (versus the artificial constructs of an RCT). By virtue of reflecting actual, real-world clinical practice, databases offer an external validity that is greater than RCTs. However, databases can be limited by missing data (particularly on such confounders as smoking, height, weight, race, over-the-counter drugs, and alcohol consumption), failures to follow up (due to turnover in healthcare plans), and data quality issues (e.g., miscoding of diagnoses). Furthermore, data reflect the healthcare delivery system and, as such, may fail to account for or capture healthcare visits outside of providers using a specific electronic health record (EHR) system, due to benefit design or other issues. The use of these data for research is also complicated by the difficulty of capturing benefits of treatment (e.g., improvement in blood pressure, quality of life) compared to the relatively

TABLE S-4 The Comparative Effectiveness Studies Inventory Project Identified 16 Candidate Topics for Comparative Effectiveness Research

Study Topic	Study Type	Age Group	Condition
Treatment of attention deficit hyperactivity disorder in children: drugs, behavioral interventions, no prescription	Comparative effectiveness treatment studies across modalities	Children	Mental diseases
Treatment of acute thrombotic/embolic stroke: clot removal, reperfusion drugs	Comparative effectiveness treatment studies across modalities	Adults	Heart and vascular diseases
Treatment of chronic atrial fibrillation: drugs, catheter ablation, surgery	Comparative effectiveness treatment studies across modalities	Adults	Heart and vascular diseases
Treatment of chronic low back pain	Comparative effectiveness treatment studies across modalities	Adults	Neurological diseases
Gamma knife surgery for intracranial lesions vs. surgery and/or whole brain radiation	Comparative effectiveness treatment studies across modalities	Adults	Neurological diseases
Treatment of localized prostate cancer: watchful waiting, surgery, radiation, cryotherapy	Comparative effectiveness treatment studies across modalities	Adults	Cancer
Diagnosis and prognosis of breast cancer using genetic tests: human epidermal growth factor receptor 2 and others	Diagnostic studies	Adults	Cancer
Over-the-counter drug treatment of upper respiratory tract infections in children	Drug–drug and drug–placebo treatment studies	Children	Respiratory disorders
Drug treatment of depression in primary care	Drug–drug and drug–placebo treatment studies	Adults	Mental disorders

TABLE S-4 Continued

Study Topic	Study Type	Age Group	Condition
Drug treatment of epilepsy in children	Drug–drug and drug–placebo treatment studies	Children	Neurological diseases
Use of erythropoiesis-stimulating agents in the treatment of hematologic cancers	Drug–drug and drug–placebo treatment studies	Adults	Cancer
Outcomes of percutaneous coronary interventions in hospitals with and without onsite surgical backup	Health services/systems studies	Adults	Heart and vascular diseases
Screening hospital inpatients for methicillin-resistant *Staphylococcus aureus* infection	Preventive interventions	Adults	Infectious diseases
Tobacco cessation: nicotine replacement agents, oral medications, combinations	Preventive interventions	Adults	Preventive interventions
Prevention and treatment of pressure ulcers	Surgical studies	Adults	Dermatological diseases
Inguinal hernia repair: open vs. minimally invasive	Surgical studies	Adults	Surgical disorders

NOTE: Study topics are categorized by study type, age group, and condition.
SOURCE: Kamerow, 2009.

high capture of treatment risks that are often also codified as "clinical diagnoses." Regardless, little is understood of how much of an effect these benefits or risks have, or the perceptions of the patients or healthcare providers regarding the benefits and risks.

Designing targeted studies within databases is a promising direction for research. For example, special data collection screens might pop up on an in-office computer when patients matching a specific set of criteria were under consideration. This idea could be extended to include the conduct of large, simple, randomized studies within the databases. The question

is whether additional aspects of data collection can be tailored (as in primary data collection efforts) within the context of an existing data collection system. These ideas are not novel, but neither have they yet been widely adopted. Other ideas discussed include capturing the data at the physician–patient interface and providing data back to clinicians. Feedback to providers could encourage them to enroll in clinical trials and could permit healthcare professionals to better understand their own treatment decisions and the impacts of those decisions.

Knowledge Synthesis and Translation That Need to Be Applied

Currently the United States lacks a single reliable source that people can use to evaluate the safety and effectiveness of medical treatments. In January 2008, the IOM published *Knowing What Works in Health Care: A Roadmap for the Nation* (IOM, 2008b), which explored the national capacity to use scientific evidence to identify highly effective clinical services. Richard A. Justman, national medical director at UnitedHealthcare served on the report committee, and he discussed key findings related to the state of knowledge synthesis and translation and opportunities to scale up national capacity to meet the anticipated demand. While there are multiple avenues available today to help consumers, physicians, and others decide which treatments are safe and effective, they all have significant limitations. Justman highlighted how the absence of a national comparative effectiveness architecture has led to an evidence base that is replete with gaps, duplications, and contradictions (Table S-5). For example, some systematic reviews of clinical evidence and some clinical practice guidelines lack scientific rigor, relying on a consensus of expert opinion rather than clinical evidence as the basis for their conclusions. The body of clinical evidence for some health services that consumers and physicians are interested in may be weak or totally lacking. Bias and conflict of interest on the part of experts further complicate the understanding of the conclusions that can be drawn from available clinical evidence. Finally, the multiple clinical guidelines available for the treatment of the same condition frequently make differing recommendations. The 2008 report urged Congress to direct the Secretary of Health and Human Services to designate a single entity to ensure the production of credible, unbiased information about what is known and what is not known about clinical effectiveness. It also recommended the appointment of a Clinical Effectiveness Advisory Board to oversee the program and the appointment of a Priority Setting Advisory Committee to identify high-priority topics. The report further prescribed the development of evidence-based methodological standards for systematic reviews, including a common language for characterizing the strength of evidence. It recommended that bias be minimized by balancing competing interests,

publishing conflict-of-interest disclosures, and prohibiting voting by members with material conflicts.

Methods That Need to Be Developed

To contend with the growing scope and scale of clinical evidence needs, work is needed to improve and refine current research methods as well as to develop innovative new approaches to ensuring the development of efficient, timely, and relevant information for healthcare decision making. Although randomized clinical trials and meta-analyses of these trials provide the best evidence for use in comparative studies of the effectiveness of clinical interventions and care, it is impossible, difficult, unethical, or prohibitively expensive to randomize all relevant factors. Eugene H. Blackstone, head of clinical investigations at the Cleveland Clinic Heart and Vascular Institute, presented five foundational methodologies that will accelerate movement from the current siloed approach to evidence generation to an approach that enables predictive and personalized medicine (Figure S-1).

Re-engineering clinical trials will require addressing six main pitfalls associated with traditional RCTs: (1) complexity, (2) data capture, (3) generalizability, (4) equipoise, (5) appropriateness, and (6) funding. For the many instances in which even RCTs are not feasible or sufficient to meet information needs, methods for conducting approximate randomized trials using balancing strategies and real-world observational clinical data have become increasingly common—although a number of their important features remain to be explored and understood. Many trials focus on early outcomes or else introduce medicines or devices that bring additional complications. Thus methods for longitudinal surveillance and long-term outcomes analysis—e.g., birth-to-death, patient-centric health records populated with discrete values for variables—are also needed.

Among the most promising methodologies emerging is the semantic representation of data. The elements of this methodology include the storage of patient data as nodes and arcs (graphs) that can seamlessly link all types of data across current medical silos, from genomics to outcomes; a rich ontology of medicine that permits natural-language queries of complex patient data without the need to understand the underlying data structure; the assembly of this ontology and the assertions that make it useful; and intelligent agents to assist in the discovery of unsuspected relationships and unintended adverse outcomes. An immediate focus should be on supporting a worldwide effort to develop the comprehensive formal ontology of medicine needed to implement semantic databases and knowledge bases.

Methods are then needed to transform the results of trials, approximate trials, and automated discovery from static publications into dynamic, patient-specific ("personalized") medical decision support tools

TABLE S-5 Duplicated Efforts by Selected Health Plans and Technology Assessment Firms, 2006

	Health Plans				Technology Assessment Firms		
Type of Service	United-Healthcare	Kaiser Permanente	Aetna	WellPoint	Hayes, Inc.	Technology Evaluation Center	ECRI Institute
Screening							
Genetic testing to predict breast cancer recurrence	√	√	√	√	√	√	√
Proteomic testing for ovarian cancer	√		√	√	√		√
Virtual (computed tomography [CT]) colonoscopy	√	√	√	√	√	√	√
Disease management							
Ambulatory blood pressure monitoring	√	√	√	√	√	√	√
Intermittent intravenous insulin therapy	√	√		√	√		√
Diagnosis							
CT angiography for suspected coronary artery disease	√	√	√	√	√	√	√
Microvolt T-wave alternans	√	√	√	√	√	√	√
Wireless capsule endoscopy	√	√	√	√	√	√	√

Treatment							
Brachytherapy for various cancers: breast, ovarian, and prostate cancer and brain tumors	√	√	√	√	√	√	√
Dysfunctional uterine bleeding and fibroids	√	√	√	√	√	√	√
Fallopian tube occlusion for permanent contraception	√	√	√		√		√
Growth factor–mediated lumbar spinal fusion	√	√	√	√	√		
Intracoronary brachytherapy	√	√	√	√	√	√	√
Minimally invasive surgery for low back pain	√	√	√	√	√	√	√
Photodynamic therapy for Barrett's esophagus and esophageal cancer	√	√	√	√	√		√
Vagus nerve stimulation for intractable depression	√	√	√	√	√	√	√
Devices							
Artificial total disc replacement for lumbar and cervical spine	√	√	√	√	√	√	√
Cochlear implants	√	√	√	√	√		√
Total artificial heart	√	√	√	√	√		√
Total hip resurfacing arthroplasty	√	√	√	√	√	√	√

NOTE: Not all reviews are comprehensive assessments. Agency for Healthcare Research and Quality evidence-based practice centers have reviewed 5 of the 20 topics listed (ambulatory blood pressure monitoring, CT angiography, proteomic testing for ovarian cancer, spinal fusion for low back pain, and uterine fibroids). The Kaiser Permanente entries represent all Kaiser regions.
SOURCE: IOM, 2008a.

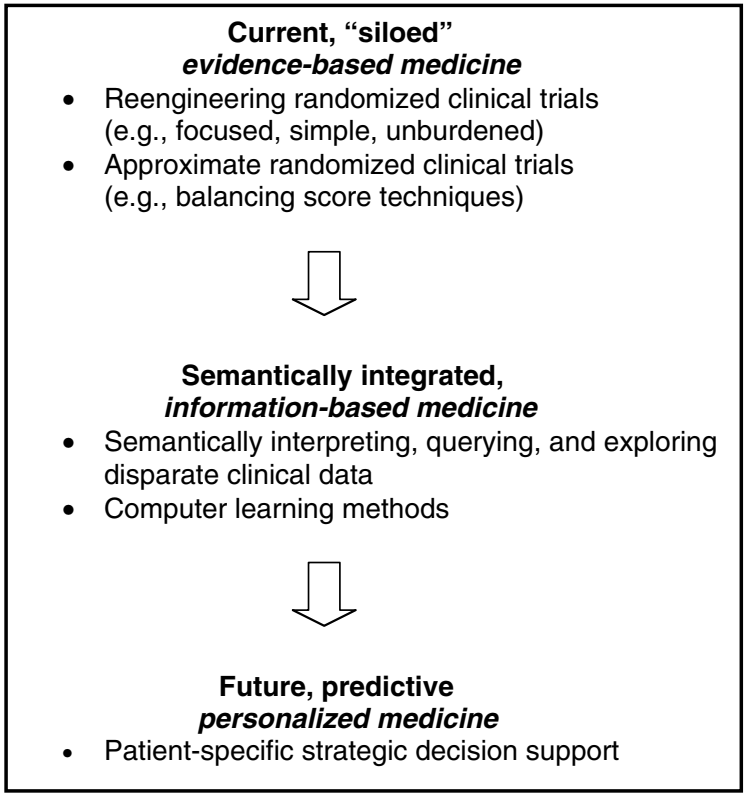


FIGURE S-1 Five foundational methodologies that need to be developed.

(simulation). Although such methodologies are widely used for institutional assessment and ranking, they need to become clinically rich and easily used real-time tools. The discrepancy between the "goodness of fit" of models to data and the minimization of prediction error needs to be addressed to enable accurate decision support. Algorithmic techniques, such as random forests–based methods, are intriguing and promise to fill the gap in accurately predicting a patient's response to treatment, but they are still in their infancy. Moving beyond randomized trials to the real world, exploiting emerging semantic technology in order to integrate currently disparate medical data, using the knowledge generated for strategic decision support, and developing the next generation of statistical tools for analysis of clinical data are but a few concrete examples of the methods that need to be developed to provide an infrastructure for learning which is the right treatment, for the right patient, at the right time.

Some preliminary estimates of the resources needed to spur methodology development are suggested.

Coordination and Technical Assistance That Need to Be Supported

AHRQ has played a leading role in promoting the evidence development, synthesis, and translation activities integral to CER. Jean R. Slutsky, director of AHRQ's Center for Outcomes and Evidence noted that CER as a concept and reality has grown rapidly over the past several years. Most of this work has built upon an appreciation for the role of technology assessment, comparative study designs, and HIT in the gathering and dissemination of best evidence to clinical practice; however, the development of the infrastructure needed for an expanded national capacity for CER has received less attention. To plan for such capacity rationally and strategically, one must have an understanding of the range of organizations currently conducting CER activities as well as some idea of which functions might benefit from either centralized or local approaches. Slutsky described lessons learned from AHRQ's work to support CER, outlined some practical realities of the current state of play, and suggested some priority areas in need of attention if the nation is to better meet the information needs of the diverse healthcare system.

Slutsky noted that priority CER needs include improvements aimed at supporting and training researchers; providing technical assistance in research design, conduct, and implementation; and developing capacities to prioritize, coordinate, fund, and engage stakeholders in CER activities. Training in research design and translation are particularly important to ensure that designs and protocols efficiently and effectively answer research questions and that findings are not used inappropriately and do not have unintended consequences. Because of the impact of CER on many different sectors (e.g., patients, industry, health plans), the research must be well designed and conducted in a fair and transparent manner, and receive adequate funding and support. Provisions in the ARRA and ACA hold promise in this respect. In addition, CER-focused public–private partnerships building on the work of other federal agencies (e.g., NIH, CMS, coverage with evidence development, and the Department of Veterans Affairs) are beginning to emerge to address some of these issues. As AHRQ discovered in developing the Section 1013 healthcare program, involving stakeholders early, listening to them, and involving them throughout the process are critically important.

The Information Networks Required

The scale of efficiencies gained through improvements to methods, coordination, and prioritization of CER will be limited by the available

capacity to capture, access, and share relevant data and information. Design and development of robust information networks and efforts to foster collaboration around common work are critical aspects of CER infrastructure. This capacity, too, has been addressed in recent legislation. The Health Information Technology for Economic and Clinical Health (HITECH) Act, enacted as part of ARRA in 2009, allocates $20 billion to be used as incentive payments to promote the adoption and "meaningful use" of electronic health records. Along these lines, presentations summarized in Chapter 3 discuss key current issues and needed capacity for networks to support the generation, synthesis, and the application of evidence, as well as for providing opportunities to support learning from clinical practice.

Information Technology Requirements

Robust, advanced clinical information systems (CIS)—including EHRs—are increasingly viewed as essential support for an evidence-based and learning health system. To provide policy makers with "order of magnitude" CIS estimates of the new or additional spending needed to speed broad adoption in care delivery organizations throughout the nation, Robert H. Miller, professor of health economics at the University of California at San Francisco described current EHR adoption, future EHR capital and operating expenditure requirements, and prospects for EHR adoption in the $648 billion hospital sector and the $447 billion physician and clinical services sector (spending levels as of 2006). EHR capabilities and estimated adoption level in hospital inpatient systems are indicated in Table S-6. Miller estimated that roughly $90 billion in new money may be needed over 8 years for robust hospital EHRs. Despite the 1.7 percent increase this represents for total hospital spending, adoption of EHRs will likely increase. There is substantial momentum in this sector, as health systems and larger hospitals increasingly see CIS as a cost of doing business, although public hospitals and unaffiliated hospitals with low or negative margins will likely lag behind.

Miller used rough estimates of the number of office-based physicians to develop an order of magnitude estimate of $40–$50 billion in new money for robust physician EHRs that may be needed over 8 years. This 1–1.25 percent average increase in physician services expenditure is feasible for most practices, and evidence suggests that the return on such an investment for physician practices could be substantial. Larger physician practices are adopting EHRs relatively rapidly, especially compared to solo/small groups (i.e., 10 physicians or fewer) and community health centers—for whom the business case is not perceived as favorable.

To achieve full EHR adoption, all types of healthcare delivery organizations need to increase CIS; however, more will be needed to achieve

TABLE S-6 Hospital Electronic Health Record Capabilities and Adoption Estimates

Stage	Description	2008
Stage 7	Medical record fully electronic; healthcare organization able to contribute continuity of care document as by-product of electronic medical record; data warehousing/mining	0.1%
Stage 6	Physician documentation (structured templates), full clinical decision support, full Radiology Picture Archiving and Communication System (PACS)	1.0%
Stage 5	Closed loop medication administration	1.3%
Stage 4	Computerized physician order entry, clinical decision support (clinical protocols)	1.9%
Stage 3	Clinical documentation (flow sheets), clinical decision support system (error checking), PACS available outside radiology	32.9%
Stage 2	Clinical data repository, controlled medical vocabulary, clinical decision support system inference engine, may have document imaging	33.2%
Stage 1	Ancillaries—lab, radiology, pharmacy	12.5%
Stage 0	All three ancillaries not installed	17.1%

SOURCE: HIMSS Analytics, Hospital IT Expenses and Budgets Related to Clinical Sophistication. Market Findings from HIMSS Analytics. (Chicago, IL: Health Information Management Systems Society, 2008).

effective EHR use for evidence-based medicine. Overall, CIS adoption will likely improve quality, but improvements are needed to EHR software, government and payer financial incentives, public performance reporting, EHR support services, and health information exchange.[3]

[3] The American Recovery and Reinvestment Act (ARRA) of 2009 provides a total of $19 billion to promote the adoption and use of HIT, particularly EHRs. The HIT-specific ARRA provisions provide $2 billion to the Office of the National Coordinator for HIT—charged with creating a strategic plan for a nationwide interoperably health information system—and allocate $17 billion for financial incentives, through Medicare and Medicaid reimbursements and to physicians and hospitals that become "meaningful users" of EHRs. The focus on "meaningful use" is a recognition of the need to not only adopt HIT but to employ HIT capabilities that improve health care through the "exchange and use of health information to best inform clinical decisions at the point of care" (HHS ONC. http://healthit.hhs.gov/portal/server.pt?open=512&objID=1325&parentname=CommunityPage&parentid=1&mode=2&in_hi_userid=10741&cached=true) [accessed September 10, 2009].

Data and Information Hub Requirements

The near-term aim to develop a comprehensive and expanded approach to CER is inherently challenging because of the lack of a controlled environment for assessing therapeutic options, the heterogeneity of patient characteristics, and the distributed nature of both the requests for and the sources of information. An evolution in the approaches to data and information hubs is needed to meet these challenges. Although large databases and clinical registries offer immediate opportunities for learning what works in health care, Carol C. Diamond, managing director of the healthcare program at the Markle Foundation, argued that the greatest promise of HIT lies in its ability to enable networked analysis—or the quick and efficient learning via a networked and distributed approach to information sharing and evidence creation. To maximize this potential, four key challenges must be addressed: (1) clearly defining the ultimate goal; (2) being open to reset definitions and assumptions about health data and research approaches; (3) articulation of new, broadly accepted working principles based on 21st-century information paradigms; and (4) developing an information policy framework that broadly addresses public hopes and concerns.

These challenges should be thought of as a jumping-off point for envisioning what is needed to move to a distributed approach to research—one characterized by connectivity, networks, and feedback loops. Rather than clinicians relying solely on large databases, centralized research centers, and analysis outside of healthcare delivery that can take months and years, Diamond presented a scenario in which clinicians access, in real time, current research and evidence syntheses as well as information provided via local networks on factors relevant to treating a particular patient (e.g., individual physician's patient outcomes versus his peers, community outbreaks, sensitivity patterns). Distributed analytic tools move research closer to practice by allowing clinicians and patients to quickly answer practical questions and to make better decisions.

This paradigm shift challenges assumptions about health data and research approaches. Diamond cited several examples to illustrate what might be achieved if research is conducted in networked environments, if information is provided when and where it is needed, and if clinicians, researchers, and patients connect the silos of clinical care and clinical research. Examples include childhood cancer networks that continuously use data to evaluate outcomes in order to improve protocols and treatments, an international and national flu surveillance network (Distribute), and the first real-world, open and nonblinded, patient-driven trial on the use of lithium for amyotrophic lateral sclerosis patients—research driven by the Web community PatientsLikeMe. As a starting point for developing

the information hubs needed, Diamond noted several examples of established distributed research models including the National Cancer Institute's Shared Pathology Information Network and Cancer Biomedical Informatics Grid and AHRQ's Distributed Research Network. Bringing distributed research networks to the scale needed to address national needs for CER will require attention to motives, standards, methods, and rules.

Integrative Vehicles Required for Evidence Review and Dissemination

The essential functions of any system dedicated to developing a robust evidence base for medical practice are synthesizing information derived from relevant trials and studies with insights emerging from clinical practice and ensuring that this information is continually updated. As clinical information systems are increasingly deployed, and as research increasingly draws upon connected and distributed data and information networks, the demand for synthesis work—to ensure studies are appropriately reviewed, vetted, and incorporated into the evolving evidence base—will also grow. Lorne A. Becker, co-chair of the Cochrane Collaboration steering group, provided an overview of the evidence synthesis development, coordination, and application within the United States and internationally and described opportunities to expand the nation's capacity to meet the anticipated demand.

Synthesis provides the necessary link between knowledge generation and its application to medical practice by identifying gaps, helping to set the research agenda, assessing the quality of individual studies, and collecting and appraising the data. Currently, evidence syntheses vary in the methods used, their complexity, and the reproducibility of results. Becker discussed large complex evidence syntheses that assess evidence over a broad domain as well as systematic reviews that have a more narrowly targeted focus. Clinical practice guidelines are also discussed.

Becker noted several advantages and disadvantages of these approaches with respect to time, cost, usefulness to decision makers, and robustness of information produced. These trade-offs suggest the need for a program that targets and supports the conduct of complex syntheses, as well as broader efforts to build a diffuse network of skilled producers to develop focused reviews. Currently, reviews in the United States are developed in a decentralized fashion involving both public and private entities. On a population basis, the United States contributes far fewer reviews than other nations (Figure S-2)—suggesting an opportunity for increased U.S. involvement in synthesis and dissemination activities as well as for gains through greater international coordination. Several notable international collaborations were noted including the Joanna Briggs Institute, the Campbell Collaboration, a consortium of health technology assessors (e.g., the

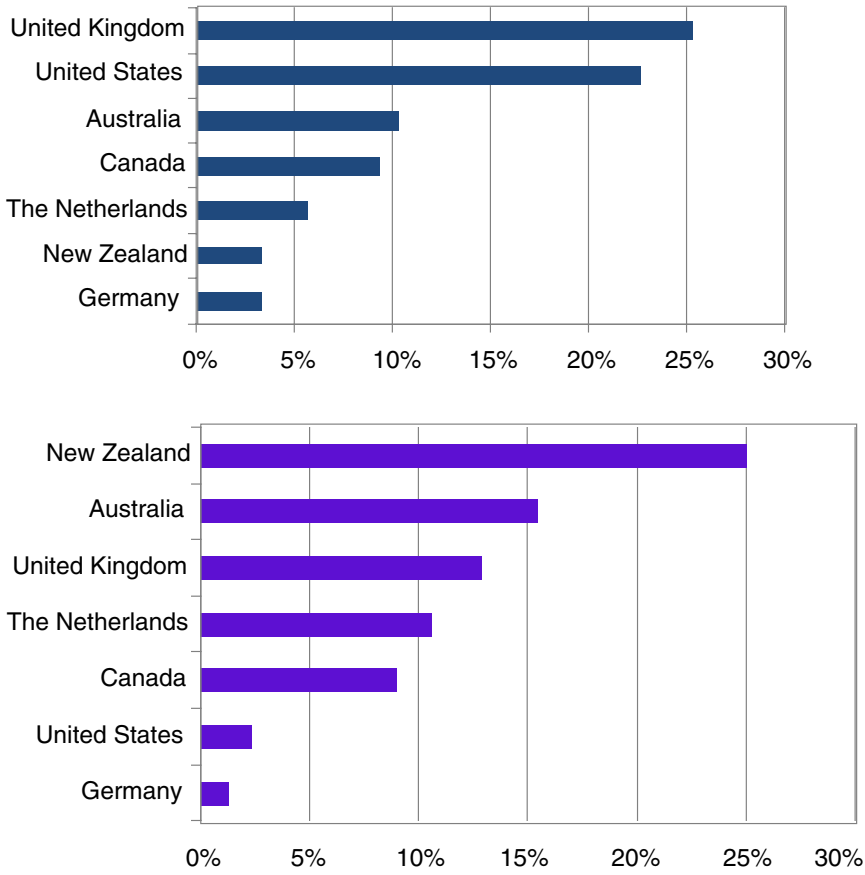


FIGURE S-2 Systematic review production by country and per capita (2004). SOURCE: Adapted from Moher et al., 2007.

European Network for Health Technology Assessment), and the Guidelines International Network.

Becker suggested that international efforts should focus on increasing efficiency of evidence syntheses through improved coordination—perhaps with the formation of a registry of systematic reviews—and fostering rapid progress in methods and standards development—for the conduct, reporting, and assessment of evidence syntheses and guidelines. Such work will benefit all healthcare decision makers, as well as accelerate the development of key infrastructure elements needed for expanded CER capacity in the United States.

The Talent Required

A comprehensive program to meet current information needs requires more than an expansion of existing programs and infrastructure. New structures, systems, and elements of HIT will need to be integrated into current practice to help improve research and practice. Furthermore, the capacity for prioritization, coordination, and conduct of CER will be increasingly interdisciplinary and involve many professions and healthcare sectors, thus requiring greater attention to human capital development. Chapter 4 provides two perspectives on what the discipline of CER and its associated workforce might look like.

Comparative Effectiveness Workforce—Framework and Assessment

The purpose-oriented nature of CER and the focus on informing practice and policy decisions suggest needed attention to how we train and develop the workforce required. Wlliam R. Hersh, professor and chair of the department of medical informatics and epidemiology at Oregon Health and Science University, and colleagues developed a framework for the CER workforce needed, made some preliminary estimates of the size of those needs, and proposed an agenda for further research. The heterogeneity of CER activities makes planning for its workforce needs challenging. Investigators and staff in CER come from many backgrounds—including clinical medicine, clinical epidemiology, biomedical informatics, biostatistics, and health policy—and work in a variety of settings, including academic units, university centers, contract research organizations, government, and industry. To simplify discussion, five key domains of CER activity were identified: (1) clinical epidemiology, (2) biomedical informatics, (3) health services research, (4) clinical guideline development and implementation, and (5) communications (Figure S-3). Several areas of significant overlap between domains were noted, including methods development and identifying information needs, suggesting required attention to interdisciplinary training and education. For each domain, the skill sets and competencies, training and education approaches, and issues related to expanding current capacity were reviewed.

The authors concluded that quantifying the needs of the overall workforce requires a better sense for the scale of expansion for the various CER activities (e.g., systematic reviews, trials, studies, guideline development, data mining). While the education and training of the current workforce can be applied to many aspects of CER, the workforce training and education needs for expanded capacity for CER will likely be substantial.

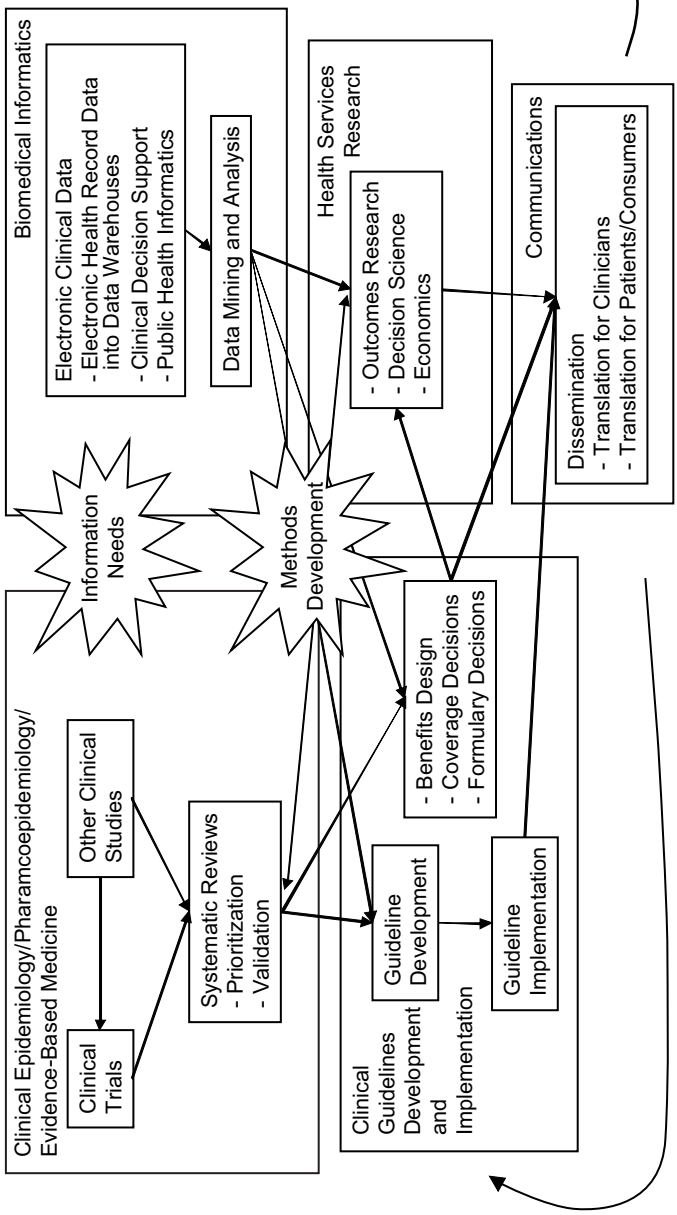


FIGURE S-3 Key activity domains for comparative effectiveness research. Workforce development will be critical to support the many primary functions within each of these domains as well as to foster the cross-domain interactions and activities identified (e.g., methods development, identifying information needs).

Toward an Integrated Enterprise—The Ontario, Canada, Case

An example of different workforce elements engaged in a system focused on developing and applying clinical effectiveness information was provided in an overview of an Ontario program aimed at ensuring that promising but unproven technologies are made available to patients for whom the risk–benefit ratio is favorable. Sean R. Tunis, director of the Center for Medical Technology and Policy, and colleagues described a system that allows purchasers (primarily hospitals) to request that a health technology be reviewed by the Ontario Health Technology Assessment Committee (OHTAC). If, after completion of this assessment there is insufficient information to recommend a coverage decision, OHTAC may request a "conditionally funded field evaluation." These studies, led by government-funded, independent research entities, are designed to address the evidence gaps necessary for policy makers to make coverage decisions. Funding this research costs approximately $8–$10 million per year (about $500,000 per field evaluation) and requires the support of Ministry of Health staff as well as hospital and university investigators with a wide variety of expertise (epidemiologists, biostatisticians, physicians, health economists, health policy experts, health services researchers, etc.). Tunis noted that the direct and explicit link between decision makers and the CER entities facilitates research timeliness and a clear focus on information satisfies the needs of decision makers and allows for evidence-based technology diffusion.

Although the U.S. healthcare system differs greatly from Ontario's in size, complexity, and design, Ontario's experiences illustrate that a significant amount of research can be achieved with very little spending if existing infrastructure is used wisely. Important lessons relevant to U.S. efforts to build CER capacity include establishing a stable funding source to support CER that, unlike the standard grant review cycle time, can fund rapidly evolving research needs; ensuring a timely process focused on the needs of decision makers to increase the likelihood that data generated by a study will be relevant; designing programs independent from government and industry, and ensuring a transparent decision-making process; increasing the efficiency and effectiveness of research by creating partnership between universities and those conducting field evaluations; and leveraging Medicare's influence on private payers to more broadly support coverage with evidence development. In addition, this analysis presents an opportunity to consider potential collaborative activities, such as international patient registries or standards of study design, which may help to globalize CER in the future.

Priorities for Implementation

Workshop discussion shaped an ambitious vision for the potential gains—in the efficiency, effectiveness, and value of health care delivered in the United States—that might be realized with a greater focus on and expanded capacity for clinical CER. Realizing this vision requires strategic and implementation priorities for the near and long term. Chapter 5 features a discussion of IT platforms, data resource and analysis improvements, the clinical research infrastructure, health professions training, and training capacity needs. Each paper summarizes suggestions presented and discussed at the workshop on staging and policies, key needs in the relevant area, and possible approaches to ramping up. Also discussed are opportunities to take advantage of existing manufacturer, insurer, and public capacities through public–private partnership mechanisms.

Information Technology Platform Requirements

A reformulation of approaches to information systems will be essential to better capture and apply clinical data important to advance care and the evidentiary basis for practice. Mark E. Frisse, professor of biomedical informatics at Vanderbilt University, noted a needed shift in informatics approaches to representing data, developing comprehensive systems, and integrating these systems into decision-making settings. IT infrastructure must be systemic, sustainable, relevant, and incremental in design and benefit; it must also be based both on principles that engender public trust and on explicit policies for use. Drawing upon his experiences with a Tennessee regional health information exchange (the Memphis Exchange), Frisse offered suggestions on IT platform requirements and approaches that will help realize significant societal benefit at a realistic marginal cost.

With proper design and integration, the current collection of databases, health record systems, health information exchanges, financing, workforce, policies, and governance can evolve into a system that can address a range of needs. Too often, the design of systems emphasizes administrative transactions and episodic care at the expense of recording data that can be used to drive care, process improvement, and promote research. A clear framework—informed by effectiveness, quality, safety, and efficiency outcomes—is needed to prioritize and support a wide range of scientific, clinical, and policy aims. The Memphis Exchange demonstrated that trust and policy—rather than technology—are the primary barriers to an integrated IT platform; that approaches can be developed to contend with issues related to combining data from disparate sources, identifying and matching data, sharing data, and protecting confidentiality and privacy; and that loosely coupled data sets from disparate resources are amenable to supporting a wide range of

research efforts. When tied to deidentification processes, these data could serve as a powerful resource for biosurveillance, public health research, quality improvement, and comparative effectiveness studies.

Information exchanges however, are just one part of a larger HIT platform. The choice and effectiveness of care delivery technologies (e.g., EHRs) is critical. If properly designed and implemented, an interconnected system will return substantial benefits at marginal costs. National investment decisions that could simplify the integration of data across disparate systems include an immediate acceleration of knowledge representations that could be applied to clinical use quickly (e.g., RxNorm, Unified Medical Language System); decisions over the extent to which payment and administration coding standards can reflect disease states and contexts required of learning health systems (e.g., International Classification of Diseases [ICD-9], Systematized Nomenclature of Medicine, ICD-10); enforcement of a few, selective standards (e.g., Logical Observation Identifiers Names and Codes, SCRIPT); promotion of efforts that make laboratory and medication history more portable in a secure and affordable way; and selection of a few, simple quality initiatives that can guide improvement of any interventions enabled by IT. Broad adoption without coupling technologies to system improvements will not produce optimal outcomes. Trials are needed to test different approaches and ensure that IT expenditures are made wisely.

Data Resource Development and Analysis Improvement

Data resources and analysis can be used to guide clinical decisions; yet, despite the many potential information resources (e.g., product developers, federal agencies, payers, practitioners, providers), a cogent framework for selecting and using these resources to ensure that care delivery is centered on patient needs is lacking. Moreover, the absence of clarity on how and for what purpose such information would be used impedes progress. Compounding these challenges is the wide variation in information resources available across provider groups. T. Bruce Ferguson from the East Carolina Heart Institute discussed the robust, procedure-focused clinical databases in the field of cardiology that have independently validated processes and outcomes linked to quality improvement. To remain effective in a practice-based, learning health system, these vertical, procedure-based clinical data will need the longitudinal, medical-condition context important for the development of quality comparative effectiveness information. Resource and analysis development work must translate into a dynamic, real-time learning infrastructure, including built-in feedback processes and a focus on the patient and the point of care.

Data resources are currently incomplete both on the patient and provider level and are inadequate for learning about key gaps. The misalign-

ment between resources and their use leads to conflicting and erroneous data and interpretation. Movement has been slow at three important levels of data resources: integrated health systems, electronic medical records, and national-level resources from providers and payers. A priority for data resource development is the accurate definition of data (both the information and context elements) to ensure appropriate use. Key needs include defining the type, source, and use of data; operationalizing data collection; defining and making change to interoperation dynamics; and standardizing data use for comparative effectiveness. Advances at the national and local levels will be important, and the National Consortium of Clinical Databases's effort to integrate databases from the Society of Thoracic Surgeons, the American College of Cardiology, and the American Heart Association is particularly promising. A better alignment of incentives for comparative effectiveness—both informational and financial—is needed, as is a better definition of the opportunity and value of clinical and research data for use in CER. Ferguson noted opportunities to link data collection with efforts to improve performance measurement and reporting at the national level and, also at the local level, to enhance provider-level evaluation for quality improvement, benchmarking, and profiling.

Although robust methods for data analysis exist, they are often limited to specific databases, are procedure—rather than outcome—based, and have limited applicability for point-of-care use. Analyses are also retrospective in scope and require expensive infrastructure, and findings are often difficult to implement. Critical to improvement is the notion that analysis must be embedded into the data context infrastructure. New, patient-centric comparative effectiveness analyses are needed, as are approaches that account for multiple procedural options or assess risk over the duration of the medical condition. Tools for clinical point-of-care application of comparative effectiveness analyses and new analytic tools for CER will affect quality, effectiveness, appropriateness, and efficiency of care.

Practical Challenges and Infrastructure Priorities for Comparative Effectiveness Research

The process of developing and completing study protocols must be efficient in order for CER to reach its full potential for improving medical care. Daniel E. Ford, director of the Johns Hopkins' Institute for Clinical and Translational Research, reflected on key barriers to efficient clinical research and outlined research infrastructure needs for improving the quality and timeliness of research. He noted that CER needs to become a common occurrence in the delivery of care in the United States, and developing the needed infrastructure will require ample and long-term support. Moreover, efficient and valuable research requires the support of multiple

stakeholders, including patients, healthcare providers, healthcare plans, and the research community. Limited participation by just one of these stakeholders can impede study progress and reduce the value of the overall investment.

To optimize the quality and value of CER, Ford suggested six priorities for infrastructure development: (1) establishing a process for timely consultation from all stakeholders—ideally, standing panels; (2) accelerating study initiation through streamlining both institutional review board and contracting mechanisms; (3) developing a standard policy on insurers' coverage of services for individuals in clinical trials; (4) enhancing capabilities to conduct research in hospitals and practices outside the academic center—perhaps building on existing practice-based research networks; (5) developing stronger partnerships between researchers and healthcare systems; and (6) developing the workforce needed for CER teams—in particular, IT professionals, database developers and analysts, and biostatisticians who are expert in the analysis of cluster-randomized designs and sophisticated observational study methodologies.

Transforming Health Professions Education

Health care is moving toward a patient-centered, evidence-based health management orientation. Computerization of health records, wider use of patient care registries, greater availability of tools that allow for tracking individuals as well as populations of patients, and information-savvy consumers will drive our current fragmented health system toward one that will emphasize greater accountability, transparency of information, and higher levels of performance. Benjamin K. Chu, Southern California regional president for Kaiser Foundation Health Plan and Hospitals, described how these changes will shape the future practice environment and suggested that health professions education should take place in environments that emulate current models of best care. Such an approach would encourage the effective use of new tools by teams of health professionals as well as use and improvement of approaches that achieve the best outcomes across the full continuum of care and patient needs.

The delivery system infrastructure needed to support the best performance of health professionals depends on clear expectations for high performance along defined and measurable dimensions of care, adoption of appropriate IT tools that provide essential information to drive performance improvement, and payment systems that value better outcomes. Key opportunities for progress include computer-assisted tools with sophisticated evidence-based decision-support protocols combined with process changes and strict adherence to bundles of care, the ability to track gaps in preventive care and chronic disease management, and payment reform that

emphasizes bundled payments for episodes of care and evidence-informed case rates, or capitation.

To correct gaps in care and to ensure safe and effective interventions, health professionals will increasingly have to work together in teams and share accountability for their patients' clinical outcomes. Acute episodes of illness will require coordination of handoffs, patient safety protocols, and checklists and other interventions designed to minimize harm and to maximize benefit to patients. Chronic disease management and adherence to preventive measures that are known to be effective will become system-wide accountability requirements. The complexity of care and the huge burden placed on shorter physician–patient interactions with a multitude of different clinicians will require that other health professionals as well as ancillary staff be used to bridge the gaps. Every touch point, enhanced with Web-based and other communication-based tools, will be an opportunity to maximize care. A new professionalism will build on the principles of lifelong learning, duty to patients, and devotion to finding best outcomes as well as emphasize teamwork and evidence-based care. Computerized simulation training will become a staple for health profession education. Team skills—the ability to lead, develop, and encourage the active contribution of other professionals in the clinical setting—will become an essential core of professionalism. Demonstrated competency both in clinical arenas and in the ability to work effectively with others will be required.

Building the Training Capacity for a Healthcare Workforce of the Future

Research holds the promise of finding and testing the answers to the challenges that face U.S. health care, but traditional approaches are inadequate. Steven A. Wartman, president of the Association of Academic Health Centers, called for the development of a new kind of research infrastructure focused on health and health care that can guide and inform decision making. Such an approach would support research to discover, disseminate, and optimize the adoption of practices that advance the health of individuals and the public as a whole. In his discussion, Wartman suggested that the key to the changes needed is expanding the continuum of medical research to ensure that discoveries ultimately serve the public. This expansion would include all aspects of health, including biomedical, public health, and multidisciplinary research on the social and environmental determinants of health. Table S-7 outlines an approach to achieving this new research vision, and among the most pressing needs is the development of a new cadre of researchers, clinicians, and health leaders—a workforce that includes, among others, health professionals, engineers, sociologists, urban planners, policy experts, and economists. The cross-cutting nature of academic health centers (AHCs) suggests an unprecedented opportunity to

TABLE S-7 An Approach to Achieving a New Vision for Health Research

New People and Skills	• Multidisciplinary teams • Strategic faculty recruitment • Expansion and training of research support staff • New partners (e.g., industry, nongovernmental organizations, faith-based organizations, payers, government, public, organizations diverse communities, patients, general public) • New venues (e.g., community-based research) • Training to provide new skills, including inter-professional training • Incentives within academia to support all types of health researchers (e.g., academic home, revised promotion, tenure criteria)
New Infrastructure	• Information technology investments (e.g., electronic health records, personal health record, regional health information organizations) • Biostatistics and data management support • Biorepositories • Streamlined clinical research approval processes • Efficient intellectual property policies • Links between academia, industry, and venture capitalists
New Investments and Incentives	• Expanded funding for clinical, translational, and social health research by the National Institutes of Health, National Science Foundation, foundations, others • Identification of new funding sources, especially for T2 and T3, behavioral, public health, and social health research • Increased organizational investment in translational research cores (e.g., informatics, clinical research nurses) • National coordination of research resources (e.g., informatics linkages, data sharing)

SOURCE: Wartman and Pomeroy, 2009.

build AHCs to foster interprofessional collaborative activity and to develop needed health research teams. These teams may reside in new departments, institutes, and centers as typical academic silos give way to more horizontal integration.

Organizational and management trends taking place in the nation's AHCs are remolding the ivory tower into a complex business enterprise. This transition is characterized by reorganization along nondisciplinary lines and a management structure that, conceptually and operationally, better aligns the entire institution. To build the needed training capacity, AHCs will need to ensure commitment of their own leaders to expand "health research," invest in new infrastructure (e.g., IT, data repositories,

biorepositories), and support curricular and training innovations to develop multidisciplinary, multisector research teams. In addition, AHC leadership can drive this new vision of health research by calling for adequate and innovative funding mechanisms, providing needed culture and infrastructure, and facilitating the partnerships with government, industry, and community groups that are needed for health research. It will be necessary to provide clear-cut career paths for health researchers along with adequate and appropriate institutional resources. Key opportunities include the provision of academic homes for translational researchers, the development of appropriate recruitment packages, and criteria for promotion and tenure.

Many healthcare sectors—industry, community, and other nonacademic organizations—have important roles in facilitating fundamental change in medical research. The involvement of community constituencies affected by research will increasingly be an essential component of health research—through contributing input into research priorities, helping build trust of community participation in research, or disseminating findings. Particularly critical are national policy makers who will drive this transformation by endorsing the importance of health research in leveraging biomedical discoveries for health improvements; by providing adequate funding for the full range of health research needed, including workforce development; and by helping to address current barriers to research (e.g., Health Insurance Portability and Accountability Act procedures).

Public–Private Partnerships

A fundamental challenge in advancing CER is developing an infrastructure that is sufficiently robust to support and nurture productive relationships among stakeholders with different perspectives and organizational missions. Without a mechanism for bringing these parties to the same table, fundamental differences in institutional cultures can impede or even preclude stakeholder-to-stakeholder communication. Public–private partnerships can bridge these gaps and remove barriers to cooperation. This mechanism not only creates space for collaboration—in which barriers to cooperation can be discussed and addressed—but also offers a structure and operational guidelines, typically tailored to a specific partnership by the participants, that help facilitate cooperative work. A value for participating entities is they can learn more and distribute new knowledge more quickly in a collaborative environment. Public–private partnerships can help link some of health care's disparate component elements and draw productively on the respective assets of participating stakeholders. Public–private partnerships are viewed by some as fundamental building blocks in the development of the CER infrastructure. A panel discussion featuring perspectives of health plans, the federal government, and industry

representatives considered current and planned public–private partnership efforts as well as how these efforts can be used in a more expansive fashion to develop infrastructure for CER.

Carmella A. Bocchino, vice president for clinical affairs and strategic planning at America's Health Insurance Plans, discussed several successful public–private partnerships in which health plans and federal agencies have partnered to create databases that are useful in identifying potential safety issues and opportunities to improve care and care delivery. An extension of these activities could contribute to the development of a national data system to serve as a central part of the nation's health research infrastructure. The United States Renal Data System, a large national data registry for end-stage renal disease patients, offers a potential model for a more comprehensive national data registry. Research and surveillance networks, such as the HMO Research Network, the HMO Cancer Research Network, and the Vaccine Safety Datalink, demonstrate the potential of distributed data networks to help address national research and public health questions. Similar models, such the National Data Aggregation Initiative (NDAI), are being explored for quality measurement and reporting. NDAI seeks to combine Medicare and private-sector data to generate physician performance measures. While these initiatives demonstrate the inherent value of developing the infrastructure and tools to aggregate and analyze these data across populations, challenges remain. Agreement is needed on a shared methodology that can facilitate comparative analyses across the broad spectrum of current clinical research. Data systems design should facilitate data mining as well as the identification and tracking of safety and effectiveness issues in real time. Progress will require the standardization and compilation of data from disparate sources as well as ensuring thoughtful and appropriate design of emerging data sources, such as EHRs, so that data is produced that can help answer questions important to understanding clinical effectiveness. Establishing governance structures will also be a key challenge, as will developing approaches for sustainable funding of these types of research and contending with issues related to ownership of data.

Rachel Behrman, associate commissioner for clinical programs and director of the office of critical path programs at the FDA, summarized two public–private partnerships housed in the FDA: the Critical Path Initiative and the Sentinel Network. The Critical Path Initiative seeks to modernize the way in which FDA-regulated products, including drugs, biological products, and medical devices, are developed, evaluated, and manufactured. The Sentinel Initiative is intended to establish a national integrated electronic structure and approach for monitoring medical product safety. These initiatives have focused on several key issues that require collaborative engagement, including research methods and data analysis tools that ensure the production of timely, reliable, and secure information, as well as

governance structures and policies that meet stakeholder needs while also putting appropriate safeguards into place. Specifically, questions related to data access, use, and stewardship need to be resolved. With respect to a CER infrastructure, attention should initially focus on developing mechanisms for priority setting, sustainable financing, and collaboration governance as well as on data transparency so that conduct and reporting of analyses result in high-quality information. Contending with issues related to proprietary data and patentable tools and processes will be essential to progress.

William Z. Potter discussed two public–private partnerships, the Biomarkers Consortium and the Alzheimer's Disease Neuroimaging Initiative (ADNI), that have productively linked pharmaceutical companies, government agencies, and other stakeholders. The Biomarkers Consortium, which aims to speed up the development of biological markers in support of drug development, preventive medicine, and medical diagnostics, demonstrates the need for careful delineation of specific areas of research focus that protect the individual interests of consortium members. Areas of collaboration were carefully selected, and research was conducted in precompetitive spaces to ensure that the work would achieve the common goals of advancing human health and improving patient care; speeding the development of medicine and therapies for detection, prevention, diagnosis, and treatment of disease; and making project results broadly available to the entire research community. The Biomarkers Consortium had to address issues related to data quality, study design variation, and data sharing—and a project on placebo response was described to illustrate how such work can inform discussions on needed improvements. The ADNI demonstrates that infrastructure can be developed to foster cross-sector communication and work. Underlying this project's initial, promising results is ADNI's ability to adequately address data transparency issues. Key barriers identified relevant to the CER infrastructure included the need for internal industry champions to drive collaborative work; the need to meet the costs of full-time equivalent and data management; skepticism by industry, NIH, and academic leadership on the value of such partnerships; and variable legal opinions on intellectual property and medicolegal risks.

Moving Forward

Although expanding CER capacity offers many potential gains for health care, the scale of needed transformation is also large and spans all healthcare sectors. A long-term strategy must appropriately incorporate existing infrastructure, prioritize and sequence needs, engage all stakeholders, and build sustained, cross-sector support. Discussed in the final workshop session were key considerations for such a strategy—roadmap

elements, quick hits, and opportunities to build support. The final chapter includes a synthesis of this session's discussion, a review of common themes heard at the workshop, and a number of possible follow-up actions to be considered for ongoing multistakeholder involvement through the IOM Roundtable on Value & Science-Driven Health Care.

The Roadmap—Policies, Priorities, Strategies, and Sequencing

Stuart Guterman, senior program director for the Commonwealth Fund's Program on Medicare's Future, outlined six broad areas discussed during the workshop that should be considered in the development of policies and strategies: data, methods, workforce, organization, translation, and financing. Clear end goals for each area, priority needs within and between categories, and key actors or existing infrastructure that could help initiate the activities needed were discussed. Suggested goals for these areas included the development of capacity to produce relevant data, ensuring maximal value of data through integration and system linkages and making data and information available to appropriate users when and where needed; development of research approaches to meet the needs of CER end users; education of a cadre of professionals—from across healthcare sectors—trained to use tools and techniques for developing and applying comparative effectiveness information; prioritization and coordination across the many organizations engaged in various aspects of evidence development—primary research, synthesis, translation—to enable more efficient information production; movement from evidence to evidence-based decision making; and sufficient and sustained funding to establish and support CER and its application as an integral part of the U.S. healthcare system.

Quick Hits—Things That Can Be Done Now

Actions that can be undertaken immediately will be essential to help accelerate progress by demonstrating in the near term the benefit of expanded CER. W. David Helms, president and CEO of AcademyHealth, noted several opportunities for collaborative efforts by stakeholders to lay the groundwork for a national capacity for CER—advocating for congressional action to establish a platform for CER, increasing federal funding for CER, articulating the case for CER, examining models for an expanded national capacity, and educating state policy representatives and Medicaid officials about the potential and needs for CER. He noted that work can also begin immediately to build up the needed workforce. The many other recommendations for immediate action offered by session respondents and throughout the workshop were also summarized. Subsequent to this meeting, Congress increased the national capacity for CER with the estab-

lishment, in the ACA of 2010, of Patient-Centered Outcomes Research Institute, previously described.

Building Support

While building upon many existing activities and infrastructure, an enhanced focus on CER is a shift in the nation's approach to clinical research and practice. Although viewed as an important element of health reform by most healthcare stakeholders, additional work is needed to build support by the public and policy makers for needed investments and potential returns from CER. An open discussion session on this topic was led by Mary Woolley, from Research!America, who noted four fundamental requirements for building support: (1) having clarity on the ultimate goal, (2) understanding the target audience, (3) ensuring all stakeholders are involved, and (4) understanding the context. This framework suggests several key opportunities to build support for the expanded development and use of CER, including finding ways to frame the many infrastructure needs in simple terms that make sense to all stakeholders, including the public and policy makers; tailoring communications to the interests and concerns of different stakeholders; and engaging in clear communication and crisp, well-tested messaging. Finally, she noted that communication should not be unidirectional, but structured to fully engage all stakeholders involved in infrastructure building. Suggestions offered by workshop participants for possible goals, for opportunities to better engage consumers and patients, and for research that might better inform communications were summarized.

Issues for Possible Roundtable Follow-Up

Throughout the course of discussions, a number of items were identified as candidates for follow-up attention by the Roundtable on Value & Science-Driven Health Care:

- ***Better characterization of the elements of the infrastructure:*** Build on the work sponsored by the Roundtable on workforce needs and IT infrastructure, continue to improve the initial estimates and pursue similar assessments related to requirements for new analytic tools and methods, establish processes for efficient and effective operation of the fields of work, and shape the strategy for attention and phasing. Include examples of effective work at institutional level.
- ***Clarification of the nature of the "prework" needed for a more systematic approach to the necessary RCTs:*** Even though a more

practical portfolio of research approaches is essential, the RCT offers the key standard for the rigor required for certain circumstances. Their most effective deployment requires attention to issues of the criteria indicating the need for an RCT, the issues and priorities to be assessed, the best structure of the research questions, and improved approaches to trial design, conduct, and data collection.

- ***More focus on the infrastructure needed for guideline development, implementation, and evaluation:*** Several issues could be productively engaged, including transparency and collaboration across professional groups on improving consistency in the methods, standards, rules, and participants in guideline development and approaches to implementation.
- ***Share meeting discussions with organizational stakeholders in elements of the infrastructure:*** Examples given included the National Quality Forum; the Association of American Medical Colleges; the Association of Academic Health Centers, the Quality Improvement Program, and CMS/Department of Health and Human Services in the context of development of the 10th quality improvement organization statement of work; the American Hospital Association Quality Forum; the International Society for Pharmacoeconomics and Outcomes Research; and provider groups.
- ***Devote additional attention to data stewardship issues:*** Because the basic resource for effectiveness research is the clinical data system, the Roundtable needs to catalyze more discussion on the integrity of this resource, including issues of maintenance, privacy, and data ownership.
- ***Identify possible incentives:*** Look at how subsidies and reimbursement regulations can stimulate increased use of HIT in medical care, increased use of HIT for application of evidence, and increased use of HIT for the development of evidence.
- ***Expand engagement of the business case and demand function for infrastructure investment:*** Give additional attention to the economic or business case for employers to appreciate the investment and its necessity to improve value from health care, the case for more attention by states, the case for deployment of the personal health record to drive more patient–provider interaction, and work on the consequences of not investing.
- ***More focus on the issues of strategies and infrastructure for implementing findings on effectiveness:*** Since evidence is virtually useless if not applied, the Roundtable could give more attention to understanding the infrastructure needs for effective guideline implementation.

- ***Sponsor discussions on training and health professions education reorientation:*** With greater appreciation for team-based, networked information stewardship roles by caregivers, the health professions groups should be recruited for collaborative consideration of the training implications.
- ***Provide information on the Roundtable's Web site:*** The resources of the workshop presentations and discussions should be posted on the Web site—slides, links, and speaker contact information.

REFERENCES

AcademyHealth. 2005. Placement, Coordination, and Funding of Health Services Research within the Federal Government. In *AcademyHealth Report*, September 2005. http://www.academyhealth.org/files/publications/placementreport.pdf (accessed September 3, 2010).

Buto, K., and P. Juhn. 2006. Can a center for comparative effectiveness information succeed? Perspectives from a health care company. *Health Affairs* 25(6):w586-w588.

Clancy, C. M. 2006. Getting to "smart" health care. *Health Affairs* 25(6):w589-w592.

Fisher, E. S., and J. E Wennberg. 2003. Health care quality, geographic variations, and the challenge of supply-sensitive care. *Perspectives in Biology and Medicine* 46(1):69-79.

Fisher, E. S., J.E. Wennberg, T. A. Stukel, D.J . Gottlieb, F. L. Lucas, and E. L. Pinder. 2003a. The implications of regional variations in Medicare spending. Part 1: The content, quality, and accessibility of care. *Annals of Internal Medicine* 138(4):273.

Fisher, E. S., J. E. Wennberg, T. A. Stukel, D. J. Gottlieb, F. L. Lucas, and E. L. Pinder. 2003b. The implications of regional variations in Medicare spending. Part 2: Health outcomes and satisfaction with care. *Annals of Internal Medicine* 138(4):288.

Health Industry Forum. 2006. *Comparative Effectiveness Forum: Executive summary*. http://healthforum.brandeis.edu/meetings/materials/2006-30-Nov./ExecBrief.pdf (accessed July 20, 2010).

Hopayian, K. 2001. The need for caution in interpreting high quality systematic reviews. *British Medical Journal* 323(7314):681-684.

IOM (Institute of Medicine). 2000. *To err is human: Building a safer health system*. Washington, DC: National Academy Press.

———. 2001. *Crossing the quality chasm: A new health system for the 21st century*. Washington, DC: National Academy Press.

———. 2007. *Learning what works best: The nation's need for evidence on comparative effectiveness in health care*. Washington, DC: The National Academies Press.

———. 2008a. *Knowing what works in health care: A roadmap for the nation*. Washington, DC: The National Academies Press.

———. 2008b. *Learning what works: Infrastructure required for comparative effectiveness research*. Washington, DC: The National Academies Press.

Kamerow, D. 2009. *Comparative effectiveness studies inventory project*. Washington, D.C.: A commissioned activity for the IOM Roundtable on Value & Science-Driven Health Care.

Kupersmith, J., S. Sung, M. Genel, H. Slavkin, R. Califf, R. Bonow, L. Sherwood, N. Reame, V. Catanese, C. Baase, J. Feussner, A. Dobs, H. Tilson, and E. A. Reece. 2005. Creating a new structure for research on health care effectiveness. *Journal of Investigative Medicine* 53(2):67-72.

McGlynn, E. A., S. M. Asch, J. Adams, J. Keesey, J. Hicks, A. DeCristofaro, and E. A. Kerr. 2003. The quality of health care delivered to adults in the United States. *New England Journal of Medicine* 348(26):2635-2645.

Moher, D., J. Tetzlaff, A. C. Tricco, M. Sampson, and D. G. Altman. 2007. Epidemiology and reporting characteristics of systematic reviews. *PLoS Medicine* 4(3):e78.

Rowe, J. W., D. A. Cortese, and J. M. McGinnis. 2006. The emerging context for advances in comparative effectiveness assessment. *Health Affairs* 25(6):w593-w595.

Wartman, S., and C. Pomeroy. 2009. Building the training capacity: Implementation priorities In *Learning what works: Infrastructure required for comparative effectiveness research*. Washington, DC: The National Academies Press.

Wennberg, J. E., E. S. Fishers, and J. S. Skinner. 2002. Geography and the debate over Medicare reform. *Health Affairs* Web Exclusive:w96-w114.

Wilensky, G. 2005. Developing a center for comparative effectiveness information. *Health Affairs* Web Exclusive:w572-w585.

1

The Need and Potential Returns for Comparative Effectiveness Research

INTRODUCTION

Understanding the comparative effectiveness of candidate interventions for similar conditions is essential for improving the development and delivery of effective health care. Many publications have highlighted the shortfalls of health care delivered in the United States, including the frequency of medical errors (IOM, 2000); wide variation in practice patterns, driven more by services available than medical needs (Fisher and Wennberg, 2003); the slow translation of research discoveries into medical practice (Balas and Boren, 2000; Woolf, 2008); the limited quality of the evidence developed to guide healthcare decision makers (Atkins et al., 2004; Califf, 2004; IOM, 2008a; Tunis and Pearson, 2006); and the adverse consequences of care administered with adequate evidence (IOM, 2001). While each highlights a different problem or concern with the current healthcare system, collectively these findings reveal systemic inadequacies in current approaches to developing evidence to help guide the health decisions of policy makers, physicians, and patients.

Underscoring the pressing need for better insights into the relative effectiveness of therapeutics and treatments are the rising and unsustainable costs of health care and the relatively low returns for those high-cost investments. In 2009, spending on health care totaled $2.5 trillion, or over 17 percent of the nation's gross domestic product. Healthcare costs are becoming increasingly burdensome, with annual out-of-pocket costs to consumers steadily increasing (KFF, 2009). Some experts suggest that medical costs due to illness and injury contribute to a significant proportion, perhaps half, of bankrupt-

cies filed by American families (Himmelstein et al., 2005). The Congressional Budget Office estimates that if left unchecked, health expenditures will rise to 25 percent of the gross national product by 2025 (CBO, 2007).

Developing and using information on which treatments work best for whom is imperative to achieving better value from national healthcare expenditures. Of the more than $2.5 trillion spent in 2009 on health in the United States, available estimates indicate that less than one-tenth of 1 percent has been devoted to such research (AcademyHealth, 2005; Moses et al., 2005). Recently, policy makers have demonstrated substantial interest in comparative effectiveness research (CER) (Jacobson, 2007), with attention and discussion focused on increased funding and on the structure, placement, and governance of an entity or agency charged with developing CER information (Kupersmith et al., 2005; Wilensky, 2006). With the passage of the American Recovery and Reinvestment Act of 2009, $1.1 billion were made available to the National Institutes of Health (NIH), the Agency for Healthcare Research and Quality (AHRQ), and the Secretary of Health and Human Services for the conduct of CER and to encourage data resource development and use for such analyses.[1] These funds provided an important down payment on efforts to move to a system focused on delivering high-value care and driven by the best evidence, and formal recommendations have been made by the Institute of Medicine (IOM) (2009) and the Federal Coordinating Council for CER (FCC, 2009). With the 2010 passage of the ACA, and establishment of the Patient-Centered Outcomes Research Institute (PCORI), the capacity for sustained investment has developed. Appendices C, D, and E offer additional background.

The infrastructure needed to expand capacity for CER extends beyond developing data resources (e.g., registries, databases, data networks). Innovative research strategies are needed to improve the efficiency and relevance of clinical research as well as to ensure the appropriate translation and use of CER information by decision makers. Consideration is also needed of how best to align the substantial promise offered by health information technologies—to gather and disseminate needed data and information—with the needs of CER. These technologies offer opportunities to reduce costs and improve the quality of health care (e.g., e-prescribing, remote monitoring, public health records, electronic health records [EHRs]) and will increase access to new types of data and modes of communication (Litan, 2008). Adopting such innovations requires infrastructure development. Careful investments in the requisite workforce, systems, and technologies can also enhance the nation's capacity to learn from health care delivered.

Consideration of such long-term strategies as well as the identification of areas where appropriate investment and coordination will enable

[1] *American Recovery and Reinvestment Act*. 2009. HR1, 111th Cong, 1st Sess.

immediate progress were the focus of the July 30–31, 2008, workshop, Learning What Works: Infrastructure Required for Comparative Effectiveness Research. The meeting's discussions were motivated by many of the issues discussed above and the resulting need to explore key elements and opportunities for infrastructure development (see Box 1-1 on p. 64).

Expanded capacity for comparative effectiveness research can provide information and insights helpful to important care decisions of patients, providers, and policy makers, but progress will require informed and careful investment in key infrastructure elements. Summarized below are possible implications of CER for healthcare stakeholders, an assessment of activities under way, and options to enhance national CER currently under consideration. The workshop's two keynote presentations offered additional perspectives on infrastructure needs by describing a long-term vision and potential returns for a healthcare system informed by CER.

Mark B. McClellan reflects on the core elements of a robust and sustainable capacity for CER and how these immediate needs might also fit into a long-term strategy to support the functions of a learning health system. Noting that form should follow function, he outlines four key evidence gaps that should inform infrastructure development: (1) baselines for evaluations, such as disease models and natural histories; (2) safety; (3) comparative effectiveness of interventions; and (4) comparative effectiveness of treatment strategies and practice patterns. Efforts should focus on all of these areas that fall short in order to develop a healthcare delivery system that provides better outcomes for each kind of patient at much lower cost. Gail R. Wilensky notes that the potential returns from increased investment in CER are enormous, and she offers some suggestions on the elements required for progress, including establishing a center charged with creating better information. It will also be important to develop and use the approaches, data resources, and analyses most useful to producing the information needed and to recognize that all stakeholders need to be a part of the decision-making process.

Background material for the workshop was assembled and prepared by staff of the IOM's Roundtable on Value & Science-Driven Health Care, founded in 2006 to provide a trusted venue for major healthcare stakeholders to consider and advance their mutual interests in the enhanced development and use of evidence in health care. The Roundtable has defined science-driven health care broadly to mean "to the greatest extent possible, the decisions that shape the health and health care of Americans—by patients, providers, payers, and policy makers alike—will be grounded on a reliable evidence base, will account appropriately for individual variation in patient needs, and will support the generation of new insights on clinical effectiveness" (IOM Roundtable on Value & Science-Driven Health Care, 2009). An expanded capacity to develop evidence on the compara-

tive benefits and risks of healthcare treatments and strategies is an essential step toward a learning health system, and it has been the principal focus of the Roundtable's working group on sustainable capacity. At the working group's request, in 2007 IOM staff authored an issue overview white paper, *Learning What Works: The Nation's Need for Evidence on Comparative Effectiveness in Health Care* (IOM, 2007). This background brief also provided context for the July 30–31 workshop, Learning What Works: Infrastructure Required for Comparative Effectiveness Research, and was included in the meeting's briefing materials. It is summarized below and included in full in Appendix A.

THE NATION'S NEED FOR EVIDENCE ON COMPARATIVE EFFECTIVENESS IN HEALTH CARE: LEARNING WHAT WORKS BEST

J. Michael McGinnis, LeighAnne Olsen, Dara Aisner, Pamela Bradley, Daniel O'Neill, and Katharine Bothner
(IOM Roundtable on Value & Science-Driven Health Care staff)

A core objective for the nation is achieving the best health outcome for every patient. This objective cannot be accomplished until better evidence is available upon which to base healthcare decisions and until existing knowledge is applied more effectively. Each need is vitally important. It is known, for example, that failure to deliver proven interventions is a substantial challenge to the quality of health care for Americans, and it is a key concern of the IOM Roundtable on Value & Science-Driven Health Care (IOM, 2007). Yet, with the current pace of change, the most rapidly growing problem is the healthcare system's inability to produce the needed evidence in a timely fashion. Medical-care decision making is now strained, at both the level of the individual patient and the level of the population as a whole, by the growing number of diagnostic and therapeutic options for which evidence is insufficient to make a clear choice. The consequences can be seen in the broad geographic variation in the intensity of services delivered for the same outcome, in the occurrence of medical errors, in patient and provider confusion about which interventions deliver the most value, and in the costs of care.

A testament to innovation is the fact that new pharmaceuticals, medical devices, biologics, and procedures are introduced constantly, and the pace is quickening. From 1991 to 2003 the number of medical device patents per year doubled, and biotechnology patents tripled. Between 1993 and 2004 there was an 80 percent increase in the number of prescriptions received by Americans. A recent review suggests that half or more of the growth in medical spending in recent years is attributable to changes in technology.

In addition to the growth in application of drugs, devices, biologics, and procedures, the world of health care is about to experience dramatic new insights into the genetic variation in individual responses to different diagnostic and treatment interventions (AdvaMed, 2004; Biotechnology Industry Organization, 2006; Foster et al., 2002; Gelijins and Rosenberg, 1994). The age of personalized medicine will soon be a reality, if the capacity can be developed to contend with these insights. Today the average clinical encounter already requires a health provider to manage more variables than would be considered reasonable given what is known about the capabilities of the human mind. Over the next decade, that same encounter will require contending with perhaps an order of magnitude more complex (IOM, 2007).

These developments hold fundamental implications for health prospects, and, to capture and use them effectively and efficiently, a proportionate commitment is required to understand their advantages and appropriate applications. It is both a capacity investment and a resource allocation problem. Of the nation's more than $2.5 trillion in 2009 health expenditures, only a tiny fraction was devoted to CER. If only 1 percent of the nation's healthcare spending were devoted to understanding the effectiveness of the care purchased, the total for effectiveness research would come to approximately $20 billion annually—about 10 times the amount in 2009. In contrast, even accounting for the support from all private and public sources, the aggregate national commitment to assessing the effectiveness of clinical interventions is far below the standard that any company would expect to invest in work to evaluate and improve its products.

Regardless of individual perspectives on reform of the many challenging issues in health policy today, there is little question about the critical need for patients and providers to have better information with which to make their decisions about the comparative advantages of healthcare options. What follows is a summary of the issues and options and is intended to inform discussions of how to proceed on this matter of central importance to health and health care. It does not provide recommendations.

Implications for Stakeholders

For patients, the stakes are very clear. Every patient should be able to feel confident that there is solid evidence that the care received is the care most appropriate to the circumstances. Yet, increasingly this is not the case. In a 2005 survey, 60 percent of Americans said they didn't believe that the United States had the best healthcare system in the world, 41 percent said they knew of a time when they or a family member had received the wrong care, and 56 percent said there should be more investment in clinical and health services research. Health providers feel similar tensions. No health

professional should be put in the position of uncertainty about the evidence in support of the care provided at his/her behest. Yet, with the pace of advances in medical procedures, pharmaceuticals, devices, and biotechnology, a sometimes confusing array of choices is presented for patients, their healthcare providers, and the healthcare organizations in which care is delivered. The integrity and reputation of healthcare delivery organizations is dependent on their ability to ensure the quality and appropriateness of the care delivered within their walls. Any decision support system is only as good as the information built into the model and should include the comparative advantages or disadvantages of different diagnostic and therapeutic options.

Healthcare manufacturers, focused as they are on returns on investment, inherently understand the importance of improving the value proposition in patient care. But their stakes go deeper. Manufacturers directly bear the economic burden of delays and inefficiencies when information is not available about the advantage of their products, not to mention the challenges of public and shareholder backlash when problems are identified too late. Without a sizable improvement in our evaluation capacity, the slower pace of understanding how and when interventions work best will retard the application of innovative treatments.

From a purchasing perspective, the need for better information is of central importance to those who pay for health care: patients, employers, insurers, and the government. Over half of the nation's health expenditures are borne by the private sector, including a sizable share by employers. For the fourth consecutive year, chief executive officers of U.S. companies have cited healthcare costs as their number one economic concern. Employers now pay 78 percent more for health care than 5 years ago, and it has been suggested by some that this increased financial burden makes it more difficult for American companies and workers to compete in the global marketplace. Often acting on behalf of employers, insurers represent the front line of the economic choices that have to be made about payment for healthcare services. This means drawing conclusions about the comparative advantages or disadvantages of proposed diagnostic or treatment interventions in the face of a paucity of such information, especially information applicable to real-world circumstances. As a payer, the government accounts for about 45 percent of health expenditures in the United States, including care that it delivers directly in its own facilities. Whether as a payer or a provider, the government has a central interest in ensuring that its clients receive the care that is most appropriate and of the greatest value.

Current Activities in Clinical Effectiveness Research

Currently, activities to assess the effectiveness of healthcare interventions are broad but underresourced and fall far short of the need (IOM, 2007). CER can be described as either primary or secondary. Primary refers to the direct generation of evidence through the use of a specific experimental methodology. Secondary refers to the synthesis of evidence from multiple primary studies in order to draw conclusions for practice. Within the overall umbrella of CER, the most practical need is for studies of comparative effectiveness, the comparison of one diagnostic or treatment option to one or more others (Wilensky, 2006).

The largest investment in CER has been made by industry, with industry-sponsored clinical trials representing a significant portion of health manufacturer investments in research and development (R&D). For example, about 40 percent of pharmaceutical R&D investments goes to the phase 3 and phase 4 trials, which have particular relevance to clinical effectiveness (PhRMA, 2006). Many of these studies are conducted with academic investigators, and others are managed by contract research organizations. Relatively few of the studies are comparative, or head-to-head, studies.

Outside of industry, several government agencies support CER, including AHRQ, which has a specific mandate and a small appropriation for CER. In 2005, the total appropriations to all federal agencies—the NIH, the Veterans Health Administration, the Department of Defense, the Centers for Medicare & Medicaid Services (CMS), the Food and Drug Administration (FDA), AHRQ, and the Centers for Disease Control and Prevention—for all health services research amounted to about $1.5 billion, and only a modest portion of this was devoted to clinical effectiveness research, far below the industry level. Additional work, also modest, is undertaken by certain of the larger healthcare delivery organizations. Evidence synthesis activity is supported by the insurance industry, professional societies, healthcare organizations, and government. AHRQ has established a network of 13 AHRQ-sponsored evidence-based practice centers that review literature and produce evidence reports, including comparative effectiveness reviews. Organizations interested in evidence reviews will often draw upon syntheses performed by several well-established technology assessment entities (IOM, 2008).

Activities and Needs Related to Comparative Effectiveness Research

Although there is a great deal of interest and activity surrounding the topic of clinical effectiveness, the aggregate research capacity is very thin, and the products fall substantially short of the need. Because of the scant resources available for the support of primary CER—head-to-head

BOX 1-1
Issues Motivating the Discussion

1. Substantial demand for greater insights into the comparative clinical effectiveness of clinical interventions and care processes to improve the effectiveness and value of health care.
2. Expanded interest and activity in the work needed—e.g., comparative effectiveness research, systematic reviews, innovative research strategies, clinical registries, coverage with evidence development.
3. Currently fragmented and largely uncoordinated selection of studies, study design and conduct, evidence synthesis; methods validation and improvement, and development and dissemination of guidelines.
4. Expanding gap in workforce with skills to develop data sources and systems, design and conduct innovative studies, translate results, and guide application.
5. Opportunities presented by the attention of recent initiatives and the increasing possibility of developing an entity and resources for expanded work on the comparative effectiveness of clinical interventions.
6. Growing appreciation of the importance of assessing the infrastructure needed for this work—e.g., workforce needs, data linkage and improvement, new methodologies, research networks, technical assistance.
7. Desirability of a trusted, common venue to identify and characterize the need categories, begin to estimate the shortfalls, consider approaches to addressing the shortfalls, and identify priority next steps.

studies—much of the work is, of necessity, secondary evidence synthesis. Yet the most pressing needs that clinicians and their patients have are for reliable studies upon which to base their decisions. The elements of the needs have been characterized in various ways, and can be grouped into the key areas indicated in Box 1-1. The key challenges that must be faced in each of these areas are summarized in Table 1-1 (Buto and Juhn, 2006; Clancy, 2006; Health Industry Forum, 2006; Hopayian, 2001; Kupersmith et al., 2005; Rowe et al., 2006).

Models for a Stronger Approach to Comparative Effectiveness Research

To narrow the rapidly growing gap between the available evidence on clinical effectiveness and the evidence necessary for sound clinical decision

TABLE 1-1 Prominent Comparative Effectiveness Research Activities and Needs—Key Challenges

Issue	Key Challenges
Head-to-head studies	Scant resources; rapidly increasing need; comparison choice
Systematic reviews	Few primary studies; inconsistent methods; uncoordinated
Comparative value insights	Little agreement on metrics or role of costs; cost fluctuation
Priority setting	Fragmentation; inefficiency; no mechanism for coordination
Study designs and tools	Clinical trial time/cost/limits; large dataset mining methods
Research life-cycle links	Efficacy–effectiveness disjuncture; postapproval surveillance
Evidence standards	Standards not adapted to needs; inconsistency in application
Practice guidance	Disparate approaches; conflicting recommendations
Coverage guidance	Narrow evidence base; limited means for provisional coverage
Application tools	Public misperceptions; incentive structures; decision support

SOURCE: IOM, 2007.

making, various organizations and recent public articles have called for the creation of a new entity and a quantum increase in spending—several billion dollars—on CER. The various approaches to building the required capacity can be grouped into four categories according to the funding patterns for their support (Box 1-2). Each of the approaches is based on an existing or recent model. Although presented as discrete models for discussion purposes, they are not mutually exclusive.

The most straightforward public-funded approach is an expanded and appropriated mandate to an existing or newly created federal agency, and the agency whose mandate most closely parallels these priorities is AHRQ. Through its Effective Health Care program, AHRQ has an existing framework into which many elements of the identified needs can easily fit. Other executive branch models include locating the primary capacity in the NIH, putting it elsewhere in the Department of Health and Human Services, or creating it as a free-standing operational federal agency.

Other possibilities include approaches that are privately funded, although this raises issues of independence and objectivity, as well as approaches with a blend of public and private funding, which could have various governing and execution structures. In the latter category are those

BOX 1-2
Models for Enhancing Capacity

Incremental funding augmentations
- Incremental model

Publicly funded entity
- Executive branch agency model
- Independent government commission model
- Legislative branch office model

Privately funded entity
- Operating foundation model
- Investment tax credit cooperative model

Public–private funded entity
- User fee public model
- Federally funded research and development center public model
- Independent cooperative model
- Independent quasi-governmental authority model

SOURCE: IOM, 2007.

approaches based on the quasi-governmental federally funded research and development centers (FFRDCs), which are funded primarily by the federal government but which are allowed to have up to 30 percent of their funding from private sources. The FFRDCs are private entities managed by nongovernmental organizations and are based on the examples of free-standing independent quasi-governmental entities such as the Federal Reserve Board, which serves as the nation's central banking system, and the IOM and the Transportation Research Board (TRB) at the National Academies. TRB, from its National Academies locus, houses publicly and privately funded work in transportation that is conceptually similar in structure to what is envisioned for CER (IOM, 2008a; Kupersmith et al., 2005; Wilensky, 2006).

Decision and Implementation Considerations

Weighing the relative strengths and weaknesses of the various models can begin with certain touchstone principles that have been suggested to

help guide their consideration. These include the characteristics of the approaches with respect to the following:

- *Scientific credibility:* ability to gain the trust and confidence of the public, the scientific community, and the other stakeholders involved.
- *Political independence:* well-insulated from the political processes that interests from all perspectives will seek to leverage.
- *Stakeholder neutrality:* ability to engage with all stakeholders—patients, providers, employers, manufacturers, and insurers—in an independent, even-handed fashion.
- *Participatory governance:* affording the opportunity for relevant stakeholders to engage as appropriate in setting priorities and agendas, while safeguarding the scientific integrity.
- *Investigator integrity:* management and conduct of the research processes, and the determination and validation of research results completely insulated from outside influence.
- *Agenda flexibility:* organizational decision making, resource allocation, and program conduct with the flexibility to respond quickly to emerging issues and changing circumstances.
- *Infrastructure efficiency:* use where possible existing capacity for the establishment of scientific standards and for the management and conduct of studies.
- *Transparency of processes and results:* specification and availability of the data on which determinations are based, and clarity as to the processes and tools used in their evaluation.

Other implementation considerations include those related to funding and program management. As noted earlier, funding estimates are in the range of several billion dollars. This is a sizable amount, although it is not particularly large in the context of the total U.S. health expenditures or in the context of the efficiencies that could be gained. Suggestions for funding mechanisms range from direct annual federal appropriation or a small set-aside from the Medicare Trust Fund to the structuring of proportionately matching contributions, including set-asides from Medicare fund expenditures, from private health insurance premiums, or from manufacturers' R&D expenditures (Health Industry Forum, 2006; Hopayian, 2001; Kupersmith et al., 2005; Wilensky, 2006). There can be many variations on these themes, but, ultimately, the source of the funds invested is not so important as the value of the return for the outcomes and efficiency of the nation's health care.

Independent Approaches Most Commonly Discussed

Because of the challenges to increasing CER primarily through a simple appropriation to an existing agency—such as the difficulty of marshaling an appropriation at a sufficient level, a lack of political independence, a limited ability to draw on other agencies—much of the recent discussion has focused on three of the independent models, often with blended public and private funding. Table 1-2 presents these as the federal agency, independent board, and hybrid models.

As independent entities, each of these approaches assumes the establishment of a governing board composed of stakeholders and charged with priority setting, broad budget allocation, and fiduciary responsibility for the program of activities. The approaches differ in their degree of insulation between the stakeholder priority setting and the conduct of the scientific studies, as well as in the ways those studies would be managed, the involvement of existing agencies, and the reporting of results (Buto and Juhn, 2006; Kupersmith et al., 2005; Wilenksy, 2006).

PCORI, established under the ACA (2010) as an independent nonprofit organization to assist in informing the health decisions of "patients, clinicians, purchasers, [and] policy-makers," fits this model. The ACA appropriated to the PCORI Trust Fund $10 million, $50 million, and $150 million for fiscal year 2010-2012. Additionally, $150 million plus $1 per Medicare part A and B enrollee has been appropriated for 2013 and $150 million plus $2 for each A/B enrollee, each year from 2014-2019. As outlined in the Act, PCORI will set a national agenda for research priorities, fund entities that conduct priority research, improve clinical effectiveness research methods, and ensure transparency and broad dissemination of its findings. It will be overseen by a Governing Board, comprised of 19 members appointed by head of the Government Accountability Office, as well as 2 ex officio representatives from the Agency for Healthcare Research and Quality and the National Institutes of Health. For more information on PCORI, see Appendix E.

Concluding Observations

As ever-increasing options evolve in health care, current gaps in knowledge and practice about which care works best will persist or worsen without the appropriate information on which to base healthcare decisions. The rate with which new interventions are introduced into the medical marketplace is currently outpacing the rate at which information is generated about their effectiveness and the circumstances of their best use. If trends continue, the ability to deliver appropriate care will be strained and may be overwhelmed. A substantially increased capacity to conduct and evaluate

TABLE 1-2 Comparative Effectiveness Research Enterprise Models

Activity	Federal	Independent Board	Hybrid
Reference model	National Institutes of Health or Agency for Healthcare Research and Quality	Federal Reserve Board	National Academy of Sciences (NAS) (Institute of Medicine [IOM]/Transportation Research Board)
Priority setting	Agency[a]/Department of Health and Human Services (HHS) Board	Governing board/staff	Governing board/ISO[b]
Budget allocation	Agency/HHS Board	Governing board/staff	Governing board/ISO
Study selection	Agency	Governing board/staff	ISO
Design/methods	Agency	Agencies/ IOM–NAS	Agencies/ISO
Agency designation	Agency	Governing board/staff	Governing board/ISO
Study management	Agency	Agencies	Agencies
Study conduct	Agency/field	Agencies/field	Agencies/field
Study certification	Agency	Governing board/staff	ISO
Study conclusions	Agency	Governing board/staff	ISO
Dissemination	Agency	Governing board/staff	ISO
Advantages	Builds on current	Independent	Independent Builds on current
Disadvantages	Politically vulnerable	No established credibility	Other missions of ISO
	Linked to one agency	Duplicate capacity	

[a] Some proposals suggest creating an agency-associated but privately operated federally funded research and development center to give the work quasi-insulated status.

[b] ISO = independent scientific organization (e.g., IOM, on the model of the NAS Transportation Research Board).

research on the clinical effectiveness of interventions brings many potential opportunities for improvement across a wide spectrum of healthcare needs. In time, the enhanced capacity to identify and apply the most appropriate care will improve health and also support innovation, by identifying the areas where it is needed most. The options reviewed above offer a sense of the possibilities and opportunities, but the need for better information is pressing.

A VISION FOR THE CAPACITY TO LEARN WHAT CARE WORKS BEST

Mark B. McClellan, M.D., Ph.D.
Director, Engelberg Center for Health Care Reform, Brookings Institution; IOM Roundtable on Value & Science-Driven Health Care

The IOM Roundtable on Value & Science-Driven Health Care has characterized the key elements of the learning health system needed to achieve the goal that by 2020, 90 percent of clinical decisions will be supported by accurate, timely, and up-to-date clinical information (IOM, 2007). Central to this vision is a sharper focus on research that compares the effectiveness of clinical and health policy interventions in real-world populations. The growing support for CER represents an important first step toward the Roundtable's long-term goal. As infrastructure needs for expanding the nation's capacity for CER are identified and prioritized, it is essential to consider how these immediate needs might support the development of a healthcare system that learns—a system that promotes innovation and improves care by efficiently identifying and disseminating knowledge about practices that lead to better outcomes and higher-value health care.

Key infrastructure for a learning health system will encompass three core elements: data networks, methods, and workforce. Outlined below are the important advances needed in each of these areas and how it might be possible to learn from existing capacity to create an effective, efficient CER infrastructure. It is crucial that this infrastructure be thoughtfully developed, as it will be central to determining the impact of an expanded CER effort on cost and quality.

The chapters that follow provide additional insights into the "how" of CER infrastructure by identifying opportunities to develop the core elements of a robust, sustainable capacity for CER; how these elements interact and reinforce one another; and how to build upon, link, and improve existing public and private system elements.

Data and Information Networks

Distributed data networks for securely and efficiently sharing relevant clinical and claims data are necessary for moving to a system that uses information captured at the point of care to influence practice patterns. These networks will be crucial to the CER infrastructure because they allow data owners—commercial health plans, the federal government, and others—to share only summary data in response to specific queries; individual-level data remains protected within the data owners' systems. This "virtual" approach to linking databases can result in statistically sound study results

while addressing concerns about patient privacy. The functioning of these networks—and the ability to incorporate findings into EHRs—depends on further development of standards. Incentives will likely need to be put in place in order to encourage the creation of linkages and the adoption of standards.

Finally, work is needed to identify and develop successful demonstration projects and pilot models—AHRQ and the FDA, for instance, are already supporting research in this area—that can be built upon and quickly brought to scale.

Research Methods

Capitalizing on the opportunities presented by emerging clinical data and information networks requires innovative approaches to clinical trials in order to allow them to be conducted under conditions of actual practice, enabling estimates of real-world effectiveness. It will also be necessary to have improved statistical and epidemiologic methods to address the limitations of nonrandomized studies employing heterogeneous but much richer and larger-scale data sources. Similarly, methods will be needed to predict patient-level responses to interventions from population-level data. The past several decades have witnessed dramatic methodological advancement in other fields, such as financial services, Web and Internet search technologies, aerospace, and flight dynamics, but such methods have not yet been fully applied in health care. As these new data resources and methodologies become more widely available, and as the challenges of finding better ways to use them are addressed, the nation is on the verge of a tremendous opportunity to improve health care.

Workforce

These critical advances must be driven by a larger and well-trained workforce, prepared to conduct studies in what might be thought of as an expanded or new field of "treatment evaluation in healthcare delivery," as both AcademyHealth and the IOM have proposed. Consideration must be given to how the many existing training and educational programs and approaches can be built upon to develop a broad, cross-disciplinary workforce with advanced capabilities in biostatistics, epidemiology, decision analysis, health economics, health services research, and program evaluation.

Learning from Existing Approaches

The pluralistic nature of the U.S. healthcare system means that careful planning and coordination is required to successfully implement new

approaches to collecting and using data, designing and conducting clinical trials, and recruiting and training a workforce for CER. Many view this heterogeneity as a barrier to the successful creation of a CER infrastructure. However, there is an upside to pluralism, as variation can be helpful in improving practice by affording opportunities to learn more quickly which interventions and policies work and which do not. Thus, a CER infrastructure should promote sharing and learning from the diverse experiences of all its stakeholders. Public–private partnerships can serve as an ideal vehicle for ensuring that all of the stakeholders within the healthcare system are represented within the CER enterprise and have the ability to coordinate with one another. Various public and private organizations would have an incentive to participate in such an effort because all would benefit from the more rapid development of evidence on the effectiveness of practices and treatment strategies.

Evidence Gaps That Inform Infrastructure Needs

In addition to identifying the critical components of a CER infrastructure, dialogue is needed on how the infrastructure can best support the learning health system. The concept of "form follows function" suggests that the development of this infrastructure should be guided by the kinds of research questions that need to be answered. The remainder of this paper will be devoted to examining four key gaps in evidence that could be closed by a learning health system and that should inform CER infrastructure development.

Establishing Baselines for Evaluations—Disease Models and Natural Histories

To move beyond evaluating the average impact of a treatment in a population and toward targeted, personalized medicine, researchers need to understand how particular types of patients are being treated. There has already been some initial progress in this area, as researchers have begun to create and use such tools as natural histories and disease models to clarify how different patients experience disease states and how they respond to different kinds of treatments and, potentially, to new additions to treatments. These models can be developed with large epidemiologic data sets—with the recognition that the data describe actual medical practices and the experiences of patients who are treated differently. Establishing such descriptive baselines is necessary for moving into more direct evaluations of the impacts of interventions on particular subpopulations of patients.

This approach has a clear resonance with the public and policy makers. For example, the public has increasingly taken advantage of the health-

care resources on the Internet, particularly Web sites and services such as PatientsLikeMe, which allow users to share experiences and gain insights on their own conditions based on the experiences of others. In a similar vein, the FDA and other scientific research groups are working to improve their understanding of exactly how different kinds of subpopulations of patients experience their conditions.

Current progress on understanding how subgroups of patients are treated can be bolstered and accelerated by appropriate infrastructure development, specifically the creation and implementation of complete standards for data collection, clinical trials, and EHRs. The potential is clear. At a recent meeting of the American Health Information Community, a number of research groups discussed how a better understanding of particular disease models with particular kinds of patients is beginning to emerge. This initial success in bridging a particular type of gap in evidence can be accelerated with appropriate infrastructure development. Expanding and improving data on treatment patterns in subgroups will also help move the biomedical research enterprise toward developing personally relevant target information that can be used to improve care for particular types of patients. Distributed data networks, as described above, are ideal for this type of work, especially since large amounts of data may be necessary in order to produce statistically significant results in studies of subpopulations.

Safety

Outstanding questions related to patient safety represent a second gap in evidence. For several reasons, some safety issues may be relatively straightforward to address as more extensive infrastructure for evidence generation is developed. First, patient safety has received broad public attention and has been backed by strong bipartisan legislation in the form of the FDA Amendments Act of 2007 (FDAAA), which endorsed the creation of a national, virtual infrastructure for quickly learning about the association between medical product use and adverse reactions and, potentially, about products' benefits in subgroups. Tracking adverse events does not have be statistically challenging, particularly in instances where safety issues should occur only very rarely, if at all. If an adverse event occurs much more frequently in a particular population of patients, it is possible to reach a responsible conclusion about excess risk without conducting a randomized controlled study. Other safety issues may be more difficult to resolve; for instance, increased adverse-event rates may be difficult to discern when medical products are used over long periods of time and when the adverse events are common healthcare problems, like heart attacks. In these situations, the difference in rates between a group treated one way and a group treated another way

may be modest. Another challenge requiring further investigation is the presence of selection bias—patients who take certain drugs are likely to have different, unobserved characteristics than those with a different medication regimen—which makes it more difficult to establish a true, causal effect of a treatment. Sometimes, follow-up studies may be necessary if observational methods provide conflicting or otherwise unsatisfactory results.

The FDA, through its Sentinel Initiative, is already well under way in its efforts to develop the needed public–private infrastructure to support post-market safety monitoring of medical products. The Reagan-Udall Foundation, a nonprofit foundation created by the FDAAA to advance regulatory and product-development science at the FDA—and to involve private sector support as well—is also working to promote the establishment of this kind of network. Initially, this network will only be used for safety monitoring. However, as it evolves, it may be able to accommodate studies that compare the safety and effectiveness of treatments in different subgroups, for example. Thus, the Sentinel Network is likely to be an important part of an infrastructure for evidence generation.

Comparative Effectiveness of Interventions

A third gap concerns comparative effectiveness; at this time, the nation does not have a sufficiently evidence-based system for deciding among treatment options. A recent study concluded that more that 40 percent of the American College of Cardiology/American Heart Association (ACC/AHA) clinical practice guidelines are based on low-quality evidence (Tricoci et al., 2009), according to the evidence-grading system of the ACC and AHA.

These shortcomings in the evidence base create certain challenges that can be addressed with infrastructure development. A primary challenge is the relatively small effect of alternative treatments on patient outcomes observed in comparative effectiveness studies. A difference of only a few percentage points in outcomes—while perhaps clinically important to some patients—might be difficult to detect except in a very large study. These differences would also perhaps be more likely to be subject to confounding if the treatments being compared are not fully randomized, as in the traditional clinical trial. Conducting carefully randomized studies, in real-world situations, of these kinds of practical treatment questions can be difficult as well as costly and time consuming. In fact, by the time a large randomized trial is completed, the information may be outdated. If the information of value to patients and providers is the impact of the treatment in particular kinds of patients, a key challenge is how to move beyond approaches that generate evidence about an overall average effect—in one population versus another—to the efficient development of information relevant to particular types of patients.

The Clinical Antipsychotic Trials in Intervention Effectiveness (CATIE), which led to significant insights about the use of alternative antipsychotic medications, suggest how future CER could focus on improving the evidence of treatment effectiveness on patient subgroups. It is important to emphasize that CATIE revealed not which treatment is better on average, but rather it advanced understanding of which kinds of patients might be treated best with one approach or another at a particular stage in their treatment. Despite the importance of the CATIE study, the evidence on comparative effectiveness that it generated has not yet resulted in an FDA labeling change. As presence on a medical product label is viewed as the gold standard for determining whether evidence is of the highest quality available, the absence of particular evidence from a label can make it difficult for such evidence to be widely accepted by physicians.

A considerable amount of work is currently under way to conduct comparative effectiveness trials and to use data that has been collected for other purposes to develop comparative effectiveness information. There are also ongoing studies to determine the limitations, the methodological challenges, and the improvement in data-collection methods that are necessary to increase the value of these data. This work to improve CER methods has the potential to help develop a truly effective learning health system, but the process will hold many challenges—foremost among them is the absence of agreement on the amount and type of evidence needed for decision-making purposes.

Practice Patterns and Treatment Strategies

The fourth type of evidence gap relates to the need to understand effective treatment strategies and policies. The lack of evidence on how to best deliver care is distinct from the need to understand the differential impacts of specific treatments in different populations, yet both are critical. Treatment strategies and policies alike must be compared in their ability to produce the best outcomes, at the lowest cost, for particular populations. The complexity of medicine is increasing exponentially in terms of the array of treatments available and how they can be used in combination for particular patients. Some estimate that there are hundreds of comparative effectiveness studies under way. Even if that number were doubled or increased tenfold, the current capacity cannot contend with the impending exponential growth in complexity of medical decision making. This complexity stems from the increasing array of medical technologies and combinations of treatments, especially for the growing number of patients with chronic diseases. When physicians have little information about how these technologies and strategies affect outcomes in their patients—particularly those with complex problems—they may provide care that is only marginally beneficial or even harmful.

The geographic variations in the kinds of care delivered to similar Medicare patient populations are at least partly the result of the lack of evidence or consensus on treatment strategies. Experts have suggested that it should be possible to reduce costs in Medicare by 20 percent or more without consequences for patient outcomes—if these variations could be addressed. These variations in costs from area to area are the result of the set of sometimes subtle differences in practice patterns, especially for chronic-disease management. Among a population of patients, for example, the rate at which they are seen for follow-up varies significantly. Thus, relevant effectiveness questions include: What proportion of patients make it to more frequent follow-up? How often are they referred to specialists, and to which type of specialists? Which diagnostic tests are done and when? What minor procedures are performed on these patients and when? In this context, the evidence generated from head-to-head comparisons of treatments in experimental settings is unlikely to address the root causes of this variation.

Resolving questions related to differences in practice will likely require other methods besides randomized clinical trials (RCTs). To compare such practices and determine which relevant policies factor into variation, research needs to account for changes in delivery systems, changes in benefit designs, and changes in payments to providers that could influence how certain practices might lead to better outcomes at a lower cost. Such assessments should be part of the science of healthcare delivery, and knowledge gained through such studies could influence practice.

These questions can be studied in real-world medical practice, where similar patient populations are exposed to different health policies and therefore may face different treatment options and strategies. These kinds of studies could help close the gap between what is known to work and what is actually delivered in medical practice. Such studies, for example, would help answer questions regarding the lack of long-term adherence to certain medications among the chronically ill. More broadly, comparisons of treatment strategies could enhance our understanding of the underlying issues related to the coordination and integration of care—the lack of which constitutes a major problem in our healthcare system today.

The infrastructure needed to address these challenges should involve broad collaboration among many stakeholders, including AHRQ, other researchers, health plans, employers, consumers and patients, and provider groups. Consensus is needed to identify the best methodological approaches for developing the kind of evidence that can show which reforms in payments, benefits, and support systems for healthcare professionals and for consumers can lead to the best results. In this case, methods development should focus on improving observational methods since strategies and policies can only be studied in real-world practice and are not amenable to the idealized academic clinical trial setting. Such studies, however, could be

very useful in uncovering key opportunities for improving outcomes while lowering costs.

Closing Observations

This paper provided examples of what an infrastructure can do to advance knowledge about which care is best as well as some insights about the different elements of infrastructure that can support a learning health system. Unless new infrastructure is informed by current gaps in evidence—and the ultimate goal of delivering care that produces better outcomes for different patient types at much lower costs—it will not be possible to close all the gaps that exist today. Fortunately, there is a tremendous opportunity to meaningfully expand the evidence base, as a result of the advances and collaborations discussed in the chapters that follow. The efforts of all stakeholders are necessary to transform health care into a system that learns much more effectively from actual practice.

THE POTENTIAL RETURNS FROM EVIDENCE-DRIVEN HEALTH CARE

Gail R. Wilensky, Ph.D., Senior Fellow, Project HOPE

Interest in the potential of comparative clinical effectiveness information as a strategy to help Americans learn to "spend smarter" has been growing among those at both ends of the political spectrum, and it can best be understood as part of the concern about healthcare quality and value, and the drive toward the increased use of evidenced-based medicine. Other countries have focused on the use of comparative effectiveness information primarily as a strategy for new drug approval in their national health systems. The potential economic gains are even greater for medical procedures where even less comparative effectiveness information has traditionally been available, since procedures account for much more of the healthcare dollar. Substantial attention has been given to the important decisions that need to be made about the appropriate structure, placement, financing, and function of an agency devoted to comparative effectiveness. It will be equally important to focus on how best to align financial incentives to encourage the use of better information in clinical decision making. The potential for better information to improve health outcomes and help moderate spending increases is enormous, and an understanding of this information is dependent upon how to capture some of the potential savings that CER could bring.

As a prelude to this discussion, I want to note that statistically significant information will not necessarily have clinical or policy importance, in terms

of guiding clinical decision making. Early in my career as a health economist, I codirected the National Medical Care Expenditure Survey (NMCES) at what was then called the National Center for Health Services Research. One of the lessons that I learned is that when very large samples are being assessed—such as 40,000 individuals, the sample size for the NMCES—almost any difference is statistically significant, but many of those differences were not relevant in terms of driving any conceivable policy decision. In the case of very small samples, on the other hand, what appear to be large differences may, in fact, not be statistically significant, which means that they should be used only with great caution in making policy decisions.

The problems that can arise from information bias are a major reason it may be important to consider new data collection, including the possibility of new prospective trials and other costly data-collection strategies, even when it looks like observed differences are substantial. This need stems from the possibility of self-selection or biased selection being introduced in analyses of observed data. An obvious example concerns the presumed advantages of hormone-replacement therapy (HRT). Prior to the Women's Health Initiative studies, relationships had been observed between the use of HRT and a variety of positive outcomes, such as improved cardiovascular health or lower rates of dementia. Unfortunately, data from the Women's Health Initiative showed these supposed advantages to be a function of selection bias related to the characteristics of the women who were using HRT. Therefore, the sizeable differences that had been observed were, in fact, not meaningful in terms of causal interpretation.

Such cases serve as reminders that it is important in designing a study not only to lay out the hypotheses and the data that will serve to support or not support the hypotheses but also to take great care in searching for correlations among the independent variables and other potential drivers of statistical bias in the data to be sampled. In addition, the differences that are likely to exist between various groups will help to determine the necessary sample size needed for the study. Finally, some determination will be needed as to whether the likely differences are ones that would be relevant at a clinical or policy level.

With that introduction, let me turn now to the kinds of data that will be relevant for comparative effectiveness analyses, remembering that the focus for these analyses is generally a medical condition and the various alternative strategies that can be used for treating that medical condition. It will be important to ensure that data are collected for various subgroups in the population that may be differentially affected by a particular medical condition. At the moment, these distinctions may be defined in terms of age and sex or other demographic characteristics, but ultimately it may be possible to differentiate probable outcomes based on an individual's genotype, phenotype, or metabolic type.

While there has been debate about the data that are most appropriate to use for comparative effectiveness, it can be argued that data need to be collected from as many sources as is possible. This includes not only the so-called gold standard of double-blinded RCTs, but the use of real-world prospective trials that Sean R. Tunis and others have been developing allow for inclusion of individuals with comorbidities, epidemiological studies, medical record analyses, registry data, administrative data, and so forth. There have been occasions where researchers have spoken as though only data reflecting the results of double-blinded RCTs should be regarded as appropriate for decision making, but comparative effectiveness analyses need to include data from many sources, although it will be important to make clear the robustness of the data collection strategies and methodologies used in the analyses. Presumably the conclusion made from the data will reflect the robustness of the data and the statistical analyses used in the assessment.

All data have limitations and are subject to error, including the results from RCTs. Specifying these limitations and biases and correcting for them wherever possible is appropriate and should be made available as part of the data release. It will also be important to find ways to reduce the costs and time required for the collection of new prospective data, given the amount of new data collection that is likely to be needed. Efforts by Bryan Luce in developing his Pragmatic Approaches to Comparative Effectiveness initiative, along with the work of Don Berry that makes use of Bayesian statistical approaches to establish shorter end points in certain types of clinical trials, represent other important efforts in this vein.

Even with these strategies to reduce the costs of new prospective trials, it is the anticipated need for a substantial amount of new data that makes the cost involved with comparative effectiveness significant. My guess is that, when fully operational, such efforts could cost several billions of dollars a year—perhaps in the neighborhood of $4 billion to $6 billion a year—although an investment of several hundred million dollars would probably be enough to make a serious start.

The first step in considering an analysis of the comparative effectiveness of various treatments for a particular medical condition is to assess the data that already exist and the analyses that have already been done. It now appears that this step may also require significant investments in time and effort. The IOM report *Knowing What Works in Health Care: A Roadmap for the Nation* served as a wake-up call that making better and more effective use of the information that exists is harder and more challenging than many of us had previously thought (IOM, 2008). Obtaining systematic reviews of existing data will also be more controversial than had previously been recognized.

Setting priorities for comparative effectiveness analyses should be informed by a two-step process: first, focus on those medical conditions—

Medicare diagnosis-related groups might be a useful proxy—that are high-cost/high-volume areas in health care; and second, focus on those medical conditions that are subject to substantial variation in terms of how they are treated. As an economist who is looking at comparative effectiveness as a way to learn how to "spend smarter," I suggest that the best place to focus early efforts would be those conditions on which a lot of money is spent and for which there is a great deal of geographic variation, since that suggests that differences in opinion exist about how best to treat the condition or, in any case, that differences exist in how the condition is actually treated. It would also be appropriate to look at issues of clinical relevance, disease burden, and the various subgroups that are particularly affected. Such considerations would certainly help determine, at a policy level, the relative importance of given interventions.

An important early step for more effective CER will be the creation of either a new center or a series of centers—my preference would be for a single center—that is part of the government or is a public–private enterprise and that is responsible for funding comparative clinical effectiveness studies. Unlike some of my colleagues, I believe it would be unwise—both at a technical level and, even more importantly, at a political level—to include cost-effectiveness analysis as part of the activities of a center for comparative clinical effectiveness. Cost-effectiveness information should be a component in reimbursement decisions made by payers and even in clinical decision making by clinicians and individuals, but these analyses should be kept separate and carried out in separate places.

One reason is that the amount of effort required to increase knowledge about comparative clinical effectiveness is of a much different magnitude in terms of the kinds of studies, the cost of the studies, and the length of time these studies will require. Also, at a technical level, some of the issues involved in cost effectiveness, particularly when it comes to such issues as discount rates, get into areas for which there are no definitive answers. Moreover, the number that one chooses to use has a significant effect on the outcome calculated. Other kinds of technical challenges that arise in cost-effectiveness analysis involve which cost to use and whose perspective to use: Should it be society's perspective? Medicare's? The employer's? Other issues are when in the lifetime of a technology the cost is measured and whose costs are being considered.

As important as the technical reasons are for keeping comparative clinical effectiveness analyses separate from comparative cost-effectiveness analyses, the political reasons are even more important. Cost-effectiveness analyses have long been held in suspicion by industry and many patient advocacy groups as a strategy to prevent them from providing or receiving the latest innovations and technologies in medical care. While this is an issue that ultimately will have to be dealt with, without better information about the likely

effects of different medical interventions in treating various medical conditions, particularly high-cost conditions, it will not be possible to effectively make use of the information on cost effectiveness. It is therefore urgent that the nation make the investment in comparative clinical effectiveness information and keep it as protected as possible, for all users—clinicians, patients, and payers, both public and private—to have available. It will be crucial to have a common data source that we can turn to that captures what is known about the likely clinical outcomes of various kinds of treatments for various subgroups of the population before we get into the next round of much more difficult decision making about how we make use of that information. Having said that, we need to have better information about cost effectiveness as well. Fortunately, that can be funded more quickly and at substantially less cost. A portion of the funding stream that is used to fund studies in comparative clinical effectiveness can be provided to CMS to fund studies in the cost effectiveness of alternative medical treatments.

In determining the best way to fund CER, it is important to take into account the issue of the preferred versus the practical. The preferred strategy would be to use a direct appropriation, as is the case with the NIH, since the information generated is clearly a public good as the economist uses the term. Unfortunately, the practical reality is that relying on a direct appropriation is likely to produce an unreliable funding stream. An alternative would be to use a combination of direct appropriations with fees, which would resemble an all-payer system, for people who are covered by private plans as well as a contribution from the public players such as Medicare. If it appeared that a direct appropriation to CMS for cost-effectiveness analysis was unreliable, a small portion of the funding stream could be diverted to CMS in order to fund the cost-effectiveness studies that are important for Medicare. It will also be important to ensure that the information on cost effectiveness that is generated is valid in terms of objectivity and credibility—just as information needs to be for comparative clinical effectiveness—or it will not be trusted. However, unlike the comparative clinical effectiveness information, payers could generate better information about cost effectiveness as long as they do so in a way that keeps the generation of the information transparent. Alternatively, the cost-effectiveness analysis could done by AHRQ or other entities in the federal government, and it could continue to be done in the private sector by private-sector payers as well.

The question then becomes how do you begin to use this information? Several important principles apply. First, the concept of expecting and allowing for different players to use this information differently is very important. If there is a sense that a single entity can and is making decisions about how information on comparative clinical effectiveness and cost effectiveness will be used, there will be a great deal of resistance by patient

advocates as well as by industry. That does not mean that such information should not be part of a realignment of financial incentives to reward clinicians and institutions in terms of how they practice and to encourage positive patient behaviors, but rather that it ought not be relegated to a single unitary decision-making entity. Such a monopoly would be inappropriate in a country as large and diverse as ours, and it would also be a political nightmare for politicians.

As part of the need to realign financial incentives so that physicians and other clinicians, as well as institutions, are rewarded for producing good clinical outcomes, a first step could be to have information available from comparative clinical effectiveness as part of a change in reimbursement policies. This would be consistent with the development of discussions on value-based insurance, where the amount of the copayment varies with the likelihood of a good clinical outcome for a particular intervention. The objective would be that procedures likely to produce good clinical outcomes for patients in particular categories or subcategories of the population would have low copayments or no copayments, while those medical procedures unlikely to have good clinical outcome would be made more expensive, although not disallowed. In this sense, comparative effectiveness information would be considered not so much as a coverage issue as a reimbursement issue.

Changes in the statutes governing Medicare would be required before we could begin to think about copayments on a variable basis, but that approach is better than some alternatives that have been proposed. Currently there is no statutory authority, when it comes to either coverage or reimbursement, to allow the agency to introduce concepts of cost or cost effectiveness. The ideas outlined here provide a way to introduce these concepts into the reimbursement process and to do so in a way that allows different private payers to use the information differently. It is one of the many changes that will need to occur.

In addition to major investments in comparative clinical effectiveness, it will be important to use a variety of strategies to improve evidence development. One of the ways this could occur would be by tying the local coverage decision making that now goes on under Medicare with evidence development. There has been a great deal of discussion over the last couple of decades about whether it is appropriate for carriers, the local payers, to have their own coverage authority, at least on an interim basis, before a national coverage decision is made. Before I had been sworn in as administrator of the Health Care Financing Administration, now CMS, I had thought that one of my goals would be to remove local coverage decision making on the grounds that this is a national program and that the benefits ought to be the same everywhere. In ensuing discussions, however, it became apparent that to do so would introduce a very conservative bias to the coverage of

new innovations and technologies in Medicare. If local carriers are going to continue making interim coverage decisions, these decisions should only occur with evidence development. Such a change would need to be statutorily driven rather than have the agency attempt to do this administratively, but it seems to be the kind of change that could be introduced to advance evidence development in as many ways as is possible. This would make it possible to harness the diversity that exists in the U.S. healthcare delivery system in order to improve the knowledge base. The concept proposed here is just one example of how to continue having the local coverage variation that already exists, but to do so in ways that still contribute to knowledge and, therefore, to improved decision making in the future.

In conclusion, the following steps need to occur. First, a center or entity should be established that is charged with creating better information on comparative clinical effectiveness. Initially, such a center could be started with an investment of perhaps a few hundred million dollars. In the long term, such a center would require funding on the order of several billion dollars to sustain its effectiveness. Second, priority setting ought to be based on both cost and geographic variation, at least as general guidance, but with allowances made to include the economic and clinical burdens of disease in making decisions. Third, it should be recognized that all stakeholders need to be a part of the decision-making process. Better to have them on the inside participating in the decision making about what is analyzed and how to treat various types of information than to have them attacking the process from the outside.

As noted earlier, it is important to generate information on cost, but the estimates should be done separately from the comparative clinical effectiveness analysis, both in the public and the private sector. It is important also to have credibility, objectivity, and transparency associated as much with the cost analysis as with the generation of comparative clinical effectiveness information. In addition, we need to recognize that as important as it is to have information on clinical or cost effectiveness out there, the necessary gains are not likely to be made unless the reimbursement system is changed to make use of the information through value-based insurance, through changing how we reimburse clinicians and institutions, and through rewarding the kind of behavior that needs to be encouraged rather than just paying more for doing more.

It will be particularly important to begin to give CMS the legislative authority to introduce what is known about clinical and cost effectiveness into its reimbursement decisions. As indicated, this would be preferable to having comparative clinical effectiveness become part of the coverage decision, which is too heavy a burden to use going forward.

And finally, it is important to understand that if the growth in medical spending is to be slowed from its current rate of 2.5 percent faster than

the economy to something more tolerable, it must be recognized that such a slowing will mean less increased cashflow over time than industry, clinicians, institutions, and patients have been used to seeing come into the system. None of them are likely to appreciate the consequences, and there will likely be charges that clinicians are being prevented from providing the best care possible to their patients. Not only may patient advocates and clinicians feel they are being denied "the best care that is out there," but industry may also feel that it is being prevented from having the opportunity to sell what could be the lifesaving or quality-improving strategy that people want. Having credible information to indicate the contrary will be a critical first step—although only a first step—if the United States is ever to learn how to treat patients better and spend smarter.

REFERENCES

ACA (Patient Protection and Affordable Care Act). 2010. Public Law 111-148.

AcademyHealth. 2005. Placement, coordination, and funding of health services research within the federal government. In *AcademyHealth Report*, September 2005. http://www.academyhealth.org/files/publications/placementreport.pdf (accessed September 3, 2010).

AdvaMed (Advanced Medical Technology Association). 2004. *The medical technology industry at a glance*. http://www.omnex.com/training/iso13485/articles/ADVAMED-Medical-Device-Industry-Facts.pdf (accessed September 3, 2010). P. 16, Chart 3.6.

Atkins, D., D. Best, P. A. Briss, M. Eccles, Y. Falck-Ytter, S. Flottorp, G. H. Guyatt, R. T. Harbour, M. C. Haugh, D. Henry, S. Hill, R. Jaeschke, G. Leng, A. Liberati, N. Magrini, J. Mason, P. Middleton, J. Mrukowicz, D. O'Connell, A. D. Oxman, B. Phillips, H. J. Schunemann, T. T. Edejer, H. Varonen, G. E. Vist, J. W. Williams, Jr., and S. Zaza. 2004. Grading quality of evidence and strength of recommendations. *British Medical Journal* 328(7454):1490.

Balas, E. A., and S. A. Boren. 2000. Managing clinical knowledge for healthcare improvements. In *Yearbook of medical informatics*, edited by V. Schattauer. Stuttgart, Germany: Schattauer Publishing. Pp. 65-70.

Biotechnology Industry Organization. 2006. *Chart: Total biotechnology patents granted per year*. http://www.bio.org/speeches/pubs/er/statistics.asp (accessed July 15, 2010).

Buto, K., and P. Juhn. 2006. Can a center for comparative effectiveness information succeed? Perspectives from a health care company. *Health Affairs* 25:w586-w588.

Califf, R. M. 2004. Defining the balance of risk and benefit in the era of genomics and proteomics. *Health Affairs* 23(1):77-87.

CBO (Congressional Budget Office). 2007 (December). *Research on the comparative effectiveness of medical treatments*. http://www.cbo.gov/ftpdocs/88xx/doc8891/12-18-ComparativeEffectiveness.pdf (accessed July 15, 2010).

Clancy, C. M. 2006. Getting to "smart" health care. *Health Affairs* 25(6):w589-w592.

FCC (Federal Coordinating Council for Comparative Effectiveness Research). 2009. *Federal Coordinating Council Report to the President and to Congress on comparative effectiveness research*. http://www.hhs.gov/recovery/programs/cer/cerannualrpt.pdf (accessed September 3, 2010).

Fisher, E. S., and J. E. Wennberg. 2003. Health care quality, geographic variations, and the challenge of supply-sensitive care. *Perspectives in Biology and Medicine* 46(1):69-79.

Foster, L., D. Nanu, and S. Alcock. 2002. Patenting in the biopharmaceutical industry—comparing the U.S. with Europe. *Drug Plus International* 1(1).

Gelijins, A. C., and N. Rosenberg. 1994. The dynamics of technological change in medicine. *Health Affairs* 13(3):28-46.

Health Industry Forum. 2006. *Comparative Effectiveness Forum: Executive summary*. http://healthforum.brandeis.edu/meetings/materials/2006-30-Nov./ExecBrief.pdf (accessed July 20, 2010).

Himmelstein, D., E. Warren, D. Thorne, and S. Woolhander. 2005. Illness and injury as contributors to bankruptcy. *Health Affairs* (Web Exclusive):W5-W63. http://content.healthaffairs.org/cgi/content/full/hlthaff.w5.63/DC1.

Hopayian, K. 2001. The need for caution in interpreting high quality systematic reviews. *British Medical Journal* 323(7314):681-684.

IOM (Institute of Medicine). 2000. *To err is human: Building a safer health system*. Washington, DC: National Academy Press.

———. 2001. *Crossing the quality chasm*. Washington, DC: National Academy Press.

———. 2007. *Learning what works best: The nations need for evidence on comparative effectiveness in health care*. http://www.iom.edu/ebm-effectiveness (accessed August 5, 2010).

———. 2008. *Knowing what works in health care: A roadmap for the nation*. Washington, DC: The National Academies Press.

———. 2009. *Initial national priorities for comparative effectiveness research*. Washington, DC: The National Academies Press.

IOM Roundtable on Value & Science-Driven Health Care. 2009. *Charter and vision statement: Roundtable on value & science-driven health care*. Washington, DC.

Jacobson, G. 2007. *Comparative clinical effectiveness and cost effectiveness research: Background, history and overview*. United States Congressional Research Service. http://www.cbo.gov/ftpdocs/88xx/doc8891/12-18-ComparativeEffectiveness.pdf (accesed July 15, 2010).

KFF (Kaiser Family Foundation). 2009. *Trends in health care costs and spending*. http://www.kff.org/insurance/upload/7692_02.pdf (accessed January 19, 2011).

Kupersmith, J., N. Sung, M. Genel, H. Slavkin, R. Califf, R. Bonow, L. Sherwood, N. Reame, V. Catanese, C. Baase, J. Feussner, A. Dobs, H. Tilson, and E. A. Reece. 2005. Creating a new structure for research on health care effectiveness. *Journal of Investigative Medicine* 53(2):67-72.

Litan, R. E. 2008. *Vital signs via broadband: Remote health monitoring transmits savings, enhances lives*. http://www.betterhealthcaretogether.org/www/docs/broadband.pdf (accessed August 5, 2010).

Moses, H., 3rd, E. R. Dorsey, D. H. Matheson, and S. O. Thier. 2005. Financial anatomy of biomedical research. *Journal of the American Medical Association* 294(11):1333-1342.

PhRMA (Pharmaceutical Research and Manufacturers of America). 2006. *Pharmaceutical industry profile*. P. 49, Table 4. http://www.phrma.org/files/attachments/2006%20Industry%20Profile.pdf (accessed August 5, 2010).

Rowe, J. W., D. A. Cortese, and J. M. McGinnis. 2006. The emerging context for advances in comparative effectiveness assessment. *Health Affairs* 25(6):w593-w595.

Tricoci, P., J. M. Allen, R. M. Califf, and S. C. Smith, Jr. 2009. Scientific evidence underlying the ACC/AHA clinical practice guidelines. *Journal of the American Medical Association* 301(8):831-841.

Tunis, S. R., and S. D. Pearson. 2006. Coverage options for promising technologies: Medicare's "coverage with evidence development." *Health Affairs* 25(5):1218-1230.

Wilensky, G. R. 2006. Developing a center for comparative effectiveness information. *Health Affairs* 25(6):w572-w585.
Woolf, S. H. 2008. The power of prevention and what it requires. *Journal of the American Medical Association* 299(20):2437-2439.

2

The Work Required

INTRODUCTION

Comparative effectiveness research (CER) is composed of a broad range of activities. Aimed at both individual patient health and overall health system improvement, CER assesses the risks and benefits of competing interventions for a specific disease or condition as well as the system-level opportunities to improve health outcomes. To meet the ultimate goal of providing information that is useful to guide the healthcare decisions of patients, providers, and policy makers, the work required includes conducting primary research (e.g., clinical trials, epidemiologic studies, simulation modeling); developing and maintaining data resources in order to conduct primary research, such as registries or databases for data mining and analysis, or to enhance the conduct of other types of clinical research; and synthesizing and translating a body of existing research via systematic reviews and guideline development methods. To ensure the best return on investment in these efforts, work is also needed to advance the development of new or refined study methodologies that improve the efficiency and relevance of research as well as reduce its costs. Similarly, to guide the overall clinical research enterprise in the efficient production of information of true value to healthcare decision makers requires the identification of priority research questions that need to be addressed, the coordination of the various aspects of evidence development and translation work, and the provision of technical assistance, such as study design and validation.

The papers that follow provide an overview of the nature of the work required, noting lessons learned about the known benefits of the country's

capacity and experience, and illustrating opportunities to improve care through capacity building. Emerging from these papers is the notion that although a number of diverse, innovative, and talented organizations are engaged in various aspects of this work, additional efforts are needed. Gains in efficiency are possible with improved coordination, prioritization, and attention to the range of methods that can be employed in CER.

Two papers provide a sense of the potential scope and scale of the necessary CER. Erin Holve and Patricia Pittman from AcademyHealth estimate that approximately 600 comparative effectiveness studies were ongoing in 2008, including head-to-head trials, pragmatic trials, observational studies, evidence syntheses, and modeling. Costs for these studies range broadly, but cluster according to study design. Challenges to develop the workforce needed for CER suggest the need for greater attention to infrastructure for training and funding researchers. Providing a sense of the overall need for comparative effectiveness studies, Douglas B. Kamerow from RTI International discusses the work of a stakeholder group to develop a prioritization process for CER topics and some possible criteria for prioritizing evidence needs. This process yielded 16 candidate research topics for a national inventory of priority CER questions. Possible pitfalls of such an evaluation and ranking process are discussed.

Three papers provide an overview of the work needed to support, develop, and synthesize research. Jesse A. Berlin and Paul E. Stang from Johnson & Johnson survey data resources for research, and discuss how appropriate use of data and creative uses of data collection mechanisms are crucial to help inform healthcare decision making. Given the described strengths and limitations of available data, current systems are primarily resources for the generation and strengthening of hypotheses. As the field transitions to electronic health records (EHRs) however, the value of these data could dramatically increase as targeted studies and data capture capabilities are built into existing medical care databases. Richard A. Justman, from the United Health Group, discusses the challenges of evidence synthesis and translation as highlighted in a recent Institute of Medicine (IOM) report (2008). Limitations of evidence synthesis and translation have led to gaps, duplications, and contradictions; and, key findings and recommendations from a recent IOM report provide guidance on infrastructure needs and options for systematic review and guideline development. Eugene H. Blackstone, Douglas B. Lenat, and Hemant Ishwaran from the Cleveland Clinic discuss five foundational methodologies that need to be refined or further developed to move from the current siloed, evidence-based medicine (EBM) to semantically integrated, information-based medicine and on to predictive personalized medicine—including reengineered randomized controlled trials (RCTs), approximate RCTs, semantically exploring disparate

clinical data, computer learning methods, and patient-specific strategic decision support.

Finally, Jean R. Slutsky from the Agency for Healthcare Research and Quality (AHRQ) provides an overview of organizations conducting CER activities and reflects on the importance of coordination and technical assistance capacities to bridge these activities. Particular attention is needed as to which functions might be best supported by centralized versus local, decentralized approaches

THE COST AND VOLUME OF COMPARATIVE EFFECTIVENESS RESEARCH

Erin Holve, Ph.D., M.P.H., Director; Patricia Pittman, Ph.D., Executive Vice President, AcademyHealth

Overview

In the ongoing discussion about CER, there has been limited understanding of the current capacity for conducting CER in the United States. This report intends to help fill this gap by providing an environmental scan of the volume and the range of costs of recent CER. This work was funded by the California HealthCare Foundation. Current production of CER is not well understood, perhaps due to the relatively new use of the term, or perhaps as a result of fragmented funding streams.

Comparative Effectiveness Research Environmental Scan

This study sought to determine whether there is a significant body of CER under way so that policy makers interested in improving outcomes can plan appropriate initiatives to bolster CER in the United States. This study does not catalog the universe of CER because existing data sources are limited by the way that research databases are developed. However, it is a first attempt to assess the approximate volume and cost of CER.

The study focused on four major objectives:

1. Identify a typology of CER design.
2. Characterize the volume of research studies that address comparative effectiveness questions.
3. Provide a range of cost estimates for conducting comparative effectiveness studies by type of study design.
4. Gather information on training to support the capacity to produce CER.

These efforts relied on three modes of data collection. The first phase included the development of a framework of study designs and topics. The second consisted of a structured search of research projects listed in two databases: www.clinicaltrials.gov and Health Services Research Projects in Progress (HSRProj),[1] a database of health services research projects in progress. The third consisted of in-person and telephone interviews with 25 research organizations identified as leaders in comparative effectiveness studies. The number, type, and costs of studies were noted, although only studies cited by funders were included in the estimates of volume because of the possibility of double-counting studies cited by both researchers and funders. Interviews were used to triangulate information on costs and the relative importance of different designs.

An important study limitation was that it was not possible to cross-reference the databases and the interviews. As a result, although these sources are comparable, they should not be aggregated.

In an initial focus group with experts on CER, a definition of CER was developed to guide the project. Though many definitions of CER have been developed,[2] this project relies on the following:

- CER is a comparison of the *effectiveness* of the risks and benefits of two or more healthcare services or treatments used to treat a specific disease or condition (e.g., pharmaceuticals, medical devices, medical procedures, other treatment modalities) in approximate real-world settings
- The comparative effectiveness of organizational and system-level strategies to improve health outcomes is excluded from this definition, as is research that is clearly "efficacy" research. This means that studies that compare an intervention to placebo or to usual care were excluded from our counts.

The expert panel also developed a framework of research designs to make it possible to categorize findings systematically. For the purposes of this study there was general agreement that there are three primary research categories[3] applicable to CER:

[1] HSRProj may be accessed at www.nlm.nih.gov/hsrproj/ (accessed September 22, 2010).

[2] In a recent report from the Congressional Budget Office, the authors state that comparative effectiveness is "simply a rigorous evaluation of the impact of different treatment options that are available for treating a given medical condition for a particular set of patients" (CBO, 2007). An earlier report by the Congressional Research Service makes an additional distinction that comparative effectiveness is "one form of health technology assessment" (CRS, 2007).

[3] During the interviews we attempted to identify research by more specific types, asking questions about pragmatic trials, registry and modeling studies, and systematic reviews.

1. head-to-head trials (including pragmatic trials);
2. observational studies (including registry studies, prospective cohort studies, and database studies); and
3. syntheses and modeling (including systematic reviews).[4]

Clinicaltrials.gov is the national registry of data on clinical trials, as mandated by the Food and Drug Administration (FDA) reporting process required for drug regulation. Clinicaltrials.gov includes more than 53,000 study records and theoretically provides a complete set of information on all clinical trials of drugs, biologics, and devices subject to FDA regulations (Zarin and Tse, 2008). While the vast majority of trials included in clinicaltrials.gov are controlled experimental studies, there are some observational studies as well.

HSRProj is a database of research projects related to healthcare access, cost, and quality as well as the performance of healthcare systems. Some clinical research may be included in HSRProj if it is focused on effectiveness. HSRProj includes a variety of public and private organizations but only a limited number of projects funded by private or industry sources.

A search was conducted in www.clinicaltrials.gov for phase 3 and phase 4 observational and interventional studies. Phase 4 studies were narrowed by searching only for studies containing the term *effectiveness*. Studies that were explicitly identified as *efficacy* studies or those that did not include at least two active comparators were excluded. The HSRProj database was also searched for studies on *effectiveness*. Because HSRProj does not differentiate between study design phases, a search was also conducted by the types of studies identified in the framework. The studies identified through the process of searching both databases were then searched by hand in order to identify projects that met the definition of CER.

The interview phase of the project included in-person and telephone interviews with research funders and researchers who conduct CER. An initial sample of individuals funding and conducting CER were contacted in response to recommendations by the focus group panel, and these initial

[4] Observational research studies include a variety of research designs but are principally defined by the absence of experimentation or random assignment (Shadish et al., 2002). In the context of CER, cohort studies and registry studies are generally thought of as the most common study types. Prospective cohort studies follow a defined group of individuals over time, often before and after an exposure of interest, to assess their experience or outcomes (Last, 1983), while retrospective cohort studies frequently use existing databases (e.g., medical claims, vital health records, survey records) to evaluate the experience of a group at a point or period in time.

Registry studies are often thought of as a particular type of cohort study based on patient registry data. Patient registries are organized systems using observational study methods to collect patient data in a uniform way. These data are then used to evaluate specific outcomes for a population of interest (Gliklich and Dreyer, 2007).

contacts in turn recommended other respondents—a "snowball" sample. In total, 35 individuals representing 25 research funders or research organizations participated in the project.[5]

Findings

For the project, 689 comparative effectiveness studies were identified in clinicaltrials.gov and HSRProj. Of these the vast majority are "interventional trials" listed on clinicaltrials.gov; specifically, they are phase 3 trials that compare two or more treatments "head to head," have real-world elements in their study design, and do not explicitly include *efficacy* in their description of the study design. Only 19 studies are phase 4 post-marketing studies that compare multiple treatments. Seventy-three CER projects were identified in HSRProj. The process of manually searching project titles confirmed that most studies in this database were observational research, although a handful of studies were specifically identifiable as registry studies, evidence synthesis, or systematic reviews.

The interviews with funders identified 617 comparative effectiveness studies, of which approximately half were observational studies (prospective cohort studies, registry studies, and database studies). Research syntheses (reviews and modeling studies) and experimental head-to-head studies also represent a significant proportion of research activities.

Neither clinicaltrials.gov nor HSRProj publish funding amounts, so interviews with funders and researchers are the sole source of data on cost. As would be expected across the range of study designs covered, there is an extremely broad array of cost for CER studies. However, despite the range, cost estimates provided in the interviews did tend to cluster, particularly by study type. While the cost of conducting head-to-head randomized trials was extremely broad, including studies as costly as $125 million, smaller trials tended to range from $2.5 million to $5 million and larger studies from $15 million to $20 million.

Likewise, the range of cost for observational studies was extremely broad but tended to cluster (Table 2-1). While large prospective cohort studies cost $2 million to $4 million, large registry studies cost between $800,000 and $6 million, with most examples falling at the higher end of this range, although a few were substantially less. Retrospective database studies tended to be less expensive and cost on the order of $100,000

[5] Though an initial group of potential respondents was identified as funders of CER, it was often necessary to speak with multiple individuals to find the appropriate person or group responsible for CER within the organization's portfolio. For this reason the response rate among individuals is lower than might be expected for a series of key informant interviews. Thirteen organizations were identified that did not suggest they received funding for CER from other sources. This subset is used as the sample of organizations that fund or self-fund CER.

TABLE 2-1 Costs of Various Comparative Effectiveness Studies

Type of Study		Cost
Head to head	Randomized controlled trials:	
	Smaller	$2.5m–$5m
	Larger	$15m–$20m
Observational	Registry studies	$2m–$4m
	Large prospective cohort studies	$800k–$6m
	Small retrospective database studies	$100k–$250k
Synthesis	Simulation/modeling studies	$100k–$200k
	Systematic reviews	$200k–$350k

to $250,000. Systematic reviews and modeling studies tended to be less expensive and have a far narrower range of cost, in part because these studies were based on existing data, with many falling between $100,000 to $200,000. It is important to note, however, that this may not include the cost of procuring data. Systematic reviews cluster around a range of $200,000 to $350,000.

There are, of course, additional activities and costs of involving stakeholders in research agenda setting as well as prioritizing, coordinating, and disseminating research on CER that are not included here.[6] These important investments will need to be considered in the process of budgeting for CER.[7]

[6] Examples of activities designed to prioritize and coordinate research activities include the National Cancer Institute's CER Cancer Control Planet (http://cancercontrolplanet.cancer.gov/), which serves as a community resource to help public health professionals design, implement, and evaluate CER-control efforts (NCI, 2007). Within the the Agency for Healthcare Research and Quality, the prioritization and research coordination efforts for comparative effectiveness studies are undertaken as part of the Effective Health Care Program. Translation and dissemination of CER findings is handled by the John M. Eisenberg Clinical Decisions and Communications Science Center, which aims to translate research findings to a variety of stakeholder audiences. No budget information is readily available for the Eisenberg Center activities.

[7] Examples of stakeholder involvement programs include two programs at the FDA focused on involving patient stakeholders, the Patient Representative Program and the comparative

Finally, the interviews also shed light on two subjects discussed at this workshop (IOM, 2008): (1) the need for additional training in the methods and conduct of CER; and (2) the need to bring researchers together to discuss the relative contributions of RCTs and observational studies in the context of CER.

While interviewees generally commented that they have some capacity to respond to an increase in the demand for CER, some noted that they have had difficulty finding adequately trained researchers to conduct such research. For the moment, training programs are limited primarily to research trainees working with AHRQ's evidence-based practice centers and the Developing Evidence to Inform Decisions about Effectiveness (DEcIDE) network. Respondents also mentioned two postdoctoral training programs designed to teach researchers how to conduct effectiveness research. The Duke Clinical Research Institute (DCRI) offers a program for fellows and junior faculty. Fellows studying clinical research may take additional coursework and receive a masters of trial services degree in clinical research as part of the Duke Clinical Research Training Program. The Clinical Research Enhancement through Supplemental Training (CREST) program at Boston University is the second program mentioned. CREST trains researchers in aspects of clinical research design, including clinical epidemiology, health services research (HSR), biobehavioral research, and translational research. Both the DCRI and CREST programs have a strong emphasis on clinical research using both randomized experimental study designs and observational designs.

Other funding for training includes the National Institutes of Health (NIH) K30 awards to support career development of clinical investigators and to develop new modes of training in theories and research methods. The program's goal is to produce researchers capable of conducting patient-oriented research on epidemiologic and behavioral studies and on outcomes or health services research (NIH, 2006). As of January 2006, the Office of Extramural Affairs lists 51 curriculum awards funded through the K30 mechanism. In November 2007 AHRQ released a Special Emphasis Notice

effectiveness research Drug Development Patient Consultant Program (Avalere Health, 2008; FDA, 2009). Other examples of stakeholder involvement programs include the National Institute for Occupational Safety and Health–National Occupational Research Agenda program, the American Thoracic Society Public Advisory Roundtable, and the National Institutes of Health director's Council of Public Representatives (COPR). These efforts can represent a sizeable investment in order to assure stakeholder involvement among the potentially diverse group of end users. For example, the COPR is estimated to cost approximately $350,000 per year (Avalere Health, 2008). From an international perspective, the UK's National Institute for Health and Clinical Excellence (NICE) allocates approximately 4 percent of their annual budget (approximately $775,000) in NICE's Citizen's Council and for their "patient involvement unit" (NICE, 2004).

for Career Development (K) Grants focused on CER (AHRQ, 2007). Four career awards are slated to support the development, generation, and translation of scientific evidence by enhancing the understanding and development of methods used to conduct CER.

The challenge of filling the "pipeline" of researchers working on comparative effectiveness is exacerbated by what many of the interviewees viewed as a fundamental philosophical difference between researchers who were academically trained in observational research and those who are trained on the job to conduct clinical trials. Several respondents noted that these differences likely arise because the majority of researchers are trained in either observational study methods or randomized trials, but rarely both. In addition, as noted by Hersh and colleagues,[8] there are many unknowns related to assessing the workforce needs for CER, including the unresolved definitional issues and scope of comparative effectiveness. As noted by Hersh, the proportion of CER that is focused on randomized trials, observational research, and syntheses has strong implications for the number of researchers (and the type of training) that will be needed. For this reason it is important to track research production and funding for CER in order to anticipate future needs.

Some respondents noted that differences in training manifest themselves in disagreements about the benefits of various observational study designs. Nevertheless, most individuals interviewed as part of this study felt that RCTs, observational studies (including registry studies, prospective cohort studies, and quasi-experiments), and syntheses (modeling studies and systematic reviews) are complementary strategies to generate useful information to improve the evidence base for health care. Furthermore, many participants agreed that, as CER evolves, it will be critical to develop a balanced research portfolio that builds on the strengths of each study type. To facilitate this balance, some of the interviews, as well as many comments at the IOM's Roundtable on Value & Science-Driven Health Care (IOM, 2008) suggest that training opportunities to bridge gaps in language and methods used by researchers may be helpful in creating a balanced portfolio of CER.

Conclusion

Findings from this study indicate that there is a greater volume of ongoing CER than may initially have been supposed. The cost of conducting this research varies greatly, although it tends to cluster by type of study.

[8] Hersh, B., T. Carey, T. Ricketts, M. Helfand, N. Floyd, R. Shiffman, and D. Hickam. *A framework for the workforce required for comparative effectiveness research*. See Chapter 4 of this publication.

The range of studies that are currently being conducted and the cost variation by study type have implications for the mix of activities that may be undertaken by an entity focused on CER. Furthermore, as identified in the interviews, assuring sufficient research capacity to conduct CER is likely to require an investment in multidisciplinary training, with an emphasis on bridging the gap between awareness of the strengths and limitations of randomized trials, observational study designs, and syntheses.

INTERVENTION STUDIES THAT NEED TO BE CONDUCTED

Douglas B. Kamerow, M.D., M.P.H.
Chief Scientist, Health Services and Policy Research, RTI International, and Professor of Clinical Family Medicine, Georgetown University

Overview

CER compares the impact of different treatments for medical conditions in a rigorous, practical manner. At the request of the IOM's Roundtable on Value & Science-Driven Health Care, IOM staff convened in 2008 a multisectoral working group to create a national priority assessment inventory. Their charge was to set criteria for choosing appropriate CER topics and then to nominate and review example topics for needed research. An abridged summary of the report is presented in Appendix B. Appendixes C and D, respectively, are the recommended priority CER studies proposed in 2009 by the IOM Committee on Comparative Effectiveness Research Prioritization and the Federal Coordinating Council for Comparative Effectiveness Research.

Introduction

CER has been defined as "rigorous evaluation of the impact of different options that are available for treating a given medical condition for a particular set of patients," (CBO, 2007) and "the direct comparison of existing healthcare interventions to determine which work best for which patients and which pose the greatest benefits and harms" (Slutsky and Clancy, 2009). Broadly construed, this type of research can involve comparisons of different drug therapies or devices used to treat a particular condition as well as comparisons of different modes of treatment, such as pharmaceuticals versus surgery. It can also be used to compare different systems or locations of care and varied approaches to care, such as different intervals of follow-up or changes in medication dosing levels for a particular condition. CER also may be used to investigate diagnostic and

preventive interventions. All of these studies may include an evaluation of costs as well as an assessment of clinical effectiveness.

Comparative assessment research is especially valuable now, in an era of unprecedented healthcare spending and large, unexplained variations in care for patients with similar conditions. The IOM's Roundtable on Value & Science-Driven Health Care set a goal for the year 2020 that 90 percent of clinical decisions will be supported by accurate, timely, and up-to-date clinical information that reflects the best available evidence. Currently, there is insufficient evidence about which treatments are most appropriate for certain groups of patients and whether or not those treatments merit their sometimes significant costs. The healthcare community requires improved evidence generation that will address how different interventions compare to one another when applied to target patient populations. CER provides the opportunity to ground clinical care in a foundation of sound evidence.

The current system usually ensures that when a new drug or device is made available, there is evidence to show its effectiveness compared to a placebo in ideal research conditions—that is, its efficacy. But there is often an insufficient body of evidence demonstrating its relative effectiveness compared to existing or alternative treatment options, especially in real-world settings. This limited scope of information increases the likelihood that clinical decisions are not based on evidence but rather on local practice style, institutional tradition, or physician preference. Although the numbers vary, some estimate that less than half—perhaps well less than half—of all clinical decisions are supported by sufficient evidence (IOM, 2007). This lack of evidence also leads to substantial geographic variations in care, further supporting the idea that patients may be subjected to treatments that are unnecessarily invasive—or not aggressive enough—for a variety of conditions. These variations in care partly explain healthcare spending differences across geographic regions that cannot be fully accounted for by price differences or illness rates. Geographic variations in treatment approach are often greater when there is less agreement within the medical community about the appropriate treatment. Variation in treatment approach for a variety of conditions is of significant concern because it has not been demonstrated that areas with higher levels of spending—where presumably patients are treated with more aggressive or expensive options or with simply more treatment—have significantly better health outcomes than areas with lower levels of spending.

The Institute of Medicine and Comparative Effectiveness

The IOM's Roundtable on Value & Science-Driven Health Care recognized the importance of furthering CER to ensure that all clinical deci-

sions are based on sound evidence. The participants at the Roundtable's July 2007 meeting of sectoral stakeholders concluded both that current resources to support head-to-head assessments of treatment options are not optimal and that a stronger consensus is needed on the priorities and approaches for assessing the comparative clinical effectiveness of health interventions. Participants at this meeting identified the need for the development of what they initially termed a "national problem list" to illustrate key evidence gaps and to prompt discussions leading to national studies.

National Priority Assessment Inventory

After initial work was done by IOM staff, the proposed project was renamed the National Priority Assessment Inventory. A working group was convened to serve as a steering committee for the project. Nominees for the working group were sought from IOM Roundtable stakeholders representing the different participating sectors—patients, caregivers, integrated care delivery organizations, insurers, regulators, employees and employers, clinical investigators, and healthcare product developers. The working group was composed of physicians and researchers representing different specialties and perspectives, coming from academia, government, private practice, medical specialty societies, and industry. The working group was given three tasks to accomplish in a series of conference calls involving either the entire group or only certain individuals:

- Review and refine selection criteria for identifying and evaluating needed research.
- Solicit and nominate candidate research topics.
- Review and comment on the final list of pilot topics.

Selection Criteria

The working group initially discussed and refined criteria that could be used to identify and evaluate candidate research topics. The final five selection criteria are listed in Box 2-1.

The first criterion selected was the *importance of the conditions* being treated or prevented. The working group wanted to concentrate on research for problems that were serious, common, or costly. Though not wanting to set quantitative cutoffs, they felt it was important that the studies chosen involved conditions that would be clearly recognized as important by clinicians, policy makers, and patients.

Second, *effective treatments or preventive interventions* needed to be available for the chosen conditions. Reasonable alternative treatments and resulting variations in practice are basic requirements for CER. Alternatives to be compared could include different drugs, different treatment modalities

BOX 2-1
Selection Criteria to Identify and Evaluate Research Topics

1. Importance of the conditions
2. Effective treatments or preventive interventions available
3. Current knowledge about the relative effectiveness not definitive
4. Research must be feasible and realistic
5. The selection process should yield a heterogeneous group of topics

(drugs, devices, surgery, etc.), different preventive interventions, or different settings and systems of care.

Third, *current knowledge about the relative effectiveness* of alternative treatments or modalities must not be definitive, so that uncertainty exists in treatment selection for different settings and populations. In short, while an existing body of research is clearly necessary, unresolved questions should still exist.

Fourth, research to answer selected questions must be *feasible and realistic*. A head-to-head trial or data analysis that could improve knowledge and guide decision making should be practical to perform, without major design or financial barriers.

Fifth, the work group agreed that this pilot selection process should yield a *heterogeneous group of topics*. The final topic list was intended to provide examples of treatment, prevention, and HSR on a variety of conditions, using several different modes of treatment, and for conditions in differing demographic groups.

Selection Process

Candidate topics were solicited in a number of ways. Working group members suggested many topics from their respective clinical and administrative experience; they also asked colleagues for suggestions. IOM staff generated topics from prior work they had done as well as from outside sources. A prioritized list of 100 Medicare research priorities generated by the Medicare Evidence Development and Coverage Advisory Committee was also reviewed, as was a list of 14 priority conditions created for AHRQ's Effective Health Care Program. A special effort was made to identify topics related to the care of children because so much effort is usually concentrated on adults and the elderly.

After this nomination process, about two dozen potential topics and studies were classified by condition, applicable population, type of treatment, setting, and other categories, and then were discussed by the working group. Staff began to evaluate the candidate topics by doing literature searches to determine available literature and feasibility. In an informal iterative process performed with the working group, the list was reduced to 16 topics. This list was then circulated to sectoral representatives for comments, and in July 2008 it was discussed and approved by members of the Roundtable on Value & Science-Driven Health Care. Staff then did further literature reviews and wrote brief summaries of the final example topics.

The 16 study topics are listed in Table 2-2 and are categorized by type of study, condition, and age group. As mentioned previously, the topics were chosen in part to provide examples from each of the categories in these tables.

Study Topic Summaries

Sixteen study topic summaries are included in Appendix B of this publication. All of the summaries are organized in the same manner:

- description of the condition or problem,
- available treatments or interventions,
- current evidence about the treatments or interventions,
- issues needing research and conclusions, and
- brief list of references.

Lessons Learned from This Project

The original, perhaps naïve, intent of this project was to produce a list of the 20 or so "best" or "most important" comparative effectiveness studies that need to be done immediately. However, evaluating and ranking studies proved to be difficult to do. Some nominations were impractical or difficult to operationalize; for others, evidence was not available. Ranking topics by potential national costs of the condition would skew the rankings toward common and expensive adult problems and leave childhood problems out entirely. Comparing topics is often an "apples vs. oranges" exercise, and common metrics are not always available. The process did, however, produce some clear lessons learned:

Much research needs to be done. It was not difficult to gather nominations and information about research that has the potential to make a real difference in costs and outcomes. It bears repeating that the topics chosen are just examples of questions that need to be answered.

Stakeholders should be consulted. A process without input and review

TABLE 2-2 The Comparative Effectiveness Studies Inventory Project Identified 16 Candidate Topics for Comparative Effectiveness Research

Study Topic	Study Type	Age Group	Condition
Treatment of attention deficit hyperactivity disorder in children: drugs, behavioral interventions, no prescription	Comparative effectiveness treatment studies across modalities	Children	Mental diseases
Treatment of acute thrombotic/embolic stroke: clot removal, reperfusion drugs	Comparative effectiveness treatment studies across modalities	Adults	Heart and vascular diseases
Treatment of chronic atrial fibrillation: drugs, catheter ablation, surgery	Comparative effectiveness treatment studies across modalities	Adults	Heart and vascular diseases
Treatment of chronic low back pain	Comparative effectiveness treatment studies across modalities	Adults	Neurological diseases
Gamma knife surgery for intracranial lesions vs. surgery and/or whole brain radiation	Comparative effectiveness treatment studies across modalities	Adults	Neurological diseases
Treatment of localized prostate cancer: watchful waiting, surgery, radiation, cryotherapy	Comparative effectiveness treatment studies across modalities	Adults	Cancer
Diagnosis and prognosis of breast cancer using genetic tests: human epidermal growth factor receptor 2 and others	Diagnostic studies	Adults	Cancer
Over-the-counter drug treatment of upper respiratory tract infections in children	Drug–drug and drug–placebo treatment studies	Children	Respiratory disorders
Drug treatment of depression in primary care	Drug–drug and drug–placebo treatment studies	Adults	Mental disorders

continued

TABLE 2-2 Continued

Study Topic	Study Type	Age Group	Condition
Drug treatment of epilepsy in children	Drug–drug and drug–placebo treatment studies	Children	Neurological diseases
Use of erythropoiesis-stimulating agents in the treatment of hematologic cancers	Drug–drug and drug–placebo treatment studies	Adults	Cancer
Outcomes of percutaneous coronary interventions in hospitals with and without onsite surgical backup	Health services/ systems studies	Adults	Heart and vascular diseases
Screening hospital inpatients for methicillin-resistant *Staphylococcus aureus* infection	Preventive interventions	Adults	Infectious diseases
Tobacco cessation: nicotine replacement agents, oral medications, combinations	Preventive interventions	Adults	Preventive interventions
Prevention and treatment of pressure ulcers	Surgical studies	Adults	Dermatological diseases
Inguinal hernia repair: open vs. minimally invasive	Surgical studies	Adults	Surgical disorders

NOTE: Study topics are categorized by study type, age group, and condition.
SOURCE: Kamerow, 2009.

by a broad set of stakeholders risks missing important topics as well as not obtaining different perspectives on all nominated topics. Clinicians ask one kind of question; payers and employers often have different concerns. All perspectives need to be considered in nominating, vetting, and ultimately deciding on research topics. That said, some topics nominated by stakeholders were not confirmed as being important or practical after literature reviews.

Research questions need to be carefully defined. Once general topics had been selected, literature reviews often found either too much or

not enough evidence to support a call for further research. Topics were often reoriented as a result of the state of the evidence, being narrowed, expanded, or abandoned according to what the evidence said.

Different types of studies are needed. Three types of comparative assessment studies are often described. The gold standard for CER is a prospective, randomized head-to-head trial comparing specific treatments or interventions for a condition. Such trials, however, can be expensive and usually take time to complete. In some cases, the extant literature is sufficient to assess comparative effectiveness, and systematic reviews with or without formal meta-analysis can be created from data taken from previous studies. This type of study is usually the fastest to complete. Finally, existing clinical and payment data systems can sometimes be analyzed to compare the effectiveness of drugs or devices without collecting new data. These studies can also be done relatively rapidly. It was often difficult to determine from the state of the literature which type of comparative effectiveness study was appropriate for each of the nominated topics.

An explicit, transparent process is important. Significant funding will be necessary to undertake these studies. Unlimited funding is never available, of course, and there will be limits on the number of projects that can be initiated simultaneously, so priorities must be assigned. While the external validity of this selection and ranking processes has not been proven, the face validity of the process will be important for continued stakeholder support.

Continuous updating is important. Advances and new data appear all the time, so frequent updates of literature reviews will help make sure that the topics selected still need research.

Prioritizing topics is very difficult. As mentioned above, ranking studies by their importance is difficult to do in a manner that is equitable. The burden of suffering caused by a condition, expressed in prevalence or mortality rates, can be determined for many conditions, but it is skewed by greater prevalence of disease in older populations. Taking into account the years of productive life lost can help to correct for some of the age bias, but such calculations do not necessarily reflect costs. Treatment cost data are biased towards conditions whose treatments involve expensive devices or drugs. And so on.

"Indication creep" is challenging to assess. Often drugs or devices that have been proven effective for one indication or condition will be used to treat other, related problems without strong evidence. Questions then arise about their effectiveness for the new indication. Are *de novo* studies required, or can results from research on similar problems be extrapolated to apply?

Systematic reviews and evidence-based guidelines are important and helpful, as are expert-written editorials. Staff found that medical journal

editorials accompanying research articles often provided a good orientation for subjects being considered. When available, systematic reviews and meta-analyses helped in the assessment of what is known and what needs to be done. Evidence-based clinical practice guidelines were also useful, especially when they included research agendas that emerged from the guideline creation process.

The distinction between efficacy and effectiveness is an important one. More often than not, clinical trials and the resulting meta-analyses report on work done in academic centers with highly selected patients. Except for some very recent trials, most do not include minority or gender diversity. This limits the ability to extrapolate from trial or meta-analysis findings to real-world populations.

Much more needs to be done. This effort is only a preliminary step in setting criteria and providing examples of CER. In several instances a single test or technology was discussed as an example of several needed study topics and areas. A much larger and more comprehensive assessment process is needed to create a truly national inventory of needed research.

CLINICAL DATA SETS THAT NEED TO BE MINED

Jesse A. Berlin, Sc.D., Vice President, Epidemiology; and
Paul E. Stang, Ph.D., Senior Director, Epidemiology,
Johnson & Johnson Pharmaceutical R&D[9]

Overview

Information is like fish: It's better when it's fresh.

Healthcare delivery, by virtue of the evolving ability of systems to capture and store data and the intense interest in analyzing those data, is poised for transformation. The key to the transformation will be how the data and analyses are used to inform decisions to improve health. If successful, the system can be transformed into a learning health system that will provide high-quality data to inform decisions made by policy makers, healthcare providers, and patients, who can then use the same data collection and analytic tools to assess the impact of those decisions. The critical elements needed to transform health care and to create a "learning organization" (Senge, 1990) include experimentation with new approaches, learning from personal and corporate past experience, learning from best practices of

[9] The authors would like to acknowledge the insightful comments of Michael Fitzmaurice on earlier drafts.

others, efficient and rapid knowledge transfer, and a systematic approach to problem solving.

This model has been recast for health care by Etheredge (2007), who espouses the value of observational research in medical care data as a laboratory for testing hypotheses and undertaking research to inform rapid learning and improve the efficiency of healthcare delivery in the United States in addition to addressing a number of other information gaps. However, the current capacity and data resources are only capable of filling the short-term needs and therefore represent a modest initial step toward informing better healthcare decision making. The contribution of electronic data collection to achieving the vision of the learning health system will be directly related to the quality, breadth, and depth of the data capture.

The basic contention underlying this paper is that appropriate use of existing data and creative uses of existing data collection mechanisms will be crucial to operationalizing the above elements, the end result of which will be improvement of healthcare decision making in the near term. With this goal in mind, the strengths and limitations of currently available administrative data to address questions of comparative effectiveness (and safety) will be explored, and additional ways in which the existing infrastructure for these databases might be used to support further data collection efforts, perhaps more specifically targeted, will be proposed.

Currently Available Healthcare Databases: Claims and Electronic Health Records

Most existing healthcare databases were created for purposes other than research. Currently, payers (including managed care organizations, insurance companies, employers, and governments) use data to track expenditures for multiple purposes. This gives rise to the so-called claims or claims-based databases that contain the coded transactional billing records between a clinician and an organization that allow the healthcare provider to be reimbursed for a given patient. In parallel, the details of the patient–clinician interaction contained within the medical record have slowly been moving from paper to computers, giving rise to EHRs, which can also be aggregated into a database. The EHR may, in some cases, represent all of the care provided to a given patient by an institution or staff health maintenance organization, but in other cases it may represent just that portion of care rendered by the individual clinician with the EHR. The Mayo Clinic, long the keeper of extensive longitudinal paper "dossier" medical records, now uses an electronic system. The General Practices Research Data (GPRD) based in the United Kingdom and records from Kaiser Permanente and the Henry Ford Health Systems are a few other examples of these

EHR databases. These data reflect not only the patient–clinician interaction but also the underlying healthcare delivery system (and its idiosyncrasies).

Existing secondary data sets, such as those listed in Box 2-2, offer a number of advantages for research. Studies are inexpensive and can be done quickly, relative to the cost of doing a clinical study, because the data have already been collected and organized into a database; the data reflect healthcare decisions and outcomes as they were actually made (vs. the artificial constructs of an RCT); and each database represents an identifiable and quantifiable source population, i.e., a "denominator" for the calculation of event rates and mean values. By virtue of reflecting actual, real-world clinical practice, these databases also offer improved external validity relative to RCTs.

Despite all of these advantages for research, secondary data have several important limitations and obvious gaps. For claims data, research would be restricted to coded diagnoses, some of which may be erroneous, may be part of an active workup (e.g., the code may be MI [myocardial infarction], but the interaction was part of a "rule-out" MI), or may omit a code altogether because it is not required for billing (e.g., one of four active problems during the visit). Some relevant data may not have been collected or consistently recorded, particularly those potential confounding factors that may be associated with the choice of therapy and with the outcome of interest. The absence of these data may be critical when making comparisons between specific therapeutic options, as they may strongly influence the study results or the conclusions drawn, but they cannot be managed in the design or analysis of these databases. The problem of confounding of comparisons between therapeutic options is why randomization is such a powerful aspect of clinical trials. Unfortunately, many existing databases lack consistent capture of such factors as smoking status, alcohol consumption, use of over-the-counter (OTC) medications; even weight and height are not always routinely captured in these databases (Box 2-2). As an example, Ilkhanoff and colleagues (2005) used a case-control study that included interview-based primary data collection from participants to show that adjustment for confounding factors not captured in electronic databases—e.g., smoking, family history, years of education—had a substantial impact on estimates of relative risk for MI associated with non-aspirin, non-steroidal anti-inflammatory drug (NANSAID) use, relative to nonusers. The inability of databases to capture OTC use of NANSAIDs and OTC aspirin use also had marked effects on study findings.

There is also a general lack of longitudinality in the patient record that reflects the current healthcare system. Loss to follow-up occurs because people are constantly changing healthcare systems or coverage, either as part of the annual choices provided by employers or because their eligibility for

BOX 2-2
Examples of Existing Healthcare Databases

Healthcare delivery systems: Kaiser Permanente, Group Health Cooperative of Puget Sound, Geisinger Health System, Henry Ford Health System, Rochester Epidemiology Project (Mayo Clinic)

Aggregated claims: PharMetrics, Marketscan (Thompson-Medstat)

Electronic medical record systems independent of a single healthcare delivery system: General Practice Research Database (United Kingdom), The Health Improvement Network (based in the United Kingdom), General Electric (Centricity), Cerner

Hospital: Premier, Cerner

Government-managed or -funded programs: Medicaid, Centers for Medicare & Medicaid Services (Medicare), Veteran's Administration, Department of Defense

U.S. government surveys: National Health and Nutrition Examination Survey, Medical Expenditure Panel

Specialty data and registries: Surveillance, Epidemiology, and End Results (Cancer Registry), Food and Drug Administration spontaneous adverse event database (voluntary drug adverse event reports), the Nordic countries linkable registries (Sweden, Denmark)

Canadian provincial databases: Saskatchewan, Manitoba, Régie de l'Assurance Maladie du Québec

Prescription-tracking databases: IMS Health, Verispan

coverage changes. About 10 to 20 percent of patients in a given insurance database may leave a given plan during a given year. Having a unique identification number that follows a given patient across all interactions with the healthcare system would alleviate this problem with longitudinality, but it is clear there are appropriate concerns about confidentiality that such a system would trigger.

Finally, there are entire segments of the population and healthcare system that are poorly represented in these data. The interactions of the 49

million Americans without healthcare coverage are essentially lost to the current system since there is no ability to link individual patients with a unique identification number. Similarly, elderly people and those in institutions are essentially overlooked in most analyses because of lack of access to their clinical data, even though these are the very groups that are most at risk for poor coordination of care and less favorable outcomes.

Studying Benefits and Risks in Existing Healthcare Data: Information Asymmetry

Because the data source represents a clinical interaction, any retrospective research will be limited to events that could be characterized as part of the coded or detailed physician interaction. In general, the benefits of treatment are more common than the risks, but unlike risk, benefits are poorly represented in these data sources, as they do not fit into the current diagnostic vernacular. Many of the benefits of treatment (e.g., reduction in blood pressure, improvement in mood or quality of life, return to full mobility, fewer number of seizures per month, reduced symptoms of schizophrenia) are not clinical diagnoses, and they are not usually captured in databases because they cannot be represented in coded claims. It is not always known what impact these other measures have on more serious events, nor is it known how their importance is perceived by patients and providers. The claims do capture use-based measures (e.g., switching drugs, changes in emergency department use, hospitalizations), and the data would include reductions in clinical events (e.g., myocardial infarction). Conversely, most potential harms from therapies *are* clinical events and would be captured in clinical encounter data (e.g., agranulocytosis, hepatitis, renal failure).

In the context of comparative effectiveness studies, differences in effectiveness (usually considered to be benefits) between two treatments, especially between two drugs in the same class, may be small in magnitude. Evaluating such small differences in effects absolutely *requires* strong control over potential confounding variables if internal validity is to be maintained. If one fails to control confounding variables, the observed differences in effects (or safety) could be misleading, as they might represent differences in the underlying characteristics of patients exposed to one or the other product rather than true differences in effect between the products. Several efforts, including the Developing Evidence to Inform Decisions about Effectiveness (DEcIDE) network and the Centers for Education and Research on Therapeutics, both of which were set up by AHRQ and the Observational Medical Outcomes Partnership from the Foundation for the NIH, are under way to better understand the strengths and limitations of these databases, the performance characteristics of the methods, and the

ability to use them as the initial line of surveillance for potential safety issues.

Because these databases reflect care as it is really delivered, generally without the consistent screening and capture of information that exists in a clinical trial, variability and uncertainty can be high. Erkinjuntti and colleagues offer a striking example in which the prevalence of dementia in a single Canadian cohort of patients varied from 3.1 to 29.1 percent simply by applying different accepted case definitions of dementia (Figure 2-1) (Erkinjuntti et al., 1997). With secondary data it is unclear what diagnostic criteria (if any) were applied in the clinical setting to arrive at a diagnosis, and there are other measures that are either inherently prone to variability, such as blood pressure, or are subjective in nature, such as the Tanner score. When measures are highly variable and taken outside the context of a carefully controlled study with standardized measurement techniques, it will often be more difficult to detect differences in these measures between treatment approaches. By engaging the clinician directly at the point of data entry, the record system itself can help standardize the capture of data, can solicit additional detail, or can push summary information and links to additional resources.

Existing databases are generally considered to be reasonably complete with respect to determining "exposure" to particular drugs (and, to a lesser

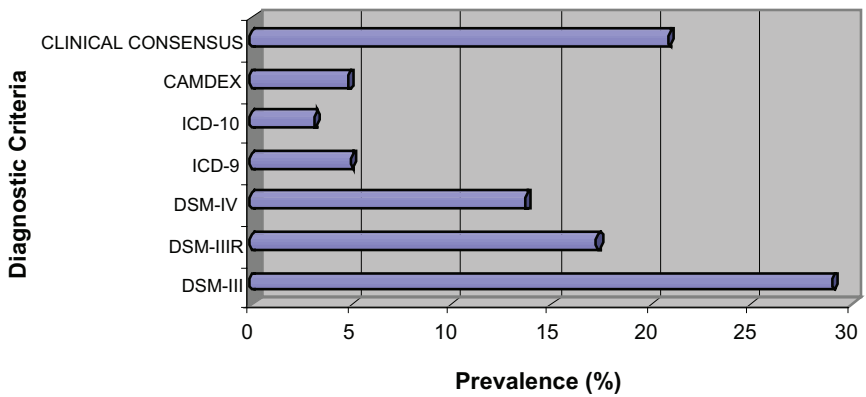


FIGURE 2-1 The prevalence of dementia across different diagnostic criteria in the same Canadian cohort.
NOTE: CAMDEX = Cambridge Mental Disorders of the Elderly Examination; DSM = Diagnostic and Statistical Manual of Mental Disorders; ICD = International Classification of Diseases.
SOURCE: Derived from data in Erkinjuntti et al. (1997).

extent, medical devices) (Strom, 2005). However, databases capture only the prescription of a medication (from EHR) or, at most, the prescription plus the dispensing of the prescription (from pharmacy records or billing). The databases do not (cannot) capture whether the drug was actually used or, if used, whether it was used correctly.

While there are a number of challenges in the use of current databases to address questions of comparative effectiveness, it is still possible to learn from current efforts to inform further designs and improvements in the capture of data, governance, and methods. The current systems can best be thought of as hypothesis generating and, potentially, hypothesis strengthening; it is unclear to what extent they can be considered definitive sources of data for confirmation. In some respects, current systems represent databases in search of a question. As the field transitions to EHRs, some of the issues with coding mentioned above will be dramatically reduced, and the value and impact of these data and the evolving methods will improve. The ability to link a patient across data sets will further strengthen the capacity of the data sets to provide more definitive answers, increasing the value of these data sets; the appearance of prompted "pop-ups" to collect or refine data at the point of entry will similarly increase the value of the data sets. Real value will come from the ability to use the EHR system as a data collection vehicle for randomized studies in the populations covered by the data sources. This concept is expanded in the following section.

Looking Ahead

Given the limitations of working with available databases, it seems likely that more robust data collection at the clinical interface would improve insights from, and the quality of, observational databases. Accumulating higher-quality data could help provide insight into current practice and help improve care. Simply adding flags to denote "rule-out" diagnoses would help researchers to better distinguish actual events from clinical workup. Similarly, with respect to general hospitalization data where much effort is focused (because of its high costs), if a separate diagnostic list of those findings was included at the time of admission (rather than the existing discharge diagnoses only), it would have a major impact on the ability to track and determine risks and effectiveness of hospital-based occurrences. Finally, as more devices are implanted, physically coding (e.g., with bar codes) the individual devices with a unique identification would facilitate future identification and the tracking of safety and outcomes. The ongoing Centers for Medicare & Medicaid Services (CMS) payments to hospitals and physicians for reporting quality data and for improved performance on

quality measures may be an appropriate mechanism to help motivate and implement such changes.

However, the innovation that holds the greatest potential for informing change in health care will be the ability to use the existing data infrastructure and healthcare delivery system to take advantage of randomization at points of clinical equipoise to generate new insights into interventions and their outcomes. There are many possible innovative approaches, but the focus here is specifically on the idea of "designing studies *into* the database."

In general terms, the goal is to increase the power of the existing data collection through the EHR by enhancing data collection with special data collection forms, e.g., screens that pop up on a computer in the physician's office for patients who match a specific set of criteria. The basic idea is to *tailor* additional aspects of data collection (as would be done in a separately designed and implemented primary study) within the context of an existing data collection system. This increases the possibility of going one step further by conducting large simple randomized studies (also called naturalistic or real-world trials) by "randomizing into the database," a concept that was first described by Sacristan and coworkers (1997) and later employed by Mosis and colleagues in the Dutch Integrated Primary Care Information database (Mosis et al., 2005a, 2005b). Their findings suggest that the technical infrastructure existed but that the requirements for recruitment in general practice for this particular study were inconsistent with the flow of patient care and were too time consuming (Mosis et al., 2006). At a high level, this would again involve the use of a computer in the clinician's office that, when a patient meeting certain criteria presents, would generate a special screen asking the clinician whether or not he or she is willing to randomize that patient.

This model is being revived across the United Kingdom in the Research Capability Programme in the Connecting for Health initiative, which includes educating the public on the importance of participating in clinical trials as a way of contributing to medical knowledge and advancement. Making such trials practical in the context of primary care may well require either structural changes in the process of care, in order to facilitate patient and physician participation, or adjustments to how these trials are conducted, so that the interruptions to the usual process of care are minimized. Various scenarios might be considered as far as when, in the course of care, computerized "prompts" would be presented to the clinician: during periodic reviews of new data, for example, vs. when a patient presents for *any* reason vs. when a patient presents with specific symptoms vs. when a prescription for one of the study drugs is written. These various scenarios have potentially very different implications for the acceptability of the trial

to the clinician. Research is ongoing in the UK GPRD to test the feasibility of "randomized trials within the database."[10]

The significance is that the *same* clinical data capture can be used as data capture for randomized (or observational) studies. The database provides part of the infrastructure for conducting a targeted study, thereby reducing time and costs while engaging the clinician in the process. This approach will require a variety of steps if it is to be feasible: modifying the existing infrastructure to produce new data collection tools and integrate those into existing systems, training clinicians to use the new systems, and perhaps even training clinicians in the principles of clinical epidemiology, so that they might better appreciate the value of research based in actual clinical settings.

The discussion so far has focused on data collection aimed at specific research questions. However, capturing the data at the physician–patient interface may have a number of other potential applications beyond contributing to larger trials. Data on individual physician practices and outcomes from specific encounters could form a foundation for single-physician-focused efforts that would allow practitioners to see and track what works and what doesn't, not just for their own patients but for all (similar) patients in the database, and to do a better job of understanding their own treatment decisions and the impact of those decisions. It is essential that data flow not be unidirectional. This concept of data flowing back to practitioners and informing future practice is captured in Figure 2-2. This is consistent with the views expressed by others (Etheredge, 2007; Stewart et al., 2007) and would be the foundation of a learning health system.

To overcome barriers generated by the structure of the healthcare system (particularly in the United States) and the dispersion of healthcare data, a broader view of data collection and integration is required. Comprehensive health records capturing *all* of the clinical encounters with a given patient will require an infrastructure to link across databases (issuing a unique patient identifier to enable this linkage), perhaps to include the personal health record maintained by the patient himself or herself. This generates more than just information technology (IT) requirements. For example, one *must* consider privacy concerns, particularly in the United States with its Health Insurance Portability and Accountability Act (HIPAA) regulations. Making links across hospitals, claims, electronic medical records, and other data sources will be best accomplished by using unique personal identifiers (e.g., social security numbers), which would require a reexamination of the country's HIPAA and protected personal information culture.

An investment in workforce and methods will also be required to realize the full benefit of these changes to the data and IT infrastructure.

[10] See http://www.gprd.com/home/default.asp (accessed September 22, 2010).

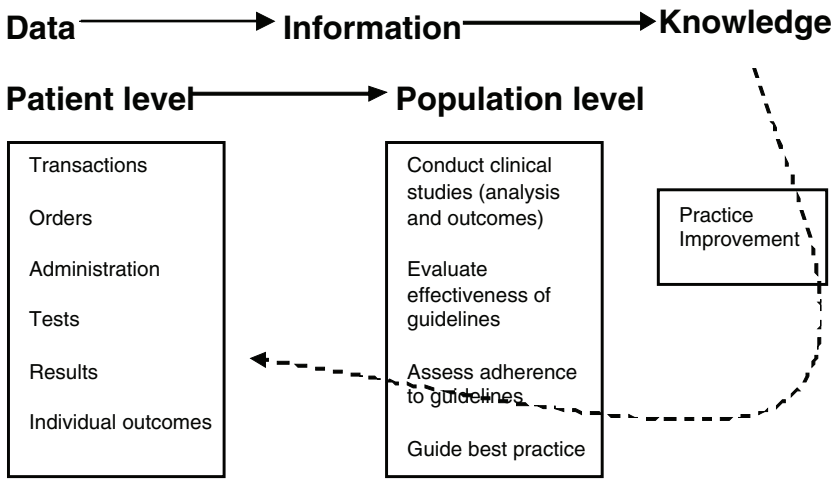


FIGURE 2-2 Data flow in a learning system.
SOURCE: Reprinted with permission from Anceta American Medical Group Association's Collaborative Data Warehouse.

Clinicians will need to be provided with more training to make sure they understand how their contribution of quality data directly affects the value of the information that they and others can retrieve. A variety of specialists, including informaticians, methodologists, and epidemiologists, will also be required to assure that there are continuing improvements to the systems.

Sharing of de-identified, individual-level, clinical trials data could provide an incredibly rich source of data to investigators. Again, however, doing so presents challenges because of privacy and informed consent concerns. Strict de-identification, required under HIPAA, is a time-consuming and costly activity (as opposed to the use of limited data sets from which certain key identifiers have been stripped). Safety analyses of erthyropoietin-stimulating agents are currently being conducted by the Cochrane Hematological Malignancies Group. For this effort, data have been provided by multiple sponsors of trials, including three pharmaceutical companies. There are other ongoing consortia of many types. For example, a Duke/FDA/Industry Cardiac Safety Research Consortium exists, the purpose of which is to create an electrocardiogram (ECG) library from clinical trials that could be used to identify early predictors of cardiac risk (Cardiac Risk ECG Library). Thus there is ample precedent for successful collaborative efforts across academic, government, and private sectors, with the success of each resting on the ability to amass data from a number of different sources.

The comments thus far have focused, at least implicitly, on pharma-

ceuticals, but it is important to remember that comparative effectiveness questions also arise with respect to medical devices. Currently, as far as is known, none of the administrative databases that are publicly available (usually for a fee) can distinguish among devices by manufacturer, so, for example, one may know a coronary stent was used, but much would be unknown, including which company made that stent and specific details around the individual stent (e.g., manufacturing lot number).

Summary

Data can be powerful and, one hopes, represent truth. Existing databases are currently most useful for paying bills and reflecting medical treatment. For research, the most useful applications of administrative data to date have been to support the analysis of potential safety issues. Although a number of statistical methods are available for controlling potential confounders, their contribution may be limited by the availability and accuracy of the required data. This lack of sufficiently detailed information on confounding factors is potentially a bigger problem than any challenges related to statistical methods when studying small-magnitude differences in the effectiveness of two therapies. Existing databases generally lack specific information on effectiveness, except when effectiveness can be represented in terms of clinical events.

Although not necessarily optimal, there are rational options that would improve the existing data collection system and its usefulness to both clinician and researcher. In particular, further development of a system at the point of care that records information, prompts the clinician for additional information, and provides insight and summarization back to the clinician should improve the quality of care by making critical data directly available when the clinician is making treatment decisions. This feedback loop is a critical link in the development of a learning health system that functions both for direct patient care and for the development of a research infrastructure that will help improve quality.

Future directions may include a mix of data quality and infrastructure efforts. Simple data-side adjustments that would improve the usefulness of these data (e.g., a "rule-out" flag that would signal more directly a clinician's intent rather than leave others to misinterpret a code as an occurrence, a hospital admission findings list, unique identification numbers on all implanted devices) would be smaller steps that could generate huge returns in the quality of patient care. The ultimate goal is to begin both building targeted studies and enhanced data capture capabilities into the framework of existing medical care databases as well as making information flow both from *and* to healthcare providers in a way that is immediately beneficial and effective in informing their care of the patient.

KNOWLEDGE SYNTHESIS AND TRANSLATION THAT NEED TO BE APPLIED[11]

Richard A. Justman, UnitedHealthcare

EBM is now the mantra of physicians, consumers, purchasers of health care, regulators, payers, and others who want to know which medical test or treatment works best. Unfortunately, an agreed-upon infrastructure to determine which treatment works best does not exist today. In fact, although everyone agrees that it is a good idea to follow the principles of EBM in deciding which treatment works best, there is no such agreement on the standards of clinical evidence to be used in a given situation. A national system for grading clinical evidence does not exist. This creates an interesting conundrum for persons who want answers to specific questions.

For example, a patient with localized prostate cancer seeking information on which treatment would be best for him is likely to obtain different answers depending upon whether he asks a primary care physician, a urologist, or a radiation oncologist. In fact, an asymptomatic 50-year-old man cannot even obtain definitive answers about whether he should be screened for prostate cancer with an inexpensive, readily available blood test for prostate-specific antigen.

Physicians seeking information based upon grades of clinical evidence face a similar dilemma. For example, microvolt T-wave alternans is an office-based test that predicts the risk of life-threatening cardiac arrhythmia. The American College of Cardiology grades the evidence supporting its use as grade 2a (Zipes et al., 2006). A physician seeking information about the use of bevacizumab to treat breast cancer will learn that the National Cancer Care Network drug compendium grades the evidence supporting its use as grade 2a. However, the definitions of grade 2a given by these two highly respected professional organizations do not match.[12]

[11] Note: This section is adapted from portions of *Knowing What Works in Health Care: A Roadmap for the Nation*, a report of the Institute of Medicine Committee on Reviewing Evidence to Identify Highly Effective Clinical Services (IOM, 2008).

[12] American College of Cardiology: "Evidence level IIa: conditions for which there is conflicting and/or a divergence of opinion about the usefulness/efficacy or a procedure or treatment. Weight of evidence/opinion is in favor or usefulness/efficacy" (Hunt, et al., 2005). National Comprehensive Cancer Networks (NCCN) Category 2A: "The recommendation is based on lower-level evidence, but despite the absence of higher-level studies, there is uniform consensus that the recommendation is appropriate. Lower-level evidence is interpreted broadly, and runs the gamut from phase 2 to large cohort studies to case series to individual practitioner experience. Importantly, in many instances, the retrospective studies are derived from clinical experience of treating large numbers of patients at a member institution, so NCCN Guidelines panel members have firsthand knowledge of the data. Inevitably, some recommendations must address clinical situations for which limited or no data exist. In these instances the congruence of experience-based judgments provides an informed if not confirmed direction for optimizing

Institute of Medicine Committee to Identify Highly Effective Clinical Services

In June 2006 the Robert Wood Johnson Foundation asked the Institute of Medicine to do the following:

- Recommend an approach to identifying highly effective clinical services.
- Recommend a process to evaluate and report on clinical effectiveness.
- Recommend an organizational framework for using evidence reports to develop recommendations on appropriate clinical applications for specified populations.

The imperative behind this request derives from the need to constrain healthcare costs, which have been rising faster than the consumer price index without a commensurate improvement in health outcomes; the need to reduce the idiosyncratic geographic variation in the use of healthcare services; the need to improve clinical quality, including health outcomes; the need to give consumers information they need to make healthcare choices; and the need to help purchasers of health care and payers to decide which services to include in their benefit designs.

Clinical Safety and Effectiveness: Current State

While there are multiple avenues available today to help consumers, physicians, and others decide which treatments are safe and effective, they all have significant limitations. In addition, the lack of a national comparative effectiveness architecture leaves the current state replete with gaps, duplications, and contradictions, as noted below (see also Table 2-3).

The FDA reviews the safety and effectiveness of medications. However, comparative effectiveness studies are not required as part of this review, so users lack information on which medication works best for a particular disease scenario. Another division of the FDA, the Center for Devices and Radiological Health (CDRH) reviews medical devices. The level of review performed by the CDRH is determined by the classification of the device. In any event, head-to-head trials and long-term outcome studies are not required in order to approve a medical device for marketing. Physicians commonly prescribe both medications and medical devices for other than the labeled indications ("off label" prescribing). This sometimes leads to

patient care. These recommendations carry the implicit recognition that they may be superseded as higher-level evidence becomes available or as outcomes-based information becomes more prevalent" (NCCN, 2008).

conflicting and confusing results. For example, bevacizumab is labeled for treatment of breast cancer. It is not labeled for treatment of age-related macular degeneration (AMD). Yet the evidence for improved vision in persons with AMD is stronger than the evidence of prolonged survival for women with breast cancer.

Multiple organizations perform systematic reviews of medical tests and treatments, including the Cochrane Collaboration; Hayes, Inc.; the ECRI Institute; and the Blue Cross Blue Shield Technology Evaluation Center. These respected organizations provide excellent systematic reviews of available clinical evidence. However, many of the topics they review are new or emerging technologies, so there is insufficient information from which valid conclusions can be drawn. Two of these organizations provide brief reports of new technologies, but the lack of available evidence makes these reports of limited use to persons who must make treatment decisions today.

Various agencies of the federal government have performed systematic reviews of available clinical evidence. Two such organizations, the National Center for Health Care Technology and the Office of Technology Assessment, did so in the past, but no longer do so. AHRQ now contracts with evidence-based practice centers to complete systematic reviews. These reports are uniformly excellent, but they address only a small number of topics about which consumers and physicians need to know. AHRQ also maintains the National Guideline Clearinghouse, a repository of clinical practice guidelines developed by other organizations. The Department of Veterans Affairs (VA), CMS, and the NIH provide information about the safety and effectiveness of medical treatments in different ways. However, they, too, lack the infrastructure to offer comprehensive information about the comparative effectiveness of health services.

Professional specialty societies and health plans, among others, develop clinical practice guidelines. Some of these guidelines provide excellent reviews of clinical evidence, noting both the strength of the evidence and the strength of the recommendations they make. Other guidelines are less fastidious in their review of evidence, relying instead on a consensus of expert opinion.

It is not clear how much is spent annually on researching the safety and effectiveness of medical treatments. However, it is likely to be less than $2 billion (IOM, 2007). This is considerably less than 1 percent of the total dollars spent annually on health care.

Deficiencies of the Current State of Comparative Effectiveness Review

As noted above, some systematic reviews of clinical evidence and some clinical practice guidelines lack scientific rigor, relying instead on a con-

TABLE 2-3 Duplicated Efforts by Selected Health Plans and Technology Assessment Firms, 2006

	Health Plans				Technology Assessment Firms		
Type of Service	United-Healthcare	Kaiser Permanente	Aetna	WellPoint	Hayes, Inc.	Technology Evaluation Center	ECRI Institute
Screening							
Genetic testing to predict breast cancer recurrence	√	√	√	√	√	√	√
Proteomic testing for ovarian cancer	√		√	√	√		√
Virtual (computed tomography [CT]) colonoscopy	√	√	√	√	√	√	√
Disease management							
Ambulatory blood pressure monitoring	√	√	√	√	√	√	√
Intermittent intravenous insulin therapy	√	√		√	√		√
Diagnosis							
CT angiography for suspected coronary artery disease	√	√	√	√	√	√	√
Microvolt T-wave alternans	√	√	√	√	√	√	√
Wireless capsule endoscopy	√	√	√	√	√	√	√

Treatment							
Brachytherapy for various cancers: breast, ovarian, and prostate cancer and brain tumors	√	√	√	√	√	√	√
Dysfunctional uterine bleeding and fibroids	√	√	√	√	√	√	√
Fallopian tube occlusion for permanent contraception	√	√	√		√		√
Growth factor–mediated lumbar spinal fusion	√	√	√	√	√		
Intracoronary brachytherapy	√	√	√	√	√	√	√
Minimally invasive surgery for low back pain	√	√	√	√	√	√	√
Photodynamic therapy for Barrett's esophagus and esophageal cancer	√	√	√	√	√		√
Vagus nerve stimulation for intractable depression	√	√	√	√	√	√	√
Devices							
Artificial total disc replacement for lumbar and cervical spine	√	√	√	√	√	√	√
Cochlear implants	√	√	√	√	√		√
Total artificial heart	√	√	√	√	√		√
Total hip resurfacing arthroplasty	√	√	√	√	√	√	√

NOTE: Not all reviews are comprehensive assessments. Agency for Healthcare Research and Quality evidence-based practice centers have reviewed 5 of the 20 topics listed (ambulatory blood pressure monitoring, CT angiography, proteomic testing for ovarian cancer, spinal fusion for low back pain, and uterine fibroids). The Kaiser Permanente entries represent all Kaiser regions.
SOURCE: IOM, 2008.

sensus of expert opinion rather than clinical evidence as the basis of their conclusions.

The body of clinical evidence for some health services in which consumers and physicians are interested may be weak or totally lacking. Two examples are the treatment of chronic wounds, an issue of vital importance to an aging population, and the nonsurgical treatment of uterine fibroids. Moreover, as noted earlier, there is no nationally agreed-upon method for rating clinical evidence or the strength of recommendations.

Bias and conflict of interest on the part of experts further complicate the understanding of the conclusions that can be drawn from available clinical evidence. Bias may be unintentional. For example, physicians trained in nonsurgical specialties may be more oriented toward medical treatments than surgical treatments. Actual conflicts of interest, such as physicians who are paid consultants to device manufacturers or pharmaceutical companies, make it difficult to discern whether the conclusions physicians draw about clinical evidence are truly independent conclusions.

Finally, the multiple clinical guidelines available for the treatment of the same condition frequently make differing recommendations. The Infectious Disease Society of America and the International Lyme and Associated Diseases Society make different recommendations concerning prolonged antibiotic use to treat neuroborreliosis (IDSA, 2006; ILADS Working Group, 2004). Similarly, different professional specialty societies related to allergies make different recommendations regarding the type of allergy immunotherapy to be used on atopic persons.

Recommendations by the Institute of Medicine Committee on Reviewing Evidence to Identify Highly Effective Clinical Services

To address the deficiencies noted above and in fulfillment of the charge to the IOM by the Robert Wood Johnson Foundation, a Committee on Reviewing Evidence to Identify Highly Effective Clinical Services (HECS) was convened. This committee was composed of persons with different perspectives on health care, including academic physicians, researchers, epidemiologists, health economists, consumers, payers, device manufacturers, physician groups, and others.

In *Knowing What Works in Health Care: A Roadmap for the Nation* (IOM, 2008), the committee made the following recommendations:

- Congress should direct the Secretary of Health and Human Services to designate a single entity to ensure production of credible, unbiased information about what is known and not known about clinical effectiveness.

- The Secretary of Health and Human Services should appoint a Clinical Effectiveness Advisory Board to oversee the Program.
- The Program should develop standards to minimize bias for priority setting, evidence assessment, and recommendations development.
- The Program should appoint a Priority Setting Advisory Committee to identify high-priority topics.
- The Program should develop evidence-based methodologic standards for systematic reviews, including a common language for characterizing the strength of evidence.
- The Program should assess the capacity of the research workforce to meet the Program's needs; if necessary, it should expand training opportunities in systematic review and CER methods.
- Groups developing clinical guidelines should use the Program's standards.
- An effort should be made to minimize bias by balancing competing interests, publishing conflict of interest disclosures, and prohibiting voting by members with material conflicts.
- Stakeholders should preferentially use clinical recommendations according to the Program's standards.

Role of a Priority Setting Advisory Committee

Anticipating significant demand for comparative effectiveness reports, a Priority Setting Advisory Committee (PSAC) will be necessary to review requests for clinical effectiveness information from stakeholders and to stratify them for review. To meet the needs of a large and varied group of stakeholders, priority setting must be open, transparent, efficient, and timely. Suggested criteria for priority setting include the improvement of health outcomes, reduction in the burden of illness and health disparities, elimination of undesirable variations, and reduction in the economic burden of disease and treatment of disease. Members of the PSAC should have a broad range of expertise and interests. Finally, the committee must be constituted in such a way as to minimize bias and to identify and address potential conflict of interest among committee members (IOM, 2008).

Conceptual Framework for Assessing Clinical Evidence

The IOM HECS committee recommended a hybrid approach to assessing clinical evidence that included developing a new infrastructure for evaluating evidence through systematic review and using existing structures for the development of clinical guidelines. Research study designs would include RCTs, cohort studies, case-control studies (identifying factors that may contribute to a medical condition by comparing a group of patients

who have that condition with a group of patients who do not), cross-sectional studies (observing some subset of a population of items all at the same time, so that groups can be compared at different ages with respect to independent variables), and case reviews. Systematic review of clinical evidence requires a common language to identify and assess the quality of individual studies, to critically appraise the body of evidence, and to develop qualitative or quantitative synthesis. Finally, clinical guidelines would be written using common language for stating the strength of clinical evidence and the strength of recommendations (IOM, 2008).

A systematic review investigates the characteristics of the patient population, care setting, and type of provider; the intervention, including the route of administration, dose, timing, and duration; the comparison group; outcome measures and timing of assessments; quality of the evidence (risk of bias, sample sizes, quantitative results and analyses, including an examination of whether the study estimates of effect are consistent across studies); and the effect of potential sources of study heterogeneity, if relevant (IOM, 2008).

Infrastructure requirements include funding, staffing, and stakeholder involvement; common agreement on the hierarchy of clinical evidence; a body of evidence sufficient to allow for systematic review; capability of satisfying the needs of different stakeholders, including consumers, physicians, purchasers of health care, and payers, among others; a repository for evidence reports; a process for systematic updating and revision; and an infrastructure sufficient to compare different treatments (IOM, 2008).

Review of Emerging Treatments

All of the foregoing assumes that questions will be directed toward health services sufficiently mature to accumulate a body of clinical evidence sufficient for systematic review. However, it is likely that many questions will arise about new or emerging treatments for which a significant body of published clinical evidence does not exist. The committee recommends that an infrastructure be developed allowing for brief reports that address what is known about emerging treatments, identify the salient questions that must be addressed, acknowledge the gaps in evidence, and articulate the opportunities for future research (IOM, 2008). Such brief reports would be also be useful in addressing rare diseases and conditions, treatments for conditions with no known effective treatment, and health services for which comparative effectiveness trials are unlikely to be completed in the short term.

Some avenues to synthesize evidence in order to address such health services include clinical trials and CMS Coverage with Evidence Devel-

opment. Other protocol-specified prospective data collection programs address similar questions.

Examples of new or emerging treatments worthy of such review are easy to identify. They include accelerated partial breast irradiation to accompany lumpectomy in the treatment of localized breast cancer, bevacizumab to treat age-related wet macular degeneration, endovascular repair of thoracic aortic aneurysm, intracerebral stenting for the prevention of stroke, endobronchial values as an alternative to lung volume reduction surgery to treat emphysema, bronchial thermoplasty to treat moderate-to-severe asthma, and injectable bulking agents to treat vesicoureteral reflux in children.

Final Comments

Given the amount spent on health care in the United States, consumers of health services, professionals who provide those services, and purchasers who pay for them are entitled to know what works and what does not. They are entitled to know which health services are definitely beneficial, which are likely to be beneficial, which have insufficient evidence supporting their use to know if they are beneficial, and which services in common use today are known to be of no benefit or, worse, that are actively harmful. Persons making choices on which treatments to use should understand the range of treatments available to them, including advantages, harms, and alternatives. However, despite the plethora of information available today, such a "single source of truth" does not exist. The foregoing comments represent one attempt at defining the knowledge synthesis necessary to answer these vital questions.

METHODS THAT NEED TO BE DEVELOPED

Eugene H. Blackstone, M.D., Douglas B. Lenat, Ph.D., and Hemant Ishwaran, Ph.D.
Cleveland Clinic

Overview

RCTs and their meta-analyses are generally agreed to provide the highest-level evidence for comparative clinical effectiveness of clinical interventions and care. However, today cost and complexity impede nimble, simple, inexpensive designs to test the numerous therapies for which a randomized trial is well justified. Further, it is impossible, unethical, and prohibitively expensive to randomize "everything."

To fill this gap, balancing-score methods coupled with rigorous study design can approximate randomized trials. They are less controlled but

use real-world observational clinical data. They may provide the only way to test therapies when it is impossible to conceive of or conduct RCTs. Although a number of their important features remain to be understood and refined, they are comparatively inexpensive and use readily available electronically stored data. Interestingly, although the intent of EBM is to reduce practice variance, this methodology draws its power from heterogeneity of care.

Unfortunately, a longitudinal birth-to-death patient-centric health record, populated largely with discrete values for variables that would be useful for both streamlined randomized and balancing-score-based clinical trials, has not been brought to fruition. Instead, clinical information remains locked in narrative, mostly within segregated institutional silos. But a new methodology is emerging both to elegantly link these silos and to provide a population-centric view of clinical data for analysis: semantic representation of data. Meaning is emphasized rather than lexical syntax. This has the promise of transforming EBM into information-based medicine. Its elements include storage of patient data as nodes and arcs of graphs that can seamlessly link disparate types of data across medical silos, from genomics to outcomes, and, in theory, across venues of care to create a virtual longitudinal health record, to say nothing of the completely longitudinal personal health record. What is required are (1) a rich ontology of medicine, the taxonomy component of which is enough to enable semantic searching, and the formal knowledge base component, which is enough to permit—even today—natural language query of complex patient data (that is, separating logical understanding of query from the need to understand underlying data schemata); (2) a worldwide effort to assemble this ontology and the assertions that make it useful; and (3) intelligent agents to assist discovery of unsuspected relationships and unintended adverse (or surprisingly beneficial) outcomes.

But if such clinically rich data were available, especially a massive amount, could they be put to effective use? Computer-learning methods such as bootstrap aggregation (bagging), boosting, and random forests are algorithmic, as opposed to the traditional model-based methods that are computationally fast and can reveal complex patterns in patient genomic and phenotypic data. These methods refocus attention from "goodness of fit" to a given set of data to prediction error for new data. Methods like this are needed to propel the country yet another step toward personalized medicine.

Thus the results of trials, approximate trials, and automated discovery need to be transformed from static publications to dynamic, patient-specific medical decision support tools (simulation). Although such methodologies are widely used for institutional assessment and ranking, they need to lead

to clinically rich, easily used, real-time tools that integrate seamlessly with the computer-based patient record.

This article highlights five foundational methodologies that need to be refined or further developed to provide an infrastructure to learn which therapy is best for which patient. They are representative of those needed for progression from current siloed EBM to semantically integrated information-based medicine and on to predictive personalized medicine. The five methodologies can be grouped into three categories:

1. Evidence-based
   - Reengineering RCTs
   - Approximate RCTs
2. Information-based
   - Semantically interpreting, querying, and exploring disparate clinical data
   - Computer learning methods
3. Personalized
   - Patient-specific strategic decision support

Reengineering Randomized Controlled Trials

Following intense preliminary work, several cardiac surgical centers began designing a randomized trial to answer a simple question: Is surgical ablation of nonparoxysmal atrial fibrillation accompanying mitral valve disease effective at preventing the return of the arrhythmia? It took a short time—weeks—to design this study, but then it had to be vetted through committees, review boards, and the FDA, leading to multiple revisions, additions, and mounting complexity. The case report form became extensive and required considerable human abstraction of information from clinical records to complete. Two core laboratories were needed and competitively bid. After more than 2 years, the trial was launched. From inception to completion, the trial is likely to take 5 years at a minimum. The cost of what was intended to be a simple, easily deployed trial will be about $2 million; large multi-institutional, multinational trials may cost upwards of 10 times this figure.

Designing and executing RCTs like this has become one of the most demanding of human feats. It may not compare with climbing Mt. Everest, but it is close. A major reason to climb this mountain is that RCTs remain the gold standard for EBM. They are purpose designed, have endured ethical scrutiny, ensure concurrent treatment, capture highest-quality data, and have adjudicated end points. Their data meet the statistical assumptions of the methods used to analyze them.

Yet, like the trek up Everest, the design and conduct of an RCT is filled with pitfalls that need to be bridged. The following six areas are among those that must be addressed if RCTs are to achieve the kind of cost effectiveness that evidence-based medical practice requires in the future: complexity, data capture, generalizability, equipoise, appropriateness, and funding.

Complexity

A deep pitfall of the current practice of RCTs is what John Kirklin, pioneer heart surgeon, called "the Christmas Tree Effect": ornamenting trials with unnecessary variables rather than keeping them elegantly simple and focused. Every additional variable increases the cost and difficulty of the trials, which reduces available resources, limiting the number of trials that can be performed. Nonessential complexity constructs a barrier to progress when instead a bridge is needed. In reengineering RCTs, data collection should be focused on the small number of variables that directly answer the question posed. A series of elegant, scientifically sound, clinically relevant, simple, focused trials will provide more answers more quickly than bloated multimillion-dollar trials that are justified as providing enormous riches of high-quality data for later (observational) data exploration.

Second, rapid development of simple pilot trials on clinically important questions should be encouraged, to be followed with simple, definitive trials. The National Heart, Lung, and Blood Institute has put into place a number of disease- and discipline-specific networks of centers devoted to simple RCTs. This is an important step forward. Two observations: (1) The trials being designed are simple only in the number of patients enrolled, not in design; funding would be better spent on highly focused, extremely simple RCTs. (2) There is no plan for funding pivotal trials based on clinical outcomes rather than surrogate and composite end points that stem from these pilot trials (Fleming and DeMets, 1996). Perhaps the focus should, therefore, shift to funding a mix of simple, inexpensive pilot trials and simple but definitive trials.

Third, adding administrative and bureaucratic complexity to many RCTs is needed for investigational device exemptions and new drug exemptions from the FDA. This introduces considerable delay by an organization that should itself promote efficient study designs focused on safety and efficacy. The heterogeneity of institutional review board requirements adds further administrative burden.

Fourth, to "survive," design and conduct of RCTs has become a "business" that is increasingly specialized and complex and distanced from the practice of medicine. Physicians with good questions believe they cannot attempt to scale the mountain. It was not always this way, and patient

recruitment suffers from it because patients' personal physicians are often no longer advocates for clinical trials. Again, simplification is key to bridging this chasm.

All four of these complexities argue for applying a kind of *symbolic sensitivity analysis* when an RCT is designed, eliminating variables that are more decorative than functional.

Data Capture

RCT technology as practiced today makes little use of discrete data elements acquired as part of clinical practice. Available computer-based clinical data could and should be used for patient screening, recruiting, and data gathering. With electronic patient records composed of "values for variables" (discrete data elements), one could electronically identify patients meeting eligibility criteria for trials, generating alerts so that healthcare providers could be on the front line of informing patients about a trial germane to their treatment. Insofar as possible, patient data, including end points, should be retrieved directly from the electronic patient record. Instead, study coordinators today laboriously fill out case report forms, translating from medical records. Reducing the data-gathering burden would not only reduce complexity and cost but also bring trials more into the sometimes messy reality of clinical practice—the very environment for which inferences about clinical effectiveness from the trial are to be made. Admittedly, redundant data abstraction, end point adjudication, and core laboratories all contribute to incrementally improving the quality of trial data, but it is questionable whether the improvements justify the accompanying costs, their impeding the climb, and permitting more climbs.

Generalizability

RCTs often focus on patient subgroups (usually the lowest-risk patients, ostensibly to reduce potential confounding and for which equipoise is unquestioned, rather than the spectrum of disease observed in the community (Beck et al., 2004). Yet results of these restrictive trials typically are extrapolated to the entire spectrum, a practice that may be treacherous no matter what the trial shows. One of the first large, costly trials sponsored by the NIH was the Coronary Artery Surgery Study of the late 1970s and early 1980s (NHLBI, 1981). About 25,000 patients were entered into a registry of patients with coronary artery disease, but only 780 were randomized (Blackstone, 2006). Yet treatment inferences from the study were applied to a broad spectrum of patients with coronary artery disease (Braunwald, 1983). Although it can be argued that pilot studies should be conducted in the patient subgroups most likely to demonstrate a treatment difference

(so-called enriched trials), these studies should be used to aid developing inclusive trials of adequate power. Just as the data acquired from clinical practice is often taken too lightly today, the data acquired from these restricted RCTs is often taken too seriously, when in truth both of these turn out in hindsight to be no more—and no less—than valuable heuristics (Ioannidis, 2005).

Equipoise

Among physicians' areas of expertise and responsibility is the task of selecting the right treatment for the right patient at the right time. Surgeons call this "indications for operation." This is the antithesis of equipoise. Thus, a number of important trials have been stopped or considerably protracted for lack of enrollment. Across time periods, nationalities, and schools of thought, each physician will follow his or her own generally consistent but somewhat idiosyncratic set of rules for deciding appropriate treatment. Thus, whenever one examines clinical practice, considerable variance is seen. This gives hope that equipoise on important medical dilemmas might be found at times. However, it also suggests the possibility of capitalizing on practice heterogeneity to conduct studies that approximate RCTs, as described later in this text, rather than seeking artificial, unnatural equipoise.

Appropriateness

Most investigators developing RCTs concentrate on efficacy. Studies are powered for anticipated (often overly optimistic) efficacy, but rarely focus on short- or long-term safety. This is even true of trials conducted for FDA approval. Indeed, for cardiovascular devices, the track record of mandated FDA safety surveillance is dismal. It usually involves a small cohort of patients for whom there is little power to detect increased occurrence of adverse events, and it generally employs a follow-up time too short to detect untoward effects of long-term device implantation. Rare adverse effects caused by long-term exposure to devices (or pharmaceuticals) may go undetected for a long time, but when they are finally detected they incite public anger, recalls, and withdrawals of effective drugs and devices (Nissen, 2006). This reaction might be avoided if a proper surveillance program were in place with impartial analysis of data, possibly assisted by the computer learning technology discussed later in this paper. The factual reporting of findings and a measured response could convince industry, the public, regulators, and even skeptics that the process is transparent and timely (Blackstone, 2005).

Are all the clinical trials that are being performed actually necessary?

Just because a trial can be mounted is no reason to initiate inappropriate trials. At the end of the appropriateness scale is the proverbial parachute trial. Not only will there not be a randomized trial of efficacy of parachutes, there is no compelling reason to do such a trial; magnitude of the effect is too large and logically obvious, although we concede that logic can trip us up. Many trials are expected at the outset to show no difference in efficacy, and yet futile trials are done, often because a regulatory body has required it. Many equivalency, nonsuperiority, and noninferiority trials could be replaced by objective performance criteria and an intense surveillance program (Grunkemeier et al., 2006).

Funding

Typically, the costs of new pharmaceutical and device trials are borne by industry sponsors, with their attendant actual and potential conflicts of interest. Relative to this, only a small number of trials are sponsored by the NIH. Yet in an evidence-based medical system, the obvious benefactors are health insurers, and to a lesser extent the pharmaceutical and device manufacturers. Shouldn't insurers be interested in sponsoring clinically relevant RCTs, including making data available from the trials to the scientific community or at least bearing the patient costs of RCTs?

Approximate Randomized Clinical Trials

What effect does chronic exposure to urban pollution have on the risk of developing pulmonary disease or cancer? What is the effect of socioeconomic status on response to therapy? What is the effect on long-term outcomes of complete versus incomplete coronary revascularization? What is the effect of chronic atrial fibrillation on stroke? Can severe aortic stenosis be managed medically rather than surgically? Is the radial artery a good substitute for the right internal thoracic artery for bypass of the circumflex coronary system? These are but a few questions for which an evidence basis is needed. Some may be answerable with cluster randomized trials (Donner and Klar, 2000). Others require epidemiologic studies, and none seem readily amenable to randomized trials. It is not possible to randomize gender, disease states, environmental conditions, choice of ancestry, or healthcare organizations in local communities. It would be unethical to randomize patients to placebo or to incomplete or sham surgery when at least knowledge at the present time, if not solid data, indicates that to do so is unsafe. Thus, there is no knowledge in the modern era about the untreated natural history of certain diseases, such as critical aortic stenosis, hypoplastic left heart syndrome, transposition of the great arteries, untreated renal failure, unset fractures, untreated acute appendicitis, or jumping out of an airplane

without a parachute. Yet clinical decisions are made on incomplete evidence or flawed logic every day. Is it possible to do better than guessing? Is there an alternative to "randomizing everything?"

When literature comparing nonrandomized treatment groups is scrutinized, the natural response is to think, "They are comparing apples and oranges" (Blackstone, 2002). This is because in real-life clinical practice there remains wide variance in practice (that is, selection bias), and this results in noncomparable groups. If it is impossible to randomize patients or impractical or unethical, or if it can be demonstrated that one cannot draw a clean, causal inference even from a randomized trial (such as a trial that inextricably confounds treatment with the skill of the person implementing the treatment), is there a way to exploit the heterogeneity of clinical practice to make better comparisons that are closer to apples to apples? Basically, the goal would be to discover within the heterogeneity of practice the elements of selection bias and account for these to approximate a randomized trial.

A quarter century ago, Rosenbaum and Rubin (1983) introduced the improbable notion that observational data can be thought of as a broken randomized trial (Rubin, 2007), with an unknown key to the treatment allocation process. They proposed that the propensity of a patient to receive treatment A versus B be estimated statistically (for example, by logistic regression) to find that key. In its simplest form, a quantitative estimate of propensity for one versus the other treatment is calculated for each patient (propensity score) from the resulting statistical analysis and used for apples-to-apples comparisons (Blackstone, 2002; Gum et al., 2001; Sabik et al., 2002).

How does a single number, the propensity score, seemingly magically achieve a balance of patient characteristics that makes it appear as if an RCT had been performed (for that is exactly—and surprisingly—what it does)? It does so by matching patients with similar propensity to receive treatment A. A given pair of propensity-matched patients may have quite dissimilar characteristics but similar propensity scores. A set of such pairs, however, is well matched (Figure 2-3). What distinguishes these patients from those in an RCT is that at one end of the spectrum of propensity scores, only a few who actually received treatment *A* match those who actually received treatment *B,* and at the other end of the spectrum, only a few patients who actually received treatment *B* match those who received treatment *A*. Thus, balance in patient characteristics is achieved by unbalancing *n* along the spectrum of propensity scores (Figure 2-4). The generic idea is called *balancing score* technology, which can be extended from two treatments to multiple treatments, or even to balance continuous variables, such as socioeconomic status or age (Rosenbaum and Rubin, 1983).

Unlike an RCT in which the allocation mechanism (randomization)

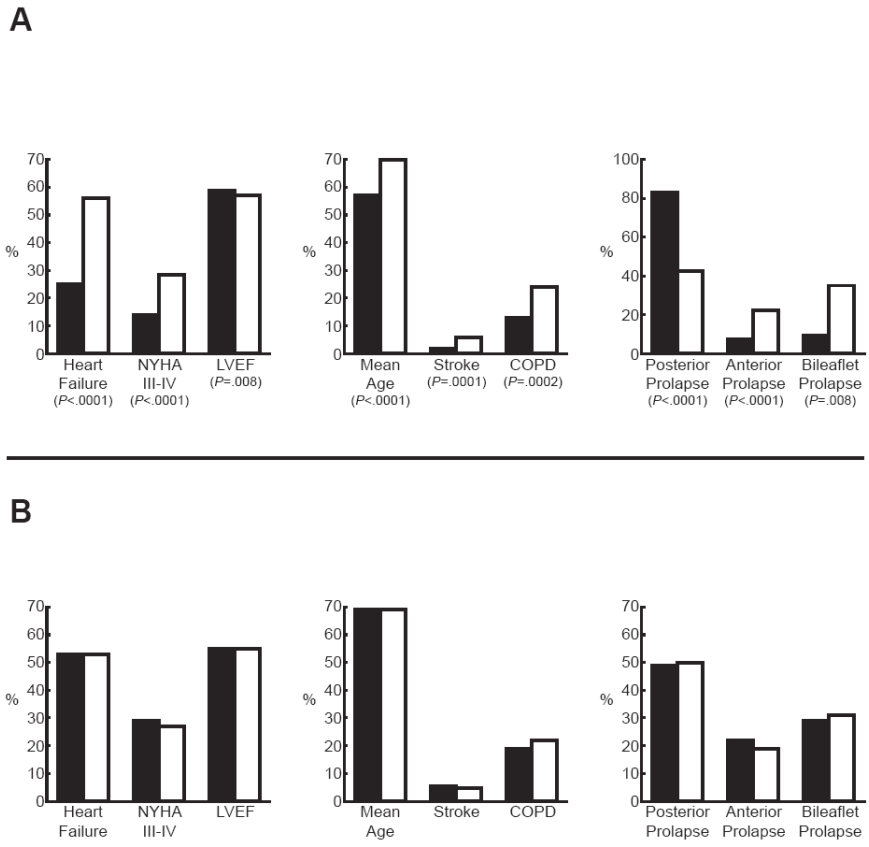


FIGURE 2-3 Comparison of patient characteristics before mitral valve repair (black bars) or replacement (unshaded bars). Unadjusted values are depicted in ***A*** and propensity-matched patients in ***B.***
NOTE: COPD = chronic obstructive pulmonary disease; LVEF = left ventricular ejection fraction; NYHA = New York Heart Association.

is known explicitly and equally distributes both known and unknown factors, propensity score methods can at best account for only those selection factors that have been measured and recorded, not for those that are unknown. Thus claims of causality, which are strong with RCTs, are weaker with propensity-based methods. This considerable disadvantage is, however, offset in a number of ways: (1) innumerable treatments can be studied at low cost based on heterogeneity of practice and availability of clinically rich data and (2) treatments or characteristics that cannot be randomized (e.g., gender, place of birth, treating facility, presence of

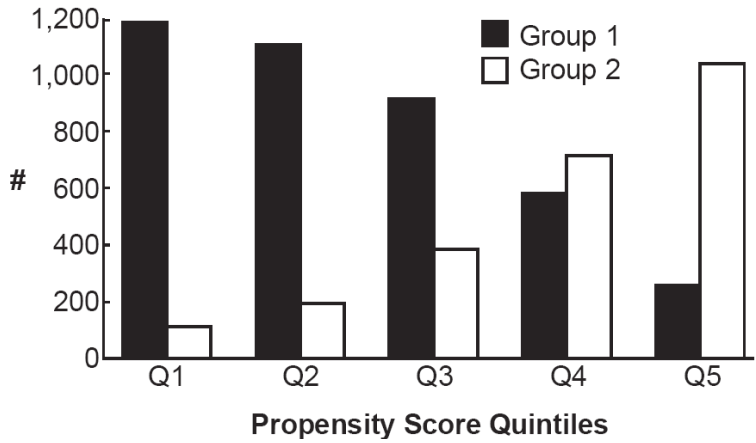


FIGURE 2-4 Achieving balance of clinical features by unbalancing *n*. Shown are two groups of patients that have been divided according to increasing quintile of propensity score. Notice at low propensity scores, the numbers of group 1 patients dominate over those of group 2, and at the other extreme, the numbers of group 2 patients dominate over those of group 1. Within each quintile, patient characteristics are well matched between groups, but these characteristics progressively change across quintiles (for example, low-risk profile in quintile 1 and high-risk profile in quintile 5).

disease) can be analyzed. Thus, there is broad applicability for a relatively inexpensive method.

It is important to say, however, that relying on clinical practice data alone is potentially irresponsible, biased, and dangerous, much like standing on untested terrain that may turn out to be thin ice, and patterns may turn out later to be artifactual "false peaks." However, these techniques may play a valuable role as a *heuristic* for helping to point RCTs in promising directions when that is possible and as better evidence than apples-to-oranges comparisons when it is not.

Taking yet another step backward, it has been claimed that traditional multivariable analysis is equally accurate in making risk-adjusted nonrandomized comparison (Sturmer et al., 2006). The problem, however, is that until now, there has been no independent support for this claim. It may be right more than 80 percent of the time, but what about the other 20 percent? Propensity-based methods provide this independent assessment. In addition, they also permit comparison when important clinical outcomes occur at a low frequency by supplying a single risk-adjustment variable: the propensity score (Cepeda et al., 2003).

Propensity methodology (and balancing scores in general) should be

elevated further. First, because propensity models are predictive ones (predicting which treatment was selected), the computer learning approach presented later in the text could be exploited to account for possibly complex interactions among selection factors. Second, comparisons based on clinically rich vs. administrative vs. electronically available laboratory databases should be tested for relative value. Third, the most appropriate method of comparing outcomes after propensity matching remains controversial and probably requires developing new statistical tests.

Semantically Interpreting, Querying, and Exploring Disparate Clinical Data

Computerized Patient Records

In 1991 the Institute of Medicine described what it called the computer-based patient record (CPR) (Barnett et al., 1993; IOM, 1991). Its creators envisioned a birth-to-death, comprehensive, longitudinal health record that contained not just narrative information but also values for variables (discrete data) to allow the record to be active, generating medical alerts, displaying trends, providing meaningful patient-level clinical decision support, and facilitating clinical research. It would not be simply an electronic embodiment of the paper-based medical record, which is what they believed the emerging electronic medical record (EMR) was.

The need for a CPR is, if anything, more acute today than it was in the early 1990s because of the increased complexity of care, the aging of the population with multiple chronic diseases, and the multiplicity of care venues from shopping malls to acute care facilities to clinics to large hospitals, to say nothing of OTC medications and a proliferation of alternative and complementary therapies, public awareness of clinical outcomes, the need to track unanticipated complications of therapy across time, and a cumbersome built-in redundancy of clinical documentation for reimbursement.

The originators recognized most of the same impediments to implementing such a system as still exist, not the least of which was that medical education would need to be altered to train a new generation of physicians how to use this new technology optimally.

What has not been clear to developers of EMRs is how discrete data might provide the underpinnings for a learning medical system. Nor did those who were willing to give up pen and ink and adopt the electronic record demand discrete data gathering as a by-product of patient care. Thus, before describing various methods that exploit discrete medical data, it is important to ask why discrete data is an asset and envision what could be done with this asset. For individual care, discrete data can be used to generate smart alerts based on the real-time assessment of data by algo-

rithms, care plans, or models developed on the basis of past experience. For informed patient consent, patient-specific predicted outcomes of therapy can be displayed based on models that are risk adjusted for individual patient comorbidities and intended therapeutic alternatives (see the later section, "Patient-Specific Decision Support"). From a population-centric vantage point, discrete data can provide outcomes and process measures for quality metrics and necessary feedback for improving patient care. This is in part because discrete data can make possible the automated compiling of quality outcomes and process measures along with variables needed for proper risk adjustment. Discrete data assist institutions in responding to clinical trials eligibility specifications to determine feasibility of studies and provide historical outcomes for estimating sample sizes. In addition, discreet data could alert physicians that a patient being seen satisfies all eligibility criteria for a clinical trial. Discrete data coupled with an intelligent query facility can be used to identify patient cohorts for observational clinical studies and approximate clinical trials. They provide the observational data for developing propensity scores, balancing scores, and conducting studies of comparative clinical effectiveness. If a true longitudinal record is created, then discrete data may identify adverse events and the substrate by which unsuspected correlated events may be identified, quite possibly with the use of artificial intelligence and computer learning techniques.

Computer-Based Patient Record Efforts at the University of Alabama at Birmingham

Kirklin and Blackstone, then at the University of Alabama at Birmingham (UAB), recognized the formidable barriers to the CPR and in October 1993 embarked on a $23 million proof-of-concept CPR in partnership with IBM. Initially, they sought an object model of medicine. Two simultaneous efforts to accomplish this resulted in the same conclusion: There is no object model of medicine because "everything is related to everything." Requirements for complex relationships coupled with the extensibility needed to keep pace with rapid medical advancement, assimilation of disparate types of data, provisions for examining data from multiple vantage points (e.g., viewing diabetes from the vantage points of genetics, anatomy, endocrinology, laboratory medicine, pharmacology, and other medical perspectives), and the feeding of computer systems without slowing patient care were huge challenges. IBM brought to the table experts in a host of different types of databases and concluded that nothing existed that would satisfy the IOM's vision of an active CPR. However, a novel vision for a system emerged from the collaboration that would be infinitely extensible, self-defining, active, secure, and fast (response time less than 300 milliseconds to those using the system clinically). It required that the container hold-

ing the data know nothing of its content and thus be *schemaless*. Rather, values for variables themselves needed to be surrounded by their context (metadata) (Kirklin and Vicinanza, 1999). Such a system was built on the IBM-transaction processing facility platform, the same as used at that time by airlines and banking. Its major unsolved problem, however, was cross-patient (population centric vs. patient centric) queries: In theory, an infinitely extensible, comprehensive, centralized data store could take an infinite time to query.

Semantic Representation of Data and Knowledge

Meanwhile, computer scientists at Stanford University (Abiteboul et al., 1997) and the University of Pennsylvania (Buneman et al., 2000) were developing methods to query semistructured (schemaless) data stored as directed acyclic graphs (DAGs). We recognized that the storage format of our UAB data could also be considered DAGs and be queried by the techniques those investigators were developing. Blackstone's move to Cleveland Clinic in late 1997 provided the opportunity to pursue development of the CPR, but in the test bed of a highly productive cardiovascular clinical research environment. Clinical researchers know, of course, that discrete data are required for statistical analysis, and for the preceding 25 years, human abstractors at the clinic had laboriously extracted data elements from narratives for every patient undergoing a diagnostic or interventional cardiac procedure, resulting in the Cardiovascular Information Registry. We also found that other investigators at the clinic had developed more than 500 clinical data registries, often containing redundant, unadjudicated, non-quality-controlled data about various aspects of medicine—even of the same patient—stored in disparate clinical silos, such as orthopedics, cardiology, oncology, and ophthalmology. For the most part, these registries did not communicate with one another.

We therefore continued our work in developing what we then called a *semantic database* that, like any DAG representation, could be extended infinitely, was self-defining, and was also self-reporting by use of intelligent agents. Some 15 years and $50 million later, we at last have a technology that can underlie an extensible multidisciplinary CPR without the need for special integration, because it is natively integrated. Each data element in such a system is a node or an arc that connects nodes (databases) in a graph resource description framework), along with context and meaning (knowledge base). Additional nodes represent medical concepts and these are all linked. Each node has an address just like an Internet in a thimble. The Internet analogy is not an empty one. The infrastructure for the World Wide Web (Cleveland Clinic is 1 of some 400+ organizations worldwide that make up the World Wide Web Consortium) is the prime example of a

container that is ignorant of content, has all the properties of a DAG, and can easily be extended to assimilate new concepts that have never before entered the mind of humankind. Our test data set for cardiovascular surgery contains 23 million nodes (terms) and 93 million relationships (statements) representing 200,000 patients.

What are the advantages of such graph structures besides infinite extensibility? First, medical taxonomies, such as those of Systematized Nomenclature of Medicine (SNOMED) (Schulz et al., 2009) or the National Library of Medicine's metathesaurus (UMLS [unified medical language system]) (Thorn et al., 2007), underlie the data model and enable semantic searches. An investigator can search for patients and their data without knowing anything about underlying data structure. Specifically, this is achieved by separating semantics from the underlying syntax, in much the same spirit as the vision for the semantic web (Berners-Lee et al., 2001). Rather than being confined to lexical searches for information, a semantic web search is based on meaning. An example of this is the contrast between a dictionary based on meaning, such as the *American Heritage Dictionary* (Pickett, 2000), and one based on lexical definitions, such as *Merriam-Webster* (*Merriam-Webster Dictionary*, 2004). Thus, a heart attack, myocardial infarction, MI, acute myocardial infarction, AMI, and the variety of ways this medical concept may be expressed in both language and specific idiosyncratic syntax in a given database, are all recognized as a meaningful single semantic concept. Conversely, when the meaning of a term (such as *myocardial infarction*) *changes*, there is no semantic confusion because at the semantic level those are separate terms (Thygesen et al., 2007). There is a many-to-many mapping between lexical terms and their semantic denotations; the latter are the loci of medical knowledge.

Second, patients' graphs are connected by a data model to both general and medical *ontologies*, not just controlled term lists or taxonomies. These ontologies are built on a skeleton of taxonomically arranged concepts, but they contain as many—and as sophisticated—assertions *about* those concepts as are needed to compose an adequate model of an area of practice (Buchanan and Shortliffe, 1984). Think of an orthopedic ontology: It contains not only a taxonomy of all the bones in the body, but also assertions about them, such as "the knee bone's connected to the thigh bone, and the thigh bone's connected to the hip bone" (Weeks and Bagian, 2000), the type of joints between them, relative sizes, and so on.

Third, because natural language queries that seem clear to human investigators are fraught with ambiguous terms and grammatical constructions (e.g., attachment of prepositional phrases), pronouns, elisions, and metaphors, the knowledge represented in rich ontologies (vs. a taxonomy) suffices—barely—to permit investigators to ask database questions in natural language rather than in the language of a database expert. For the last

few years, Cleveland Clinic has collaborated with Douglas B. Lenat and his group in Austin, Texas, who, for the last 24 years, have built a top-down ontology of general concepts that starts with "thing" at the top and goes all the way down to such domain-specific concepts as "kidney" and "dialysis," and millions of general rules and facts that interrelate and, therefore, partially define those terms and model a portion of human knowledge (Lenat and Guha, 1990). Not surprisingly, to cope with divergence across humans' models of the world, that ontology—Cyc—required its knowledge base to be segmented into locally consistent (but *only* locally consistent) contexts. Since 2007 a group of us from Cleveland Clinic and Cycorp have worked together to tie low-level medical ontology concepts to the general Cyc ontology of things.

An investigator can now type into a Semantic Research Assistant™ a simple English sentence such as "Find patients with bacteremia after a pericardial window." Although complete automatic parsing of realistically large and complex investigator queries is still far beyond today's state-of-the-art artificial intelligence software (Lenat, 2008), one thing that is possible today, and which the current system does, is to successfully extract entities, concepts, and relations from the text as it understands the meaningful fragments of the query. These fragments are understood as logical clauses (in the system's formal representation), each of which is translated into a short, comprehensible English phrase and presented to the investigator. The investigator selects those fragments believed to be relevant, at which time an amazing thing happens almost every time: There is only a single semantically meaningful *combination* of those fragments, and only a single query that makes sense, given common sense constraints, domain knowledge constraints, and discourse pragma. *Combining* the fragments entails, for example, deciding which variables from each fragment unify with variables from other fragments, or whether they represent separate entities, and deciding whether each variable should be quantified existentially or universally, and in what order. The full query is then assembled, an English paraphrase of it is presented to the investigator, and a SPARQL translation of it is presented to the semantic database, which returns answers that are displayed to the investigator. Often, in the course of this process, some clauses that were not explicitly included by the investigator can be suggested; at other points in the process, the investigator may tweak the query by replacing a term with one of its generalizations or siblings or descendants in the ontology.

Fourth, a semantic-ontology approach also permits truly intelligent patient search of medical concepts. This is becoming increasingly important as patients seek out information about their medical conditions. A patient might type into a medical semantic search engine, "I have a racing heart." The semantic search engine produces a number of hits that don't include

NASCAR racing but rather tachycardias, such as atrial fibrillation, presenting the patient with definitions and treatment options.

What now needs to be developed to implement semantic databases and knowledge bases for intelligent search of all of medicine is a comprehensive formal ontology of medicine. This will require a worldwide effort. Already some of this is going on. For example, the Cardiovascular Gene Ontology provides full annotation for genes associated with cardiac disease processes.[13]

In the future such systems may actively ask relevant questions about correlations and trends within longitudinal records by means of true artificial intelligence. Automated intelligent agents could assist in discovering unsuspected relationships, unintended adverse outcomes, and surprising beneficial effects (AAAI, 2008). It could be central to realizing a learning medical system, a key component of what 21st-century medicine must become.

Computer Learning Methods

As much as one can dream of a longitudinal database that might permit innovative research for information-based medical care, it is important to ask, "If we actually had these data, would we know what to do with them?" One useful way to look at the issue is to use a "trees and woods" analogy in which individual patients, their data, and their genes are like the individual trees, and groups of patients or populations are the woods (Blackstone, 2007). The expression "Ye can't see the wood for the trees" (Heywood, 1546) implies that there may be patterns in the wood that can be discerned by overview that are not visible by attention paid only to individual trees. Here is an example: If one sits on the sidelines of an Ohio State University football game, one can only see individual band members playing at half-time and their feet moving around. But from an aerial view, one can see the band is in formation spelling the word *Ohio*. Patterns in medical data represent the general ways that patients react to their disease or treatment. They are the incremental risk factors, the modulators, or the surrogates for underlying disease and treatment mechanisms (Kirklin, 1979).

The rapidly developing science of computer learning promises methods far more robust than traditional statistical methods for discovering these patterns (Breiman, 2001). Many of them are based on multiple bootstrap samples (Diaconis and Efron, 1983; Efron, 1979, 1982; Efron and Tibshirani, 1986), each of them analyzed and aggregated (Breiman, 1996). This can be illustrated by analyzing 15 potential risk factors for death after

[13] See http://www.geneontology.org/GO.cardio.shtml (accessed September 8, 2010).

mitral valve surgery. These are designated A through O in panel A in Figure 2-5, which shows the first five bootstrap analyses. The tall vertical bars designate variables identified in each analysis. Note that no analysis yields identical risk factors. But now consider a running average of these results (Figure 2-5, panel B). Notice the running average of these unstable results progressively reveals a clear pattern: Variables A, C, D, I, and J are signal, and the rest are noise (Figure 2-5, panel C).

Imagine extending this concept. For example, at each iteration the algorithm could average the contribution of a predictor based on its appearance in previous iterations (boosting)—an adaptive weighted average (Bartlett et al., 2004; Freund and Schapire, 1996; Friedman, 2001, 2002). Bagging produces an average, but unlike boosting, it uses the same weight for each iteration.

Other computer learning techniques are being developed, such as Bayesian analysis of variants for Microarray methodology, which is being used to discover empiric gene expression profiles that are highly predictive for colorectal cancer recurrence (Ishwaran and Rao, 2003; Ishwaran et al., 2006). Unsupervised hierarchical bootstrap clustering almost completely separates patients experiencing cancer recurrence from those whose cancer has not recurred. What is important to recognize is that these methods solve the problem of having a large number of parameters (P) compared to number of individuals (n), a key factor in genomic analysis and research.

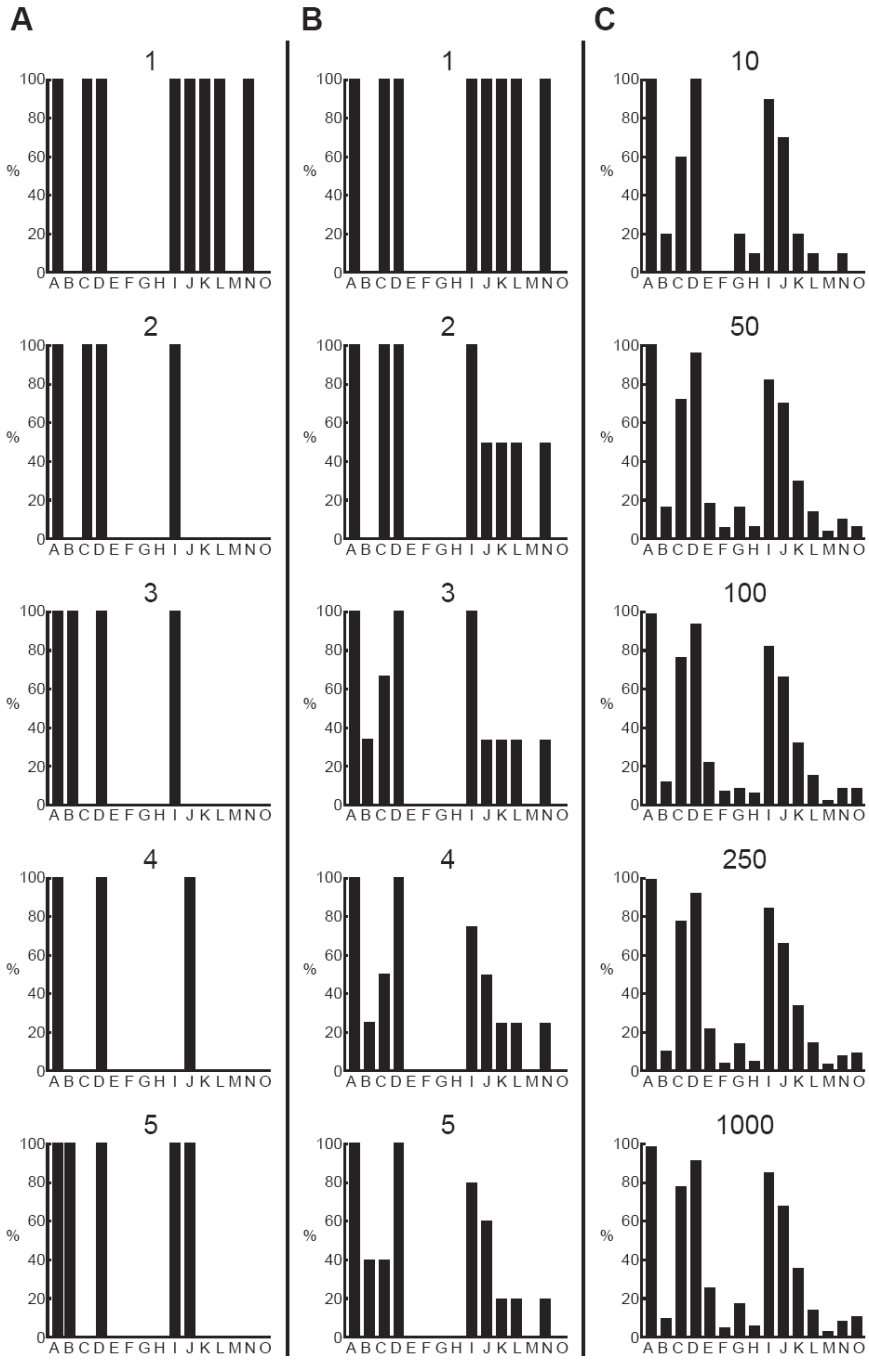
These methods are still in their infancy; many are based on computer-intensive methods such as bootstrap sampling or random forest technology. Variables may be selected by importance value (Breiman, 1996, 2001; Ishwaran, 2007) or by signal-to-noise ratios rather than by traditional P values, which become progressively less useful as n becomes large. Prediction error is minimized rather than maximizing goodness of fit.

An important feature of all ensemble learners is that they are computationally highly parallelizable—either for large-scale parallel computers or for grid computing. This may become important as researchers start looking at a huge number of patients, when speed of computation for clinical inferencing may be of the essence.

Patient-Specific Strategic Decision Support

Finally, to come full circle, consider personalized medicine. Joel Verter once said that RCTs are "sledgehammers, not fine dissecting instruments." Medicine needs to head toward fine dissecting instruments, toward personalized strategic decision support. With $n = 1$, a new paradigm of RCT needs to be developed for genomic-based personalized medicine (Balch, 2006).

Consider a 59-year-old man with ischemic cardiomyopathy and anterior MI resulting in left ventricular aneurysm. He has an ejection fraction



of 10 percent; 4+ mitral valve regurgitation; extensive coronary artery disease, including 90 percent left anterior descending coronary artery stenosis; and multiple comorbidities. Should the recommended therapy be continued medical treatment, coronary artery bypass graft (CABG), CABG plus mitral valve anuloplasty, a Dor operation, or cardiac transplantation? This complex information is too multidimensional for assimilation by the human mind. It calls for a cognitive prosthesis (Reason, 1999). Ideally, this patient's data would be entered automatically by a CPR into a strategic decision aid and the long-term expected survival would be depicted for multiple alternative therapies along with uncertainty limits, although not all therapies may be applicable.

Locked in the medical literature even today are static risk factor equations that could be used in dynamic mode for strategic decision support for a patient such as this (Levy et al., 2006). Random forest technology also can generate outcome risk estimates for individual patients by "dropping" their characteristics down a forest of trees, where they will land at a specific node in each tree with patients having similar characteristics and known outcome. Results of all patients at each node become the average ensemble predicted outcome for an individual patient. Thus, it is possible to imagine that in the future there will be methods by which patient-specific prediction of outcomes are generated and alternative therapies compared for patient decision support.

A library of modules must be developed for constructing strategic decision aids such as this. These in turn must be coupled to values for variables in a CPR so that no human intervention is required to depict comparable predictions of results. Then it must be prospectively verified that the simulated results match actual outcomes. The medical record thus becomes an active revealing and learning tool.

FIGURE 2-5 Example of automated variable selection by bootstrap aggregation (bagging). Fifteen variables labeled A through O are depicted as potential predictors of death after mitral valve surgery. In column A, analyses of five bootstrap samples are shown. Tall bars indicate the variable was selected at $P < .05$, and gaps represent variables not selected. Variables A and D were selected in all cases, but otherwise the analyses appear to be unique. Panel B shows a running average of these five analyses. Variables A, D, I, and J were selected more often than others. Panel C shows averages of 10, 50, 100, 250, and 1,000 bootstrap analyses. Notice that no variable was selected 100% of the time, and all 15 were selected at one time or another. But if variables appearing in 50% or more analyses are considered reliable risk factors, then variables A, C, D, I, and J fit that criterion of "signal," and the rest are "noise."

Summary

Moving beyond today's Mt. Everest level of difficulty, RCTs need to become more nimble and simple to better reflect the real world and to have their financing restructured. Heterogeneity in practice facilitates approximate randomized trials via propensity score methods that are inexpensive and widely accessible but which require patient-level clinical data stored as discrete values for variables. Emerging semantic technology can be exploited to integrate currently disparate, siloed medical data—responding to investigators' complex queries and patients' imprecise ones—and in the near future holds the promise to automate discovery of unsuspected relationships and unintended adverse or surprisingly beneficial outcomes. A next generation of analytic tools for revealing patterns in clinical data should build on successful methods developed in the discipline of machine learning. Both new knowledge learned and resulting algorithms should be transformed into strategic decision support tools. These are but a few concrete examples of methods that need to be developed to provide an infrastructure to determine the right treatment for the right patient at the right time.

Resources Needed

What resources are needed to develop this infrastructure?

Reengineering Randomized Controlled Trials

The cost of an NIH-sponsored simple trial appears to be in the range of $2 million, but multi-institutional, multinational large trials driven by clinical end points can consume 10 times that figure. If one uses $100 million as a metric, this means 5 to 50 such trials of therapy can be supported. Considering all the therapies of medicine for which the evidence base is weak, it is clear that demanding gold-standard RCTs for everything is unaffordable. The cost of RCTs that are highly focused, ethically unambiguous, and feasible could be brought down to a quarter, perhaps even a tenth, of this figure based on practical experience. This will require maximum use of electronic patient records, consisting of values for variables, and quite specifically longitudinal surveillance data to study the long-term side effects of therapies.

Approximate Randomized Controlled Trials

The NIH and National Science Foundation (NSF) should join forces and solicit 3-year methodology grants of approximately $250,000 per year,

10 per year. For this $7.5 million investment, a strong understanding of how best to use nonrandomized data would emerge. With this would come production of publicly available statistical software.

If rich discrete clinical data were available for analysis, a typical study using these methods for nonrandomized comparison would cost approximately $75,000. The cost would double if extensive integration of data was necessary, possibly over healthcare networks. For $100 million, it would be possible to conduct more than 1,000 such approximate randomized trials. This could have a major impact on acquiring what might be called "silver-level" evidence for practice.

Semantically Integrating, Querying, and Exploring Disparate Clinical Data

Based on several years of work, it seems that developing a comprehensive ontology of medicine—a new framework for analysis across disparate medical domains—will cost about 1 hour of time per term for an analyst, programmer, and clinical expert. One need not start from scratch, but can exploit SNOMED, UMLS, and other term lists and ontologies to start the process. Assuming that 100,000 terms would need to be defined in this fashion, that the wages would be $300 per hour, and that 25 ontologists would be needed, this work could be completed in 2 years at a cost of $36 million. This would include the software that must be programmed to implement a global effort in rallying medical experts to this task.

Computer Learning Methods

Knowledge discovery in medicine involves both methodologic development and applications. These should go hand in hand in this new field because it would accelerate the development of methods as they encounter problems requiring further methodologic work. The NSF has begun an initiative called Cyber-Enabled Discovery and Innovation (Jackson, 2007). This began with a $52 million first-year budget and is intended to ramp up $50 million per year and finish within 5 years for a total of $750 million. It would be useful to add $10 million per year for direct application to biomedicine, for a total sustained level for these activities within 5 years of $50 million.

Patient-Specific Strategic Decision Support

Costs in this area are largely for developing software, including the interfaces to EMR systems. This could be done for approximately $10 million. One could envision every study of clinical effectiveness having a

patient-specific prediction component built into it. Again, based on experience doing this, approximately $25,000 per study would be required to adapt and test the software and couple it with EMRs for decision support. It is also likely that at some point, the FDA may become involved with tools such as this and would introduce regulations that are more costly to meet than those of performing the studies.

COORDINATION AND TECHNICAL ASSISTANCE THAT NEED TO BE SUPPORTED

Jean R. Slutsky, Director, Center for Outcomes and Evidence, Agency for Healthcare Quality and Research

Overview

CER as a concept and reality has grown rapidly in the past 5 years. While it builds on an appreciation for the role of technology assessment, comparative study designs, and the increased role of health information technology to gather evidence and distribute it to the point of care, the capacity and infrastructure for this research has received less targeted attention. Understanding the landscape of organizations and health systems undertaking CER is challenging but essential. Without knowing what capacities and infrastructure currently exist, rational strategic planning for the future cannot be done. It is also important to address which functions can be most effective if they are centralized, which are most effective if they are local or decentralized, and how different activities relate to each other in a productive way. This paper will explore the practical realities of what exists now, what is needed for the future, and how the needs of the country's diverse healthcare system for CER can best be met.

The Agency for Healthcare Research and Quality Perspective

AHRQ plays a significant role in CER. Under a mandate included in Section 1013 of the Medicare Prescription Drug, Improvement, and Modernization Act of 2003, AHRQ is the lead agency for CER in the United States. AHRQ conducts health technology assessment at the request of CMS and analyzes data and suggests options for coverage with evidence development (CED) and post-CED data collection. AHRQ also provides translation of CER findings, promotes and funds comparative effectiveness methods research, and funds training grants focused on comparative effectiveness. AHRQ has an annual budget of over $300 million ($372 million for 2009), and received funds specifically for work on CER ($30

million and $50 million in 2008 and 2009, respectively).[14] AHRQ has built a flexible, dynamic infrastructure for CER that includes 41 research centers nationwide with more than 160 researchers. The program includes K awards for career development such as the Mentored Research or Clinical Scientist Development awards, and Independent Scientist awards. AHRQ has also funded methods research that heretofore had not been funded except on an ad hoc basis. AHRQ has also put concentrated funding into the translation of CER findings.

Given the pressing need for evidence, it is important to keep in mind the high costs of precision. It costs money to conduct precise studies that answer detailed questions, which speaks to the importance of not only priority setting but also understanding that this is an important and difficult task.

In the United States, the landscape for CER is, for the most part, very well intentioned. All parties engaged in this work want to do the right thing and see what works best for patients in the United States and elsewhere. Nonetheless, current efforts are too ad hoc in nature. There are no adequate organizing principles, except for those outlined in Section 1013, where language in the legislation focuses on setting priorities, having transparent processes, involving stakeholders, and having a translation component. The effect is that in the United States there is essentially only a very limited capacity to conduct CER and to translate that research into meaningful and useful applications. The United States is not accustomed, nor organized effectively, to conduct this type of research. In part, this is because the system tends not to grow researchers who have the capacity to move beyond what might be described as a parochial mind set regarding the types of research study designs, a mind set which has limited the capacity to readily generate hypotheses and study designs appropriate for CER. Generally, researchers do not involve stakeholders to the extent that is required for research aimed at generating information to guide end users such as patients and physicians. A key shift needed in the current approach to research is to involve patients and other key stakeholders, such as industry and health plans, in the formulation of questions for investigation and in study design.

As mentioned above, AHRQ is currently conducting CER under legislative mandate. Other federal agencies also conducting CER include the NIH, CMS (CED), and the VA; some have done so for decades. In the private sector, health plans and industry are also engaged in CER. In addition, CER-focused public–private partnerships are starting to form, as are private–private partnerships. Unlike other forms of research, CER will most certainly require the kinds of partnerships that are now emerging.

[14] See http://www.ahrq.gov/about/budgtix.htm (accessed September 8, 2010).

There are a number of common pitfalls in CER. One of the most significant failures is that comparative effectiveness studies are not designed in ways that capture meaningful end points as well as longitudinal and relevant outcomes—end points that would be meaningful not only to patients but also to decision makers. There are also significant issues about the applicability of the CER studies that are conducted, and the work we do for CMS is reflective of the need for research conducted in patients representative of the Medicare population. The elderly tend not to be studied in rigorous trials to the extent that children are, although AHRQ has begun to address this discrepancy through, for example, work for the Medicaid and State Children's Health Insurance programs. There is also a failure to clinically address relevant heterogeneity and biological heterogeneity as well. Finally, there are also responsibility issues, as evidenced in discussions of "who pays and who stays" when the need for an important study has been noted. There is currently discussion of such issues on Capitol Hill.

Today's reality in CER, therefore, can be summarized as follows. There is general sentiment that CER can be a positive thing if it is done fairly, is well designed, and is transparent. This is important because of the potential impact of CER on many different sectors—not just patients, but also industry and health plans. If CER is not conducted in a way that stakeholders can understand—and, importantly, in a way by which they have input into the process—it could happen that CER does not have the impact, in terms of improving health outcomes, that everyone hopes it will have. As AHRQ discovered in developing the Section 1013 healthcare program, involving stakeholders early, listening to them, and involving them throughout the process through to the end and implementation of the findings, is critically important.

Another issue with CER today is that there is no agreement on the best methods for setting priorities. Everyone tries to set priorities finds that reaching consensus is hugely difficult. In part that is because the process of setting such priorities tends to revert to personal or narrowly defined considerations. If a consensus is to be reached, rather than thinking about individual priorities, it will be important to instead learn to focus on national priorities. Experience shows, however, that such priority shifting is difficult.

Another aspect of the reality of CER today that has been disappointing is that there has been less emphasis on designing good studies than on just the concept of CER. A great deal of time is spent talking about where CER should live, what it should look like, and how it should be funded, but there has tended not to be adequate discussion about how to conduct CER most effectively from a methodological and implementation standpoint. Further discussion is needed on both sides of this issue—discussion not only about which box CER lives in, if you will, but about what's inside the box.

Once a decision to perform a CER study is made, finding a payer or a funder for the research component is difficult. Obviously, with a budget of $30 million AHRQ cannot be the sole funder of many of these studies. Regardless, the reality is that the challenges related to funding lead to less rigorous and less innovative study designs. Somehow the issue of how different players collaborate to fund these studies must be addressed. Otherwise, the rigorous study designs needed to move this field forward will never be developed.

Successes, however, make the fundamental concept of CER worthwhile. Among the successes, for example, are cases where CER uncovered findings that exceeded expectations. Successes also include what might be considered negative findings, where research results proved to be not good news for certain subpopulations but still provided findings that were not previously known. Whether findings from CER are viewed as positive or negative, the overarching consideration is that they inform how care is provided.

Conclusions

In closing, it will be important to be mindful of several critical factors while developing the infrastructure necessary to advance CER. First, more coordination is needed in setting priorities. Much has been learned from individual efforts that have taken place both within government and outside government. What is needed now is to capitalize on these lessons learned and to begin moving forward together in a more coordinated way to reach consensus in the setting of priorities for CER. A more systematic approach to the conduct of CER is also warranted, if only because CER tends to receive a smaller slice of research funding and so it will be important to be systematic and strategic in spending limited funds effectively. Coordination is imperative.

A stronger emphasis on training, methods, and translation is also needed. These three factors are separate, but they are not inseparable. Enhanced education is necessary to train the next generation of researchers on the methodologies of new research designs and on methods of translating research findings in ways that are actionable, understandable, and not leading with blunt-edge decisions. At the same time, there needs to be more robust training targeted to help next-generation investigators work effectively with all relevant stakeholders. More funding is needed, specifically for well-designed studies that meet priorities, that are not underpowered and that address meaningful health outcomes. Further, more public–private partnerships are needed to move CER forward from an implementation standpoint and to resolve some of the funding problems that heretofore have hindered CER. Finally, more training is needed on the use of findings to avoid inappropriate or unintended consequences. Too often, the defini-

tion of success is whether the results of a given study were published in a peer-reviewed journal. That focus needs to be shifted to the true heart of the matter, which is how the findings can best be used in practice and how they are relevant to decision makers.

As to the future of CER, public–private funding and participation is a critical necessity for CER to go forward. More effort is needed to develop designs and protocols that more efficiently and effectively answer CER questions. This would encompass not merely conducting new methodological research but also working with stakeholders and users of research as well as people affected by the research. Public–private funding and participation is a necessity. Finally, there are a number of important issues that will take a global approach that need to be addressed. Training on research design and translation must become an accepted use of healthcare dollars. Wide attention to vastly improved priority setting, at macro and micro levels, is also necessary. Transparency across participation is important, so no one gets unequal access and everyone is at the table. Improved technical assistance for conducting and implementing CER will also be critical to success.

REFERENCES

AAAI (Association for the Advancement of Artificial Intelligence). 2008. *Fall symposium on automated scientific discovery*. November 7-9, 2008. Arlington, VA.

Abiteboul, S., D. Quass, J. McHugh, J. Widom, and J. L. Wiener. 1997. The Lorel query language for semistructured data. *International Journal of Digital Libraries* 1:68-88.

AHRQ (Agency for Healthcare Research and Quality). 2007. *Special emphasis notice: AHRQ announces interest in career development (K) grants focused on comparative effectiveness research*. http://grants.nih.gov/grants/guide/notice-files/NOT-HS-08-003.html (accessed August 5, 2010).

Avalere Health. 2008. *Patient and clinician participation in research agenda setting: Lessons for future application*. Washington, DC: Avalere Health.

Balch, C. M. 2006. Randomized clinical trials in surgery: Why do we need them? *Journal of Thoracic and Cardiovascular Surgery* 132(2):241-242.

Barnett, G. O., R. A. Jenders, and H. C. Chueh. 1993. The computer-based clinical record—Where do we stand? *Annals of Internal Medicine* 119(10):1046-1048.

Bartlett, P. L., P. J. Bickel, P. Buhlmann, and Y. Freund. 2004. Discussions of boosting papers, and rejoinders. *Annals of Statistics* 32:85-134.

Beck, J., P. J. Esser, and M. B. Herschel. 2004. Why clinical trials fail before they even get started: The "frontloading" process. *Quality Assurance Journal* 8(1):21-32.

Berners-Lee, T., J. Hendler, and O. Lassila. 2001. The semantic web: A new form of web content that is meaningful to computers will unleash a revolution of new possibilities. *Scientific American* May 17:1-18.

Blackstone, E. H. 2002. Comparing apples and oranges. *Journal of Thoracic and Cardiovascular Surgery* 123(1):8-15.

———. 2005. Could it happen again? The Bjork-Shiley convexo-concave heart valve story. *Circulation* 111(21):2717-2719.

———. 2006. Consort and beyond. *Journal of Thoracic and Cardiovascular Surgery* 132(2): 229-232.

———. 2007. From trees to wood or wood to trees? *Journal of Surgical Research* 147: 59-60.

Braunwald, E. 1983. Effects of coronary-artery bypass grafting on survival. Implications of the randomized coronary-artery surgery study. *New England Journal of Medicine* 309(19):1181-1184.

Breiman, L. 1996. Bagging predictors. *Machine Learning* 24:123-140.

———. 2001. Statistical modeling: The two cultures. *Statistical Science* 16:199-231.

Buchanan, B. G., and E. H. Shortliffe, eds. 1984. *Rule based expert systems: The mycin experiments of the Stanford heuristic programming project*. Reading, MA: Addison-Wesley.

Buneman, P., M. Fernandez, and D. Suciu. 2000. UnQL: A query language and algebra for semistructured data based on structural recursion. *Very Large Database Journal* 9:76-110.

CBO (Congressional Budget Office). 2007. *Research on the comparative effectiveness of medical treatments*. Publication No. 2975. Washington, DC: CBO. http://www.cbo.gov/ftpdocs/88xx/doc8891/12-18-ComparativeEffectiveness.pdf (accessed September 3, 2010).

Cepeda, M. S., R. Boston, J. T. Farrar, and B. L. Strom. 2003. Comparison of logistic regression versus propensity score when the number of events is low and there are multiple confounders. *American Journal of Epidemiology* 158(3):280-287.

Congressional Research Service (CRS). 2007. *Comparative clinical effectiveness and cost-effectiveness research: Background, history, and overview.* Washington, DC: CRS.

Diaconis, P., and B. Efron. 1983. Computer-intensive methods in statistics. *Scientific American* June 1:116-130.

Donner, A., and N. Klar. 2000. *Design and analysis of cluster randomization trials in health research*. New York: Oxford University Press.

Efron, B. 1979. Bootstrap methods: Another look at the jackknife. *Annals of Statistics* 7:1-26.

———. 1982. *The jackknife, the bootstrap and other resampling plans*. Monograph. Philadelphia, PA: Society for Industrial and Applied Mathematics.

Efron, B., and R. Tibshirani. 1986. Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statistical Science* 1:54-77.

Erkinjuntti, T., T. Ostbye, R. Steenhuis, and V. Hachinski. 1997. The effect of different diagnostic criteria on the prevalence of dementia. *New England Journal of Medicine* 337(23):1667-1674.

Etheredge, L. 2007. A rapid-learning health system. *Health Affairs* 26:w107-w118.

FDA (Food and Drug Administration). 2009. Patient participation in FDA regulatory issues. http://www.fda.gov/ForConsumers/ByAudience/ForPatientAdvocates/PatientInvolvement/default.htm (accessed August 5, 2010).

Fleming, T. R., and D. L. DeMets. 1996. Surrogate end points in clinical trials: Are we being misled? *Annals of Internal Medicine* 125(7):605-613.

Freund, Y., and R. E. Schapire. 1996. Experiments with a new boosting algorithm. *Proceedings of the International Conference on Machine Learning* July:148-156.

Friedman, J. H. 2001. Greedy function approximation: A gradient boosting machine. *Annals of Statistics* 29:1189-1232.

———. 2002. Stochastic gradient boosting. *Computational Statistics & Data Analysis* 38: 367-378.

Gliklich, R. E., and N. A. Dreyer, eds. 2007. *Registries for evaluating patient outcomes: A user's guide*. Rockville, MD: Agency for Healthcare Research and Quality.

Grunkemeier, G. L., R. Jin, and A. Starr. 2006. Prosthetic heart valves: Objective performance criteria versus randomized clinical trial. *Annals of Thoracic Surgery* 82(3):776-780.

Gum, P. A., M. Thamilarasan, J. Watanabe, E. H. Blackstone, and M. S. Lauer. 2001. Aspirin use and all-cause mortality among patients being evaluated for known or suspected coronary artery disease: A propensity analysis. *Journal of the American Medical Association* 286(10):1187-1194.

Heywood, J. 1546. *A dialogue conteynyng the nomber in effect of all the prouerbes in the englische tongue*.

Hunt, S. A., W. T. Abraham, M. H. Chin, A. M. Feldman, G. S. Francis, T. G. Ganiats, M. Jessup, M. A. Konstam, D. M. Mancini, K. Michl, J. A. Oates, P. S. Rahko, M. A. Silver, L. W. Stevenson, C. W. Yancy, E. M. Antman, S. C. Smith, Jr., C. D. Adams, J. L. Anderson, D. P. Faxon, V. Fuster, J. L. Halperin, L. F. Hiratzka, A. K. Jacobs, R. Nishimura, J. P. Ornato, R. L. Page, B. Riegel, American College of Cardiology, American Heart Association Task Force on Practice Guidelines, American College of Chest Physicians, International Society for Heart and Lung Transplantation, and Heart Rhythm Society. 2005. ACC/AHA 2005 guideline update for the diagnosis and management of chronic heart failure in the adult. *American Heart Association* 112:145-235.

IDSA (Infectious Diseases Society of America). 2006. Frequently asked questions about Lyme disease. http://www.idsociety.org/lymediseasefacts.htm (accessed August 5, 2010).

ILADS (International Lyme and Associated Diseases Society) Working Group. 2004. Evidence-based guidelines for the management of Lyme disease. *Expert Review of Anti-Infective Therapy* 2(1):S1-S13. http://www.ilads.org/files/ILADS_Guidelines.pdf (accessed August 5, 2010).

Ilkhanoff, L., J. D. Lewis, S. Hennessy, J. A. Berlin, and S. E. Kimmel. 2005. Potential limitations of electronic database studies of prescription non-aspirin non-steroidal anti-inflammatory drugs (NANSAIDS) and risk of myocardial infarction (MI). *Pharmacoepidemiological Drug Safety* 14(8):513-522.

Ioannidis, J. P. 2005. Why most published research findings are false. *PLoS Medicine* 2(8): e124.

IOM (Institute of Medicine). 1991 *The computer-based patient record: An essential technology for health care*. Washington, DC: National Academy Press.

———. 2007. *Learning what works best: The nations need for evidence on comparative effectiveness in health care*. http://www.iom.edu/ebm-effectiveness (accessed August 5, 2010).

———. 2008. *Knowing what works in health care: A roadmap for the nation*. Washington, DC: The National Academies Press.

Ishwaran, H. 2007. Variable importance in binary regression trees and forests. *Electronic Journal of Statistics* 1:519-537.

Ishwaran, H., and J. S. Rao. 2003. Detecting differentially expressed genes in microarrays using Bayesian model selection. *Journal of the American Statistical Association* 98:438-455.

Ishwaran, H., J. S. Rao, and U. B. Kogalur. 2006. Bamarray: Java software for Bayesian analysis of variance for microarray data. *BMC Bioinformatics* 7(1):59.

Jackson, A. 2007. Cyber-enabled discovery and innovation. *Notices of the American Mathematical Society* 54:752-754.

Kamerow, D. 2009. *Comparative effectiveness studies inventory project*. Washington, DC: A commissioned activity for the IOM Roundtable on Value & Science-Driven Health Care.

Kirklin, J. W. 1979. A letter to Helen (Presidential address). *Journal of Thoracic and Cardiovascular Surgery* 78(5):643-654.

Kirklin, J. W., and S. S. Vicinanza. **1999. Metadata and computer-based patient records.** *Annals of Thoracic Surgery* 68(3 Suppl.):S23-S24.

Last, J., ed. 1983. *A dictionary of epidemiology*. New York: Oxford University Press.

Lenat, D. B. 2008. The voice of the turtle: Whatever happened to AI? *Artificial Intelligence Magazine* 29(2):11-22.

Lenat, D. B., and R. V. Guha. 1990. *Building large knowledge-based systems: Representation and inference in the Cyc project*. Reading, MA: Addison-Wesley.

Levy, W. C., D. Mozaffarian, D. T. Linker, S. C. Sutradhar, S. D. Anker, A. B. Cropp, I. Anand, A. Maggioni, P. Burton, M. D. Sullivan, B. Pitt, P. A. Poole-Wilson, D. L. Mann, and M. Packer. 2006. The Seattle heart failure model: Prediction of survival in heart failure. *Circulation* 113(11):1424-1433.

Merriam-Webster Dictionary. 2004. Springfield, MA: Merriam-Webster.

Mosis, G., B. Koes, J. Dieleman, B. Stricker, J. van der Lei, and M. C. Sturkenboom. 2005a. Randomised studies in general practice: How to integrate the electronic patient record. *Informatics in Primary Care* 13(3):209-213.

Mosis, G., A. E. Vlug, M. Mosseveld, J. P. Dieleman, B. C. Stricker, J. van der Lei, and M. C. Sturkenboom. 2005b. A technical infrastructure to conduct randomized database studies facilitated by a general practice research database. *Journal of the American Medical Informatics Association* 12(6):602-607.

Mosis, G., J. P. Dieleman, B. Stricker, J. van der Lei, and M. C. Sturkenboom. 2006. A randomized database study in general practice yielded quality data but patient recruitment in routine consultation was not practical. *Journal of Clinical Epidemiology* 59(5):497-502.

NCCN (National Comprehensive Cancer Networks). 2008. *About the NCCN clinical practice guidelines in oncology*. Available at http://www.nccn.org/professionals/physician_gls/about.asp (accessed September 8, 2010).

———. 2009. *NCCN drugs & biologics compendium*. http://www.nccn.org/professionals/drug_compendium/content/contents.asp (accessed September 8, 2010).

NCI (National Cancer Institute). 2007. Cancer care outcomes research and surveillance consortium. http://outcomes.cancer.gov/cancors (accessed May 30, 2009).

NHLBI (National Heart, Lung, and Blood Institute). 1981. National Heart, Lung, and Blood Institute coronary artery surgery study. A multicenter comparison of the effects of randomized medical and surgical treatment of mildly symptomatic patients with coronary artery disease, and a registry of consecutive patients undergoing coronary angiography. *Circulation* 63(6 Pt. 2):I1-I81.

NICE (National Institute for Health and Clinical Excellence). 2004. *Summary budget statement*. www.nice.org.uk/niceMedia/pdf/boardmeeting/brdsep04item12.pdf (accessed July 31, 2004).

NIH (National Institutes of Health). 2006. K30 clinical research curriculum award. http://grants.nih.gov/training/k30.htm?sort=inst#list (accessed January 17, 2006).

Nissen, S. E. 2006. Adverse cardiovascular effects of rofecoxib. *New England Journal of Medicine* 355(2):203-205.

Pickett, J. P., ed. 2000. *The American heritage dictionary of the English language*, 4th ed. Boston, MA: Houghton Mifflin.

Reason, J. 1999. *Human error*. Cambridge, MA: Cambridge University Press.

Rosenbaum, P. R., and D. B. Rubin. 1983. The central role of the propensity score in observational studies for causal effects. *Biometrika* 70:41-55.

Rubin, D. B. 2007. The design versus the analysis of observational studies for causal effects: Parallels with the design of randomized trials. *Statistics in Medicine* 26(1):20-36.

Sabik, J. F., A. M. Gillinov, E. H. Blackstone, C. Vacha, P. L. Houghtaling, J. Navia, N. G. Smedira, P. M. McCarthy, D. M. Cosgrove, and B. W. Lytle. 2002. Does off-pump coronary surgery reduce morbidity and mortality? *Journal of Thoracic and Cardiovascular Surgery* 124(4):698-707.

Sacristan, J. A., J. Soto, I. Galende, and T. R. Hylan. 1997. A review of methodologies for assessing drug effectiveness and a new proposal: Randomized database studies. *Clinical Therapeutics* 19(6):1424-1517.

Schulz, S., B. Suntisrivaraporn, F. Baader, and M. Boeker. 2009. SNOMED reaching its adolescence: Ontologists' and logicians' health check. *International Journal of Medical Informatics* 78(Suppl. 1):S86-S94.

Senge, Peter M. 1990. *The fifth discipline: The art and practice of the learning organization*. New York: Doubleday.

Shadish, W., T. Cook, and D. Campbell. 2002. *Experimental and quasi-experimental designs for generalized causal inference*. Boston, MA: Houghton Mifflin.

Slutsky, J., and C. Clancy. 2009. AHRQ's effective health care program: Why comparative effectiveness matters. *American Journal of Medical Quality* 24:67-70.

Stewart, W. F., N. R. Shah, M. J. Selna, R. A. Paulus, and J. M. Walker. 2007. Bridging the inferential gap: The electronic health record and clinical evidence. *Health Affairs* 26(2): w181-w191.

Strom, B. L. 2005. Overview of automated data systems in pharmacoepidemiology. In *Pharmacoepidemiology,* 4th ed., edited by B. L. Strom. Chichester, UK: John Wiley and Sons.

Sturmer, T., M. Joshi, R. J. Glynn, J. Avorn, K. J. Rothman, and S. Schneeweiss. 2006. A review of the application of propensity score methods yielded increasing use, advantages in specific settings, but not substantially different estimates compared with conventional multivariable methods. *Journal of Clinical Epidemiology* 59(5):437-447.

Thorn, K. E., A. K. Bangalore, and A. C. Browne. 2007. The UMLS knowledge source server: An experience in Web 2.0 technologies. *AMIA Annual Symposium Proceedings* 2007:721-725.

Thygesen, K., J. S. Alpert, and H. D. White. 2007. Universal definition of myocardial infarction. *Journal of the American College of Cardiology* 50(22):2173-2195.

Weeks, W., and J. Bagian. 2000. Developing a culture of safety in the Veterans Health Administration. *Effective Clinical Practice* 3:270-276.

Zarin, D. A., and T. Tse. 2008. Medicine: Moving toward transparency of clinical trials. *Science* 319(5868):1340-1342.

Zipes, D. P., A. J. Camm, M. Borggrefe, A. E. Buxton, B. Chaitman, M. Fromer, G. Gregoratos, G. Klein, A. J. Moss, R. J. Myerburg, S. G. Priori, M. A. Quinones, D. M. Roden, M. J. Silka, C. Tracy, S. C. Smith, Jr., A. K. Jacobs, C. D. Adams, E. M. Antman, J. L. Anderson, S. A. Hunt, J. L. Halperin, R. Nishimura, J. P. Ornato, R. L. Page, B. Riegel, J. J. Blanc, A. Budaj, V. Dean, J. W. Deckers, C. Despres, K. Dickstein, J. Lekakis, K. McGregor, M. Metra, J. Morais, A. Osterspey, J. L. Tamargo, and J. L. Zamorano. 2006. ACC/AHA/ESC 2006 guidelines for management of patients with ventricular arrhythmias and the prevention of sudden cardiac death: A report of the American College of Cardiology/American Heart Association task force and the European Society of Cardiology committee for practice guidelines (Writing committee to develop guidelines for management of patients with ventricular arrhythmias and the prevention of sudden cardiac death). *Journal of the American College of Cardiology* 48(5):e247-e346.

3

The Information Networks Required

INTRODUCTION

The scale of efficiencies that might be gained through developing prioritization and coordinating capacities and improving the methods used for comparative effectiveness research (CER) will be limited by the infrastructure available to support the capture, access, and sharing of relevant data and information. Design and development of robust information networks, and efforts to foster collaboration around common work, will therefore be a critical aspect of creating the infrastructure for expanded CER—necessary for the generation and application of evidence alike, as well as for providing opportunities to support learning from clinical practice. In addition to the federal efforts to increase the adoption and use of electronic health records as described previously, many organizations have developed such capacities, and drawing upon these and other resources through systematic, linked, and coordinated networks would greatly enhance the nation's fundamental capacity to generate evidence. Papers included in this chapter describe what was known about capacity in 2008, give a rough estimate of the necessary capacity, and offer initial suggestions on policies or activities for progress. These issues are considered in more depth in the Institute of Medicine (IOM) workshop summary publication on *The Digital Infrastructure for the Learning Health System: The Foundation for Continuous Improvement in Health and Health Care* (IOM, 2011).

Clinical information systems (CISs)—including electronic health records (EHRs)—hold particular promise, given their emerging prominence at the nexus of clinical research, clinical practice, and decision

making. Appropriately designed EHRs not only serve as a means for practitioners to access best practices and evidence guidelines, but they also capture a broad array of information important to the diagnosis and treatment of individual patients. To provide policy makers with "order of magnitude" estimates of the spending needed to speed broad adoption of CISs in care delivery organizations throughout the nation, Robert H. Miller from the University of California at San Francisco describes current EHR adoption, future EHR capital and operating expenditure requirements, and prospects for EHR adoption in the hospital and in physician and clinical services sectors.

Work is also needed to develop the technical capacity, methods, standards, and policies for the efficient exchange of information from EHRs and other data sources (e.g., administrative databases, clinical registries) and to disseminate evidence syntheses and other resources to guide practice. Although large databases and clinical registries offer immediate opportunities for learning what works in health care, Carol C. Diamond from the Markle Foundation argues that the greatest promise of health information technology (HIT) lies in its ability to enable quick and efficient learning via a networked and distributed approach to information sharing and evidence development. To maximize this potential, approaches to data and information hubs will need to evolve to address four key challenges: (1) clearly defining the ultimate goal; (2) being open to reset our definitions and assumptions about health data and research approaches; (3) articulating new, broadly accepted working principles based on 21st-century information paradigms; and (4) developing an information policy framework that broadly addresses public hopes and concerns. Diamond illustrates how these challenges are a jumping-off point for moving to a distributed approach to research—one characterized by connectivity, networks, and feedback loops.

Finally, an essential function of any system dedicated to developing a robust evidence base for medical practice is the synthesis of information derived from relevant trials, studies, and insights emerging from clinical practice. As data resources and networks expand, demand will also grow for synthesis work to ensure studies are appropriately reviewed, vetted, and incorporated into the evolving evidence base. Lorne A. Becker from the Cochrane Collaborative provides an overview of current approaches to evidence review, synthesis, coordination, and dissemination—internationally and within the United States—and offers some suggestions on key opportunities for expanding capacity to meet the anticipated demand.

ELECTRONIC HEALTH RECORDS: NEEDS, STATUS, AND COSTS FOR U.S. HEALTHCARE DELIVERY ORGANIZATIONS

Robert H. Miller, Ph.D., Professor of Health Economics in Residence, University of California at San Francisco, Institute for Health & Aging

Introduction

Implementing ubiquitous evidence-based medicine (EBM) requires robust CISs, especially EHRs. As a result, policy makers want to understand what EHR capabilities are needed, to what extent they've been implemented, likely costs for further adoption and maintenance, and prospects for full implementation of those capabilities.[1] Despite keen policy-maker interest, as of mid-2008 few "global" cost estimates for ubiquitous CISs in the healthcare delivery system had been generated (CBO, 2008), in large part because the U.S. healthcare system is so large and diverse, while usable CIS cost and benefit data have been so scarce.

In 2006, of the $2.1 trillion in total healthcare expenditures, the $648 billion hospital sector and $447 billion "physician and clinical services" sector were the most intensive users of CISs, incurring the bulk of CIS capital and operating expenses (Catlin et al., 2008). Other, smaller healthcare delivery system sectors that had less intensive CIS use together accounted for another $400 billion or so in expenditures in 2006. Using National Health Expenditure Accounts terminology, these sectors included dental services, "other" professional and personal healthcare services, nursing home care, and home health care. Healthcare sectors least relevant to this analysis accounted for $600 billion in spending; they included retail outlet sales of medical products, administration and government public health, research, construction, and equipment.

Because this report is derived from a presentation given in July 2008, it does not include a description or analysis of recent economic developments or of the 2009 economic stimulus legislation on CIS adoption and cost. It does include a handful of updated references for studies that had been in manuscript form in July 2008 and which were subsequently published.

Methods

For this overview, two main types of data sources were used: (1) data from peer-reviewed articles and non-peer-reviewed reports, and (2) interview data from research conducted on behalf of the 2007–2008 California Governor's Health Information Technology Financing Advisory Commis-

[1] Personal health records constitute a separate set of capabilities that consumers could use to view and act on data from various sources, including from the EHRs that hospitals and physician organizations use.

sion, for which researchers obtained information on CIS adoption, CIS business case and value proposition, and ability to finance CISs from large health systems, rural hospitals, public hospitals, medical groups, independent practice associations, and community health centers—albeit only for one state, California.

For the July 2008 IOM workshop, the author generated rough order-of-magnitude estimates of new or additional spending on CISs needed to implement EBM, over and above current spending on CIS capital projects and operating costs. Such ballpark cost estimates could be useful to policy makers that want to know whether the likely new spending on CISs is closer to (say) $50 billion than it is to $500 billion dollars, whether most healthcare delivery system organizations can afford the new CIS spending, and what public policies are needed to achieve ubiquitous CIS adoption. Given that any CIS cost estimates would be rough, we aimed to create cost estimates that were more likely to err on the high than on the low side—if conservative (worst-case) estimates of CIS costs were "manageable" for delivery system organizations, then, likely CIS costs would be even more manageable.

Hospital Sector Clinical Information Systems

What's Needed and What's Been Adopted

Table 3-1 contains a brief description of hospital CIS (EHR) capabilities and adoption, using a stages-of-CIS adoption schema used by the Health Information Management Systems Society (HIMSS) and data that HIMSS obtained for early 2008 (HIMMS Analytics, 2008). The schema shows a hierarchy of CIS adoption, with organizations at a higher stage of adoption typically having capabilities found at lower stages. Most hospitals had new or old (i.e., "legacy") stage 1 ancillary systems that manage basic information on radiology orders, laboratory orders, and pharmacy prescriptions. The most basic systems depend on orders that providers first write out by hand, and that are made electronic at some point prior to test/prescription processing.

Stage 2 CIS capabilities can pull patient data together from many (often isolated and disparate) information systems into a central data repository that enables managers to generate reports and providers and staff to more easily view more demographic, test result, prescription, and other data. Stage 3 capabilities enable improved data presentation and capture, some checking for errors in prescription and test ordering, as well as digital imaging.

As of 2007, CIS capabilities at stage 4 and beyond were still relatively rare—according to HIMSS data, less than 5 percent of hospitals had com-

TABLE 3-1 Hospital Electronic Health Record Capabilities and Adoption Estimates

Stage	Description	2008
Stage 7	Medical record fully electronic; healthcare organization able to contribute continuity of care document as by-product of electronic medical record; data warehousing/mining	0.1%
Stage 6	Physician documentation (structured templates), full clinical decision support, full Radiology Picture Archiving and Communication System (PACS)	1.0%
Stage 5	Closed loop medication administration	1.3%
Stage 4	Computerized physician order entry, clinical decision support (clinical protocols)	1.9%
Stage 3	Clinical documentation (flow sheets), clinical decision support system (error checking), PACS available outside radiology	32.9%
Stage 2	Clinical data repository, controlled medical vocabulary, clinical decision support system inference engine, may have document imaging	33.2%
Stage 1	Ancillaries—lab, radiology, pharmacy	12.5%
Stage 0	All three ancillaries not installed	17.1%

SOURCE: HIMSS Analytics, 2008. For more information see www.connectingforhealth.org/resources/CCEndorser.pdf (accessed September 8, 2010).

puterized physician order entry (CPOE), where the ordering physician (rather than support staff) did the data entry. CPOE systems are considered more likely to affect decision making than simpler ordering systems, as they can generate patient safety/quality reminders and alerts when physicians enter data at the point of care, rather than when staff enter order data later in the ordering process. At the far end of the spectrum are the least implemented capabilities, such as "closed loop" medication administration (with bar coding), physician documentation, and robust capability for health information exchange.

American Hospital Association (AHA) data from 2006 also provided information on the CIS adoption in U.S. hospitals, although they likely overstated CIS adoption because executives and staff in hospitals with advanced CISs were more likely to respond to a survey on CIS adoption than were respondents in hospitals without advanced CISs (AHA, 2007a). The AHA survey findings indicated that 66 percent of hospitals had results

viewing for lab and radiology, and between 46 percent and 66 percent of hospitals had lab/radiology/pharmacy order entry by staff; however, in only 10 percent of hospitals did more than half of physicians routinely use CPOE capabilities for medications ordering. Similarly, while the survey results also showed some progress in implementing CIS capabilities with built-in EBM rules. For example, 31 percent of hospitals provided real-time drug alerts at the point of order data entry (by staff or providers), 37 percent provided "back-end" (not real-time) drug alerts, but only about 10 percent of hospitals offered providers with suggested clinical guidelines and pathways for patient care.

Order-of-Magnitude Cost Estimates

Generating even rough estimates of future CIS costs throughout the U.S. healthcare delivery system is a perilous endeavor due to a lack of high-quality evidence about CIS adoption and cost. Nevertheless, it is possible to show how rough CIS capital and operating costs estimates could be generated and describe the pitfalls of any one estimate.

A crude estimate of hospital sector CIS capital costs can be created by multiplying the number of staffed U.S. community hospital beds by the estimated cost per hospital bed of implementing robust CIS capabilities, and then subtracting the proportion of the CIS capital cost already incurred. If one was to take a CIS cost per bed estimate of $57,000 found in a 2005 RAND report (which the RAND researchers believed was very rough) and increase it to $100,000 in order to account for inflation and to create a distinctly conservative (high) bias to the cost estimate (Girosi et al., 2005), robust CISs in all U.S. hospitals would cost $90 billion in capital costs, given about 800,000 community hospital beds, and perhaps another estimated 100,000 hospital beds with similar characteristics in federal and state hospitals (AHA, 2007b).

Hospitals already have spent some portion of this hypothetical $90 billion in hospital sector CIS capital cost; however, how much they've spent is unclear due to limitations in evidence on the CIS adoption, cost, and spending. Suppose that hospital-sector organizations have already incurred 25 percent of the capital cost of robust CISs on average, and so need to incur nearly $70 billion in additional capital expenditures (75 percent of the hypothetical $90 billion). We assume an 8-year time horizon for achieving robust CISs for nearly all hospitals, since implementing CISs can take years, even in a large health system with substantial information systems staffing, while U.S. hospitals and health systems are at varying stages of implement-

ing CISs. In that case, CIS capital spending per year would amount to $8.5 billion on average.[2]

Hospitals already are spending some portion of the hypothetical $8.5 billion per year on new CIS capabilities—but again, we don't know how much. For example, while one 2008 survey projected that U.S. hospitals would spend about $10 billion per year on all HIT-related capital projects, it provided no estimates of spending only on CISs or on only new CIS capabilities (rather than on replacements for old capabilities) (HIMMS Analytics, 2008).

To show how calculating "new" CIS expenditures might work, suppose for example that CISs accounted for half of the $10 billion per year in hospital HIT-related capital projects, and that new CIS capabilities (not just replacement of old capabilities) accounted for 60 percent of that $5 billion per year. In that case, hospitals would already be spending about $3 billion of the $8.5 per year billion for needed CISs, leaving about $5.5 billion per year in additional "new" CIS capital expenditures per year, or about $42 billion over 8 years.[3] In one of many simple simulations, and assuming that each dollar of new capital spending creates $0.25 in new operating expenses, hospitals would incur about $48 billion for additional operating costs over the 8 year period, for a total of $90 billion in new CIS expenditures—about $11 billion total in CIS spending per year.

Given the generally poor quality of the evidence behind the assumptions, a much lower (e.g. $50 billion) or much higher (e.g., $130 billion) amount over 8 years is equally plausible for the hospital sector. The important point is that these amounts can be seen as a rough, order-of-magnitude range for additional hospital CIS spending.

Prospects for New Clinical Information System Deployment

Is the needed CIS spending over 8 years "feasible" for the U.S. hospital sector? In 2006, $90 billion for new CIS spending would translate into an average 1.7 percent increase in hospital spending per year. Assuming no financial benefit from CIS investment, and given median hospital net margins of around 5 percent in 2007 (MedPAC, 2008), such CIS spending might be feasible for many hospitals only if some major construction projects were delayed. Obviously, such spending would not be feasible for

[2] To keep the exposition simple, we do not include spending to replace some of the additional CIS hardware (a shrinking part of total CIS expenditure) and ignore interest and discount rates for future costs and benefits.

[3] In fact, the total "new" CIS costs would be somewhat higher, since any new CIS capital expenditures create new capital replacement (depreciation) costs and new operating costs—for software maintenance, additional information systems, clinical staffing, and so on.

financially weaker hospitals or if hospital margins deteriorated due to a difficult economy.

Stakeholders care about "net" CIS costs—that is, total CIS costs *less* cost savings and new revenues. From the hospital's perspective, the good news is that some CIS-related financial return to the hospital is likely, since hospitals will use EHRs to create some efficiency savings from operations, and generate some revenues from higher reimbursement coding and from new services—benefits that potentially could make a substantial contribution towards paying for new CIS costs. Any government subsidies would contribute an additional amount. Moreover, despite a likely unfavorable measurable financial return on investment (as of mid-2008), many executives and boards—especially in larger hospital systems and larger nonsystem hospitals—appeared to view advanced CISs as a cost of doing business. That is, while a CIS investment might not be justified based on a measurable return on investment analysis, many health system leaders see it as necessary expense in order to compete successfully in a market place that increasingly will compare organizations by the quality of care provided (Miller et al., 2009a).

Physician Practice Clinical Information Systems

What's Needed and What's Been Adopted

Physician offices typically can use two different types of information systems. The best chronic disease management systems (CDMSs) for chronic and preventive care (e.g., for diabetics, asthmatics, women needing cervical cancer screening) use electronic data from billing, scheduling, registration, and lab systems, plus manually entered data, to create paper patient data summaries and reminders for visits, along with lists of patients needing services and provider performance reporting. In this setting, CDMS software coexists with the paper medical record. Simpler CDMS software imports no electronic data. While such systems are useful, policy attention has focused on ambulatory care EHR software that typically includes a suite of capabilities that physicians can use in day-to-day care, including electronic viewing, documenting, prescribing, lab order entry, care reminders, and messaging, as well as the capabilities found in CDMS software (see Table 3-2). Here we focus only on EHRs.

Order-of-Magnitude Cost Estimates

A rough estimate of overall "new" CIS capital spending on physician EHRs can be created by multiplying the number of active office-based phy-

TABLE 3-2 System and Capabilities in Chronic Disease Management Systems and Electronic Health Records

System and Capabilities	Explanation and/or Benefits
Chronic Disease Management Systems (for chronic/preventive care patients)	Best products use electronic data from billing, scheduling, registration, and lab systems, plus some manually inputted data; keeps paper chart
Patient data summaries (paper)	Provides relevant data at the point of care
Reminders (paper)	On the paper data summaries
Lists of patient needing services	Permits outreach to patients overdue for tests or visits
Provider performance reporting	Enables managers and providers to understand Quality Improvement performance
Electronic Health Records	Replaces paper chart; best ones also replace chronic disease management systems
Prescribing	Permits drug–drug/allergy interaction alerts; reduces input errors
Lab ordering	Reduces input errors
Documenting	Best products have templates for types of patients
Messaging with providers	Improves provider communication
Messaging with patients	Improves patient–provider communication; best products enable patients to view data, order prescriptions, make appointments
AND	
Patient data summaries	Provides relevant data during visit; enables customizable views
Reminders	Typically built into documenting and ordering
Lists of patient needing services	Permits outreach to patients overdue for tests or visits
Provider performance reporting	Enables managers and providers to understand Quality Improvement performance

SOURCE: Derived from Miller et al., 2009b.

sicians times the EHR capital cost per physician, less EHR costs already incurred (based on an estimate of physician EHR adoption), and adding new operating and depreciation costs over time. In 2005–2006, between 310,000 and 500,000 physicians practiced in office settings (depending on the data source), while the EHR capital cost per physician was around $40,000.[4] If only 10 percent of the higher estimate of office-based physicians had robust EHRs, the total EHR capital cost over 8 years would be around

[4] The lower estimate physician cost is based on data from E. Hing and C. Burt, *Characteristics of Office-Based Physicians and Their Medical Practices: United States, 2005-2006,* (Hyattsville, MD: National Center for Health Statistics, 2008), while the higher estimate is

$20 billion, or $2.5 billion per year, plus some hardware replacement cost. Since physicians have already spent funds on EHR capital projects, and are already spending funds on EHR capital expenses, "new" spending on EHRs might amount to about $15 billion over an 8-year period.

Using the same approach for office-based physician offices as for hospitals, and taking the higher office physician estimate, an order-of-magnitude estimate for new physician EHR spending would amount to roughly $40 billion to 50 billion over 8 years, including new operating expenses—or about a 1 percent to 1.25 percent average increase in physicians-services sector expenditures per year averaged over each of the 8 years.

Here, too, CIS cost over 8 years is a feasible expenditure for most physician practices, even in a worst-case scenario of no offsetting savings or increased revenues or subsidies. In fact, some evidence suggests that the financial return to office-based physicians could be substantial (even without subsidies), which could greatly reduce the net CIS expenditure figure (Miller et al., 2009a).

Prospects for New Clinical Information System Deployment

Large medical groups have been implementing EHRs at a good pace (DesRoches et al., 2008), because some groups face a favorable CIS business case, and because some large groups consider CISs a cost of doing business for reasons similar to the hospital sector. Solo/small groups (i.e., 10 physicians or fewer) were adopting EHRs more slowly, because the EHR business case was not perceived as favorable enough to physician practice owners, EHRs were disruptive and stressful to implement, and adequate technical support was typically hard to find. The pace of CIS implementation for all types of practices should increase to some extent, due to anticipated or actual patient pressure and greater reimbursement rewards for EHR-enabled performance, and (possibly) some subsidies from hospitals seeking to bind physicians to their organizations. Obviously, any new government subsidies or new support services could substantially increase EHR adoption. In the physician sector, absent any special CIS subsidies, financially weaker organizations would fall behind in CIS adoption, a special concern when some of those organizations also serve the disadvantaged and underserved.

---

based on D. Smart, *Physician Characteristics and Distribution in the U.S.: 2006 Edition*, (Chicago, IL: American Medical Association, 2006).

Is the Clinical Information System Expense Worthwhile?

Again, these estimates of new CIS costs for hospitals and physicians are very rough. They are intended as order-of-magnitude estimates to put the overall potential cost in the perspective of the overall spending in the hospital and physician healthcare sectors.

Will benefits justify the substantial cost for hospital and physician EHRs? We know that simply implementing EHR capabilities doesn't mean they will be used for EBM. Much out-of-the-box EHR software lacks easy-to-use and useful evidence-based templates with reminders and alerts, reports on patients needing services, and reports on provider performance; meanwhile, the information systems staff expertise that hospitals and large groups can tap to compensate for software limitations has been unavailable to the majority of physicians in solo and small groups. Moreover, without special performance incentives, many healthcare providers won't use even easy-to-use and useful software since practicing EBM requires difficult changes in workflow and sometimes additional staff. Even if providers did use the software, it could take years for comparative effectiveness researchers to obtain truly comparable data from many different organizations' practice settings. Obtaining such data requires promulgating precise definitions of measures and methods of obtaining data, but enforcing such standards would be especially difficult to achieve given the wide variation in EHR and billing software, physician documentation and data validation practices, and data from health information exchange—all of which could affect the quality of CER measures.

Increasing EHR use and especially EHR use for quality improvement will depend on a series of substantial changes in out-of-the-box EHR software, government and payer financial incentives, public performance reporting, EHR support services, and improved health information exchange. While each can contribute to increasing EHR adoption and especially use for quality improvement, appropriate financial (dis-)incentives and public reporting are the most important policy carrots and sticks that can encourage providers to practice EBM.

DATA AND INFORMATION HUB REQUIREMENTS

Carol C. Diamond. M.D., Ph.D., Managing Director, Healthcare Program, Markle Foundation

Overview

The vision set forth by the IOM's *The Learning Healthcare System* is compelling, and it has been clearly articulated in that workshop summary (IOM, 2007). In a learning health system,

- the "best evidence for the collaborative healthcare choices of patient and provider" is generated and applied,
- the "process of discovery [is] a natural outgrowth of patient care," and
- the system encourages and ensures "innovation, quality, safety, and value in health care."

Today's healthcare system falls very short of realizing these goals. As outlined in the workshop summary, the healthcare system continues to suffer from escalating costs, poor quality and outcomes compared with other industrialized nations, large variations and inconsistencies in the delivery of evidence-based care, and a chronic failure to apply even the currently available research and evidence base to the actual care that is delivered. While research and innovation are rapidly accelerating the development of new treatments and diagnostics, progress in using and applying evidence in healthcare decision making lags far behind.

HIT holds great promise in accelerating both the research needed and its dissemination in order to bring about a learning health system. Indeed, it is arguable that the greatest promise of HIT lies in its ability to enable *networked* analysis—or the rapid learning via a networked and distributed approach to information sharing and evidence development about what works and what does not work in clinical care. To maximize this potential, it is critical to address four key challenges. The following things must be done:

1. Clearly define the ultimate goal.
2. Be open to reset definitions and assumptions about health data and research approaches.
3. Articulate new, broadly accepted working principles based on 21st-century information paradigms.
4. Develop an information policy framework that broadly addresses public hopes and concerns.

Clearly Defining the Ultimate Goal

What is the goal? First and foremost, the goal should be to generate and use information to improve actual healthcare decision making for the many and varied participants in health care, from patients and clinicians at the point of clinical decisions to policy makers charged with creating new financing approaches to public health experts responsible for detecting emerging public health risks.

Too often clinical research efforts fall short of these critical end goals. For example, a great deal of time and money are allocated to discussing the

specific technologies, data collection efforts, methodologies, and analytic protocols and standards, as well as to controlling for confounders. Rarely is the same effort or priority placed on putting "research into practice" (Woolf, 2008). All too often no one considers what it is that will have the biggest impact on the quality of everyday health decisions.

What is the alternative? Since 2002 the Markle Foundation has brought together over 100 leading experts in the fields of research, medicine, policy, and business in a public–private collaborative called *Connecting for Health* to tackle the thorny and practical issues involved in getting critical information into wider use throughout the healthcare system. As a way of painting a picture of the goals we are all working toward, we created a series of hypothetical future scenarios that demonstrate the power of networked analysis (Connecting for Health, n.d.a).[5]

In one example, we describe a physician in a small practice in the suburbs. In preparation for his visit with his next patient, Theresa, he clicks on a button and is able to access information that tells him the patient is coming in to determine whether she might need to switch to a new diabetes medication. At the same time, the information system he uses allows him to compare his own patients' outcomes to the outcomes of his peers' diabetic patients. He uses an information *network* to get the answers that he needs.

The patient in this scenario also has a respiratory infection. The literature suggests that a particular antibiotic is recommended, but taking into account recent community outbreaks and their sensitivity patterns, the system helps the physician determine that a different antibiotic will likely produce better results.

In this future scenario, the clinician's decision making is based upon networked analysis of research evidence and local information in real time. Using these approaches, a different kind of "research" emerges. Instead of relying solely on the assembly of research data in large databases in centralized research centers where the data are analyzed over months and years by scientists outside of everyday healthcare delivery, networked information and distributed analytic tools make it possible for clinicians and patients to answer their real, practical questions in real time in order to make better decisions.

This is a very different paradigm from the one people live with today, and it is a reminder to challenge assumptions about health data and research approaches.

[5] For more information see www.connectingforhealth.org/connectivity/ (accessed September 8, 2010).

Be Open to Resetting Definitions and Assumptions About Health Data and Research Approaches

A comprehensive approach to clinical CER is inherently challenging today because of the lack of a controlled environment for assessing therapeutic options, the heterogeneity of patient characteristics, and the distributed nature of both the requests for, and the sources of, information. These challenges should encourage us to think in new terms. How can these "problems" become the jumping-off point for a new and much more distributed approach to research that is characterized by connectivity and networks and not just databases?

The current world of biomedicine is best characterized as a collection of information "islands" containing an unprecedented explosion of data, with more peer-reviewed articles published every year. The problems with the field lie not so much in the generation of new data or evidence but rather in the slow and uneven dissemination of innovation and information. While the findings may be cutting edge, the approach for sharing them and getting them into practice has been likened to "methods recognizable to Gutenberg" (Buetow, 2008).

Looking around at the environment today, it can be seen that most efforts are focused on data collection, data cleaning, and then compartmentalizing the data across a highly fragmented system. This approach is slow and costly and often falls short. Very often the data needed for a particular research question are hard to collect, of uneven quality, and ultimately poorly suited to provide the answers sought.

The current model also places a huge burden on providers because they are being asked to supply the same information to different requestors in different ways. This results in redundant repositories of information created in response to many different research questions or purposes. The model is inefficient in that it is difficult to repurpose or reuse specific pieces of data. Furthermore, large aggregate data sets create greater privacy and security risks. Most important to consider, however, is that the model lacks what is essential: connectivity and feedback loops. Only with networks to connect the fragmented knowledge base and the capacity to use feedback loops will it be possible to meet the goals of a learning health system: to get information when and where it is needed to make better decisions.

What does it look like when the gap between clinical research and clinical delivery is overcome? Childhood cancer is often used as an example of an area where the silos of clinical care and research have been connected. Clinicians and researchers are part of a unique community that has been able to use clinical data continuously to evaluate outcomes to improve protocols and treatments. The result has been a dramatic improvement in treatment and survival rates. Today 75 to 90 percent of children 10 years

old and younger with cancer are in a formal protocol (NCI, 2008a). This compares to only 3 percent of adults with cancer who are in formal protocols (Murthy et al., 2004). The result has been a significant improvement in childhood cancer survival rates over the past decade and a half.

So what makes this community different? It is not that they were early adopters of HIT per se. Rather, it is that the gap between researchers and clinicians has been narrowed, both in terms of the lag between evidence creation and its use in clinical care and in terms of the blurring of the line between the roles of researcher and clinician.

But what if the lines were to be blurred even farther? What would it mean if not just clinicians but also patients were involved in driving research? What if patients could bring their own very real and pressing questions and unique information about treatments, symptoms, and disease progression to networked health information? There are examples of this already happening today. The Web community PatientsLikeMe is an example of a highly evolved patient social networking site that actively conducts research (PatientsLikeMe, 2008a). According to the site's founder, the site was "built to accelerate the transfer of knowledge about what works and what doesn't" (PatientsLikeMe, 2008b). Its community includes more than 1,600 amyotrophic lateral sclerosis (ALS) patients, which is twice the number of patients in the largest ALS trial ever. While patients openly discuss a wide range of symptoms and side effects, which may include some very personal information, the most important point is that this is a research community. The data collection is highly structured, with patients using validated tools to collect a significant amount of data. The participants on the site have come together to share their data because they believe that, by doing so, they will learn more about their conditions in return. They are involved in the first real-world, open, and nonblinded patient-driven trial on the use of lithium for ALS.

It is fascinating to see what patient involvement and activation can do for research. For instance, one of the ALS research scales in use was not appropriate for patients with a forced vital capacity below 50 percent. A patient suggested a change to the scale, which was accepted, and the scale was modified so that it could better accommodate patients with low forced vital capacity.

Observing this network and seeing its potential makes it clear how valuable it can be to begin to shift current ways of thinking about data collection and patient involvement, and it encourages an opening of the aperture on how we think about data creation and collection.

Articulate New, Broadly Accepted Working Principles Based on 21st-Century Information Paradigms

As illustrated in the examples above, it is important to acknowledge and leverage the characteristics of the 21st-century environment in which the needs for sharing and accessing information are increasingly distributed. As consumers, physicians, and others increasingly use the Internet to create, access, and use health information, the traditional paradigms are changing dramatically.

Connecting for Health has articulated a set of first principles that can serve as a guide or set of working principles to characterize the current environment (Connecting for Health, n.d.b).[6] These principles suggest that as the need for information is increasingly distributed, it follows that it should be possible to leave the data distributed as well. The professional and ancillary sources of data—or, to use a network term, the "nodes"—are becoming more sophisticated in terms of analytic capabilities. It should not always be assumed that knowledge creation starts with collecting the data from the source. New paradigms should assess the merits of *pushing* the question to the data across a network and only analyzing the answers centrally. It is important to see research as occurring in a connected environment where clinicians, researchers, and patients are networked and where the strict roles of researcher and clinician are necessarily blurred.

A compelling example of a distributed network being used for national and international flu surveillance is the Distribute model.[7] This model uses summarized counts of influenza-like-illness (ILI) syndrome reported from existing syndrome surveillance systems. Rather than centralizing the data collection or requiring standardized definitions across data sources, the system relies on local definitions of ILI and collects only aggregate counts of ILI by age band. The result is an ability to track flu trends more quickly and cost effectively than ever before. During the 2007–2008 flu season, the total weekly volume of the 8 regions participating in the current Distribute network (about 250,000–350,000 visits per week) was comparable to the volume of total visits in the nationwide sentinel reporting system (about 200,000–400,000 visits per week) funded and operated for decades by the Centers for Disease Prevention and Control (CDC). Furthermore, the very low risk of a privacy breach has encouraged voluntary participation by both national and international participants.

This example demonstrates what can be achieved when a project focuses on only the minimum information required to inform public safety. By reducing the time and effort required to centrally collect, clean, and

[6] For more information see www.connectingforhealth.com/resources/first_principles.pdf (accessed September 8, 2010).

[7] See http://www.syndromic.org/index.php (accessed August 5, 2010).

analyze data from multiple sources, and by reducing privacy risks that can discourage participation, a distributed model allows for information to be shared with decision makers—from local public health departments to local hospitals or school systems—more quickly and cost effectively.

Develop an Information Policy Framework That Broadly Addresses Public Hopes and Concerns

The purpose of this meeting is to conceive a new body or mechanism to address CER. Many suggest that government has a critical role to play in funding and establishing such a center. But to establish a powerful and distributed network to improve health and health care, the investment should focus on four things: motives, standards, methods, and rules.

The *motive* to share information is critical, and at present research communities still remain silos of information and data. Clearly government has an opportunity to link research funding to a requirement to participate in a more collaborative, distributed, and networked approach. The *standards* will determine how to make information shareable and useable without having to first collect it in one place. New *methods* are needed—and are emerging—for doing research in a more networked and distributed way. Examples of distributed models used in clinical, quality CER efforts show that networked models for handling composite data analysis can be flexible in order to address a range of research questions. For example, Shared Pathology Information Network was a research initiative of the National Cancer Institute (NCI) designed to allow cancer researchers access to a virtual database to locate appropriate human tissue specimens across pathology laboratories and institutions (NCI, 2004). The model allows authorized researchers greater access to data from multiple institutions while preserving local control of the data by those institutions. The distributed research network is an example of a distributed model for comparative effectiveness research.[8] In this model, supported by the Agency for Healthcare Research and Quality (AHRQ), multiple types of information sources, including administrative data, EHRs, inpatient data, and disease registries, are leveraged across a federated system that will allow for composite data analysis without requiring the aggregation of all the raw data in a single centralized database. Lastly, Cancer Biomedical Informatics Grid, another project of the NCI, is an attempt to create a networked system that allows multiple users from the cancer community to access large data sets in standard formats without creating a centralized database of all of the raw data (NCI, 2008b).

[8] See http://sites.google.com/site/phgrid/Distributed-Research-Network (accessed August 5, 2010).

Finally, and most critically, new *rules* are needed to promote public trust in information sharing. The *Connecting for Health* common framework offers a 21st-century privacy approach (Connecting for Health, 2008a). Since 2002 the Markle-led *Connecting for Health* collaborative has brought together key organizations from all sectors to develop a common approach to information policies. Our own efforts have demonstrated the value and importance of establishing information policies, rules, and technologies that satisfy the following three requirements:

1. core privacy principles,
2. sound network design, and
3. oversight and accountability.

Core Privacy Principles

The nine core privacy principles, summarized below, are based on U.S. Fair Information Practices. Meaningful safeguards will be achieved by using both policy and technology tools to achieve the core privacy principles and by ensuring that these nine principles are applied together.

1. **Openness and transparency:** Policies for information use and sharing are clearly communicated to participants.
2. **Purpose specification:** The purpose of the data collection effort is clearly specified and narrowly suited to the need.
3. **Collection limitation and minimization:** Only data needed for specified purposes are collected and shared.
4. **Use limitation:** Data are used only for the agreed upon and stated purposes.
5. **Individual participation and control:** Individuals can find out what data have been collected and who has access, exercise meaningful control over data sharing, have access to information about them, request corrections, and see audit logs.
6. **Data integrity and quality:** Mechanisms to ensure the data are relevant, accurate, complete, and up to date.
7. **Security safeguards and controls:** Tools and mechanisms are in place to ensure that data are secured against breaches, loss, or unauthorized access and improper authentication.
8. **Accountability and oversight:** Mechanisms and accountable parties are established for monitoring compliance with policies and procedures for handling a breach.
9. **Remedies:** Mechanisms for handling complaints and remedies for affected parties are established in the event of a breach.

Sound Network Design

Sound network design helps ensure that information is protected while it is shared. Sound network design should do the following:

- Incorporate technical tools that facilitate trusted use: audit, access, authorization, authentication, and accuracy.
- Promote technological choices that limit the potential for abuse and mitigate the risks of large breaches, including distributed architecture and use of de-identified information.
- Enable interoperability and flexibility, supporting a diversity of applications, using secure, open Web standards.
- Support and encourage networked approaches to information sharing that are consumer accessible, including through the Internet and mobile devices.

Oversight and Accountability

HIT efforts must establish oversight and accountability, including critical governance and enforcement mechanisms. These mechanisms should do the following:

- Include all affected in the development of approaches and policies.
- Ensure that the framework and its attributes are adopted.
- Include clear mechanisms of enforcement appropriate to the specific activity, such as through contractual agreements or regulatory mechanisms.
- Designate responsibility for monitoring and oversight.

These attributes have guided the work in developing detailed policies and technology approaches for health information exchange and for services that enable consumers to access their own health information. The *Connecting for Health* common framework has been developed and adopted by providers, insurers, e-health companies, consumer groups, and privacy experts (Connecting for Health, n.d.c).[9]

The public understands the opportunity but needs to trust the system in order to fully participate. While individuals do not want their data to be misused, our most recent survey shows that three-fourths of those surveyed see the value in sharing their personal information to look for

[9] For more information see www.connectingforhealth.org/resources/CCEndorser.pdf (accessed September 8, 2010).

disease outbreaks or to improve information for research (Connecting for Health, 2008b). Patients won't accept an either/or proposition: safeguard my data or use it to improve my health. They, like us, urgently want (and need) both.

INTEGRATIVE VEHICLES REQUIRED FOR EVIDENCE REVIEW AND DISSEMINATION

Lorne A Becker, M.D., Emeritus Professor, SUNY Upstate Medical University, and Co-Chair, Cochrane Collaboration Steering Group

Overview

To implement what the IOM terms a *learning health system,* in which the most effective clinical practices reflecting the best available evidence are naturally embedded in patient care, necessary infrastructure must be established to efficiently develop and disseminate knowledge about what works best in health care. Evidence to support clinical decision making can come from multiple sources, but primary studies, whether randomized trials, cohort studies, case studies, cross-sectional studies, or studies using other designs rarely provide adequate answers to questions of clinical effectiveness as individual pieces of evidence. Research studies comparing the effectiveness of different treatment strategies must be combined using valid methods into evidence syntheses that show the combined results of all relevant research on a given topic to inform healthcare decisions.

Many countries have well-established national mechanisms for producing evidence syntheses. In addition, a growing number of international collaborative efforts have been developed to work together on various components of the evidence synthesis process in a way that spreads the load, reduces duplication, and benefits all. This paper outlines some of the opportunities available for the United States in these international collaborative activities.

Types of Evidence Synthesis

Many different approaches have been taken in designing evidence syntheses—with significant variations in the methods used, their complexity, and the reproducibility of their results. Several examples exist in the United States. AHRQ conducts and supports CER through its Effective Health Care program. Part of this program involves evidence-based practice centers (EPCs) established (1) to facilitate the synthesis of knowledge from data generated by a network of research organizations and (2) to then translate that knowledge into patient-targeted information. AHRQ

also sponsors the U.S. Preventive Services Task Force, a private panel of experts in primary care and prevention that reviews evidence and generates recommendations for clinical preventive services. Other federal programs include the Centers for Medicare & Medicaid Services (CMS) Medicare Evidence Development and Coverage Advisory Committee, established to conduct reviews of clinical effectiveness to advise on medical topics under evaluation at CMS; the Drug Effectiveness Review Project, a public–private collaboration reporting on the comparative effectiveness of drugs within and between drug classes; and the National Institutes of Health (NIH) Consensus Development Program conferences that convene independent panels of experts to collect information and develop consensus statements on a clinical topic selected by NIH staff. The Drug Effectiveness Review Project and the Consensus Development Program both consider the evidence reviews from EPCs in developing their reports.

In the private sector, Blue Cross and Blue Shield Association's Technology Evaluation Center assesses clinical effectiveness for both public and private entities and is a designated EPC; the ECRI Institute, another EPC, is nonprofit and conducts cost-effectiveness analyses and technology assessments for both public and private health-sector organizations; and Hayes, Inc., is a for-profit organization that develops technology assessments for healthcare organizations and networks.

In virtually all of these U.S. examples, the focus is on relatively complex evidence syntheses, which attempt to synthesize evidence over a broad domain. By contrast, the majority of evidence syntheses produced in the United States and elsewhere are systematic reviews with a much more narrowly targeted focus. As an example, a focused systematic review might examine the effectiveness of a single intervention, such as inhaled corticosteroids (Nannini et al., 2007) or influenza vaccination (Poole et al., 2006) in preventing episodes of chronic obstructive pulmonary disease (COPD). Complex reviews, such as those produced by the organizations listed above, typically take on a much broader question that can only be addressed by synthesis of evidence about multiple interventions of different sorts. For example, the evidence synthesis prepared for the U.S. Preventive Services Task Force, in examining the question of whether asymptomatic individuals should be screened with spirometry to detect undiagnosed COPD (Lin et al., 2008), outlined eight different "key questions" addressing diverse issues that included prevalence and risk factors for COPD, accuracy of spirometry, smoking cessation rates, and the effectiveness and potential harms for a wide range of potential interventions for those individuals newly diagnosed as having COPD. Many of the individual key questions were sufficiently complex to each require a complex rather than a focused evidence synthesis—an example being the single key question addressing the effectiveness of pharmacologic treatments, oxygen therapy, or pulmonary

rehabilitation in reducing morbidity and mortality. A practice guideline is typically broader still, addressing and making recommendations about screening, diagnosis of symptomatic individuals, and management at all levels of the disease (Qaseem et al., 2007).

Because of their complexity and costs, only a handful of these complex evidence syntheses are produced each year. For example, only 167 such syntheses were produced in 2006 (IOM, 2008). Clinical practice guidelines are produced in much larger quantities with much less consistency in their quality or rigor in their methods. Hundreds of new or revised guidelines are added each year to www.guideline.gov, with a variety of organizations involved in their production, including government agencies, professional organizations, and health systems.

Annual production of focused systematic reviews is much larger than that of the more complex syntheses. In contrast to the 167 complex syntheses produced in the United States each year, it has been estimated (Moher et al., 2007) that approximately 2,500 systematic reviews were published in 2004, and this number is growing rapidly each year. Authors from the United States contributed about a quarter of these systematic reviews. However, taking into account all reviews done by other countries normalized to population, the United States contributes fewer reviews than other nations, such as New Zealand, Australia, Netherlands, United Kingdom, and Canada, on a per capita basis. This suggests not only an opportunity for increased United States involvement but also for significant gains possible through greater international coordination.

In addition to their narrow focus, systematic reviews tend to be selective about the types of research that they include. Approximately 60 percent of the articles reviewed by Moher et al. (2007) included only evidence from randomized controlled trials (RCTs). A further 25 percent included data from quasi-randomized studies or non-RCTs, and only 12 percent included data from cohort or case-control studies or studies with other observational designs. The focus on randomized trials is primarily based on concerns about methodological quality and the risks of bias that are inherent in studies that employ other research designs.

Because the RCT is the study design least susceptible to bias (particularly selection bias), the methods for finding and evaluating risks of bias in systematic reviews based on RCTs have been well studied and the best approaches to dealing with these risks have been outlined in detail (Higgins and Greene, 2008). There is much less agreement on the most appropriate methods for combining results from studies with other designs—primarily because of the difficulty in assessing the probability and magnitude of bias. Oxman et al. (2006) have summarized the types of approaches used by various organizations in addressing this issue in the preparation of clinical guidelines and concluded that "as the range of study designs that are

included is broadened, an increasing amount of work is required to drive decreasingly reliable estimates of the effects of interventions."

Advantages of Focused Systematic Reviews

While focused systematic reviews are sometimes seen as less useful for decision makers because of their narrow scope and reliance on randomized trials, they have a number of advantages over more complex reviews. Because of their focus, systematic reviews are less expensive and require less effort to produce than complex evidence syntheses, which require review and processing of larger bodies of literature and often need a larger team that includes individuals with a broader set of methodological skills and content expertise. Systematic reviews may also have an advantage over complex syntheses in their generalizability, since complex reviews are more likely to include consideration of factors, such as costs, availability, and other issues that vary from setting to setting.

Because evidence syntheses of all sorts may become invalidated when new evidence appears, there is general recognition of the need for periodic updating of these documents. This is a larger problem for complex syntheses than for focused systematic reviews for two reasons. Because they summarize more evidence across a broader area of content, complex reviews are at higher risk that an important new study will appear that is relevant to one of the many topics they include. Thus complex reviews are likely to require updating more frequently than focused systematic reviews. Also, because of their complexity, updating is likely to be a more difficult task, because changes in the evidence for one of the subquestions of the synthesis may have important implications for other subquestions or their combined interpretation. Because of their narrow focus, methodological issues in systematic reviews have been subjected to much greater study than the methods for more complex syntheses. As noted above, much of this work is focused on reviews that combine results from RCTs. However, methods development continues for the incorporation of other designs, particularly for questions that may not be well addressed using RCTs alone. Thus focused systematic reviews that restrict themselves to RCTs are likely to give the most unbiased estimates of effects of the studies being grouped in the evidence synthesis, and are likely to have the greatest ability to assess and estimate the possible effects of a variety of risks of bias. Focused reviews that extend their scope to studies other than RCTs are more susceptible to bias, and complex reviews that combine different study designs for a variety of related questions are most susceptible to bias and have the least developed methods.

In contrast to complex reviews, for which funding is available from a variety of sources, there is very little grant or other funding available in the

United States to support the production of focused systematic reviews. The relatively large number of reviews produced despite this handicap reflects the fact that they are increasingly gaining recognition as valid pieces of scholarly work. Systematic reviews are frequently compact enough to be submitted as journal articles; those that are published are often highly cited. Some journals are de-emphasizing the traditional narrative review article in favor of this more scientific and disciplined approach to synthesis. In addition, the recent recommendation that planning for new research studies should always begin with identification of an up-to-date systematic review, or performance of one if no such review is available (Clarke et al., 2007), may further accelerate the increase in production of systematic reviews.

Priority Setting

Because of their complexity and costs, it is clear that a prioritization process will be needed to direct the efforts of those producing complex evidence syntheses to be certain that the limited number that can be funded will address the most appropriate topics, and any prioritization process will of necessity exclude many important but lower priority questions. However, even individuals with lower-priority conditions, or disorders requiring less expensive interventions, will be best served by a clear knowledge of the evidence available about comparative effectiveness. For the many decision makers, patients, clinicians, policy makers, and others who find themselves needing evidence on questions not covered by the small available set of complex evidence syntheses, focused systematic reviews can fill important gaps. In fact, with the rapid growth in the production of systematic reviews, it is intriguing to speculate that it may someday be possible to find all relevant high-quality studies in the literature summarized using this technique. An initial analysis of the number of systematic reviews needed to synthesize the evidence from all RCTs has already been completed (Mallett and Clarke, 2003).

Thus the need for prioritization is very different when viewed from the perspective of focused systematic reviews than when considering complex evidence syntheses. For the latter, it is important to concentrate resources on a relatively small set of high-priority, large-impact questions of comparative effectiveness. For systematic reviews, however, the priority is to create a comprehensive resource that provides syntheses that cover as much as possible of the existing evidence terrain. A second priority for those producing systematic reviews needs to be the coordination of their efforts, so that individuals wishing to find a relevant evidence synthesis will be able to do so without being confused by the availability of multiple overlapping and possibly conflicting systematic reviews.

Inclusion of Focused Systematic Reviews in Complex Reviews

Another advantage of focused systematic reviews is their ability to serve as building blocks for guidelines or more complex evidence syntheses. Because they restrict themselves to narrowly focused questions and tend to exclude broader contextual issues, their results may be less restricted to a specific country, population, or care setting. Results of a focused systematic review can then be combined with additional information from a variety of different sources (using complex syntheses, guidelines, or a variety of other techniques) to address broader questions—thus fitting with the strategy of "globalize the evidence: localize the decision" outlined by Eisenberg (2002).

This approach of using systematic reviews as building blocks is explicitly used by many producers of guidelines and other complex evidence syntheses. Detailed recommendations for the process of finding, evaluating, and incorporating systematic reviews have been made based on the experiences of producing complex evidence syntheses in U.S. EPCs (Whitlock et al., 2008).

A recent survey by the Appraisal of Guidelines, Research and Evaluation collaboration (Burgers et al., 2003) found that most guideline developers make use of systematic reviews in their guideline development process. In their series of background papers commissioned by the World Health Organization (WHO) to improve the use of research evidence in WHO guidelines and other products, Oxman and colleagues (2006) have also noted the widespread expectation that systematic reviews should be used to inform the development of guidelines and have produced recommendations about how this should be done.

Even if an existing review does not address the exact question needed for the complex synthesis, or if it is in need of updating, the detailed specification of search strategies and results included in high-quality systematic reviews, along with their critical appraisals of the studies included in the review, can give developers of complex reviews a head start and decrease the time needed to produce their reviews.

A Combined Approach to Focused and Complex Evidence Syntheses

Given these advantages, plans to increase support for evidence syntheses in the United States should recognize the need for both components—a targeted program of complex syntheses supported in a very directive fashion accompanied by more general efforts to build a diffuse network of skilled producers of focused systematic reviews that can be used as building blocks for guidelines and complex syntheses. A number of countries have

taken this approach and have developed mechanisms to actively support the production of both types of syntheses.

As an example, in the early 1990s, an explicit decision was made in the United Kingdom to set up two different types of centers to address these two functions (Sheldon and Chalmers, 1994). The National Health Service (NHS) Centre for Reviews and Dissemination was charged with either proactively commissioning or carrying out focused reviews for the NHS. The UK Cochrane Centre was given the task of participating in an international collaboration to build, maintain, and disseminate a database of focused systematic reviews of RCTs and to keep these reviews up to date. In subsequent years, a number of other countries (including Canada and Australia) have adopted a similar approach by funding activities of the Cochrane Collaboration as a mechanism for production of focused systematic reviews by participating in an organized international effort, in addition to setting up other mechanisms for the production of more complex syntheses.

The Cochrane Collaboration

The Cochrane Collaboration is an international network with the aim of improving healthcare decision making, by producing and regularly updating systematic reviews synthesizing the results of these controlled clinical trials. Approximately half of the 16,000 people in over 90 countries who work with the Cochrane Collaboration are review authors, but there are many additional roles that are filled by organized subgroups of the collaboration known as "entities," of which there are four different types: Cochrane review groups provide the editorial role for Cochrane reviews and support authors in a variety of ways, including their search for evidence to include in the review. Each review group focuses on a particular area of health (e.g., colorectal cancer, infectious diseases, schizophrenia, tobacco addiction). Cochrane Centres and their branches support Cochrane contributors and entities located in their geographic and linguistic area—providing coordination, training, help with translations, and networking. Methods groups are made up of individuals interested in advancing the still young and rapidly evolving science of research synthesis. These groups develop methodological standards and advise the collaboration on how the validity and precision of systematic reviews can be improved. Each Cochrane methods group focuses on a specific area such as statistics, adverse effects, bias, or information retrieval. Networks (or "fields") focus on dimensions of health care other than specific health problems, such as the setting of care (e.g., primary care), the type of consumer (e.g., older people), or the type of intervention (e.g., vaccines).

The majority of funding for Cochrane activities is directed to these entities—thus supporting the infrastructure required to produce reviews.

Cochrane authors for the most part volunteer their time, or find their own funding to support the authoring of their reviews. Funding for Cochrane entities comes from a large variety of national governments, international governmental and nongovernmental organizations, universities, hospitals, private foundations, and personal donations. The collaboration has made the decision not to accept funding from conflicted organizations, such as pharmaceutical companies, and this is clearly spelled out in organization-wide policy limiting uses of funds from corporate sponsors. Although much of the funding comes from national government sources, or from nongovernment funders based in a single country, the funds are used to support efforts in other countries as well. For example, Cochrane review groups based in the United Kingdom each provide support for authors in multiple countries in their area of content focus.

The editorial process for Cochrane reviews is quite different from that used by medical journals. It has been designed to not only ensure high-quality systematic reviews but to do so in a way that builds the skills of authors, and thus it has served to increase the workforce available to perform syntheses of this type. The process involves frequent iterative interactions between authors and editors at every stage of review production, beginning with the selection of a topic. No review can begin until the title has been approved by the relevant Cochrane review group. This avoids the duplication of effort that would result from different teams unknowingly working on different Cochrane reviews on overlapping topics and ensures that the planned scope of each review fits well with others in the Cochrane Library. Authors must next submit a protocol outlining in detail the approach they will take and the methods they will use for their review. Cochrane editors send this protocol to peer reviewers who have content, methodological, and statistical expertise, and who provide authors with detailed feedback. Completed reviews are again sent for peer review and are often extensively edited following reviewers' comments. Because "involving and supporting people of different skills and backgrounds" is one of the collaboration's key principles, involvement of relatively inexperienced authors is not discouraged.

Because Cochrane reviews have multiple authors, experienced reviewers can be mentors for their less experienced coauthors. In some cases, this mentorship process has been formally supported. One example is the human immunodeficiency virus (HIV)/acquired immunodeficiency syndrome (AIDS) Mentoring Programme (Oliver et al., n.d.). This collaborative project of the South African Cochrane Centre and the Cochrane HIV/AIDS review group (CRG) was established in 2000, when the CRG found that few of its HIV/AIDS systematic reviews were relevant to sub-Saharan Africa, and identified a need for first-time authors from the region to have ongoing support. The effort has been quite successful with 20 authors having

received mentorship. Three reviews have been published in the Cochrane Library, 12 are in progress, and the initiative has now been extended to include authors from South Asia.

The primary product of the collaboration is the Cochrane Database of Systematic Reviews (CDSR).[10] From a small beginning in 1995 when it contained only 36 full systematic reviews, this database has grown with each issue and now contains almost 4,000 systematic reviews covering the whole range of healthcare interventions. Initially, Cochrane reviews addressed only questions of effectiveness of interventions, using primarily evidence from controlled clinical trials. While this is still largely true, many Cochrane reviews now also include results of studies using other designs, such as controlled before/after and interrupted time series designs, for questions that have not been studied using randomized trials. Because of the international scope of the Cochrane Collaboration, the reviews cover a broad range of topical areas, with applicability for both developing and developed countries. Beginning in 2008, the database was expanded beyond its initial coverage of only systematic reviews of interventions and now includes reviews of diagnostic test accuracy (Leeflang et al., 2008) and reviews that synthesize research on issues relevant to systematic review methodology.

The Cochrane approach of producing a coordinated database of focused systematic reviews using an international collaborative process has a number of advantages and has been an effective way to build capacity for evidence synthesis in the countries participating in this effort. In addition to the mechanisms for author development noted above, the Cochrane process has contributed to capacity building by advancing systematic review methods, developing tools to help in the evidence synthesis process, and forging important partnerships with universities and other academic institutions.

Working in a Cochrane methods group provides opportunities for methodologists to further develop the methods used in systematic reviews and also provides them with a large set of reviews and protocols to serve as a substrate for their research. Detailed guidance on systematic review production from these groups has been incorporated into two Cochrane Handbooks (Diagnostics Test Accuracy Working Group, 2009; Higgins and Green, 2008). The groups have also developed the Cochrane Methodology Register[11] which is continuously updated and now includes more than 11,000 citations to journal articles, book chapters, conference proceedings, conference abstracts, and reports of ongoing methodological research. The aim of the register is to include all published reports of empirical method-

[10] Available from http://www.thecochranelibrary.com (accessed June 15, 2009).

[11] Available at http://www3.cochrane.org/access_data/cmr/accessDB_cmr.asp (accessed June 16, 2009).

ological studies that could be relevant for inclusion in a Cochrane methodology review, along with comparative and descriptive studies relevant to the conduct of systematic reviews of healthcare interventions.

In order to facilitate the production of high-quality systematic reviews by a widely dispersed international group of authors, the collaboration has developed a number of tools to help authors identify studies for inclusion in their reviews, use appropriate and standardized methods in conducting their reviews, and produce their reviews in the format specified for CDSR.

The literature identification tool is the Cochrane Central Register of Controlled Trials (CENTRAL). Each Cochrane review group maintains a database of relevant studies that includes references to both published and unpublished reports. These individual study registers are assembled quarterly into CENTRAL, which is then published as part of the Cochrane Library to make it available for broader public use by health researchers and others wishing to perform evidence syntheses. Approximately two-thirds of the references in CENTRAL are derived from specially designed searches of MEDLINE and Excerpta Medica Database (EMBASE). The remainder consist of references uncovered by authors or by organized efforts of Cochrane review groups, centers, or fields to find additional studies through activities, such as handsearching of journals or conference proceedings, follow-up of references from other studies, or contact with trialists or others who may have knowledge of additional studies not included in MEDLINE or EMBASE. A recent assessment showed that searching beyond CENTRAL found only a very small number of trials (Royle and Milne, 2003).

The collaboration's authoring tool is a complex piece of software known as RevMan. The software has been continually refined over many years. It is structured to guide authors through the appropriate steps in conducting and writing up their review, and links are provided to the relevant section of the Cochrane handbooks at each step. RevMan also incorporates a number of statistical tools, developed in conjunction with the Cochrane methods groups, that allow authors to perform meta-analyses of their data in a standardized way.

In addition to serving as essential infrastructure components for the collaboration, these resources have been made freely available and are now in widespread use in the production of evidence syntheses by others. For example, the majority of systematic reviews include a search of CENTRAL as one method of identifying studies, many reference the methods outlined in one of the Cochrane handbooks, and a large number are prepared using RevMan.

Many of the aims and activities of the Cochrane Collaboration fit well with the missions of universities and academic institutions. This has led to a number of fruitful collaborations. Most Cochrane editorial groups and many centers and fields are located in university settings or have close ties

with specific university departments or other units. Both parties benefit from this sort of collaboration—the Cochrane entity is supported and finds colleagues and collaborators within the university, while the university faculty not directly employed in the Cochrane entity become linked with an international network of individuals with unique methods and content skills and knowledge.

The ability of the Cochrane Collaboration to simultaneously increase the available number of systematic reviews and to improve and build the infrastructure required to support production of evidence syntheses of all types has led several countries to build support for the collaboration into their budgets. The majority of this support has been directed toward funding of Cochrane infrastructure (for example, support for Cochrane Review Groups) and has not been tied to the production of systematic reviews on specific topics. The United Kingdom has been the leader in this regard. Cochrane activities in the United Kingdom have been funded continuously since 1992, and the funders continue to feel that this approach of funding the infrastructure to support the production of methodologically sound systematic reviews on topics chosen by review authors is an important component of their approach to supporting evidence synthesis.

The United States has provided support to some Cochrane groups as well, using a variety of funding mechanisms. One of the first Cochrane editorial groups to be formed, the Neonatal Review Group has had funding from the National Institute of Child Health and Human Development (USA) for the support of its infrastructure since its inception. This has allowed the preparation and continuous updating of a classified bibliography of virtually all reports of randomized trials in the field of neonatology and of systematic reviews (incorporating meta-analysis) of the results of this body of research.[12] This group currently has 253 completed reviews with 65 reviews in progress. The HIV/AIDs review group has had CDC/Global AIDS Program (2008–present) and National Institute of Mental Health (2007–2008) support and has produced 52 reviews, with 51 reviews in progress. The Prostatic Diseases and Urologic Cancers has had Veterans' Administration (1998–2003) and National Institute of Diabetes and Digestive and Kidney Diseases (2005–present) support and has produced 29 reviews, with 23 in progress. The Cochrane Eyes and Vision Group (CEVG), started in 1997, has 73 completed reviews and 54 reviews in progress. The CEVG U.S. Satellite, with support from the National Eye Institute from 2002 to 2009, has produced 18 completed reviews and has 36 additional reviews in progress, to date. One of the most active of the collaboration's Fields/Networks is

[12] The Cochrane Neonatal Group. Available at http://neonatal.cochrane.org (accessed December 14, 2008).

the U.S.-based Complementary Medicine Field,[13] which is supported by the National Center for Complementary and Alternative Medicine. Unlike Cochrane Review Groups, Fields do not have a direct editorial role in the production of Cochrane reviews but identify health issues of importance to specific populations and/or intervention types and support CRGs in their production of relevant reviews in a variety of ways. One important function of Fields is their contribution of trials to CENTRAL, and the Complementary Medicine Field has been a major contributor in this regard, with a database that includes over 7,000 reports of clinical trials. The field has had a very active role in the identification, translation, and critical appraisal of reports of complementary medicine interventions published in Chinese journals, and also has assembled an organized list of all of the Cochrane reviews addressing complementary medicine interventions.

Advantages of a Collaborative International Approach to Evidence Synthesis

The example of the Cochrane Collaboration demonstrates the many advantages of an organized international approach to the production of evidence syntheses, and the benefits it brings in terms of prioritization, methods development, and capacity development. This section explores these issues in more detail and also notes other international collaborative approaches, some of which involve more complex evidence syntheses.

Prioritization

Because of the many disparate stakeholders involved and the large variations in morbidity, resources, and other factors from country to country, it is difficult to conceptualize an international process that would adequately represent all of the relevant perspectives in prioritizing questions in order to decide on a relatively small number of questions to be addressed by complex evidence syntheses. However, as noted above, in the case of focused systematic reviews, prioritization is more about coverage of the entire evidence terrain and coordination of efforts to avoid duplication. Thus a model in which local actors produce focused evidence syntheses, in accord with their own local priorities but also in a coordinated collaborative fashion, can result in a comprehensive set of syntheses that addresses multiple disparate priorities and that can serve as building blocks for more specific analyses performed at a local level.

[13] Available at http://medschool.umaryland.edu/integrative/cochrane_about.asp (accessed September 8, 2010).

Methods Development

An international approach can be helpful in developing the rigorous methods needed for evidence syntheses of all sorts. The required methodological expertise is frequently not available within a single country or region, and many of the issues are sufficiently complex that only a handful of individuals around the globe have a good grasp of them and are at the leading edge in their development. In addition to the work done within the Cochrane Collaboration, international collaborative groups have advanced the science of evidence synthesis in a number of ways.

One example is the family of standards for reporting of research studies and of evidence syntheses. The complete set has been assembled by the Enhancing the Quality and Transparency of Health Research Network—an international initiative that seeks to enhance the reliability of medical research literature by promoting transparent and accurate reporting of research studies.[14] The Quality of Reporting Meta-analyses (recently renamed Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines address reporting of systematic reviews of randomized trials (Moher et al., 2000), while the Meta-Analyses and Systematic Reviews of Observational Studies guidelines address reporting of systematic reviews of observational studies (Stroup et al., 2000).

A second international collaborative activity has addressed the methods to be used to assess the quality of evidence and the strength of evidence-based recommendations in a standardized way. Such methods have been developed by groups such as the Grading of Recommendations Assessment, Development, and Evaluation Working Group alliance (Guyatt et al., 2008).

Additional Examples of International Collaboration in Evidence Synthesis

A number of groups other than the Cochrane Collaboration have now begun to organize evidence syntheses using an international collaborative model. The Joanna Briggs Institute publishes a library of systematic reviews relevant to nursing.[15] The Campbell Collaboration, an international research network modelled after the Cochrane Collaboration, produces systematic reviews of the effects of social interventions involving areas such as education, crime and justice, and social welfare (Campbell Collaboration, n.d.). Some groups have had a more narrow content focus, such as the consortium of guideline development organizations and profes-

[14] See http://www.equator-network.org/about-equator (accessed June 12, 2009).

[15] See http://www.joannabriggs.edu.au/pdf/JBI_LibSR_info.pdf (accessed June 12, 2009).

sional societies formed to produce joint guidelines for the management of COPD (Schünemann et al., 2009).

In other cases the international efforts have been directed at organizing a standard format for specific components of complex evidence syntheses so as to allow portions to be shared and used in other countries or settings. Developers of health technology assessments in the European Union have taken this approach. Through a consortium organization (European Network for Health Technology Assessment) with 63 partners from 32 countries and that they are developing a "core model" that defines 9 different domains of a health technology assessment and that defines standard elements for each domain.[16] The model currently under development addresses only medical and surgical interventions, but a similar effort directed at diagnostic technologies is planned for the future. Similar work on standardization is being undertaken by the Guidelines International Network (GIN)—an international not-for-profit association of organizations and individuals involved in the development of clinical practice guidelines. GIN has defined a minimum data set that should be included in all evidence tables summarizing interventions (GIN, 2009) and is in the process of formulating a second to address evidence tables relating to diagnostic test accuracy. The aim is to have data in these tables presented in a consistent format that would allow guideline developers to use the efforts of others in developing their own evidence tables.

Future Directions

While the examples just listed show some beginning steps toward international integrated vehicles for evidence synthesis, there is much still to be done. Some of the additional needs include continued advancement in methods, particularly for complex evidence synthesis or syntheses involving designs other than RCTs; continued improvements in the quality of evidence syntheses; and improved coordination so as to decrease unnecessary duplication of effort.

Methods for focused systematic reviews that combine data from controlled clinical trials are well advanced. There is much less agreement, however, on the most appropriate methods for combining results from studies with other designs—primarily because of the difficulty in assessing the probability and magnitude of bias in these studies. Methods development for complex reviews and guidelines is even less advanced. As noted, international collaborative efforts have already made some beginnings in this area. On a national level, both the Centre for Reviews & Dissemination in

[16] See http://www.eunethta.net/upload/Founding%20Partners/EUnetHTA%20Collaboration%20Work%20Plan%202009_June292009_FINAL.pdf (accessed July 20, 2010).

the United Kingdom and AHRQ through its EPCs in the United States have made major contributions. Both have produced publications that address some of the methodological issues involved in complex syntheses (AHRQ, 2008; CRD, 2009)

Methods development for guidelines is also needed. In an international survey of 18 clinical guideline programs, Burgers et al. (2003) found little consistency in methods, although they did note that all respondents intended to develop their guidelines using rigorous methods. They also reported a trend toward increasing use of evidence-based methods, such as the use of electronic database searches and systematic reviews. Most guideline processes also incorporated consensus procedures of some sort. A recent literature review commissioned by the WHO (Schünemann et al., 2006) to inform its guideline production process found no experimental research or studies that compared components of guideline methods advice. They did note, however that many organizations that produce guidelines have a "guidelines for guidelines" document to guide their processes, and they found empirical evidence that organizations that publish their guidelines for guidelines produce more methodologically sound guidelines. The authors of the review were able to recommend a set of 19 principles for use by the WHO in guideline development.

Given the current state of methods development for evidence syntheses, it is clear that at least some of the funding for comparative effectiveness studies in the United States should be directed to promotion of further advances in methodology. While this funding could take the form of increased support for existing organizations within the United States, there would be clear advantages to align these efforts with international groups that are performing similar work.

Avoiding Duplication of Effort

Currently many different groups perform evidence synthesis of various sorts, and do so in a relatively uncoordinated way—leading to much needless duplication of effort. Greater organization of these efforts on an international scale would be helpful. One useful first step would be the formation of a registry of systematic reviews, analogous to the registries of clinical trials currently being set up. Prospective registration would allow any individual contemplating the performance of a systematic review to determine if a relevant review were already available or in progress, and it would also simplify the process of searching for systematic reviews. A registry could also be effective in improving the quality of systematic reviews—particularly if it included a mechanism for ensuring that review protocols are always produced. Prospective registration and publication of protocols, as done by the Cochrane Collaboration and the Joanna

Briggs Institute, have been suggested as a way to reduce the possibility of selective outcome reporting bias in systematic reviews (Schünemann et al., 2006) and to address the possibility of nonpublication bias.

Conclusions

In summary, a number of different private and public entities in the United States currently conduct and disseminate evidence syntheses of various types to support clinically effective medical practice, and an expansion of these efforts would be welcome and beneficial. While there is clearly a need for complex evidence syntheses to address the highest-priority topics, the number of these produced is likely to be relatively small (as at present) and to leave many gaps. These gaps can and will continue to be filled by focused systematic reviews and other evidence syntheses produced in the United States and in other countries. Clinicians, policy makers, and the public will benefit from these efforts regardless of the degree of direct U.S. involvement. However, increased participation by the United States in international collaborative efforts such as those discussed in this paper would bring a number of benefits in addition to increasing the number of high-quality evidence syntheses produced. These include additional opportunities for workforce training in the United States, as well as participation in international efforts to develop the tools, methods, and standards for evidence synthesis.

REFERENCES

AHA (American Hospital Association). 2007a. *Continued progress: Hospital use of information technology.* Chicago, IL: AHA.

———. 2007b. *Fast facts on U.S. hospitals.* Chicago, IL: AHA.

AHRQ (Agency for Healthcare Research and Quality). 2008. Effective healthcare program. *Research Reviews*. http://effectivehealthcare.ahrq.gov/healthInfo.cfm?infotype=rr&ProcessID=60 (accessed July 20, 2010).

Buetow, K. 2008. *Presentation to AHIC.* http://archive.healthit.hhs.gov/portal/server.pt/gateway/PTARGS_0_869092_0_0_18/all_materials.pdf (accessed July 20, 2010).

Burgers, J. S., R. Grol, N. S. Klazinga, M. Makela, J. Zaat, and AGREE Collaboration. 2003. Towards evidence-based clinical practice: An international survey of 18 clinical guideline programs. *International Journal of Quality Health Care* 15(1):31-45.

Campbell Collaboration. n.d. http://www.campbellcollaboration.org/about_us/index.php (accessed July 20, 2010).

Catlin, A., C. Cowan, M. Hartman, S. Heffler, and the National Health Expenditure Accounts Team. 2008. National health spending in 2006: A year of change for prescription drugs. *Health Affairs* 27(1):14-29.

CBO (Congressional Budget Office). 2008. *Evidence on the costs and benefits of health information technology*. http://www.cbo.gov/ftpdocs/91xx/doc9168/05-20-HealthIT.pdf (accessed July 20, 2010).

Clarke, M., S. Hopewell, and I. Chalmers. 2007. Reports of clinical trials should begin and end with up-to-date systematic reviews of other relevant evidence: A status report. *Journal of the Royal Society of Medicine* 100(4):187-190.

Connecting for Health. 2008a. *We need a 21st-century privacy approach allowing Americans to protect and share health information to improve quality*. http://www.connectingforhealth.com/resources/20080822_policy_brief.pdf (accessed July 20, 2010).

———. 2008b. *Americans overwhelmingly believe electronic personal health records could improve their health*. http://www.connectingforhealth.org/resources/ResearchBrief-200806.pdf (accessed July 20, 2010).

——— n.d.a *Connectivity in the 21st century*. http://www.connectingforhealth.org/connectivity/ (accessed July 20, 2010).

——— n.d.b *Connecting for health decision-making for population health "first principles."* http://www.connectingforhealth.com/resources/first_principles.pdf (accessed July 20, 2010).

———. n.d.c *Connecting for health common framework for networked personal health information: Statement of endorsement*. http://www.connectingforhealth.org/resources/CCEndorser.pdf (accessed July 20, 2010).

CRD (Centre for Reviews and Dissemination). 2009. *Systematic reviews: CRD's guidance for undertaking reviews in health care*. http://www.york.ac.uk/inst/crd/SysRev/!SSL!/WebHelp/SysRev3.htm (accessed July 20, 2010).

DesRoches, C. M., E. G. Campbell, S. R. Rao, K. Donelan, T. G. Ferris, A. Jha, R. Kaushal, D. E. Levy, S. Rosenbaum, A. E. Shields, and D. Blumenthal. 2008. Electronic health records in ambulatory care—A national survey of ohysicians. *New England Journal of Medicine* 359(1):50-60.

Diagnostic Test Accuracy Working Group. 2009. *Handbook for DTA reviews*. http://srdta.cochrane.org/en/authors.html (accessed June 12, 2009).

Eisenberg, J. M. 2002. Globalize the evidence, localize the decision: Evidence-based medicine and international diversity. *Health Affairs* 21(3):166-168.

GIN (Guidelines International Network). 2010. *Evidence Tables Working Group*. http://www.g-i-n.net/activities/etwg/members-of-the-etwg (accessed July 15, 2010).

Girosi, F., R. Meili, and R. Scoville. 2005. *Extrapolating evidence of health information technology savings and costs*. Santa Monica, CA: RAND Corporation.

Guyatt, G. H., A. D. Oxman, G. E. Vist, R. Kunz, Y. Falck-Ytter, P. Alonso-Coello, and H. J. Schünemann. 2008. GRADE: An emerging consensus on rating quality of evidence and strength of recommendations. *British Medical Journal* 336(7650):924-926.

Higgins, J. P. T., and S. Green, eds. 2008. *Cochrane handbook for systematic reviews of interventions*. Version 5.0.1. http://www.cochrane-handbook.org (accessed June 15, 2009).

HIMSS (Health Information Management Systems Society) Analytics. 2008. *Hospital IT expenses and budgets related to clinical sophistication. Market findings from HIMSS Analytics*. Chicago, IL: HIMSS.

IOM (Institute of Medicine). 2007. *The learning healthcare system: Workshop summary*. Washington, DC: The National Academies Press.

———. 2008. *Knowing what works in health care: A roadmap for the nation*. Washington, DC: The National Academies Press.

———. 2011. *Digital infrastructure for the learning health system: The foundation for continuous improvement in health and health care: Workshop summary*. Washington, DC: The National Academies Press.

Leeflang, M. M., J. J. Deeks, C. Gatsonis, P. M. Bossuyt, and Cochrane Diagnostic Test Accuracy Working Group. 2008. Systematic reviews of diagnostic test accuracy. *Annals of Internal Medicine* 149(12):889-897.

Lin, K., B. Watkins, T. Johnson, J. A. Rodriguez, and M. B. Barton; U.S. Preventive Services Task Force. 2008. Screening for chronic obstructive pulmonary disease using spirometry: Summary of the evidence for the U.S. Preventive Services Task Force. *Annals of Internal Medicine* 148:535-543.

Mallett, S., and M. Clarke. 2003. How many Cochrane reviews are needed to cover existing evidence on the effects of health care interventions? *American College of Physicians Journal Club* 139(1):A11.

MedPAC (Medicare Payment Advisory Committee). 2008. *Medicare payment policy report to the Congress, March 2008*. Washington, DC: MedPAC.

Miller, R. H., K. D'Amato, N. Oliva, C. E. West, and J. W. Adelson. 2009a. California's digital divide: Clinical information systems for the haves and have-nots. *Health Affairs* 28(2).

———. 2009b. *Barriers to financing clinical information systems in California healthcare delivery system organizations: Report to the Governor's Health Information Technology Financing Advisory Commission*. Sacramento: California Council on Science and Technology.

Moher, D., D. J. Cook, S. Eastwood, I. Olkin, D. Rennie, and D. F. Stroup. 2000. Improving the quality of reports of meta-analyses of randomised controlled trials: The QUOROM statement. *British Journal of Surgery* 87(11):1448-1454.

Moher, D., J. Tetzlaff, A. C. Tricco, M. Sampson, and D. G. Altman. 2007. Epidemiology and reporting characteristics of systematic reviews. *PLoS Medicine* 4(3):e78.

Murthy, V. H., H. M. Krumholz, and C. P. Gross. 2004. Participation in cancer clinical trials: Race-, sex-, and age-based disparities. *Journal of the American Medical Association* 291(22):2720-2726.

Nannini, L. J., C. J. Cates, T. J. Lasserson, and P. Poole. 2007. Combined corticosteroid and long-acting beta-agonist in one inhaler versus long-acting beta-agonists for chronic obstructive pulmonary disease. *Cochrane Database System Review* 4:CD006829.

NCI (National Cancer Institute). 2004. *Shared Pathology Informatics Network (SPIN)*. http://spin.nci.nih.gov (accessed August 5, 2010).

———. 2008a. *NCI Cancer Bulletin*. http://www.cancer.gov/ncicancerbulletin/NCI_Cancer_Bulletin_031808/page4/print (accessed December 15, 2008).

———. 2008b. *About caBIG*. https://cabig.nci.nih.gov/overview (accessed August 5, 2010).

Oliver, J., N. Siegfried, G. Kennedy, and T. Horvath. n.d. *Putting Africa first: Supporting novice authors of HIV/AIDS reviews in Africa*. http://www.imbi.uni-freiburg.de/OJS/cca/index.php/cca/article/viewArticle/1334 (accessed June 12, 2009).

Oxman, A., H. Schünemann, and A. Fretheim. 2006. Improving the use of research evidence in guideline development: 7. Deciding what evidence to include. *Health Research Policy and Systems* 4(1):19.

PatientsLikeMe. 2008a. http://www.patientslikeme.com/ (accessed December 15, 2008).

———. 2008b. *The value of openness*. http://blog.patientslikeme.com/ (accessed December 15, 2008).

Poole, P. J., E. Chacko, R. W. Wood-Baker, and C. J. Cates. 2006. Influenza vaccine for patients with chronic obstructive pulmonary disease. *Cochrane Database System Review* 1:CD002733.

Qaseem, A., V. Snow, P. Shekelle, K. F. Sherif, T. J. Wilt, S. Weinberger, and D. K. Owens. 2007. Diagnosis and management of stable chronic obstructive pulmonary disease: A clinical practice guideline from the American College of Physicians. *Annals of Internal Medicine* 147(9):633-638.

Royle, P., and R. Milne. 2003. Literature searching for randomized controlled trials used in Cochrane reviews: Rapid versus exhaustive searches. *International Journal of Technology Assessment in Health Care* 19(4):591-603.

Schünemann, H., A. Fretheim, and A. Oxman. 2006. Improving the use of research evidence in guideline development: 1. Guidelines for guidelines. *Health Research Policy and Systems* 4(1):13.

Schünemann, H. J., M. Woodhead, A. Anzueto, S. Buist, W. Macnee, K. F. Rabe, and J. Heffner. 2009. A vision statement on guideline development for respiratory disease: The example of COPD. *Lancet* 373(9665):774-779.

Sheldon, T., and I. Chalmers. 1994. The UK Cochrane Centre and the NHS Centre for reviews and dissemination: Respective roles within the information systems strategy of the NHS R&D programme, coordination and principles underlying collaboration. *Health Economics* 3(3):201-203.

Stroup, D. F., J. A. Berlin, S. C. Morton, I. Olkin, G. D. Williamson, D. Rennie, D. Moher, B. J. Becker, T. A. Sipe, and S. B. Thacker. 2000. Meta-analysis of observational studies in epidemiology: A proposal for reporting. Meta-analysis of Observational Studies in Epidemiology (MOOSE) group. *Journal of the American Medical Association* 283(15):2008-2012.

Whitlock, E. P., J. S. Lin, R. Chou, P. Shekelle, and K. A. Robinson. 2008. Using existing systematic reviews in complex systematic reviews. *Annals of Internal Medicine* 148(10):776-782.

Woolf, S. H. 2008. The meaning of translational research and why it matters. *Journal of the American Medical Association* 299(2):211-213.

4

The Talent Required

INTRODUCTION

Most of the activities integral to comparative effectiveness research (CER) have been conducted on a small scale over the past several decades; yet, meeting an increased demand for CER and the efficient translation and application of CER findings requires more than simply expanding existing programs and infrastructure. In addition to incorporating the new structures, systems, and elements of health information technologies (HITs) into current practice, innovative new approaches will be needed to drive improvements in both research and practice. Work will be increasingly interdisciplinary—requiring coordination and cooperation across professions and healthcare sectors. One of the key themes of workshop discussion was the need for increased funding and support for training a workforce to meet the unique needs of developing and applying comparative effectiveness information.

Papers in this chapter were presented in draft form at the workshop to begin to characterize the workforce needs for the emerging discipline of CER.[1] William R. Hersh and colleagues explore the heterogeneous set

[1] Comments of workshop reactor panel participants guided the development of the manuscript by Hersh and colleagues presented in this chapter. Sector perspective panelists included Jean Paul Gagnon (sanofi-aventis), Bruce H. Hamory (Geisinger Health System), Steve E. Phurrough (Centers for Medicare & Medicaid Services), and Robert J. Temple (Food and Drug Administration). Panelists commenting on training and education needs included Eric B. Bass (Johns Hopkins University), Timothy S. Carey (University of North Carolina at Chapel Hill), Don E. Detmer (American Medical Informatics Association), David H. Hickam (Eisenberg Center), and Richard N. Shiffman (Yale University).

of activities that contribute to the field of CER and define key workforce components and related training requirements. CER will draw its workforce from a variety of backgrounds—clinical medicine, clinical epidemiology, biomedical informatics, biostatistics, and health policy—and settings, including academic units, university centers, contract research organizations, government, and industry. A key challenge will be developing programs to foster interdisciplinary and cross-sector approaches.

To provide an example of how different workforce elements might be best organized and engaged in a system focused on developing and applying clinical effectiveness information, Sean R. Tunis and colleagues present an overview of a program for health interventions assessment in Ontario, Canada. A direct link between decision makers and CER entities facilitates research timeliness and a clear focus on the information needs of decision makers. Ontario's experience provides insights on how the United States might best expand CER capacity, offers a model for developing an integrated workforce that addresses important organizational and funding issues, and suggests some possible efficiencies to be gained through international cooperation.

COMPARATIVE EFFECTIVENESS WORKFORCE—FRAMEWORK AND ASSESSMENT

William R. Hersh, M.D., Oregon Health and Science University; Timothy S. Carey, M.D., M.P.H., University of North Carolina; Thomas Ricketts, Ph.D., University of North Carolina; Mark Helfand, M.D., M.P.H., Oregon Health and Science University; Nicole Floyd, M.P.H., Oregon Health and Science University; Richard N. Shiffman, M.D., M.C.I.S., Yale University; David H. Hickam, M.D., M.P.H., Oregon Health and Science University[2]

Overview

There have been increasing calls for a better understanding of "what works" in health care (IOM, 2008), driven by a system that allows for learning and improvement based on such an understanding (IOM, 2007).

[2] We thank the following individuals who provided comments, critiques, and additions to early versions of this report: Mark Doescher, M.D., M.P.H., University of Washington; Erin Holve, Ph.D., AcademyHealth; Marian McDonagh, Pharm.D., Oregon Health & Science University; Lloyd Michener, M.D., Duke University; Cynthia Morris, Ph.D., Oregon Health & Science University; LeighAnne Olsen, Ph.D., Institute of Medicine; Robert Reynolds, Sc.D., Pfizer Corp.; Robert Schuff, M.S., Oregon Health & Science University; Carol Simon, Ph.D., The Lewin Group; Brian Strom, M.D., M.P.H., University of Pennsylvania; Jonathan Weiner, Dr.P.H., Johns Hopkins University.

One of the means for assessing what works is CER. The AcademyHealth Methods Council defines CER as "research studies that compare one or more diagnostic or treatment options to evaluate effectiveness, safety, or outcomes" (EHR Adoption Model, 2008). The goals of this report are to define the many components of CER, to explore the necessary training requirements for a CER workforce, and to provide a framework for developing a strategy for future workforce development.

The objective of CER is to provide a sustainable, replicable approach to identifying effective clinical services (IOM, 2008). However, although the term *CER* is widely used, there is no consensus on how best to achieve this objective, and there is little understanding of the challenges required to meet it. There is, for example, wide disagreement about the importance of its different components. The Institute of Medicine (IOM) committee on "knowing what works in health care" emphasizes the central role of *comparative effectiveness reviews* as a critical linkage between evidence-based medicine (EBM) and practice guidelines, coverage decision making, clinical practice, and health policy (IOM, 2008), whereas Tunis views the knowledge of CER as deriving from *practical clinical trials* that compare interventions head to head in real clinical settings (Tunis, 2007). The IOM Roundtable on Value & Science-Driven Health Care expands the notion of CER to include other forms of learning about health care (IOM, 2007), such as the growing amount of data derived from secondary sources, including electronic health record (EHR) systems, which feeds other analyses, such as health services research (HSR). This knowledge in turn drives the development and implementation of clinical practice guidelines, benefits coverage decisions, and allows the general dissemination of knowledge to practitioners, policy makers, and patients. The ideal learning health system will feed back knowledge from these activities to inform continued CER.

While some organizations take an optimistic view of the benefits that CER can bring to improving the quality and cost-effectiveness of health care (Swirsky and Cook, 2008), others sound a more cautionary note. CER will not occur without political and economic ramifications. For example, the Congressional Budget Office notes that CER might lower the cost of health care, but only if it is accompanied by changes in the incentives for providers and patients to use new, more expensive technologies even when they are not proven to be better than less expensive ones (Ellis et al., 2007). A report from the Biotechnology Industry Organization raises concerns that population-based studies may obscure benefits to individual patients or groups and that even in the absence of statistically significant differences among interventions, some individuals may benefit more from some treatments than others (Buckley, 2007). Finally, many argue that CER could turn out to be ineffective unless it is funded and conducted independently of the

federal executive branch by a dedicated new entity (Emanuel et al., 2007; Kirschner et al., 2008; Wilensky, 2006).

In the United States, a clear leader in CER has been the Agency for Healthcare Research and Quality (AHRQ). The AHRQ research portfolio includes evidence-based practice centers (EPCs) (Helfand et al., 2005), which perform comparative effectiveness reviews—that is, syntheses of existing research on the effectiveness, comparative effectiveness, and comparative harms of different healthcare interventions (Slutsky et al., 2010). The work of the EPCs feeds AHRQ's Effective Health Care Program,[3] which also supports original CER through the Developing Evidence to Inform Decisions about Effectiveness network and via dissemination through the John M. Eisenberg Clinical Decisions and Communications Science Center (Eisenberg Center). AHRQ has also made a substantial investment in funding HIT projects to improve the quality and safety of healthcare delivery. The agency also funds health services research as well as pre-and postdoctoral training and career development (K awards) in all of these areas.

Another potential venue for increased CER is the effort by the National Institutes of Health (NIH) to promote clinical and translational research (Zerhouni, 2007). While many think of clinical and translational research as "bench to bedside" (i.e., moving tests and treatments from the lab into the clinical setting), the NIH and others have taken a broader view. With the traditional bench-to-bedside translational research labeled as "T1," other types of translation are defined as well, such as "T2" (assessing the effectiveness of care shown to be efficacious in controlled settings, or bedside to population) and "T3" (delivering care with quality and accountability) (Woolf, 2008). NIH has sponsored many trials that qualify as CER, and although this type of research is not a primary focus for the agency, the training needed to conduct CER overlaps that of T2 and T3 translation. Thus the Clinical and Translational Science Awards (CTSA) initiative greatly expands the clinical research training needed to conduct CER.[4] As CER absorbs researchers and staff, however, it may also compete with other types of research programs in T1 and some T2 areas. Over the past 3 years, the NIH has awarded funding to 38 CTSA centers, with a goal for an eventual steady state of 60 centers. These centers aim to speed the translation of research from the laboratory to clinical implementation and to the community. The work of CER, examining the effectiveness of treatments in real-world settings, including watching

[3] For more information, see http://effectivehealthcare.ahrq.gov (accessed September 8, 2010).

[4] Since this paper was originally authored, the 2009 American Reinvestment and Recovery Act provided $1.1 billion of funds for activities related to CER—including $400 million to the Office of the Secretary of the Department of Health and Human Services, $600 million to AHRQ, and $400 million to the NIH.

for harms to patients with multiple comorbidities, is highly relevant to the CTSA initiative.

One challenge for CER is that it currently exists as a heterogeneous field rather than a specific discipline. While this heterogeneity is probably appropriate to the status of CER as an emerging field of study and effort, it also makes planning for its workforce needs challenging. Investigators and staff in CER come from many backgrounds, including clinical medicine, clinical epidemiology, biomedical informatics, biostatistics, and health policy. They work in a number of settings, including academic units, university centers, contract research organizations, government, and industry. It is not known how well the capacity of the current workforce would absorb any sort of marked increase in demand for CER activities. Finally, there is no specific entity that funds CER, despite calls for there to be so (Wilensky, 2006).

Nonetheless, a variety of stakeholders must have access to the best comparative information about medical tests and treatments (Drummond et al., 2008). Physicians need to be able to assess the benefits and harms of various clinical decisions for their patients, who in turn themselves are becoming increasingly involved in decision making. Likewise, policy makers must weigh the evidence for, and against coverage of, increasingly expensive technologies, especially when marginal costs vastly exceed marginal benefits.

Therefore this report was approached with the assumption that CER should be encouraged as part of the larger learning health system. The authors of this report, leaders with expertise in major known areas of CER, were recruited to define the scope of CER, answer a set of questions concerning the workforce, and work together to develop a framework and a plan for future work. The first task was to achieve consensus among ourselves for defining the components of CER. The next task was to develop a framework for enumerating the workforce and to propose an agenda for defining its required size, skill set, and educational requirements. A draft of this report was presented at the workshop described in this proceedings on July 30–31, 2008. A reactor panel provided some initial feedback, and subsequently more experts were contacted, all of whom are listed in the footnotes on pp. 191 and 192. This led to finalization of the framework and agenda for further research and policy making related to the CER workforce.

Framework for Comparative Effectiveness Research Workforce Characterization

The scope of CER was defined by developing a figure that depicts the subareas of CER and that is organized around the flow of information and knowledge. Next a preliminary model was developed for how workforce

needs might be quantified. The knowledge and challenges in each area were elaborated, followed by a discussion of the issues that will arise with efforts to expand the scope and capacity of CER.

As illustrated in Figure 4-1, information and knowledge originate from clinical trials and other clinical research studies, particularly studies using registries, EHRs, practice network data, and pharmacoepidemiologic studies. This information is synthesized in comparative effectiveness reviews and technology assessments, sometimes including meta-analyses, decision analyses, or economic analyses, which inform the development of evidence-based clinical guidelines and decisions about coverage. HSR evaluates the optimal delivery and the societal health and economic effects of the corresponding changes in the health system. Finally, the information and knowledge are disseminated to both patients and professionals. Each of these components cycles back to its predecessors, and the continuously learning health system maintains a constant interaction among them.

It was also recognized that there are many areas of overlap among the components. For example, experts in biomedical informatics can work synergistically with clinical epidemiologists to determine data requirements and information needs for CER studies. Likewise, clinical guideline developers and implementers can collaborate with health services researchers in technology assessment.

Characterization of Specific Components of the Workforce

The next task was to develop a framework for enumerating the workforce and to make some estimates of its necessary size. Each author was assigned one of the major components of Figure 4-1 and asked to address the workforce needs in that particular area, taking into account the following questions:

1. What are the issues and problems for the workforce at present?
2. What skill set is needed to address current issues and problems?
3. Where are these skills currently developed or obtained?
4. What will be the projected needs as CER scales up in healthcare settings? Do we need more people? Do we need to further develop current capacity? What are the training needs?
5. What are the recommendations for assessing and measuring the needs for the current and future workforce?

Clinical Epidemiology

A core concept underlying CER is that there is a continuum that begins with research evidence, then moves to systematic review of the overall body

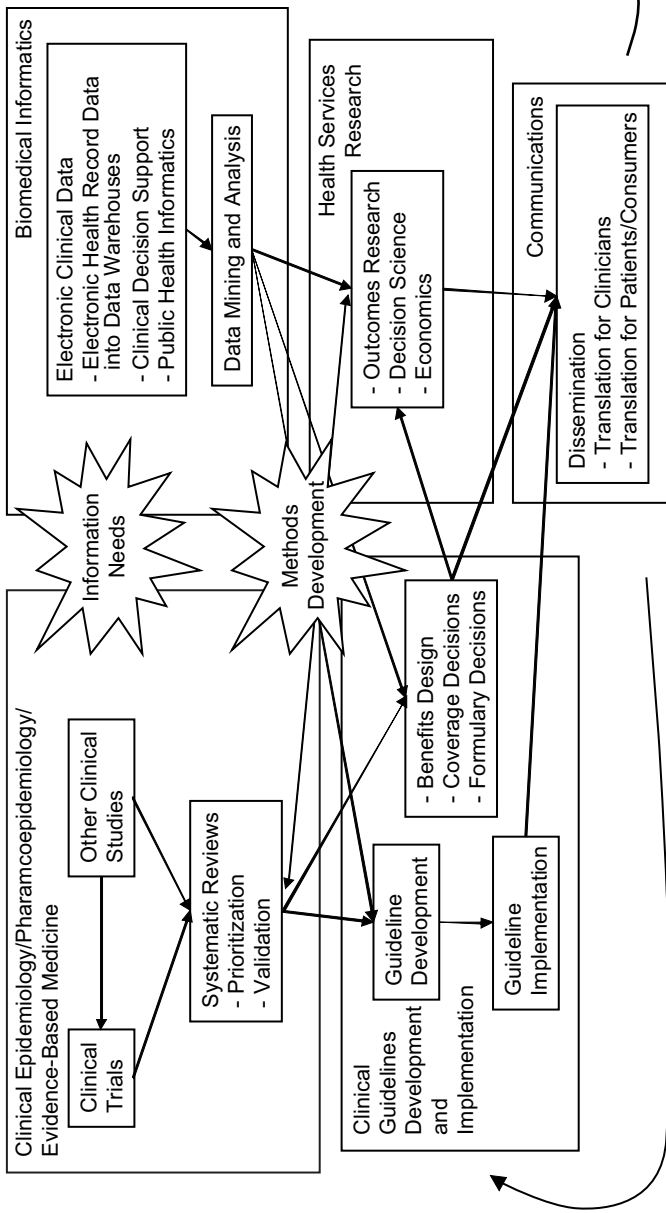


FIGURE 4-1 Key activity domains for comparative effectiveness research. Workforce development will be critical to support the many primary functions within each of these domains as well as to foster the cross-domain interactions and activities identified (e.g., methods development, identifying information needs).

of evidence, and then to the interpretation of the strength of the overall evidence that can be used for developing credible clinical practice guidelines (IOM, 2008). While they overlap with other disciplines, the skills required to conduct CER are not widely taught. This section focuses on the four types of research involved in CER analyses as well as the personnel needed to conduct those analyses: (1) practical clinical trials and conventional clinical research, (2) systematic evidence reviews and technology assessment, (3) pharmacoepidemiologic research, and (4) clinical epidemiology methods research.

Practical Clinical Trials and Conventional Clinical Research

A wide variety of studies are useful in CER (Chou et al., 2010). Most would agree, however, that increasing the amount of CER will require expanding the capability for conducting practical, head-to-head "effectiveness" trials. Such trials are distinct from the so-called efficacy or explanatory clinical trials performed in the regulatory approval process. Explanatory trials, which focus on comparison with placebo treatments in highly selected subjects, are a necessary step in evaluating new therapies, but they are usually not an adequate guide for clinical practice. It can be difficult to determine from such trials—and from the systematic reviews that aggregate them—what the "best" treatments are. In contrast, effectiveness trials, such as practical clinical trials, compare treatments in a head-to-head manner in settings that can be applied to real-world clinical practice. The characteristics that distinguish effectiveness from explanatory (efficacy) studies are listed in Box 4-1 (Gartlehner et al., 2006).

Tunis and colleagues note a number of disincentives to perform-

BOX 4–1
Characteristics Distinguishing Effectiveness from Explanatory Studies

1. Populations in primary care or general population
2. Less stringent eligibility criteria
3. Health outcomes
4. Long study duration; clinically relevant treatment modalities
5. Assessment of adverse events
6. Adequate sample size

SOURCE: Gartlehner et al. (2006).

ing head-to-head comparisons of treatments, such as the disease-oriented nature of the NIH and the commercial motivations of pharmaceutical and other companies (Tunis et al., 2003). Indeed, few such trials have been performed. In a recent survey, Luce and colleagues were able to identify fewer than 20 such trials in the literature (Luce et al., 2008). A frequently stated goal for comparative effectiveness is for the number of effectiveness trials performed each year to grow to 50 trials. As discussed below, accomplishing this goal will require methodological advances in designing and conducting studies as well as training programs devoted to this new type of clinical trial research.

Because so few effectiveness trials have been performed, training in how to design and conduct them is not widely available. While there is overlap, the expertise and the team composition required for practical clinical trials differ from what is required for smaller efficacy trials. For example, practical clinical trials will need to use streamlined, more efficient procedures for recruitment and monitoring than large efficacy trials use (Califf, 2006). They should take advantage, for instance, of Web-based tools for trial management and the potential for using EHR systems to identify, recruit, and allocate subjects to treatment arms within and across health systems (Bastian, 2005; Langston et al., 2005; Reboussin and Espeland, 2005). They also need to develop methods for involving consumers and, for trials conducted in practice networks, office-based clinicians in the design and conduct of trials. Finally, some practical trials require specialized statistical skills (Berry, 2006).

Comparative Effectiveness Reviews and Technology Assessments

Comparative effectiveness reviews are a cornerstone of evidence-based decision making (Helfand, 2005). These reviews follow the explicit principles of systematic reviews, but they are more comprehensive and multidisciplinary, requiring a wider range of expertise. As noted in the EPC *Guide to Conducting Comparative Effectiveness Reviews*, comparative effectiveness reviews "expand the scope of a typical systematic review, which focuses on the effectiveness of a single intervention, by comparing the relative benefits and harms among a range of available treatments or interventions for a given condition. In doing so, [comparative effectiveness reviews] more closely parallel the decisions facing clinicians, patients, and policy makers, who must choose among a variety of alternatives in making diagnostic, treatment, and healthcare delivery decisions" (Methods Reference Guide, 2008). While some technology assessments are similar in scope to a comparative effectiveness review, most are smaller, more focused reviews that require a narrower range of expertise.

Within the emerging, somewhat poorly defined field of CER, conduct-

ing comparative effectiveness reviews and technology assessments is the most developed component. In contrast with other components of CER, guiding principles and explicit guidance for the conduct of comparative effectiveness reviews are available and are widely used. Examples include guidance tools from the UK National Institute for Health and Clinical Excellence (NICE)[5] and the recently released EPC Guide (Methods Reference Guide, 2008).

The underlying disciplines for conducting CER are clinical epidemiology and clinical medicine. Individual comparative effectiveness reviews are usually conducted by project teams led by a project principal investigator under the oversight of a center director. The center director must have exceptional, in-depth disciplinary knowledge and skills in the underlying core disciplines of clinical epidemiology, clinical medicine, and medical decision making. The director should also have applied experience in addition to theoretical knowledge of these areas. For example, it is essential that the director have experience working with guideline panels, coverage committees, health plans, consumer groups, and other bodies that use evidence in decision making. Without such leadership, comparative effectiveness reviews may miss the mark, failing to address the information needs of the target audiences.

It is also important that the director, or other senior investigators, have experienced conducting clinical research studies and not just appraising them. Qualifications for center directors generally include an M.D. degree with additional training leading to a master's degree plus a record of academic productivity representing outstanding contributions in a field such as clinical research design, literature synthesis, statistics, pharmacoepidemiology, or medical decision making. The most important competencies of the project leader are an understanding of clinical research study designs and clinical decision making. Collectively, the project leader and other investigators and staff must have expertise in various areas, such as interviewing experts (including patients) to identify important questions for the review to address, protocol development, project management, literature retrieval and searching, formal methods to assess the quality and applicability of studies, critical appraisal of studies, quantitative synthesis, and medical writing.

This workforce can be characterized based on the experience of the AHRQ EPCs. Through the Effective Health Care Program, the EPCs have completed 15 CERs over a period of approximately 3 years. The average cost of an AHRQ CER is $250,000 to $350,000, depending on its complexity. In these centers, investigators usually have a Ph.D. in epidemiology, pharmacoepidemiology, or biostatistics, or an M.D. with research fellowship training and a master's degree in a pertinent field. Ideally, all

[5] See http://www.nice.org.uk/ (accessed September 8, 2010).

participants should have experience conducting systematic reviews and an understanding of methodological research in the area of systematic reviews, clinical epidemiology, meta-analysis, or cost-effectiveness analysis. Most importantly, they should have the ability to work with healthcare decision makers who need information to make more informed decisions; they should be able to formulate problems carefully, often working with technical experts (including patients and clinicians) to develop an analytic framework and key questions addressing uncertainties that underlie controversy or variation in practice; they should have a broad view of eligible evidence, one that has recognized that the kinds of evidence included in a review depends on the kinds of questions asked and on what kinds of evidence are available to answer them; and they should understand that while systematic reviews do not in themselves dictate decisions, they can play a valuable role in helping decision makers clarify what is known as well as unknown about the issues surrounding important decisions and, in that way, affect both policy and clinical practice (Helfand, 2005).

Also required for systematic reviews are research librarians who have skills in finding evidence for systematic reviews through using electronic bibliographic databases, citation-tracking resources, regulatory agency data repositories, practice guidelines, unpublished scientific research, Web sites and proprietary databases, bibliographic reviews, expert referrals, and publications of meeting proceedings, as well as hand-searching of key journals. Statisticians are needed who have skills in providing advice and critique on the statistical methods used in published and unpublished clinical studies; in conducting statistical analyses, including meta-analysis and other standard analysis and computation; and in preparing statistical reports, including figures and tables. EPCs also require editors who can improve the readability and standardization of evidence reports. In addition, EPCs require research support staff. Research associates must have the ability to critically assess the effectiveness and safety of medical interventions; experience with systematic reviews of the medical literature; knowledge of the fundamentals of epidemiology, study design, and biostatistics; facility in conceptualizing and structuring tasks; and experience with clinical research methods. Research assistants need skills in maintaining bibliographies; coordinating peer review contacts and documents; and assisting in the development of summary reports, figures, tables, and final reports using particular style guidelines. Table 4-1 shows the typical staffing for a CER evidence report funded by AHRQ for a 1-year period.

Although the number of systematic reviews that is necessary may be among the easier of the "how much" questions to ask, there is no clear answer. The Cochrane Collaboration[6] originally estimated a need

[6] See http://www.cochrane.org/ (accessed September 8, 2010).

TABLE 4-1 Required Staffing for a Comparative Effectiveness Research Evidence Report

Role	Activity	Training	Full-Time Equivalent
Center director	Leadership	Clinical epidemiology, clinical medicine, decision making	0.05–0.3
Principal investigator	Leadership	Clinical epidemiology and clinical medicine	0.4
Co-investigator	Domain expertise	Clinical	0.2
Co-investigators	Methods expertise	Clinical + fellowship or master's or Ph.D.	0.4 to 0.6, depending on scope
Research associate	Critical appraisal	M.S./M.P.H./other master's	1.0
Research assistant	Data management	B.S. or M.S.	0.5
Librarian	Literature searching	M.L.S.	0.05
Statistician	Statistical analysis	M.S. or Ph.D.	0.1

for 20,000 reviews; to date, it has completed 3,539 and developed 1,868 protocols for reviews that are proposed or under way. The AHRQ EPCs have produced 168 evidence reports and 16 technical reviews.[7] The Drug Effectiveness Review Project[8] produced 28 original reports and updated 45 reports in its first 3 years.

Of course, systematic reviews are not static documents and, as such, require updating when new evidence becomes available. As increasing numbers of reports are completed, the workforce needs will shift from producing reports to updating them. Shojania et al. have noted that systematic reviews published in the medical literature have a half-life of about 5.5 years, with about 23 percent requiring updating within 2 years of publication (Shojania et al., 2007). Moher et al. surveyed the literature on signals that updates are required and noted that few robust methods exist for detecting them. It is clear, however, that the growing number of systematic reviews being performed will require updating as new evidence from CER and related work becomes available (Moher et al., 2007).

As the existence of the Cochrane Collaboration, an international effort, indicates, efforts going on in other countries may be useful in various areas

[7] See http://www.ahrq.gov/clinic/epcix.htm (accessed September 8, 2010).
[8] See http://www.ohsu.edu/drugeffectiveness/ (accessed September 8, 2010).

of CER—and especially useful in systematic reviews, which are based on scientific literature. For example, NICE produces reports of evidence for health care. Programs in Canada and Sweden provide these as well.

Current training programs are probably adequate to absorb a moderate increase in demand for systematic reviews, but significant expansion of systematic reviews will require more capacity, which will inevitably lead to competition with other clinical research needs. The current training pathways are heterogeneous, and schools of public health and medicine could be much more explicit in developing tracks and certificate programs in systematic review and related areas. These could exist within degree programs in clinical effectiveness, epidemiology, informatics, and so forth. There is a substantial need, however, for biostatisticians and methodologists (who may be in epidemiology or other disciplines) to advance meta-analytic methods in systematic reviews.

Pharmacoepidemiology

An additional area where particular expertise will be needed is pharmacoepidemiology. Recent efforts of the Food and Drug Administration (FDA) to expand drug safety monitoring will require the employment of dozens of pharmacoepidemiologists, creating a competition for their services with the increasing CER activities. The recent FDA Administration Amendments Act of 2007[9] calls for expanding the Prescription Drug Use Fee Act to devote more effort to drug safety, including in areas such as pharmacoepidemiology (Kirschner et al., 2008). It has been estimated that this could require the additional need for 80 to 100 pharmacoepidemiologists (Mullin, 2007). This will create competition for pharmacoepidemiologists who could perform CER work.

An even more challenging problem is that the number of Pharm.D.'s and Ph.D.'s specifically trained in pharmacoepidemiology in North America is small and inadequate to meet growing needs. To meet new additional demand will take time and a several-fold increase in graduates. This will require expanding existing programs and establishing new programs. No one knows how easily individuals trained in other subsets of epidemiology (infectious disease, cardiovascular, environmental, and so on) can be retrained into pharmacoepidemiologists. Device safety and CER will be even more challenging, given their specialized nature and the paucity of high-quality randomized controlled trials (RCTs). Not all pharmacoepidemiology programs examine devices; it will be necessary to expand the field at the same time that training is being expanded.

[9] Food and Drug Administration Amendments Act of 2007. 2007. HR 3580, 110th Cong.

Clinical Epidemiology Methods Research

Clinical epidemiology integrates epidemiologic methods and knowledge of clinical practice and decision making in order to develop clinical research methodology and to appraise clinical research (Fletcher and Fletcher, 2005; Haynes et al., 2005). Its purpose is to develop and apply methods to observe clinical events that will lead to valid conclusions. The availability of senior clinical epidemiologists is limited. This limitation is important because it will affect the capacity to train clinical researchers, conduct practical clinical trials and comparative effectiveness reviews, and develop new methods for clinical research.

To sum up the areas considered so far, analysis of the relevant data in the areas of clinical epidemiology, clinical research, pharmacoepidemiology, and EBM show that the workforce required is likely to be substantial, not available solely based on those who are currently trained, and dependent on the amount of systematic reviews, clinical trials, and other work related to CER that policy makers and others believe must be funded. Furthermore, for all categories of workers, and especially physicians, CER will find itself in a competition both for various types of clinical researchers and also for clinical practitioners, for which there is already a looming shortage (Dall et al., 2006), especially in primary care (Goodman, 2008), which is the area where the need is greatest and from which CER physician researchers are likely to be drawn.

Biomedical Informatics

Another discipline with many contributions to make in CER is biomedical informatics (BMI). This discipline is focused on the acquisition, storage, and use of information in health care and biomedical research, usually assisted by information technology (IT) (Hersh, 2002). The use of BMI for CER is one of a number of "re-uses" or secondary uses of clinical data derived from the EHR and other sources of patient information (Safran et al., 2007). An example of how this has been done is provided in the learning health system workshop summary (Weissberg, 2007). Other potential areas for reuse of EHR data include public health surveillance, health information exchange, clinical and translational research, and personal health records.

The reuse of clinical data currently accounts for a negligible portion of the effort that healthcare delivery organizations devote to clinical IT implementation. Most of the effort goes to deploying systems and is focused on their optimal use for direct clinical care. Furthermore, many individuals who work in the collection or storage of data potentially useful for CER also have other, and sometimes more prominent, roles in the workforce.

However, new training and skills will be required as additional reuse of clinical data is undertaken.

It must be noted that much data in EHR systems are not research-quality data. Clinical documentation is often not a high priority for clinicians. Forms and other types of clinical data capture can be cumbersome and time consuming for busy clinicians to use, and clinicians often do not appreciate the importance of entering high-quality data as part of routine clinical care. BMI workers must be well attuned to the needs of CER and related disciplines if they are to meet the informatics needs of CER.

Of course, implementing EHRs and reusing their data are not the only areas of BMI that are of importance to CER. Biomedical informaticians have skills that are needed in a variety of other areas, including the following:

1. information needs assessment;
2. data mining, text mining, and other forms of knowledge discovery (e.g., tools that help streamline the production of systematic reviews) (Cohen et al., 2006); and
3. ontology development and knowledge management (e.g., projects like The Biomedical Research Integrated Domain Group (BRIDG) Model and other efforts to improve BMI in clinical research) (Fridsma et al., 2008).

Before getting into the details of the informatics workforce for CER, let us take a broader look at that workforce more generally. Most research assessing the HIT workforce has looked only at specific settings or professional groups. In developed countries, there are generally three categories of professionals who make up the HIT workforce:

1. IT professionals—usually with a technical background, such as computer science or management information systems,
2. health information management professionals—the allied health profession historically focused on medical records, and
3. biomedical informatics professionals—working at the intersection of IT and health care, usually with a formal background in one or both.

Probably the most comprehensive assessment of the HIT workforce was carried out in England (Eardley, 2006). This analysis estimated that the HIT workforce employed 25,000 full-time equivalents (FTEs) out of 1.3 million workers in the National Health Service, or about 1 IT staff per 52 non-IT workers. Studies done in the United States have generally focused on one group in the workforce, such as IT or health information management

professionals. Gartner Research assessed IT staff in integrated delivery systems of varying size (Gabler, 2003). Among 85 such organizations studied, there was a consistent finding of about 1 IT staff per 56 non-IT employees, which was similar to the ratio noted above in England.

More recently, Hersh and Wright used the Healthcare Information and Management Systems Society (HIMSS) Analytics Database[10] to analyze hospital IT staff (Hersh and Wright, 2008). This database contains self-reported data from about 5,000 U.S. hospitals, including elements such as number of beds, total staff FTEs, total IT FTEs (as well as broken down by major IT job categories), applications, and the vendors used for those applications. A recent addition to the HIMSS Analytics Database is the Electronic Medical Record (EHR) Adoption Model, which uses eight stages to rate hospitals on how far they have gone toward creating a paperless record environment (EHR Adoption Model, 2007). "Advanced" HIT is generally assumed to be stage 4, which includes computerized physician order entry (CPOE) and other forms of clinical decision support that have been shown to be associated with improvements in the quality and safety of health care (Chaudhry et al., 2006).

Hersh and Wright found the overall IT staffing ratio to be 0.142 IT FTE per hospital bed. Extrapolating to all hospitals beds in the United States, this suggests a total current hospital IT workforce size of 108,390 FTEs. They also found that average IT staffing ratios varied based on the EMR Adoption Model score. Average staffing ratios generally increased with adoption score, but hospitals at stage 4 had a higher average staffing ratio than hospitals at stages 5 or 6. If all hospitals in the United States were operating at the same staffing ratios as stage 6 hospitals (0.196 IT FTE per bed), a total of 149,174 IT FTEs would be needed to provide coverage—an increase of 40,784 FTEs from the current hospital IT workforce.

No studies have quantified the numbers of BMI professionals, although some studies have qualitatively assessed certain types, such as chief medical information officers (Leviss et al., 2006; Shaffer and Lovelock, 2007). The value of BMI professionals is also hinted at in the context of studies showing flawed implementations of HIT leading to adverse clinical outcomes (Han et al., 2005), which may have been preventable with application of known best practices from informatics (Sittig et al., 2006), and other analyses showing that most of the benefits from HIT have been limited to a small number of institutions with highly advanced informatics programs (Chaudhry et al., 2006). Others have documented the importance of "special people" in successful HIT implementations (Ash et al., 2003).

With this general framework, it is possible to discuss the needs of the

[10] This database is derived from the Dorenfest IDHS+ Database, see http://www.himssanalytics.com (accessed September 22, 2010).

informatics workforce for CER. One place to start is with the institutions funded by the CTSA program. Many institutions funded under the CTSA initiative have developed research data warehouses for clinical data that can be used for reuse of clinical data. Workforce needs include those with IT skills in deploying EHR systems, relational databases, and networked applications, as well as those with a more clinically focused orientation who will actually carry out CER activities. The skill set for CER varies depending on the job. Table 4-2 lists the job titles, job responsibilities, and degrees and skills required for various HIT positions. There is unfortunately very little standardization in these jobs. There is also minimal overlap in the skill sets. The jobs can be broadly divided into those that are IT (more technical and less requiring of clinical expertise) and BMI (less technical and more requiring of clinical expertise).

Where will these BMI skills be developed or obtained? Although some technical and clinical skills are obtained through one's formal education, much skill development in BMI currently takes place on the job. In addition, with the rapidly changing nature of IT and BMI, many skills must be learned on the job because some applications did not exist during the individual's education or training. A repeated statement heard from employers of IT and BMI personnel is that "soft skills" are essential. These include the ability to work in groups as well as to communicate effectively orally and in writing. BMI personnel in particular are often viewed as functioning in a "bridge" capacity among IT and clinical personnel.

What are the projected needs as CER is scaled up in healthcare settings? The research by Hersh and Wright cited above indicates that the need for IT personnel increases with the increasing sophistication of EHR adoption, perhaps leveling off after the implementation of CPOE and clinical decision support are reached. The estimates by Hersh and Wright do not include any BMI personnel because the resource they used for their work did not include data on these personnel. In addition, because the data resource did not include data on those who specifically do CER activities, the researchers also do not explicitly include any of these activities. CER informatics work will require both IT and BMI personnel. One common assertion concerning BMI personnel is that there should be one physician and one nurse trained in BMI in each of the 5,000+ hospitals in the United States (Safran and Detmer, 2005). This has led to the 10×10 ("ten by ten") program of the American Medical Informatics Association, which aims to provide a detailed introduction in BMI to 10,000 individuals by the year 2010 (Hersh and Williamson, 2007).

Of course, there are other needs for BMI professionals and researchers as well. Areas described above, such as information needs assessment, data and text mining, and ontology and knowledge management will require even more personnel. Indeed, the BMI field is rapidly evolving, with grow-

TABLE 4-2 Job Titles, Responsibilities, and Training Required for Health Information Technology Professionals

Job Title	Job Responsibilities	Degrees and Skills Required
Information Technology (IT)		
Chief information officer	Oversees all IT operations of organizations	IT, computer science (CS), or management information systems (MIS)
Director, clinical research informatics	Oversees clinical research applications, including comparative effectiveness research (CER)	biomedical informatics (BMI)
Data warehouse manager	Oversees development of research data warehouse	IT, CS, or MIS
Web designer	Designs Web front end for data access systems	IT, CS, or MIS
Web engineer	Deploys Web back end for data access systems	IT, CS, or MIS
Research applications programmer	Develops CER and other applications	IT, CS, or MIS
Database administrator	Administers research data warehouse	IT, CS, or MIS
Project manager	Manages CER and other projects	IT, CS, or MIS
Biomedical Informatics		
Chief medical information officer	Oversees clinical IT applications, including research data warehouse	BMI
Physician leads	Provide leadership in implementation and use of electronic health records	BMI, formally or informally
Medical informatics researcher	Oversees data mining activities for CER	BMI
Medical informatics researcher	Oversees information needs assessment for CER	BMI
Research analyst	Works with medical informatics researchers to collect data and carry out analysis with medical informatics researchers	Variety

ing attention paid to the need for professional development and recognition (Hersh, 2006, 2008).

There is a need for more research to better characterize the optimal IT, health information management, and BMI personnel for general operational EHR and related systems as well as for CER activities specifically. Such research must measure not only health IT practices as they are now, but also what they may become as the implementation agenda advances. Research must sample a wide variety of health organizations to determine quantitatively (e.g., number of people and their skills and education required) as well as qualitatively (e.g., notions of what they need that is not measured by surveys) what they do now, plan to do in the future, and should be doing to achieve an optimal learning health system.

Clinical Guidelines Development and Implementation

Practice guidelines represent an effector arm of the comparative effectiveness process (IOM, 1990). Once scientific studies are performed and their outcomes are systematically reviewed, multidisciplinary guideline development panels are convened to transform the summarized knowledge into recommendations about appropriate care. Those recommendations are then disseminated and presented to many different types of teams for implementation. Because there is some confusion about the use of these terms, *guideline authoring* will be defined as the translation of scientific evidence and expert consensus into policy statements. *Dissemination* refers broadly to the publication and spread of those policy statements. *Guideline implementation* refers to the operationalization of policies in clinical settings with the goals of improving specific processes and outcomes of care and of addressing specific barriers and challenges to uptake.

Ideally, guidelines are produced by multidisciplinary teams that together provide a complete skill set. Unfortunately, these teams are often convened for a single purpose, and skills and knowledge accumulated by team members are not reused in subsequent guideline development efforts.

Guideline development requires topic (domain) expertise that varies from one topic to the next. To create evidence-based guidelines, knowledge must be distilled from the scientific literature and combined with expert judgment. Authors typically work from evidence tables, meta-analyses, and systematic reviews to summarize the facts that are known about a topic. Such evidence summaries may be sought externally or produced by experts within the team itself. Often, however, there are "holes" in the evidence base that must be addressed, either by eliciting expert opinion and experience or by developing an agenda for further research.

Even within a single guideline topic there are often multiple clinical perspectives that should be represented, such as primary care and specialty

care, medical and surgical approaches, and the insights of paraprofessionals. Guideline authoring teams also require two types of methodologists: those who can help topic experts to understand the evidence and its quality, and those who understand the guideline development process, including how evidence quality and benefits and risks should be weighed in order to create statements about the strength of a recommendation. Furthermore, the perspectives of patients who suffer from the condition of interest are often invaluable in formulating recommendations that accommodate the values of that group of people with the greatest stake. The process of formulating policy from scientific evidence requires yet another skill set. Finally, skills in teambuilding, mediation, project management, and leadership are essential to ensure that a well-designed product emerges from the process in reasonable time.

Where are these skills currently developed or obtained? Expertise in clinical care comes most often from clinical training and experience. Expertise in judging evidence may come from coursework in epidemiology and study design. Expertise in policy development usually derives from experience for clinicians, but it may be obtained in formal health policy and public policy studies. Skill in the implementation of guideline recommendations is developed in different ways depending on the specific intervention. Expertise in education, evidence-based decision making, marketing, psychological conditioning, informatics, social and organizational behavior, regulation, financial analysis, and healthcare administration may all be useful in various situations, depending on the implementation strategies selected. Also, as noted above, the experience gained in authoring or implementing a specific guideline is often wasted when not reused.

Because guideline authors tend to focus on policy creation, they pay little attention to how those policies will be implemented. In many situations, authors deliberately introduce vagueness and underspecification because they are unsure how to address such things as gaps in evidence, lack of consensus, and potential legal implications. These limitations in clarity must be identified and resolved before the guideline recommendations can be implemented. The American Academy of Pediatrics is piloting a program called the Partnership for Policy Implementation in which a pediatrician-informatician is made a part of the guideline development team to help ensure that the guideline product can be implemented effectively.

What are the projected needs in this area? Currently, AHRQ's National Guidelines Clearinghouse contains more than 2,000 "evidence-based" guidelines. Based on observations of how soon they become outdated (Shekelle et al., 2001), these guidelines should be reviewed and reaffirmed, revised, or retired every 5 years. Even allowing for no growth, 400 guideline review teams must reassemble each year. If the advice of the IOM is followed and a central agency is developed to help standardize guideline devel-

opment efforts, this could decrease the number of teams creating guidelines and would probably—through "certification"—result in improved reuse of guideline development skills.

The guideline authoring process needs more individuals who are skilled at evidence searching, extraction, and filtering, as well as in policy development. The EPCs may be capable of meeting the current needs of requesting organizations, but additional numbers for performing updates, horizon scans, and filling in holes in the evidence base will be necessary. It will also be valuable to have additional staff within the national professional organizations who can coordinate and lead guideline development initiatives.

The lack of a central organization complicates the estimation of workforce needs. There is little opportunity for guideline developers and implementers to interact—there are no national organizations or national meetings that are attended by both groups. This contributes to the poor communication described above. The Guidelines International Network provides such a venue, but it is composed mostly of Europeans. AHRQ and the National Guidelines Clearinghouse might consider convening such an activity. The quantity of workers required in this area would depend on the number of guidelines necessary, followed by an assessment of what organizations require in order to implement them in their local settings.

Health Services Research

HSR is currently a robust and growing field. It draws from a number of disciplines. Health services researchers regularly participate in CER, and organizations and departments using the label "HSR" successfully compete for grants and contracts in this area. Informal conversations with training program directors indicate that graduates today have multiple job opportunities.

There are many programs claiming to train health services researchers. Whether they evolved from a conscious assessment of what the health system required in terms of research or whether they evolved from other beginnings is not necessarily important; what is important is that the field is established, recognized in formal policy, supported by institutions and professions, and capable of guiding its own destiny. The future of HSR is open to speculation—whether it will evolve as a specialized profession with a coherent and formally bounded sphere of influence or remain more an informally defined "point of view" is an open question. The text in this section is derived from a monograph authored by Ricketts (2007).

As CER gains more attention from scientists, practitioners, and funders, health services researchers will likely adapt their skills and content expertise toward issues of CER. There will, however, be some particular challenges to health services researchers as they become involved in CER studies. CER

requires detailed knowledge of randomized trial design, and many HSR studies use observational studies. CER will also require knowledge of the clinical conditions under study and the practice contexts in which the treatments are applied. Nonclinicians will need either to acquire this knowledge or to develop detailed collaborations with clinicians who have symmetric training in methods.

As a field rather than a specific discipline, HSR has skill sets that cut across multiple domains. In the past, there have not been specifically defined core competencies within HSR. Over the past several years AHRQ partnered with AcademyHealth to define doctoral-level "core competencies" for HSR through a series of white papers and meetings (Forrest et al., 2005, 2009). Table 4-3 shows the current list of competencies.

Some of these competencies extend beyond CER. For example, most CER does not involve primary data collection. There is a substantial overlap between the core competencies in Table 4-3 and the skills needed to conduct CER, including a knowledge of study designs, the ability to develop conceptual models, responsible conduct of research, secondary data methods, study design, implementation of protocols, clear scientific communication, and collaboration with stakeholders.

HSR is recognized in universities and research institutes as a pathway for development of advanced inquiry. Academics are recognized by the vocational cognomen "health services researcher" as often as by the more academic discipline titles of "economist" or "sociologist." There is now an extensive infrastructure in universities, research institutes, and centers for training health services researchers. AcademyHealth, in its inventory of training programs, lists 127 graduate programs in HSR in the United States and Canada. The complexity of the field is reflected in the variety and scope of programs that identify themselves as preparing health service researchers. To complicate matters more, the practicing health services researcher may not actually be formally trained in a program called "HSR."

Some of the programs will likely adapt relatively easily to the need to incorporate skill sets important to CER. As noted elsewhere in this paper, meeting future demands for literature synthesis and meta-analysis will probably not be difficult. Other components of CER, such as pharmacoepidemiology and the assessment of treatment harm through the merging of disparate secondary data, may need to invest in additional doctoral-level training positions in order to meet rising demand. One risk of not meeting increased demand will be the need to use less-well trained individuals to conduct CER.

AHRQ and AcademyHealth have recently completed a general assessment of the training issues and needs for HSR professionals (Ricketts, 2007). Those authors recognized that CER represented one aspect of HSR,

TABLE 4-3 Competencies for Health Services Research

Core Competency	Educational Domains
Breadth of health services research (HSR) theoretical and conceptual knowledge	Health, financing of health care, organization of health care, health policy, access and use, quality of care, health informatics, literature review
In-depth disciplinary knowledge and skills	(Variable depending on the discipline or interdisciplinary area of specialization)
Application of HSR foundational knowledge to health policy problems	Health, financing of health care, organization of health care, health policy, access and use, quality of care, health informatics, literature review
Pose innovative HSR questions	Scientific method and theory, literature review, proposal development
Interventional and observational study designs	Study design, survey research, qualitative research
Primary data collection methods	Health informatics, survey research, qualitative research, data acquisition and quality control
Secondary data acquisition methods	Health informatics, HSR data sources, data acquisition and quality control
Conceptual models and operational measures	Scientific method and theory, measurement and variables
Implementation of research protocols	Health informatics, survey research, qualitative research, data acquisition and quality control
Responsible conduct of research	Research ethics
Multidisciplinary teamwork	Teamwork
Data analysis	Advanced HSR analytic methods, economic evaluation and decision sciences
Scientific communication	Proposal development, dissemination
Stakeholder collaboration and knowledge translation	Health policy, dissemination

SOURCE: Data derived from Forrest et al., 2005, 2009.

but that the field is currently very labile. Three approaches to workforce planning for CER and HSR can be recommended:

1. Researchers conducting CER as well as funders and policy makers planning new initiatives should regularly communicate with educators so that new needs for training can be incorporated in a timely fashion. While training programs should not alter curriculums with each new federal request for applications, they should be responsive to changes in the research and policy environment.
2. AHRQ and AcademyHealth should continue to conduct periodic surveys (every 2 to 3 years) and key informant interviews to assess

the state of the workforce for health services and CER, communicating with researchers in industry, contract research, and academe regarding the quality and availability of personnel at multiple levels of training.

3. AHRQ and AcademyHealth should also regularly examine the number and type of training programs in HSR and CER and communicate with funders regarding adequacy of supply and congruence of curriculums with the expressed need of the research organizations conducting CER. Modifications to predoctoral, postdoctoral, and career development (K series) programs can be based on these evaluations.

Dissemination

The purpose of evidence translation and dissemination is to develop practical tools to improve decision making by end users. The group of end users is broadly defined and includes people who have medical problems (patients), their families and caregivers, clinicians, healthcare administrators, governmental policy makers, and employers. Evidence translation is the process of extracting key messages from evidence summaries (systematic reviews or technology assessments) and placing those messages into the context of the decisions made by end users. To be useful, translation needs to lead to the creation of products (such as summaries tailored for particular audiences) that can be accessed by those end users. Dissemination is then the step of making those products accessible for the end users. Dissemination can occur through various avenues, including the distribution of printed products, making the products available on Web sites, and other modes of electronic distribution (such as interactive decision aids, podcasts, and e-mail).

To help individuals use clinical evidence in their decisions, the summaries of that evidence must be unbiased. The evidence sources are often complex scientific documents that provide detailed and highly nuanced explanations of the body of evidence. The process of evidence translation requires careful analysis and summarization to avoid errors or oversimplification. Evidence summaries are only useful if they are applicable to the actual decisions made by the people involved in health care. This activity requires a thorough knowledge of the methodologies used in clinical research and systematic reviews. It also requires a clear understanding of the clinical context in which the evidence will be applied. Thus, the technical skills required to perform translation are similar to those required for the development of systematic reviews. Individuals must be able to understand the methodologies in order to translate the reports without introducing bias. The process of developing key messages often involves

simplification, and this requires a careful deconstructing of the information in a systematic review.

After the process of evidence translation has determined the messages that will be useful for decision makers, the next step is to summarize this information in products that can serve as tools to aid decisions. Because such tools can take various forms, the skills are multidisciplinary and include the ability to provide clearly written documents and to design interactive content for electronic distribution. Developing effective decision tools can also benefit from the input of stakeholders and opinion leaders. This input helps to ensure that the evidence translation will meet the critical needs of decision makers. After the decision tools are developed in draft form, testing them with end users provides valuable insight into how they can be modified and improved. Both of these steps (obtaining formative input and performing testing with end users) require qualitative research skills. Finally, after the decision tools are developed and tested with end users, they are ready for public release and dissemination. Dissemination is a specialized activity that requires skills related to public relations, journalism, and communications.

The multidisciplinary team required to perform evidence translation and dissemination includes a variety of individuals who have different skills and who commonly have diverse educational backgrounds. While some individuals may play more than one role (e.g., a clinician with skills in clinical research methodologies), the required skills are so diverse that a multidisciplinary team is needed. Table 4-4 summarizes the individuals who compose the team.

What are the projected needs in this area as CER is scaled up? At present there are relatively few groups doing state-of-the-art evidence translation and dissemination in the United States. As programs to increase CER move forward, there will be a growing need for individuals to perform translational work. Some members of the multidisciplinary team (particularly clinicians and methodologists) will be drawn from the same pool as those performing the work of information synthesis. Thus, the need for translation will increase the need for an infrastructure to train such individuals. For the other team members, the necessary training will come from programs that provide training in informatics, qualitative research, and communication.

The need for additional workers in dissemination-related areas of CER is based on the amount of activity deemed required to most effectively distribute such knowledge. This in turn is a function of the output from the other activities shown in Figure 4-1 that feed into dissemination. Another factor in quantifying the amount of dissemination required is the different types of healthcare professionals (e.g., physician specialists, physician

TABLE 4-4 Roles, Skill Sets, and Backgrounds of Personnel Involved in Comparative Effectiveness Research Dissemination

Role	Skill Set	Contribution	Educational Background
Clinician	Understanding of clinical issues	Defining key decisions to which the evidence will be applied	Medicine, nursing, pharmacy
Research methodologist	Understanding sources of bias	Defining key messages derived from systematic reviews or technology assessments	Clinical epidemiology
Writer	Synthesizing contextual information and key messages	Creating plain language explanations	Health communication
Qualitative researcher	Qualitative data collection and synthesis	Interviewing key informants and end users	Qualitative methods
Computer programmer	Creating Web-based and other interactive content	Developing electronic tools to aid decision making	Biomedical informatics and/or computer science
Dissemination specialist	Understanding audiences and avenues for dissemination	Developing effective dissemination strategies	Communication, public relations, journalism

generalists, nonphysicians) and patients (e.g., those with varying levels of health and general literacy).

Overlapping Areas

Figure 4-1 shows two areas of explicit overlap, and there are likely to be more. Each of these areas is likely to generate additional needs, such as people who work, perform research, and teach at these margins. The first of these areas of overlap is information needs assessment. Trialists and systematic reviewers must, for example, work with domain experts to determine the key aspects of their research questions to be studied. Likewise,

guideline implementers and developers must be driven by an information needs assessment process.

The second area of overlap is methodology. As with most research in general, CER is driven by a diverse set of methods applied in many medical domains. Furthermore, the intersection of these areas may create the need for new methodology, such as the best methods for approaching, if possible, research-quality data in an EHR. It will take additional workforce to conceptualize and develop this methodology, followed by practitioners to implement it and professors to teach it. Technology assessment is one specific area of methodology that will require all of this.

Summarizing and Quantifying

The above analysis of workforce components has identified the broad range of activities that make up CER. These will come not only from research in traditional clinical research and related areas but also through analysis of the growing amount of data in EHR systems, experiences with clinical guidelines development and implementation, aspects of HSR, and more widespread dissemination of knowledge. Current research and other work in these areas remains productive, but a substantial scaling up of funding will require better policy coordination, more funding, and—what is assessed in this paper—understanding and planning for workforce needs. There are several challenges to achieving the vision and goals for CER related to workforce needs.

The first challenge in defining the CER workforce is to grapple with the larger question of the quantity of CER that is necessary for the learning health system. To quantify the needs, these questions must be answered:

1. What quantity of comparative clinical trials and other clinical research will be required?
2. What quantity of CER systematic reviews will need to be performed?
3. What amount of pharmacoepidemiological and related analysis will be required, or even possible, especially given the small number of pharmacoepidemiologists?
4. How many medical centers will be willing or able to use their EHR systems or local guideline implementation to provide data for CER?
5. What quantity of clinical practice guidelines will need to be produced?
6. What types and amounts of HSR will be necessary for CER?
7. What types and quantities of dissemination will be required for CER? At how many levels will the content require reformatting?

These various aspects of CER do not exist in isolation. In this analysis,

a number of areas have been identified that require an interaction of activities and skills across various areas of CER. Therefore it will not be enough simply to plan in discrete areas such as clinical epidemiology, BMI, or HSR. Furthermore, this analysis focused only on the actual work to be done and not on the leadership required to guide CER work. As in all fields, leadership will be necessary to develop and advocate for the vision of CER and the learning health system, to manage its deployment and training, and to interact with the leadership of related and separate disciplines.

Recognizing that there are different areas of CER and diverse skill needs within them, a total summation of the workforce will quantify the need in each area and then sum those needs across all the areas. How much CER, and how much of each component of CER, must be quantified by policy makers and others who must take into account the demand for each type of CER, the supply of the workforce, competition for other tasks these workers might do, and funding available for CER.

The original intention in this report was to provide a quantitative first approximation of the workforce needs for CER. However, as the authors developed the framework and explored the issues more deeply, it was apparent that there are too many unanswered questions about the scope, breadth, and quantity of CER that need to be clarified in order to achieve the larger goals for a learning health system. This view was validated by many of the experts listed in the acknowledgements, who advised against attempting to quantify needs as long as there is such an unclear picture of, and future for, CER.

There are a number of reasons why a quantitative assessment of the CER workforce is not possible. The main one is that the true scope of CER is not known. For example, even in the area of clinical epidemiology, which has probably the most clarity about needs of any of the areas that were assessed, there is no clear answer about how many systematic reviews, practical and other clinical trials, and pharmacoepidemiological analyses are required. While the number of personnel required for systematic reviews is relatively well understood, the requirements for the other categories of clinical trials and pharmacoepidemiological analyses are much less clear.

Beyond clinical epidemiology, the picture becomes even less certain. While BMI, development and implementation of clinical guidelines, and dissemination could become a major part of CER, the amount of each that will need to be done—or that even falls under the rubric of CER—is not clear. Furthermore, in all of these areas, CER would be secondary to the larger tasks of maintaining IT systems for clinical care; using guidelines to improve the quality, safety, consistency, and cost effectiveness of operational clinical care; and disseminating all types of clinical knowledge. How much of the work would actually be CER is not known or easy to determine. Even in HSR, the amount of research to be done that could be

classified as CER is not certain. Clearly, in an analysis held under the rubric of "EBM," there is little evidence to make sound judgments about specific workforce needs.

It is certain, however, that CER will require a diverse array of skilled workers to meet its agenda. Any effort to undertake CER in a major way, such as through the establishment of a centralized public agency or private institute, should have the quantification, efficient deployment, and required education of the workforce as an early research agenda item. Determining how to implement and scale up CER will be a major challenge if its size and scope is to be seriously increased.

This leads to a number of larger policy questions. For example, how will CER be financed? In the case of research studies, who will fund the work comprehensively, especially in light of a national research enterprise that focuses on disease-based and investigator-initiated research? Likewise, who will fund the development of clinical practice guidelines? For studies derived from the secondary use of clinical data, which medical centers and health systems will be required to participate, and how will they be funded? Will it become an expected part of healthcare delivery? Finally, how will the knowledge generated from CER be disseminated? What amount of dissemination will be required, and how will it be funded?

It is also worth noting that work in CER will face competition from other areas for researchers and their staffs, especially among physician researchers. As the baby boomer generation enters the Medicare age group, there is a growing need for physicians and other clinical practitioners. Likewise, there are also demands for physicians to enter non-CER research areas, such as those encouraged by the CTSA program, which could be a help or a hindrance to CER research. This is true for other areas of CER as well, such as the need for pharmacoepidemiologists because of the growing amount of drug monitoring and safety called for in recent FDA-related legislation. As such, any policy or funding that increases CER will need to recognize the competition for workers and skills from related areas, which will drive up the salaries of researchers and their staffs.

There is also competition for workers even within CER work. This analysis has mostly focused on academic settings, but there are others who have an interest in performing CER and related work. This includes government agencies, nonacademic healthcare systems, and manufacturers of drugs, devices, and other medical tests and treatments. One possible silver lining in this competition is the potential to partner with international organizations engaged in similar work, such as NICE. While not all CER work transfers easily across borders, populations, and cultures, there is likely some amount that can.

Research and policy development will also need to be provided to the locations where leadership in CER will be required to pilot the leading edge

of research, especially in methodologies, and to train the next generation. The best sites to establish these centers of excellence will probably be those that house EPCs, CTSA centers, AHRQ HSR training, and informatics research and educational programs. As in all of these specific initiatives, national consortiums should be established to share the vision, best practices, and policy for developing the pipeline of new researchers.

The determination of the scope and amount of workforce required for CER is a research agenda itself. Particularly for the workforce required, and for the components within it (such as clinical epidemiology, BMI, and HSR), research should be undertaken to identify not only the skills required now, but also how the workforce will be best organized in the future for maximum efficiency, best-quality output, and the anticipated expansion. This should be done through a variety of methods, including estimates of quantitative needs (e.g., amount of research, number of researchers and their staff, existing capacity, how much expansion is required) as well as qualitative understanding (e.g., people and organizational challenges, academic homes, career advancement).

Once this research agenda identifies the workforce and the skills it needs, it will also be necessary to determine the types of educational programs required to train those individuals, such as the competencies and curriculums of such programs. This will require policy on how to fund such education, especially for those with increasing burdens of educational debt already acquired for their basic education. This may be another area where international partnerships may be helpful.

CER promises an exciting approach to improving the quality of health care while reducing its cost though more efficient use of the most effective approaches to clinical care. A major part of achieving this system will be a coordinated and adequately funded approach. There are many challenges to reaching that goal, including the provision of a workforce that can bring the requisite knowledge and skills to CER problems and solutions. Much further research, policy, and funding will be required to achieve this vision.

TOWARD AN INTEGRATED ENTERPRISE—THE ONTARIO, CANADA, CASE

Danielle Whicher, Kalipso Chalkidou, Irfan Dhalla, Leslie Levin, and Sean R. Tunis
Center for Medical Technology Policy, Baltimore, Maryland, USA
National Institute for Health and Clinical Excellence, London, UK
Department of Medicine, University of Toronto, Ontario, CA
Medical Advisory Secretariat (MAS), Ontario Ministry of Health and Long Term Care, Toronto, Ontario, CA

Overview

In an effort to ensure that important emerging health technologies are not used indiscriminately but are available to patients for whom the risk–benefit ratio is favorable, the Ontario provincial government has recently expanded its capacity to conduct comparative effectiveness studies. The new system allows purchasers (primarily hospitals) to request that a health technology be reviewed by the Ontario Health Technology Assessment Committee (OHTAC), an arms-length advisory committee to the Ministry of Health and Long-Term Care (MOHLTC). If, after completion of this assessment, which generally includes a systematic review produced by a governmental agency, OHTAC decides there is insufficient information to recommend a coverage decision, it may request a "conditionally funded field evaluation." These studies, led by government-funded independent research entities, are designed to produce the evidence necessary for policy makers to make coverage decisions. Funding this research requires approximately $8 million to $10 million in incremental spending per year and the support of Ministry of Health staff, as well as of hospital and university investigators with a wide variety of expertise, including epidemiologists, biostatisticians, physicians, health economists, health policy experts, and health services researchers.

The direct and explicit link between the decision makers and the CER entities facilitates research timeliness and helps ensure a clear focus on generating information that is carefully designed to satisfy the needs of decision makers. Because purchasers contact OHTAC prior to investing in medical technologies, this system encourages evidence-based technology diffusion.

Although the healthcare system in the United States differs greatly from Ontario's in size, complexity, and design, Ontario's experiences provide insights in terms of workforce issues, organization, and funding that are relevant to U.S. efforts to build comparative effectiveness capacity. In addition, U.S. efforts may benefit from various collaborative activities with Ontario and other international evidence-generating entities, such as clini-

cal trials, patient registries, and standards of study design, which may help to globalize CER in the future.

Background

Healthcare delivery in Canada is primarily a provincial responsibility, but the federal government does provide substantial funding contingent on the provinces adhering to the conditions set forth in the Canada Health Act of 1984. According to the act, provinces must provide all "medically necessary" hospital and physician care free of charge (Iglehart, 2000; Lewis et al., 2001). Hospital services include diagnostic tests, inpatient medicines, and medical devices and equipment, as well as inpatient medical and nursing care. Outpatient physician care is also covered. In contrast to inpatient medication use, the public coverage of outpatient prescription medications varies by province, and most provinces provide universal coverage only for elderly and poor people. However, most non-elderly Canadians have private drug insurance through their employer. Provinces also differ in their coverage of home care, nursing homes, and other community-based care (Iglehart, 2000; Lewis et al., 2001).

Although the proportion of health care paid for by the federal government decreased through the 1990s, increased fiscal capacity has allowed the federal government to increase its provincial transfers over the last decade. Nevertheless, provincial governments still pay for the majority of healthcare costs, and the recent surge in the development of promising but often unproven medical technologies has placed added pressure on both public payers (i.e., provincial ministries of health) and providers (i.e., hospitals) throughout Canada (Levin et al., 2007). Hospital chief executive officers throughout Ontario have expressed frustration with the public pressure to adopt new technologies in the absence of objective information regarding their benefits and risks in comparison to those currently available. The Canada Health Act obliges hospitals to provide interventions that are "medically necessary," but the evidence base necessary to make this determination is often incomplete, not unlike the situation in the United States and most other countries working to develop evidence-based policy decisions (Levin et al., 2007). Furthermore, because hospitals in Canada negotiate their budgets with provincial ministries of health or regional health authorities, they are limited in their ability to raise funds to pay for new technologies (Iglehart, 2000).

Ontario, like other Canadian provinces and the United States, faces the challenge of improving healthcare quality while simultaneously limiting increases in healthcare spending. In an effort to manage the diffusion of novel health technologies and to ensure that they are received by patients in whom the risk–benefit ratio is favorable (consistent with the manner in

which the term *medically necessary* is interpreted in the Canadian context), the MOHLTC in Ontario has developed a unique network to carry out CER on nondrug healthcare interventions (Goeree and Levin, 2006; Levin et al., 2007). There is no similar network for the assessment of emerging drug technologies. For a brief explanation of comparative effectiveness entities for medical drug technologies, see Box 4-2.

This paper begins with a description of the various entities involved in CER in Ontario. Following a brief overview of the workforce and funding requirements, it discusses potential lessons for policy makers in the United States.

Developing a Comparative Effectiveness Capacity for Nondrug Medical Technologies

Establishing an Agenda and Making Policy Recommendations

As has happened in most jurisdictions, nondrug healthcare interventions have historically rapidly diffused into the Ontario healthcare system, even in the absence of definitive clinical evidence of benefit. One reason for this is that technologies can enter the healthcare system through a variety of "portals," including hospitals and other healthcare providers, community programs, and nursing homes (Goeree and Levin, 2006; Levin et al., 2007). To manage the diffusion of new health technologies and improve care, the Ontario MOHLTC established the Medical Advisory Secretariat (MAS) in 2001 and the OHTAC in 2003. These two entities work in concert to determine which technologies should be used in Ontario, which should not, and which require further research (Figure 4-2). In making this decision, the Ministry of Health considers many factors, including not only clinical effectiveness and safety but also cost effectiveness and budget impact.

Although any interested party can ask OHTAC to assess a new health technology, most requests are made by hospitals or the Ministry of Health. These requests are initially processed by the Medical Advisory Secretariat (MAS), a unit within the ministry that is staffed by an information specialist, 2 policy analysts, 10 clinical epidemiologists, and 3 administrative staff.[11] During this initial processing, MAS completes a template, which includes information on the potential clinical effect size, public or professional pressure to use a new technology, and a preliminary comparison with alternative healthcare interventions (Goeree and Levin, 2006; Levin et al., 2007).

The results of the initial analysis are presented to OHTAC, which is composed of at least 12 individuals (currently, there are 25 members on the

[11] Personal communication, L. Levin, June 13, 2008.

BOX 4-2
Comparative Effectiveness Entities for Drug Evaluation

At the national level, new drugs undergo a systematic review (via what is called the Common Drug Review process) by the Canadian Agency for Drugs and Technologies in Health. These systematic reviews are completed in 4 to 6 weeks and focus on publicly available clinical trials data as well as pharmacoeconomic evaluations submitted by pharmaceutical manufacturers who wish to have their drug listed on formularies. The evidence is then assessed by the Canadian Expert Drug Advisory Committee, an advisory committee of 12 members from a variety of fields, including clinical trial methodologists, experts in health technology assessment, drug policy, or health economics, and two public representatives, plus a chair (Tierney and Manns, 2008). Based on this committee's assessment, the agency either recommends that a drug be covered and added to public formularies or not. The committee does not have the authority to request additional research if the current literature fails to address policy makers' concerns.

In Ontario, the Committee to Evaluate Drugs, composed of 16 members, including 2 lay people and 14 physicians and pharmacists (Committee to Evaluate Drugs, 2007b), uses the recommendations from the Canadian Agency for Drugs and Technologies in Health as a basis for its own recommendation for the province's publicly funded drug programs. A final coverage decision is made by the Executive Officer of the Ontario Public Drugs Program in the Ministry of Health. These decisions are not binding on private drug insurance plans, which provide employment-based pharmaceutical coverage for most working Ontarians, though the decisions are considered by private decision makers.

In April 2007 the Ministry of Health launched the Drug Innovation Fund, which provides CA$5 million annually to pay for independent research projects relating to health outcomes research, providing information to support decision making, and supporting the development of an independent research capacity in academic institutions throughout Ontario to provide information for decision makers on the impact of drug access and use, optimal use of drugs, and drug adherence (Committee to Evaluate Drugs, 2007b).

(For more information on specific studies paid for by the fund, visit http://www.health.gov.on.ca/english/providers/program/drugs/drug_ innov_fund/pdf/funding_successful_proposals.pdf [accessed September 22, 2010]).

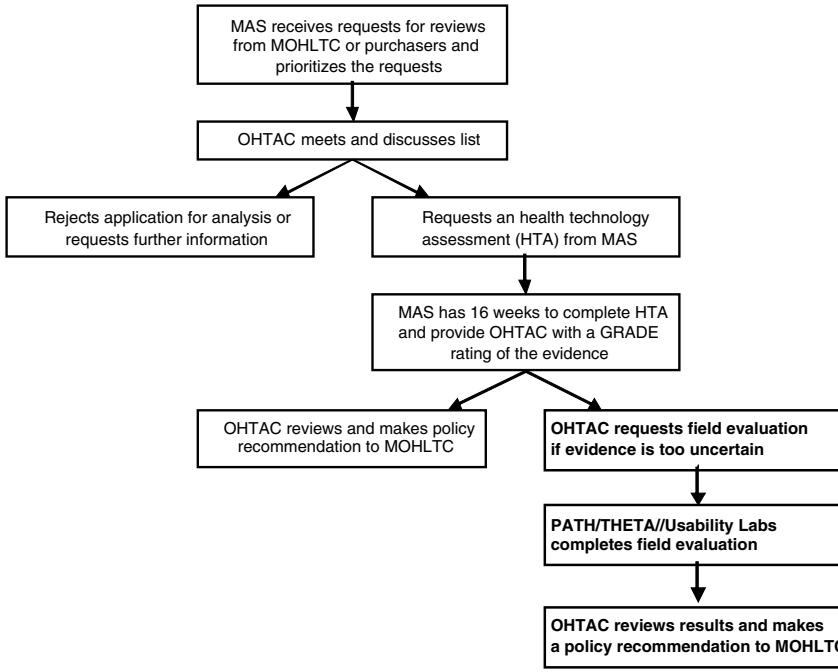


FIGURE 4-2 Schematic of explicit link between decision makers and comparative effectiveness entities.
NOTE: GRADE = Grading Recommendations Assesment, Development, and Evaluation Working Group; MAS = Medical Advisory Secretariat; MOHLTC = Ministry of Health and Long Term Care; OHTAC = Ontario Health Technology Assessment Committee; PATH = Program for the Assessment of Technologies in Health; THETA = Toronto Health Economic and Technology Assessment Collaboration.
SOURCE: Whicher, D. M., K Chalkidou, I. Dhalla, L. Levin, and S. Tunis. 2009. Comparative effectiveness research in Ontario, Canada: Producing relevant and timely information for health care decision making. *The Milbank Quarterly* 87(3): Figure 1, page 589. Reprinted with permissions from John Wiley and Sons.

committee). The committee includes representation from the Ontario Medical Association and the Ontario Hospital Association as well as from the community and long-term care sectors. Individual members have expertise in nursing, medicine, health economics, epidemiology, ethics, and technology assessment (Goeree and Levin, 2006; Ontario Health Technology Advisory Committee, n.d.). The members meet monthly to provide feedback to MAS and to provide policy recommendations to the deputy minister of health (Goeree and Levin, 2006). Based on the initial analysis produced by MAS,

OHTAC can choose to request a systematic review, commonly referred to as a health technology assessment (HTA), reject the application, or request further information before making a final decision (Figure 4-2).

If OHTAC requests an HTA, then MAS, in collaboration with academic partners at University of Toronto and McMaster University, produces a health technology assessment within 16 weeks of the initial request. The review includes evidence relating to the technology's safety, clinical effectiveness, and efficacy (produced by MAS) as well as its cost effectiveness (produced through academic collaborations); and the quality of evidence is assessed by a rating consistent with Grading of Recommendations Assessment, Development and Evaluation Working Group (GRADE) guidelines.[12] Evidence can be given a GRADE quality rating of high (future research is unlikely to have an impact on the estimate of the effect), moderate, low, or very low (any estimate of the effect is very uncertain). The GRADE framework depends on a number of factors, including the type of evidence available (e.g., randomized controlled trials vs. observational studies), the quantity of evidence (e.g., the number of studies), and the consistency of the evidence, as well as an assessment of any potential biases (Atkins et al., 2004). Using the GRADE framework provides consistency and transparency in MAS recommendations.

Based on the HTA and estimates for cost effectiveness, OHTAC can either make a recommendation to the MOHLTC or, if the evidence of clinical effectiveness of the technology is lacking, OHTAC can request that a "conditionally funded field evaluation" be performed (Figure 4-2). Following the formation of any draft recommendation, there is a period of public comment when relevant stakeholder groups are targeted and the draft document is posted on the OHTAC website.

OHTAC has determined that a multifaceted approach to public engagement is preferable to programs that conduct field evaluations to generate an evidence base for important high-demand health technologies.

There are several programs in Ontario, all of which are independent and largely government funded, with the capacity to carry out field evaluations for promising technologies. Since the requests for field evaluations are coming directly from decision makers, studies can be designed specifically to address the evidence that is critical to making coverage or purchasing decisions. Field evaluations can take many forms, including (Levin et al., 2007)

- RCTs,
- observational studies (e.g., registries or cohort studies, with or without contemporaneous controls),

[12] See http://www.ohsu.edu/drugeffectiveness/ (accessed September 8, 2010).

- time-series studies,
- chart reviews, and
- multisite safety assessments.

OHTAC may also use information from polling studies or from the development of "Ontariorised" microeconomic policy models to inform policy decisions.

Appropriate study design is discussed by these independent entities in collaboration with experts knowledgeable about the health technology under investigation. While field evaluations are being conducted, the Ontario Health Insurance Plan covers physician costs for the medical technologies used for patients participating in the study as long as a technology is already insured (Goeree and Levin, 2006; Levin et al., 2007). If there is no fee code for the technology, alternative funding arrangements may be made. This arrangement gives patients access to emerging technology—and allows manufacturers to generate some revenue—before a long-term policy decision has been reached. Field evaluations are apportioned to one of the first two agencies described below based on existing capacity, and funding to carry out these evaluations is provided by the Ministry of Health. Each field evaluation also includes an assessment of the technology's cost effectiveness.

Program for the Assessment of Technology in Health

The Program for the Assessment of Technology in Health (PATH), located at McMaster University and St. Joseph's Healthcare Centre, is the longest-standing entity for government-funded field evaluations. The program has 20 staff, including 4 graduate students, research associates, a biostatistician, and university faculty (PATH Research Institute, 2008a). Four or five staff work on a given field evaluation with the help of various project consultants. In addition, PATH has been actively involved in developing master's and doctoral-level degree programs at McMaster University in the field of HTA (PATH Research Institute, 2008b).

Over the past couple years, PATH has completed several studies, the results of which have had a significant impact on decision making. One such study is described in Box 4-3.

Toronto Health Economics and Technology Assessment Collaboration

The Toronto Health Economics and Technology Assessment Collaboration (THETA) was established in July 2007 at the University of Toronto (Toronto Health Economics and Technology Assessment Collaborative, 2007a). The group now has 28 investigators from a variety of backgrounds,

BOX 4-3
Comparative Effectiveness Study Comparing Drug-Eluting Stents to Bare Metal Stents for Treatment of Coronary Artery Disease

In 2002, the Medical Advisory Secretariat completed a gray literature-based health technology assessment on the clinical effectiveness of drug-eluting stents (DESs) compared to bare-metal stents (BMSs) and concluded that randomized controlled trial (RCT) evidence was imminent and would likely show that a DES was more effective than a BMS, following which there would be a steep diffusion curve for DESs. However, when the initial RCT results on DESs were published later that year, they demonstrated that there was uncertainty regarding generalizability. Due to this, the Ontario Health Technology Assessment Committee (OHTAC) recommended that the Ministry of Health commission a field evaluation from the Program for the Assessment of Technology in Health (PATH). The study proposed by PATH was a prospective observational study, which took advantage of both an existing provincewide registry set up by the Cardiac Care Network (CCN) of Ontario and the ability to link this registry to administrative databases housed at the Institute for Clinical Evaluative Sciences. Additional fields were added to a preexisting CCN database to facilitate a study comparing different stent designs. The objective of the study was to estimate the reduction in risk of revascularization within 2 years of treatment with a DES compared to a BMS. The study was also intended to estimate the cost effectiveness of DESs compared to BMSs. During the course of the study, hospitals were able to provide DESs free of charge to patients enrolled in the study (Bowen et al., 2007; Tu et al., 2007).

Interestingly, the study in part provided evidence for the use of DESs in some "off-label" indications and suggested that DESs may be no more effective than BMSs for many "on-label" indications. The results demonstrated an incremental benefit for DESs only in high-risk patients, defined as those patients who have two of three risk factors for restenosis (diabetes, small vessels, or long lesions) (Bowen et al., 2007; Tu et al., 2007). Based on these results, OHTAC recommended that DESs be used only in patients at high risk for restenosis. Data continue to be collected on patients who receive DESs, and initial estimates suggest this controlled diffusion of DESs led to a cost savings of about $20 million dollars in 2007 and 2008 (Bowen et al., 2007). This has resulted in a conversion rate from BMSs to DESs in Ontario currently estimated at 25 percent, compared to a conversion rate of 90 percent as reported in the *New York Times* on October 21, 2006 (Feder, 2006).

including health economists, decision analysts, biostatisticians, and health services researchers. In addition to designing and executing field evaluations, THETA is actively involved in developing classes on HTA and field evaluations at the University of Toronto (Toronto Health Economics and Technology Assessment Collaborative, 2007b). THETA is currently in the process of implementing an RCT for deep brain stimulation (DBS) for the treatment of resistant depression. Study participants (a maximum of 20 patients per year) selected for the study will be randomized to receive DBS for a 12-week period or not. This randomization will be repeated 1 additional time, leaving 4 treatment groups: patients who do not receive DBS, patients who receive DBS for the first 12 weeks, patients who receive DBS for the second 12-week interval, and patients who receive DBS for the entire 24 weeks. The main outcome measures include the effect of deep brain stimulation on depressive symptoms, physical and mental health functioning, work and social adjustment, quality of life, and cognitive functioning. The study will be completed in 2011.

Institute for Clinical Evaluative Sciences

PATH and THETA are often able to increase the impact of their studies by collaborating with researchers at the Institute for Clinical Evaluative Sciences (ICES). ICES is an independent, nonprofit organization that receives core funding for its activities from the Ministry of Health. In addition to direct, project-specific funding from various provincial and national organizations, ICES faculty compete for peer-reviewed research grants (ICES, 2007). The organization has about 75 faculty members and nearly 200 staff. ICES faculty are able to link large data sets to monitor patterns of use for various drugs and medical technologies as well as patterns in quality of care (Center for Global eHealth Innovation University Health Network, n.d.). Information from these large data sets has proven invaluable to conducting various field evaluations. For example, ICES researchers played a significant role in the PATH study comparing drug-eluting stents to bare metal stents described in Box 4-3.

The skills of the ICES faculty plus the institution's extensive existing privacy controls make ICES the ideal repository for registry creation and subsequent data analysis. For example, ICES has created several registries to collect data related to positron emission tomography (PET) scanning. There are currently six cancer-related PET registry studies being conducted through this arrangement as well as a registry for implantable cardiac defibrillators. In addition to PET registry studies, there are five prospective clinical trials being conducted by the Ontario Clinical Oncology group for cancer indications. Three of these trials are ongoing, and two RCTs have

recently completed patient accrual (Ministry of Health and Long Term Care, 2008).

University Health Network Usability Laboratories

The University Health Network Usability Laboratories have 15 employees, including human factors analysts and engineers, and are primarily concerned with assessing the safety of medical technologies, which is an important consideration for policy makers and purchasers (Center for Global eHealth Innovation University Health Network, n.d.). The laboratories handle requests from OHTAC for information relating to the ease of use of the technology, qualifications necessary to manage the technology, or risks to hospital staff or patients (Levin et al., 2007). Several topics currently under review from the usability laboratories include safety concerns regarding computed tomography radiation, magnetic resonance imaging, and smart infusion pumps.

Workforce Analysis for Comparative Effectiveness Network in Ontario

Personnel

The activities described above require staff from a variety of backgrounds, including health policy experts, health economists, clinical epidemiologists, biostatisticians, health services researchers, human factors analysts, and engineers, as well as physicians, nurses, hospital representatives, and information specialists. In addition, the success of this network is dependent on the willingness of university faculty and clinical experts to assist in the development of study designs and the collection of necessary data. Therefore, although there is a limited number of core staff, as described above, the system itself includes a far greater range of human resources working collaboratively to fill evidence gaps of importance to decision makers.

In addition, PATH and THETA are involved in developing workshops, classes, and degree programs at, respectively, McMaster University and the University of Toronto to meet future workforce needs. For example, McMaster University has the Center for Health Economics and Policy Analysis,[13] which is funded by McMaster University and the Ontario Ministry of Health. The center offers classes in health economics and policy analysis to students from a variety of degree programs (Centre for Health Economics and Policy Analysis, n.d.). The University of Toronto offers

[13] Centre for Health Economics and Policy Analysis. Available at www.chepa.org/Whoweare/Centre/tabid/59/Default.aspx (accessed July 15, 2008).

degree programs in health technology assessment and management, HSR, and clinical epidemiology through the Department of Health Policy, Management, and Evaluation (Department of Health Policy, Management, and Evaluation, 2008).

Provincial Government Funding for Field Evaluations

Currently, the Ministry of Health spends CA$8 million to CA$10 million a year on field evaluations for high-demand, emerging medical technologies. Technology costs are generally excluded from this figure, but they are also paid for by the Ministry of Health. This figure also excludes the cost of university and hospital-based researchers whose salaries are paid for by their employers or by external granting agencies. Approximately CA$5 million of this funding is invested in the PET registries, leaving CA$3 to CA$5 million for additional field evaluations. The higher cost of the PET registries is primarily due to the costs of the PET radioisotope being paid for from the OHTAC budget. For most conditionally funded field evaluation projects, other government departments cover the clinical costs.

Policy Implications for the United States

Establish a Stable Funding Source to Support Comparative Effectiveness Research

Government funding for the comparative effectiveness programs established in Ontario is critical, because product manufacturers often lack the incentives and hospitals usually lack the resources to support this research. Studies to address important unanswered questions identified by OHTAC are designed and implemented in a short time frame, primarily because a pool of resources is available to support this work. It is also worth noting that the time frame for funding decisions is extremely short, which is essential when attempting to evaluate promising emerging technologies on a time frame that is meaningful for clinical and health policy decision making. To create a similar capacity for conducting research aimed at addressing issues of importance to healthcare decision makers in the United States, it is important to identify a continually available, renewable source of funding. Since there is a mix of public and private health insurers in the United States, it would be beneficial to adopt a system where all health insurers were required to contribute funds to the programs. Furthermore, there will need to be a capacity for rapid decisions about allocation of these funds to support prospective studies. Standard grant review cycle times are unlikely to be adequate to support a productive comparative effectiveness enterprise in the United States or elsewhere.

Ensure That the Process Is Timely and Directed and That Evidence Generation Is Directed at Questions of Importance to Decision Makers

The process of generating evidence described in this paper is both timely and directed at the evidence needs of healthcare decision makers. Once OHTAC requests an HTA from MAS, a full systematic review is returned within 16 weeks, at which time OHTAC can decide to request a full field evaluation. This close and ongoing contact between the Ministry of Health, OHTAC, MAS, and the various programs that conduct field evaluations and economic analyses ensures that studies are responsive to the questions of importance to policy makers and potential purchasers. In Ontario, studies are designed collaboratively with input from government officials, hospital representatives, physicians, health economists, and health services researchers. Keeping decision makers involved in this process increases the likelihood that the data generated by the study will be relevant. In the United States, it will be necessary to establish efficient mechanisms for considering input from a broad range of experts and stakeholders in priority setting, protocol development, and study implementation. The methods and strategies for achieving this are not fully developed or well documented, and considerable work will be necessary in order to achieve functioning mechanisms to obtain broad input and to achieve consensus around priorities and methods.

Design Programs That Are Independent from Government and Industry and Ensure That the Decision Making Process Is Transparent

Although the government is the main source of funding for CER in Ontario, programs conducting the various field evaluations have remained independent. This independence from the Ministry of Health allows these programs to design and implement studies without unmanageable political influence and to more freely engage with consultants and experts. In addition, the fact that OHTAC is a board at "arm's length" from the Ontario Ministry of Health keeps the recommendation process independent from the ministry, thereby separating it from the actual decision-making process.

Efforts have been made by the Ontario government and OHTAC to ensure that the entire process is open to the public. Any Ontario citizen is welcome to submit a request for an assessment of an emerging nondrug medical technology, stakeholder engagement and feedback are solicited via targeted approaches, and all decision and reasons for those decisions are made available via the Internet. Transparency in healthcare decision making is critical to establishing trust from the general public. Decision makers in Ontario continue to look for and adopt new methods to ensure that the public is engaged in the process. When developing a system in the United

States, efforts should be made to ensure that citizens are not only aware of these efforts but also encouraged to engage in the process. Public engagement processes also need to be designed so that those with vested interests do not unduly influence decision making.

Create Partnerships Between Universities and Programs Responsible for Conducting Field Evaluations

The Ontario technology assessment network relies on partnerships between programs conducting field evaluations and various universities, such as the University of Toronto (THETA) and McMaster University (PATH). This partnership allows these programs to draw on the expertise of academics and physicians working at these universities when designing and implementing various studies. Furthermore, this connection has led to the development of classes and degree programs that will help to fill future workforce and expertise requirements. The maintenance of ongoing relationships between the Ontario Ministry of Health and academic programs that specialize in comparative effectiveness studies appears to be important for the efficiency and effectiveness of this work. This bears some similarity to the network of EPCs in the United States and a number of similar academically based networks that develop focused expertise and relationships in order to conduct particular types of projects. It may be sensible to explore the establishment of a network of centers with expertise in conducting comparative effectiveness studies that maintain ongoing relationships with CMS, private payers, and a broad network of stakeholders with an interest in this subject.

Leverage Medicare's Influence on Private Payers

It may be argued that one reason for the effectiveness of Ontario's system is that decision making is relatively centralized compared to the situation in the United States. The payer (the MOHLTC) decides how new nondrug technologies are used in Ontario. In the United States, the existence of a large number of decision makers makes it more difficult to control the diffusion of emerging medical technologies because the technologies can enter the healthcare system through any number of private as well as public payers.

Still, although there is not one central decision maker in the United States, private payers are often influenced by Medicare's coverage decisions, though it is increasingly common that large private payers make decisions that differ from those of Medicare. The influence that Medicare wields on private coverage decisions could be leveraged to develop a comparative effectiveness network, especially if Medicare were to use the existing

Medicare Evidence Development and Coverage Advisory Committee or to establish a new multistakeholder board to perform a function similar to OHTAC. Another factor to consider is that the United States has a much larger HSR capacity than Ontario; this domestic network could be leveraged to review the evidence necessary for the production of coverage recommendations. Where uncertainty remained after a thorough review of all available evidence, Medicare could commission a "coverage with evidence development" (CED) study using government funding, a policy option already used in a number of cases (Tunis and Pearson, 2006). There has been increasing interest in private payer models of CED as well, and it would be particularly effective to have public and private payers supporting the same studies using this policy mechanism.

Methodology Implications for the United States

Draw on Existing Capacities to Support Comparative Effectiveness Research

Government funding for CER in Ontario is relatively small because MAS, PATH, and THETA are able to make use of existing capacities within the province, such as ICES and university researchers and clinicians, to help support their projects. Once these programs receive requests from OHTAC, they are able to launch studies fairly quickly and efficiently, which is critical given the rapid evolution of high-demand, emerging medical technologies.

Unlike in Ontario, where only a small number of clinical research programs are capable of performing the research needed by the Ministry of Health, in the United States there are many HSR organizations as well as an extensive network of universities and teaching hospitals that could help support a CER agenda. The mechanism used in Ontario of assigning individual projects to research programs may not be scalable to the United States, and a competitive procurement process may be more suitable.

With the strong focus on EBM that currently exists, now is an ideal time to choose a high-demand medical technology and implement pragmatic studies in order to demonstrate how CER can be used to inform medical decisions. In addition, initial studies are necessary to refine current methods and inform discussions about the additional capacity necessary to build a comparative effectiveness network.

Invest in a Centralized Capacity to Set Up and Collect Information from Patient Registries

The Ontario network takes advantage of the existence of a separate, larger program (ICES) responsible for creating registries and cross-linking

databases. Although these databases serve to address a range of policy questions other than coverage decisions, the databases and various ICES analyses are used to support many of the field evaluations designed by PATH and THETA. In addition, the ICES databases allow PATH and THETA to implement studies more quickly and at a lower cost than would otherwise be possible if these databases did not exist.

In the United States there are a number payers, including Medicare, United Healthcare, and Blue Cross Blue Shield, that routinely collect patient information through administrative databases and registries. To make this information useful to researchers and decision makers, it would be beneficial to develop greater coordination in the work of collecting and analyzing administrative and registry data.

Use a Combination of Research Approaches to Inform Decision Makers

The technology assessment system in Ontario relies on a number of different study designs to assess emerging technologies and address critical evidence gaps. Decision makers in Ontario rely on information from a number of sources, including systematic reviews, cost-effectiveness modeling, and (if necessary) field evaluations. In addition, when field evaluations are deemed necessary, they are designed to be responsive to the questions of policy makers and care providers and are focused on the costs and effects of the medical technology in real-world practice.

Adopting a similar approach in the United States would help to ensure that studies are directed at the decision-making process and will likely reduce the number of studies concluding that more evidence is needed before a decision can be reached.

"Globalizing" Comparative Effectiveness

Many of the evidence gaps relating to emerging technologies in Ontario have also been identified as important evidence gaps in the United States and abroad. This overlap suggests that there is an opportunity to facilitate linkages and collaboration for activities of mutual benefit. There are lessons to be learned not only from the Ontario experience but also from those in other countries. For example, a government-funded, centralized HTA program in the United Kingdom commissions studies on topics where the evidence base is limited. This program could serve as a useful model for a commissioned-research CED program housed within Medicare.

With respect to individual studies, international partnerships may be helpful, particularly for rare diseases where the number of patients eligible for a study in any single country is small. However, international studies also have disadvantages: they may take longer to initiate; the collection,

assessment, and integration of data may be complicated; and the data may not be generalizable. Furthermore, in order for an international collaboration to be successful, there must be agreement about appropriate study design and outcome measures.

Conclusion

There is currently great interest internationally in both comparative effectiveness and coverage with evidence development. The Ontario experience demonstrates that a significant amount of research can be achieved for a relatively small amount of money if researchers, clinicians, and decision makers work together and make use of existing infrastructure. In the United States and throughout the world, there is a high demand for information on comparative effectiveness for emerging medical technologies, not only for payers and hospitals but also for individual clinicians and patients as well. Beginning to improve the capacity to make evidence-based medical decisions requires immediate action because the pace of medical technology innovation continues to increase, and, as it does, so does the list of questions that need to be answered in order to inform decision makers.

REFERENCES

Ash, J., P. Stavri, R. Dykstra, and L. Fournier. 2003. Implementing computerized physician order entry: The importance of special people. *International Journal of Medical Informatics* 69:235-250.

Atkins, D., D. Best, P. A. Briss, M. Eccles, Y. Falck-Ytter, S. Flottorp, G. H. Guyatt, R. T. Harbour, M. C. Haugh, D. Henry, S. Hill, R. Jaeschke, G. Leng, A. Liberati, N. Magrini, J. Mason, P. Middleton, J. Mrukowicz, D. O'Connell, A. D. Oxman, B. Phillips, H. J. Schunemann, T. T. Edejer, H. Varonen, G. E. Vist, J. W. Williams, Jr., and S. Zaza. 2004. Grading quality of evidence and strength of recommendations. *British Medical Journal* 328(7454):1490.

Bastian, H. 2005. Consumer and researcher collaboration in trials: Filling the gaps. *Clinical Trials* 2:3-4.

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

Bowen, J. M., R. Hopkins, M. Chiu, G. Blackhouse, C. Lazzam, D. Ko, J. Tu, E. Cohen, K. Campbell, Y. He, A. Willan, J.-E. Tarride, and R. Goeree. 2007. *Clinical and cost-effectiveness analysis of drug eluting stents compared to bare metal stents for percutaneous coronary interventions in Ontario: Final report* (Report no. Hta002-0705-02). Hamilton, ON: Program for the Assessment of Technology in Health, St. Joseph's Healthcare Hamilton/McMaster University.

Buckley, T. 2007. *The complexities of comparative effectiveness*. Washington, DC: Biotechnology Industry Organization.

Califf, R. 2006. Clinical trials bureaucracy: Unintended consequences of well-intentioned policy. *Clinical Trials* 3:496-502.

Center for Global eHealth Innovation University Health Network. n.d. *Healthcare human factors group*. www.ehealthinnovation.org/?q=hhf (accessed July 3, 2008).

Centre for Health Economics and Policy Analysis. http://www.chepa.org/Whoweare/Centre/tabid/59/Default.aspx (accessed July 15, 2008).

Chaudhry, B., J. Wang, S. Wu, M. Maglione, W. Mojica, E. Roth, S. Morton, and P. Shekelle. 2006. Systematic review: Impact of health information technology on quality, efficiency, and costs of medical care. *Annals of Internal Medicine* 144:742-752.

Chou R, N. Aronson, D. Atkins, A. S. Ismaila, P. Santaguida, D. H. Smith, E. Whitlock, T. J. Wilt and D. Moher. 2010. AHRQ Series Paper 4: Assessing harms when comparing medical interventions: AHRQ and the Effective Health-Care Program. *Journal of Clinical Epidemiology* 63:502-512.

Cohen, A., W. Hersh, K. Peterson, and P. Yen. 2006. Reducing workload in systematic review preparation using automated citation classification. *Journal of the American Medical Informatics Association* 13:206-219.

Committee to Evaluate Drugs. 2007a. *Terms of reference and administrative guidelines*. Ontario: Ministry of Health and Long-Term Care.

———. 2007b. *Drug innovation fund to advance research into value of medicines. Approved funding for research proposals from 2007/08 review cycle*. Ontario: Ministry of Health and Long-Term Care.

Dall, T., A. Grover, C. Roehrig, M. Bannister, S. Eisenstein, C. Fulper, and J. Cultice. 2006. *Physician supply and demand: Projections to 2020*. Washington, DC: Health Resources and Services Administration.

Department of Health Policy, Management, and Evaluation. 2008. *Course descriptions*. http://www.hpme.utoronto.ca/about/gradprograms/msc-htam/courses.htm (accessed July 15, 2008).

Drummond, M., J. Schwartz, B. Jönsson, B. Luce, P. Neumann, U. Siebert, and S. Sullivan. 2008. Key principles for the improved conduct of health technology assessments for resource allocation decisions. *International Journal of Technology Assessment in Health Care* 24:244-258.

Eardley, T. 2006. *NHS Informatics Workforce Survey*. London, UK: The Association for Informatics Professionals in Health and Social Care.

EHR (Electronic Health Record) Adoption Model. 2008. *The EHR adoption model*. Chicago, IL: Healthcare Information Management and Systems Society.

Ellis, P., C. Baker, and M. Hanger. 2007. *Research on the comparative effectiveness of medical treatments: Issues and options for an expanded federal role*. Washington, DC: Congressional Budget Office.

Emanuel, E., V. Fuchs, and A. Garber. 2007. Essential elements of a technology and outcomes assessment initiative. *Journal of the American Medical Association* 298:1323-1325.

Feder, B. 2006. Doctors rethink widespread use of heart stents. *New York Times*. October 21. http://www.nytimes.com/2006/10/21/business/21stent.html (accessed July 2, 2008).

Fletcher, R., and S. Fletcher. 2005. *Clinical epidemiology: The essentials*, 4th ed. Baltimore, MD: Lippincott Williams & Wilkins.

Forrest, C., A. Millman, J. Hines, and E. Holve. 2005. *Health services research competencies: Final report*. Baltimore, MD: Johns Hopkins Bloomberg School of Public Health.

Forrest, C. B., D. P. Martin, E. Holve, and A. Millman. 2009. Health services research doctoral core competencies. *BMC Health Services Research* 9(1):107.

Fridsma, D., J. Evans, S. Hastak, and C. Mead. 2008. The BRIDG project: A technical report. *Journal of the American Medical Informatics Association* 15:130-137.

Gabler, J. 2003. *2003 integrated delivery system IT budget and staffing study results*. Stamford, CT: Gartner Corp.

Gartlehner, G., R. Hansen, D. Nissman, K. Lohr, and T. Carey. 2006. A simple and valid tool distinguished efficacy from effectiveness studies. *Journal of Clinical Epidemiology* 59:1040-1048.

Goeree, R., and L. Levin. 2006. Building bridges between academic research and policy formulation: The PRUFE framework—An integral part of Ontario's evidence-based HTPA process. *Pharmacoeconomics* 24(11):1143-1156.

Goodman, D. 2008. Improving accountability for the public investment in health profession education: It's time to try health workforce planning. *Journal of the American Medical Association* 300:1205-1207.

Han, Y., J. Carcillo, S. Venkataraman, R. Clark, R. Watson, T. Nguyen, H. Bayir, and R. Orr. 2005. Unexpected increased mortality after implementation of a commercially sold computerized physician order entry system. *Pediatrics* 116:1506-1512.

Haynes, R., D. Sackett, G. Guyatt, and P. Tugwell. 2005. *Clinical epidemiology: How to do clinical practice research,* 3rd ed. Baltimore, MD: Lippincott Williams & Wilkins.

Helfand, M. 2005. Using evidence reports: Progress and challenges in evidence-based decision making. *Health Affairs* 24:123-127.

Helfand, M., S. Morton, E. Guallar, and C. Mulrow. 2005. Challenges of summarizing better information for better health: The evidence-based practice center experience. *Annals of Internal Medicine* 142(12, Pt. 2).

Hersh, W. 2002. Medical informatics: Improving health care through information. *Journal of the American Medical Association* 288:1955-1958.

———. 2006. Who are the informaticians? What we know and should know. *Journal of the American Medical Informatics Association* 13:166-170.

———. 2008. Health and biomedical informatics: Opportunities and challenges for a twenty-first century profession and its education. In *IMIA yearbook of medical informatics 2008,* edited by A. Geissbuhler and C. Kulikowski. Stuttgart, Germany: Schattauer. Pp. 138-145.

Hersh, W., and J. Williamson. 2007. Educating 10,000 informaticians by 2010: The AMIA 10 × 10 program. *International Journal of Medical Informatics* 76:377-382.

Hersh, W., and A. Wright. 2008. *What workforce is needed to implement the health information technology agenda? An analysis from the HIMSS Analytics database*. Paper read at AMIA Annual Symposium Proceedings, Washington, DC.

ICES (Institute for Clinical Evaluative Sciences). 2007. *Knowledge igniting change: 2007 annual report*. http://www.ices.on.ca/file/Annual_Report_2007.pdf (accessed July 3, 2008).

Iglehart, J. K. 2000. Revisiting the Canadian health care system. *New England Journal of Medicine* 342(26):2007-2012.

IOM (Institute of Medicine). 1990. *Clinical practice guidelines: Directions for a new program*. Washington, DC: National Academy Press.

———. 2007. *The learning healthcare system: Workshop summary*. Washington, DC: The National Academies Press.

———. 2008. *Knowing what works in health care: A roadmap for the nation*. Washington, DC: The National Academies Press.

Kirschner, N., S. G. Pauker, and J. W. Stubbs. 2008. Information on cost-effectiveness: An essential product of a national comparative effectiveness program. *Annals of Internal Medicine* 148:956-961.

Langston, A., M. McCallum, M. Campbell, C. Robertson, and S. Ralston. 2005. An integrated approach to consumer representation and involvement in a multicentre randomized controlled trial. *Clinical Trials* 2:80-87.

Levin, L., R. Goeree, N. Sikich, B. Jorgensen, M. C. Brouwers, T. Easty, and C. Zahn. 2007. Establishing a comprehensive continuum from an evidentiary base to policy development for health technologies: The Ontario experience. *International Journal of Technology Assessment in Health Care* 23(3):299-309.

Leviss, J., R. Kremsdorf, and M. Mohaideen. 2006. The CMIO: A new leader for health systems. *Journal of the American Medical Informatics Association* 13:573-578.

Lewis, S., C. Donaldson, C. Mitton, and G. Currie. 2001. The future of health care in Canada. *British Medical Journal* 323(7318):926-929.

Luce, B., L. Paramore, B. Parasuraman, B. Liljas, and G. deLissovoy. 2008. Can managed care organizations partner with manufacturers for comparative effectiveness research? *American Journal of Managed Care* 14:149-156.

Methods Reference Guide. 2008. *Methods reference guide for effectiveness and comparative effectiveness reviews*. Rockville, MD: Agency for Healthcare Research and Quality.

Ministry of Health and Long Term Care. 2008. *Bulletin: Accessing positron emission tomography (PET) studies*. http://www.health.gov.on.ca/english/providers/program/ohip/bulletins/4000/bul4477.pdf (accessed July 3, 2008).

Moher, D., A. Tsertsvadze, A. Tricco, M. Eccles, J. Grimshaw, M. Sampson, and N. Barrowman. 2007. A systematic review identified few methods and strategies describing when and how to update systematic reviews. *Journal of Clinical Epidemiology* 60:1095-1104.

Mullin, T. 2007 (April 17). Statement of Theresa Mullin, Ph.D., Assistant Commissioner for Planning, Food and Drug Administration. *Congressional Record*, D507-D508. Washington, DC: Subcommittee on Health, Committee on Energy and Commerce, U.S. House of Representatives.

Ontario Health Technology Advisory Committee. n.d. *OHTAC membership*. http://www.health.gov.on.ca/english/providers/program/ohtac/committee.html (accessed July 10, 2008).

PATH (Programs for Assessment of Technology in Health) Research Institute. 2008a. *Meet our Team*. http://www.path-hta.ca/team.htm (accessed July 2, 2008).

———. 2008b. *HTA educational learning program*. http://www.path-hta.ca/help.htm (accessed July 1, 2008).

Reboussin, D., and M. Espeland. 2005. The science of Web-based clinical trial management. *Clinical Trials* 2:1-2.

Ricketts, T. 2007. *Developing the health services research workforce*. Washington, DC: AcademyHealth.

Safran, C., and D. Detmer. 2005. Computerized physician order entry systems and medication errors. *Journal of the American Medical Association* 294:179.

Safran, C., M. Bloomrosen, W. Hammond, S. Labkoff, S. Markel-Fox, P. Tang, and D. Detmer. 2007. Toward a national framework for the secondary use of health data: An American Medical Informatics Association white paper. *Journal of the American Medical Informatics Association* 14:1-9.

Shaffer, V., and J. Lovelock. 2007. *Results of the 2006 Gartner-AMDIS survey of CMIOs: Bridging healthcare's transforming waters*. Stamford, CT: Gartner.

Shekelle, P., E. Ortiz, S. Rhodes, S. Morton, M. Eccles, J. Grimshaw, and S. Woolf. 2001. Validity of the Agency for Healthcare Research and Quality clinical practice guidelines: How quickly do guidelines become outdated? *Journal of the American Medical Association* 286:1461-1467.

Shojania, K., M. Sampson, M. Ansari, J. Ji, S. Doucette, and D. Moher. 2007. How quickly do systematic reviews go out of date? A survival analysis. *Annals of Internal Medicine* 147:224-233.

Sittig, D., J. Ash, J. Zhang, J. Osheroff, and M. Shabot. 2006. Lessons from "unexpected increased mortality after implementation of a commercially sold computerized physician order entry system." *Pediatrics* 118:797-801.

Slutsky, J., D. Atkins, S. Chang and B. A. Collins Sharp. 2010. AHRQ Series Paper 1: Comparing medical interventions: AHRQ and the Effective Health-Care Program. *Journal of Clinical Epidemiology* 63(5):481-483.

Swirsky, L., and L. Cook. 2008. *Comparative effectiveness: Better value for the money?* Washington, DC: Alliance for Health Reform.

Tierney, M., and B. Manns. 2008. Optimizing the use of prescription drugs in Canada through the common drug review. *Canadian Medical Association Journal* 178(4):432-435.

Toronto Health Economics and Technology Assessment Collaborative. 2007a. *Research*. http://theta.utoronto.ca/research (accessed July 1, 2008).

———. 2007b. *Education*. http://theta.utoronto.ca/static/education (accessed July 1, 2008).

Tu, J. V., J. Bowen, M. Chiu, D. T. Ko, P. C. Austin, Y. He, R. Hopkins, J. E. Tarride, G. Blackhouse, C. Lazzam, E. A. Cohen, and R. Goeree. 2007. Effectiveness and safety of drug-eluting stents in Ontario. *New England Journal of Medicine* 357(14):1393-1402.

Tunis, S. 2007. *Comparative effectiveness: Basic terms and concepts*. San Francisco, CA: Center for Medical Technology Policy.

Tunis, S. R., and S. D. Pearson. 2006. Coverage options for promising technologies: Medicare's "coverage with evidence development." *Health Affairs* 25(5):1218-1230.

Tunis, S. R., D. Stryer, and C. Clancy. 2003. Practical clinical trials—Increasing the value of clinical research for decision making in clinical and health policy. *Journal of the American Medical Association* 290:1624-1632.

Weissberg, J. 2007. Use of large system databases. In *The Learning Healthcare System: Workshop Summary,* edited by L. Olsen, D. Aisner, and J. McGinnis. Washington, DC: The National Academies Press. Pp. 46-50.

Wilensky, G. 2006. Developing a center for comparative effectiveness information. *Health Affairs* 25:w572-w585.

Woolf, S. 2008. The meaning of translational research and why it matters. *Journal of the American Medical Association* 299:211-213.

Zerhouni, E. 2007. Translational research: Moving discovery to practice. *Clinical Pharmacology and Therapeutics* 81:126-128.

5

Implementation Priorities

INTRODUCTION

Significant gains in the efficiency, effectiveness, and value of health care delivered in the United States are possible with a greater system focus on developing and applying insights on what works best for whom. The near-term needs for an expanded and broadly supported capacity for comparative effectiveness research (CER) include infrastructure for the requisite work (e.g. methods, technical support, coordinating capacities), information networks, and workforce. Identification of the highest-priority implementation needs will guide strategic and coordinated development of needed capacity. Consideration is also needed of how infrastructure development might best build upon existing capacity. Papers in this chapter focus on five key areas for work: (1) information technology (IT) platforms, (2) data resource and analysis improvement, (3) clinical research infrastructure, (4) health professions training, and (5) building the training capacity. Each paper offers suggestions for prioritization and staging of policies, as well as possible approaches to increasing the scale of activities. Also discussed are opportunities to take advantage of existing manufacturer, insurer, and public capacities through public–private partnership.

The first three papers focus on developing information acquisition and exchange tools as well as the research approaches essential to speeding evidence development. Based on his experiences developing a regional health information exchange in Tennessee (the Memphis Exchange), Mark E. Frisse of Vanderbilt University suggests several implementation priorities for the development of an IT platform that will realize significant

societal benefit at a realistic marginal cost. With appropriate design and integration, current collections of databases, health record systems, health information exchanges, financing, workforce, policies, and governance, it can be evolved into a system that addresses a range of needs in care delivery, process improvement, and research. T. Bruce Ferguson from the East Carolina Heart Institute discusses clinical database work in the field of cardiology and identifies key opportunities to apply data resource and analysis infrastructure toward the development of dynamic, real-time learning systems, centered on the patient and decisions at the point of care. Finally, Daniel E. Ford of Johns Hopkins University discusses opportunities to improve the efficiency and effectiveness of clinical research by streamlining and standardizing processes and policies, increasing investments in practice-based networks and training and retaining research support personnel. Two papers focus on the workforce at the front lines of evidence application and development—health professionals and clinical researchers. Benjamin K. Chu from Kaiser Permanente describes changes to the healthcare delivery system that will shape the future practice environment and illustrates how training and practice environments for health professions education should seek to emulate and improve upon current models of best care. Steven A. Wartman of the Association of Academic Health Centers describes a needed expansion of medical research to a multidisciplinary approach that addresses all aspects of health. He offers some suggestions on how the training capacity might be developed to accelerate a shift to research focused on the discovery, dissemination, and optimized adoption of practices that advance the health of individuals and the public.

This chapter concludes with discussion highlighting opportunities to take best advantage of existing infrastructure elements—such as data resources, expertise, and technology platforms. Speaking from key sector perspectives, Carmella A. Bocchino from America's Health Insurance Plans, Rachael E. Behrman from the Food and Drug Administration (FDA), and William Z. Potter from Merck Research Laboratories, discuss how public–private partnerships can create needed space for cross-sector collaboration around common areas of interest and expertise.

INFORMATION TECHNOLOGY PLATFORM REQUIREMENTS

Mark E. Frisse, M.D., M.Sc., M.B.A., Professor of Biomedical Informatics, Vanderbilt University

Overview

The overarching intent of this publication is to better understand the requirements necessary to transform our fragmented healthcare infrastruc-

ture into a learning health *system*. This system must be structured in a way that draws on the best evidence, delivers the best value, adds to learning throughout the system of care, leads to improvements in the nation's health, and ensures that "each patient receives the right care at the right time" (IOM, 2007, 2008).

Where IT platform requirements are concerned, with thought and cautious action, it is possible to realize the aims of a learning health system through an evolution of our current collection of databases, health record systems, health information exchanges, financing, workforce, policies, and governance. Properly designed and integrated, the composite system would be able to address a wide range of needs at a manageable marginal cost for each. However, the *status quo* without thoughtful attention to the ends and means may actually impede long-term progress at the expense of short-term expedience.

A recent report by the National Research Council provides some guidance. Among the principles for change espoused in this report is the assertion that health technologies should "record available data so that today's biomedical knowledge can be used to interpret them to drive care, process improvement, and research" (NRC, 2009). All too often, the design of current systems emphasizes administrative transactions and episodic care at the expense of other priorities. Data are often embedded into specific applications and not represented in a way that clarifies their context or allows reinterpretation as both our analytic techniques and our needs change (NRC, 2009).

An Infrastructure Framework

IT platforms should be based on a clear framework that enables progress toward a wide range of scientific, clinical, and policy aims, while allowing for these aims to evolve over time. The framework should be guided by the analysis and prioritization of initiatives according to their value, difficulty, and requirements for data sharing. The framework should identify potential outcomes according to their impact on effectiveness, quality, safety, and efficiency. In practice, this framework would provide a means of assembling governance, policy, technology, and processes into a series of components that work with one another and that can evolve incrementally over time toward the primary goal of supporting and improving our ability to create and use healthcare knowledge. Such an infrastructure focuses on components that must be assembled to realize specific outcomes. It is these components that should be the focus of activity. Instances of component collections—including various forms of electronic health records (EHRs), personal health records, and health information exchanges—should be viewed not as monolithic products but instead in terms of what their com-

ponents contribute separately and collectively to meeting a specific clinical need.

There are many discrete components and functions, including digital connectivity, source identification, data integrity checking, record location, data aggregation, audits, data collections, and computer–human interfaces. A *system* is composed of multiple instances of each component (e.g., databases and record locator services) originating in a diverse array of local and national settings and designed for different primary purposes. Each instance of a component can in theory be funded through different means and managed under different governance and operational controls. Each component's means of representing data can differ as long as two characteristics are met: (1) ways to combine data in order to achieve practice aims must be implemented, and (2) original data elements must be maintained in their original format and, to the greatest extent possible, coupled with the context in which they were obtained.

What unites the disparate instances of components and creates a true system is a clear separation of data from application, a retention of source and context, and a common minimal set of governance structures and policies that address appropriate uses, performance, financing, and responsibility. Governance, policy, and standards are coordinated only to the minimal extent necessary to achieve results, to gain trust, to demonstrate value, and to support incremental progress. *System* value is recognized not through successful implementation but rather through the impact the system and its components have on measurably improved outcomes.

Lessons from Memphis

The work necessary for developing a regional health information exchange in Memphis, Tennessee (the Memphis Exchange), demonstrates the feasibility of applying these principles and the practicality of this approach. The Memphis Exchange is based on technologies and practices in use for over a decade at the Vanderbilt University Medical Center and described elsewhere (Stead, 2006; Stead and Starmer, 2007). This system produces short-term system-based results, supports incremental improvements, and fosters evolutionary change (Frisse et al., 2008; Johnson et al., 2008). Many lessons have been learned during its 3 years of use and operation.

First, trust and policy—not technology—are the primary barriers to realizing a desired IT platform. Developing data-sharing agreements governing use and oversight was arguably the most challenging initial task. This effort was accelerated considerably by efforts made through the Markle Foundation's *Connecting for Health* initiative (Connecting for Health, 2006).

Second, information from many different systems and encoded in many

different acceptable standards can be combined inexpensively. These data are "liquid" and are not tied to a specific application but instead to a source, a context, and a unique individual. Each clinical or administrative data element is "wrapped" with a meta-level tag that provides a general description while the original data element—in whatever format it is received—is retained. Currently, the exchange receives data from multiple systems at over 20 major healthcare institutions. Some data elements—like laboratory results—can be presented in a uniform format using Logical Observation Identifiers Names and Codes (LOINC) (Porter et al., 2007). Such an approach can be generalized and can provide intermediate results while the long-term process of standards convergence takes place.

Third, identification and matching of data can be achieved with a degree of precision if attention is devoted to measuring performance using a "gold standard" data set of 5,000 to 10,000 patients. Such a matching approach is not a master patient index in a traditional sense because no unique patient identifier is generated and linkages are represented as data clusters rather than as absolute mappings.

Fourth, perceptions of ownership are more important than the locality often embodied in the "centralized vs. decentralized" debate. In the Memphis Exchange, each participating institution publishes its data to its own "vault." A vault in this context is a logical database that may be housed in a central or distributed cluster of databases. What is important is that each institution providing data maintains control of its data until they are combined and used to treat an individual patient. When data are used, actual use is recorded in logs, and efforts to assure nonrepudiation are enforced. Our contention is that no system is completely centralized, and many significant queries can only be answered through a collection of loosely coupled systems.

Fifth, confidentiality and privacy can be achieved through a relatively absolute "opt in" or "opt out" decision made at each institution. The primary focus of our confidentiality efforts is on developing a network of trust that is heavily audited and rigorously enforced. This approach ensures that the only individuals examining data are those who have rights (by law or consent). Emphases on selective data, drugs, or other disorders are not easily manageable and cannot be absolutely enforced unless all free-text documents are excluded. Unfortunately, these text documents (e.g., transcribed medical histories) often provide the most meaningful information both for patient care and for chart review.

Finally, based on the Vanderbilt experience, loosely coupled data sets from disparate resources seem capable of supporting a wide range of research efforts. Using technologies and methods similar to those of the Memphis Exchange, Vanderbilt researchers have developed a deoxyribonucleic acid (DNA) biobank linked to phenotypic data derived from the

Vanderbilt EHR (Roden et al., 2008). Employing an opt-out consent model, these researchers have developed a statistically de-identified mirror image of the electronic medical record (EMR) called a "synthetic derivative." These records are linked to DNA extracted from discarded blood samples. In one test, the de-identification algorithm removed 5,378 of the 5,472 identifiers, with an error rate for complete Health Insurance Portability and Accountability Act (HIPAA) identifiers of less than 0.1 percent. The aggregate error rate—which includes any potential error, including non-HIPAA items, partial items, and items that are not inherently related to identity—was 1.7 percent. The ability of these de-identification procedures to discover and suppress identifiers was sufficient for institutional review boards to judge the research done with this system to be consistent with an Office of Human Research Protections "nonhuman subjects" designation.

It should be possible to apply such a process equally well to health information exchanges or other ways of accessing information from disparate sources. Such applications will be powerful tools in biosurveillance, public health research, quality improvement, and comparative effectiveness studies.

Applicability to Information Technology Platform Requirements

This approach is very affordable. The total operational costs for a region of 1 million people are under $3 million a year. Even with additional expense incurred by increasing connectivity to smaller care settings and enhancing data-analytic capabilities, the overall cost will be less than $5 million (or $5 dollars per capita per year). This expense should be compared with overall healthcare expenditures, which are estimated at $7.4 billion, or $7,400 per capita, per year. Thus the expense would amount to less than 0.07 percent of per capita healthcare expenditures. Because the costs are largely offset by reductions in duplicate testing, efficiencies in quality metrics, public health reporting, and other functions, the costs that could be allocated to knowledge management and development of a learning health system are insignificant by almost any degree. Extrapolating to a population of 350 million, our cost estimates ($1.7 billion) are less than estimates provided in Chapter 3 of this publication, but our cost models may be based on different assumptions (Miller, 2008).

The Role of Electronic Health Records

The Memphis Exchange is but one part of a larger health information technology (HIT) platform. Clearly, the choice and effectiveness of care delivery technologies (such as EHRs) are critical. Using Miller's estimates, marginal annual operating expenditures (per capita per year) would be in

the range of $50 (Miller, 2008). As expected, the costs for systems to deliver the details of care exceed the cost estimates for integrating EHRs into a broad IT platform. EHR costs will likely be offset by efficiencies or driven by other practice imperatives, so the question is not so much what a system costs but the extent to which such a system improves practice performance and the extent to which it can send and receive data from other sources to achieve desired results. If the systems are properly designed, their marginal cost to achieve broader aims is very low.

Properly designed, the marginal benefit of a connected system is quite substantial, and the marginal cost of creating such a system (in context to overall healthcare technology costs or to healthcare expenditures overall) can be very low. Thus the greatest risk to realizing great benefit at low financial and societal cost is likely to be the inclination to create monolithic systems that overengineer and promise more than they can deliver.

Additional Initiatives and Decisions

Some national investment decisions can be made that would simplify the integration of data across disparate systems. Although the Memphis Exchange argues that much can be done without the monolithic standardization efforts and privacy initiatives espoused by many, much more can and must be done to make this experience more applicable. Among the most valuable steps that could be taken are an immediate acceleration of knowledge representations that could be quickly applied to clinical use (e.g., RxNorm, unified medical language system), decisions about the extent to which payment and administration coding standards can reflect disease states and contexts required of learning health systems (e.g., International Classification of Diseases [ICD]-9, Systematized Nomenclature of Medicine, ICD-10), enforcement of a few—and only a few—selective standards (e.g., LOINC, SCRIPT), promotion of efforts that make laboratory and medication history more portable in a secure and affordable way, and selection of a few simple high-quality initiatives that can guide improvement of any interventions enabled by IT (Frisse, 2006).

Focused trials with immediate findings are essential to ensure that IT expenditures are made wisely. Proposed legislation to accelerate the adoption of HIT does not assure an optimal outcome. Applying more funds to technologies that are not coupled to system improvements may help, may hurt, or may do both.[1]

[1] U.S. Senate Committee on Finance. 2009. *American Recovery and Reinvestment Act of 2009*.

DATA RESOURCE DEVELOPMENT AND ANALYSIS IMPROVEMENT

T. Bruce Ferguson, Jr., M.D., Chairman,
Department of Cardiovascular Sciences, East Carolina Heart Institute
and Brody School of Medicine at East Carolina University; and
Ansar Hassan, M.D., Ph.D., Brody School of Medicine at ECU

Overview

Enormous challenges face U.S. healthcare stakeholders if the 2020 goal of the Roundtable on Value & Science-Driven Health Care—that 90 percent of clinical decisions will be supported by accurate, timely, and up-to-date clinical information that reflects the best available evidence—is to be met. Among the most complex of these challenges is the issue of the data and data analysis that will be used to drive those clinical decisions. Knowledge about the comparative effectiveness of (1) diagnostics and treatments, (2) providers choosing and administering diagnostics and treatments, and (3) the direct value and benefit to individual patients of (1) and (2) is what must be assembled from data and data analysis going forward. Within the context of CER, using cardiovascular disease as an example, this paper will address the data resource development and the data analysis improvement necessary for the migration of health care toward these 2020 goals.

Data as Knowledge

Despite a multiplicity of potential information resources, there is no cogent framework for selecting and using these resources. Within cardiovascular disease, each of the major stakeholder groups has independently developed, financed, and extensively used data generated from systems that are mostly perceived to be proprietary. These data types include the following:

- Data from the medical product (pharmaceutical and device) companies, which are incentivized to collect safety and efficacy data from pivotal randomized clinical trials (RCTs) for FDA approval of their technologies. The knowledge generated from these studies is critical to the regulatory process. Because equipoise is necessary to randomize patients, particularly in noninferiority trial designs, this body of knowledge is scientifically valid but limited in its applicability to overall care delivery evaluation of effectiveness. Controversy surrounds the application of these trial findings to patients beyond the trial design and beyond the FDA labeling for

the technologies or pharmaceuticals. Investment in postmarket data collection and analysis, except as required for physician and hospital reimbursement (e.g., Centers for Medicare & Medicaid Services [CMS] Pay with Evidence Development program), has generated an important data void in our healthcare system (Bach, 2007).

- Healthcare data available from the public domain and through federal agencies such as CMS, Centers for Disease Control and Prevention (CDC), Agency for Healthcare Research and Quality (AHRQ), and the Social Security Administration require analytical expertise and may be expensive. These data provide knowledge on the administrative, financial, and quality characteristics of care delivery based on claims and administrative data that may be somewhat limited in describing actual clinical care delivery.
- Payers have developed robust administrative and claims-based proprietary systems that extend up to—but as yet do not include—whether a patient actually ingested the medication that was prescribed and filled. These systems are relatively unique in that they give a longitudinal documentation of care with data, some of which have been risk adjusted. These data provide knowledge about longitudinal care processes delivered by multiple providers but are confined to specific payer groups for defined periods of time.
- Practitioners in cardiovascular disease have developed robust clinical observational databases, such as the Society of Thoracic Surgeons' National Adult Cardiac Surgery Database (Ferguson et al., 2002), the American College of Cardiology Foundation's National Cardiovascular Data Registry (ACCF, 2008), and the American Heart Association's Get with the Guidelines (Giugliano and Braunwald, 2007). In addition, regional databases, such as the Northern New England Cardiovascular Consortium (Malenka et al., 2005) and the New York State Cardiac Surgery and Percutaneous Coronary Intervention Registries, have been collecting data for over 15 years. These clinical registries have developed methods to describe risk-adjusted outcomes that, along with processes of care, describe care delivery specific to the procedure-based episode of care. They have independently validated the processes and outcomes of care that are linked to quality improvement. These systems provide knowledge about those care episodes that is clinically relevant but limited in its scope.
- Providers have also devoted considerable effort to the development of guidelines to direct clinical care (ACC, 2008). This is a resource-intensive effort, and much of the data available for guideline development falls short of class I data. The knowledge contained in

the guidelines represents what expert consensus suggests should be done in clinical scenarios that fit into the guideline construct; however, this may limit their usefulness in comparative effectiveness analyses. More recently, the specialty societies have developed guidelines for appropriateness of care, which may become more useful (Douglas et al., 2008).

The fifth stakeholder—the patients and their families—in part desires that this knowledge be integrated in such a way that care delivery centered on the needs and medical conditions of the patient is always available. This requires knowledge about processes and preferably risk-adjusted outcomes of care, as well as administrative and financial data. This cannot be accomplished by using data from just one stakeholder's system or by employing just one type of knowledge data.

Figure 5-1 illustrates the reason for this. For a patient with a medical condition for which there are two potentially applicable therapies, clinical trials data are unlikely to differentiate between the two therapies because of trial design issues (panel A). A more accurate representation of potential therapeutic effectiveness for that patient is derived from the pool of "application" data, or knowledge gained from data describing the ongoing application of health care to patients. In fact, this is the data domain in which most patients and providers reside and which represents the real challenge regarding data resources and data analysis for comparative effectiveness. A slightly different way of looking at this is represented in panel B of Figure 5-1. Wennberg et al. (2002, 2007) have described a recommendation for Medicare reform based upon three categories of medical services and their direct links to health care spending in the Medicare program. In fact, the majority of health care delivered is either preference- or supply-sensitive care, where the knowledge for these decisions comes from application data. For example, in the United States over 75 percent of patients currently undergoing coronary artery bypass grafts (CABGs) wouldn't have been eligible for enrollment in the surgical arms of the major randomized trials of percutaneous coronary interventions (PCIs) vs. CABGs based on National Adult Cardiac Surgery Database data (Taggart, 2006), while at the same time an estimated 70 percent of drug-eluting stent (DES) use in this country is currently presumed to be "off-label" (Tung et al., 2006). In terms of comparative effectiveness between these two therapies, in a recent systematic review of PCI vs. CABG by an AHRQ-sponsored evidence-based practice center, observational analyses were excluded from the principal meta-analysis of trials, which concluded that survival at 10 years was similar between the two therapies (Bravata et al., 2007). Until recently, data from RCTs of PCIs with or without DESs vs. CABGs have not demonstrated any difference in outcomes at 1- or 5-year follow-up (Daemen et

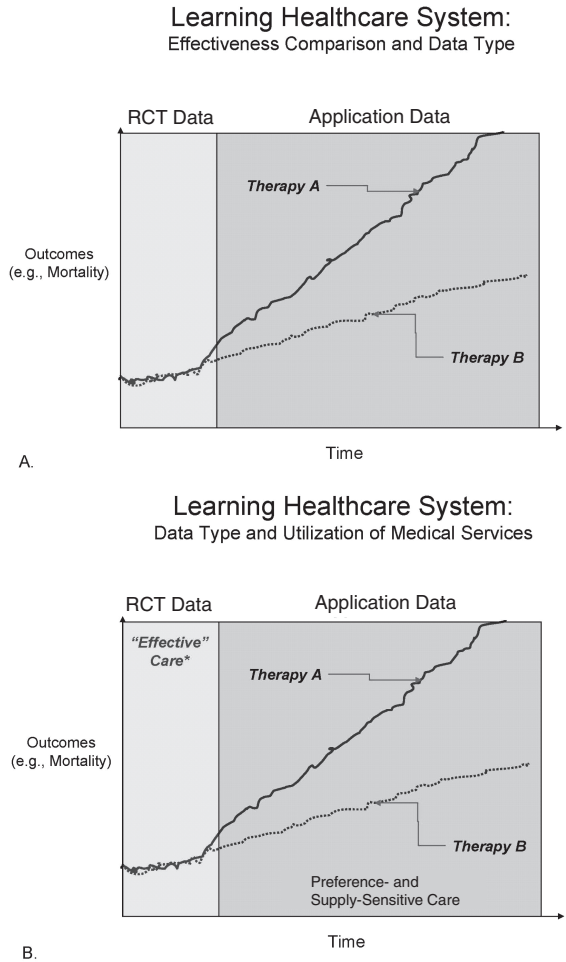


FIGURE 5-1 Panel A shows the hypothetical relationship between information generated from RCT data and application data (data generated through the application of health care to patients) on two different therapeutic interventions. As a result of trial design and equipoise for randomization, an outcome such as mortality is unlikely to be measured as discernibly different. Over time, however, application data may highlight differences in that outcome. Some controversy exists as to whether data from RCTs is appropriate for making decisions in the application data space, and vice versa. Panel B relates this construct to the utilization of medical services as described by Wennberg et al. (2002). The majority of service utilization is in the preferences- and supply-sensitive categories; these activities fall under the application data categorization and constitute the primary target area for comparative effectiveness research going forward.
NOTE: RCT = randomized controlled trial.

al., 2008; Hlatky et al., 2004). If a patient met the enrollment criteria for the trial, these data could be applied and characterized as "effective" care. In contrast, multiple large observational analyses have consistently demonstrated an increasingly significant survival benefit from CABGs as compared with PCIs, beginning at 1 year postintervention (Smith et al., 2006). From the perspective of a patient whose medical condition places him or her in the "gray box," the recommendation for therapy A vs. therapy B would be based on preference- or supply-sensitive care considerations and application data. Recently, the Synergy between PCIs with Taxus and Cardiac Surgery "all comer" trial of PCIs with DESs vs. CABGs demonstrated a mortality difference that was similar to the large observational studies. This example illustrates the complexity of the data requirement for a comparative effectiveness study. It also emphasizes the need for data resources that come from all stakeholders, and the need for taking into account all stakeholder's perspectives. Califf and colleagues (2007) have also emphasized that there needs to be partnership development among these stakeholders to address the cardiovascular disease epidemic. Their argument that an ongoing risk–benefit balance of technology needs to be derived in part from its ongoing use by providers is important because the information from this use becomes a component for comparative effectiveness analyses in the learning health system of the future.

Current Demand Shortfalls for Data Resources and Data Analysis

Data Resources

It is important that the framework in which to assess the current demand shortfalls for data resources and analysis for a learning health system be synchronous with the framework necessary to turn these shortfalls into solutions. This in turn emphasizes principles outlined by the Institute of Medicine (IOM) in 2006 regarding the nature of healthcare information, which must become more aligned with the IOM's six tenets of health care and within the context of comparative effectiveness.

With respect to data resources, these demand shortfalls can be categorized into structure, administration (process), and organization (Figure 5-2). As outlined above, these available data fall short in providing patient-level data that are complete across the medical condition for that patient, as defined by Porter and Teisberg (2006). The provider-level data necessary to address quality of care delivery are also incomplete, both per patient and across the medical condition. As outlined by Califf and others, the data infrastructure is not designed as a resource to generate data where gaps in information for comparative effectiveness exist (in part, the application data in Figure 5-1), namely health policy and quality improvement

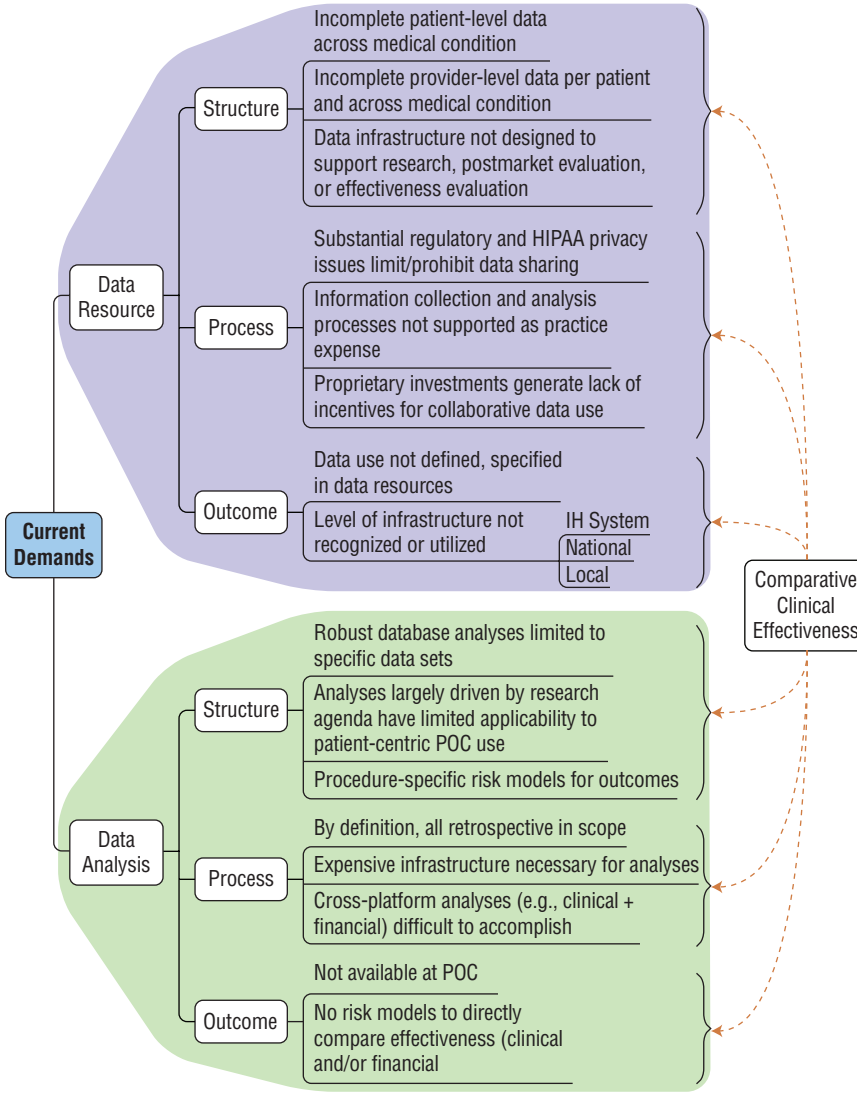


FIGURE 5-2 An analysis of current demand shortfalls in terms of potential comparative effectiveness data resource and data analysis. Individual points are discussed in the text. Each one of these levels, however, relates directly to moving the comparative clinical effectiveness and learning health system agenda forward.
NOTE: HIPAA = Health Insurance Portability and Accountability Act; IH = international health; POC = point of care.

(QI) research, postmarket evaluation, and effectiveness (both medical and financial) on an ongoing basis.

Administratively, the current demands have highlighted a number of obstacles as well. There are substantial regulatory and HIPAA privacy issues that limit or even prohibit data sharing across the patient's medical condition. In terms of financial support for these data resources, information collection and analysis processes by providers for quality improvement have not been supported as a recognized practice expenditure worthy of specific reimbursement. Significant proprietary investments in data resources have resulted in few incentives for collaborative data use among and across these stakeholders.

In terms of organization, there remains an important disconnection between data resources and data uses that are not defined and specified in these data resources; this disconnection sometimes produces conflicting and erroneous data and interpretation from these otherwise important resources. Finally, we have been slow to recognize that there are in fact at least three important levels of data resources that must be used fully in order to move this agenda forward. Integrated health systems, with major financial commitments to EMR systems, bring a unique and important experience to the table, but one that is still very limited in its applicability to most provider systems (James, 2007). The national-level resources from providers and payers have a much broader potential impact. In addition, it is important to recognize that local and regional resources are making investments in the learning health system; these entities may be able to address certain of these data resource and analysis issues without some of the obstacles and shortfalls present at these other levels.

Data Analysis

Analysis shortfalls can be grouped into structure, administration (process), and implementation (outcome) categories (Figure 5-2). In terms of structure, robust analyses that are available from clinical data sets such as the Society of Thoracic Surgeons and the American College of Cardiology (ACC) are now largely confined to those data sets, which mostly capture data from procedural ("vertical") but not longitudinal ("horizontal") episodes of care. These robust systems have developed risk models to facilitate cross-site comparisons of outcomes and within-site comparisons of observed vs. expected (or predicted) outcomes, using national-level populations to generate and validate the risk models. These models, however, are procedure based and are not based on medical outcomes, nor do they incorporate increasingly important additional data such as epidemiologic, socioeconomic, and long-term survival data. Finally, a majority of these analyses have been driven by a health policy and clinical research agenda

because of the design of the data system; they have had limited applicability for point-of-care use that is patient-centric. The use of administrative data sets for outcomes analysis, with the limited clinical information available, has been a challenge (Hall et al., 2007; Krumholz et al., 2006).

For data analysis administration, there continues to be a temporal discontinuity between the data sets and the analyses, with all of these analyses being retrospective in scope. Overall, where these large data set analyses are concerned, analysis activities have required rather expensive infrastructure to manage the collection and analysis of the data. Finally, in part because of the structural nuances of the data sets, cross-platform analyses (clinical + clinical, clinical + administrative, clinical + financial) have so far been difficult to accomplish.

The implementation shortfalls highlight the principle that outcomes are more important to patients than structure or process of care. The fact that these analyses and their outcomes are not generally available is an important concern in a learning health environment. It is important not only to learn what the most effective care is, but it is also important to be able to make that decision and apply it as close to the point of care as possible. For example, there have not yet been risk models developed that directly compare, for patients, the effectiveness of therapeutic options, although work in this area is beginning (Ferguson, 2008; Singh et al., 2008). Financial cost and effectiveness data need to be part of this point-of-care implementation.

Clinical comparative effectiveness assessment as part of a learning health system will, to a varying degree, affect each of these resource and analysis demand shortfall issues. The stress that these demands place on existing data resources is substantial. The possible opportunities for migrating to more operationally sustainable platforms in the future become somewhat clearer when coupled with the IOM criteria for future health information.

An Overview of Next Steps

Data Resource Development

The source of much of these data in the future will be the EMR infrastructure, which is still mostly site specific. However, the focus of this paper is on the resource development steps that are critical but generic to a functional data infrastructure for CER going forward.

First, it will be necessary to better define the type, source, and use of data for comparative effectiveness (Figure 5-3). This includes the classic Donabedian triad, but in this case outcomes include both clinical and financial data. In addition, data and metrics for efficiency, effectiveness, and appropriateness need to be available. In terms of data sources, the use

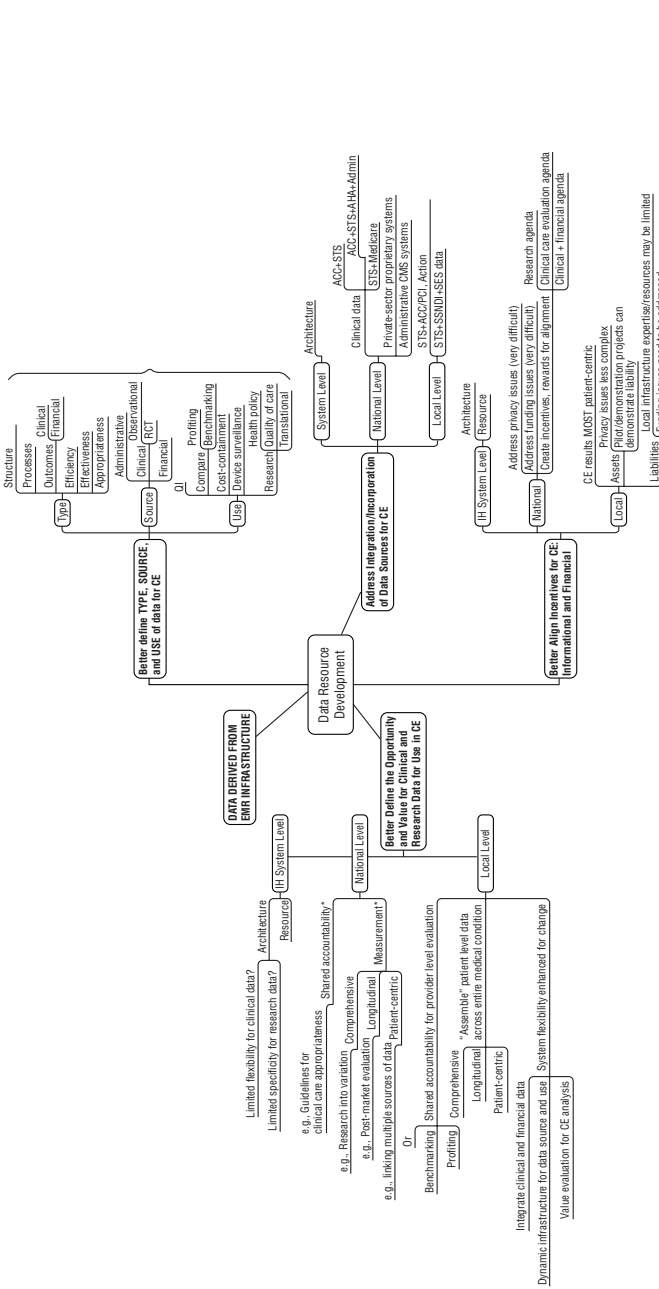


FIGURE 5-3 A relational "map" suggesting major areas of data resource development for comparative effectiveness studies.
NOTE: ACC = American College of Cardiology; AHA = American Hospital Association; CE = comparative effectiveness; CMS = Centers for Medicare & Medicaid Services; EMR = electronic medical record; IH = international health; PCI = percutaneous coronary intervention; QI = quality improvement; RCT = randomized controlled trial; SES = socioeconomic status; SSNDI = Social Security National Death Index; STS = Society of Thoracic Surgeons.
SOURCE: Ferguson, T. B., IOM workshop presentation, July 30-31, 2008.

of administrative, financial, and clinical data, including both RCTs and observational information, needs to be agreed upon, as does the resource sharing between the components of these data sources. Finally, there needs to be substantive agreement on how the data from these data resources will be used, with the correct data applied to the correct use. Indeed, much of these data are already available; resource development efforts need to be focused on how to operationalize data collection, how to define and make change and interoperation dynamic, and how to standardize data use for comparative effectiveness analyses. In examining how to address the integration and incorporation of data sources for CER, it becomes clear that progress can be made at all three levels of healthcare delivery systems, as suggested above. At the integrated health system level, the extent to which this is possible is largely defined by individual system architecture. At the national and local levels, different resources and opportunities available at one level are not available at the other level; however, success at either or both levels moves the agenda forward. Integration of data at the patient level across administrative and clinical data platforms can and has been accomplished locally. At the national level, the National Consortium of Clinical Databases (NC2D) is examining how the Society of Thoracic Surgeons, the ACC, and the American Hospital Association clinical database activities can address these integration and incorporation challenges from the data resource perspective much the same way that these societies have partnered to create guidelines for clinical care and appropriateness. Other national-level data integration projects are under way, facing different sets of challenges with respect to privacy and technology than perhaps exist at the local level.

An important third area for data resource development is the need to better align incentives for CER, both informational and financial. At the integrated health system level, this alignment would be dependent on both architecture and resource factors. Again, differential opportunities between national and local settings become apparent. Currently, both privacy and funding issues for these resource development activities are very difficult at the national level. As mentioned, specialty society efforts in QI, funded by provider contributions to support and participate in society-led databases, have yet to be recognized as valid practice expenses despite their substantive contributions to an improved quality of care (Ferguson et al., 2003). On the other hand, the national platform allows for the creation of new and important incentives and rewards for alignment; this has implications for the research, clinical care evaluation, and clinical plus financial data agendas going forward. It is in this area where perhaps the differentiation between national and local activities might result in the greatest development achievements early on. First and foremost, the local level is where comparative effectiveness implementation and results will be the most

patient-centric. Potential assets include the facts that privacy issues are currently far less complex at the local level than at the national level and that the opportunity for pilot projects to demonstrate feasibility is substantial. Potential liabilities include the fact that local infrastructure expertise and information resources may be limited, although as healthcare systems move into the EMR environment, this is less likely to be the case. Additionally, funding for these activities still remains an issue to be addressed, because they remain an expensive investment. Better alignment of these incentives for CER will reduce the overall cost, while making sustainable comparative effectiveness studies a part of everyday clinical care delivery.

An additional data resource development area is to better define the opportunity and value of clinical and research data (part of application data, Figure 5-1) for use in CER. The value of these data is referenced to administrative data, largely through the major payer resource mechanisms, which are ubiquitous and applicable to all providers. It is not meant to diminish the importance or utility of these data, but only to acknowledge that it likely has limited usefulness for evaluation of comparative effectiveness.

At the national level, this comparative effectiveness agenda ultimately must intersect with the design principles for a national system for performance measurement and reporting (IOM, 2006). It is in part through this intersection that the comparative effectiveness agenda and six healthcare aims articulated by the IOM in 2001 (IOM, 2001) can be pursued simultaneously. Specifically, in each of the areas a comprehensive, longitudinal, and patient-centric measurement can be linked to clinical and research data resources. In addition, shared accountability can be linked to guidelines for clinical care and appropriateness of care.

At the local level substantial opportunity exists as well. Provider-level evaluation for quality improvement, benchmarking, and profiling can be most easily extended to the shared accountability criterion at this level. Migration from data with a provider-centric focus to assembling patient-level data across the entire medical condition is beginning to occur at the local level, while meeting the comprehensive, longitudinal, and patient-centric parameters outlined above. Finally, the local level provides for increased system flexibility for change. With these anticipated dynamic developments in data resources, the integration between clinical and financial data can keep pace with the annual financial reassignment process. The data resource infrastructure must equally be dynamic and rapidly amenable to changes in data definitions and measurement specifications. Perhaps most importantly, it is at this local level where a patient-centric healthcare value can be measured, where value is the quality of patient outcomes relative to the dollars expended (Porter and Teisberg, 2007). Again, from the patient's perspective, comparative effectiveness evaluation is most important and has its greatest impact at this level.

Data Analysis Improvement

As important as the data resource development process will be, the evolution of data analysis will be a key feature for making the comparative effectiveness agenda operational over the long term. As mentioned, most robust data analyses available today are by definition retrospective; they involve harvesting, aggregating, and then analyzing existing data. As useful as this information can be, there remains too great a gap between these analysis outcomes and point of care.

The first challenge is to incorporate these new data resource developments into analyses (Figure 5-4). Thus at the national level, long-term clinical and financial outcomes analyses are of critical importance. Additionally, the data resource developments will require the generation of new risk models to assess outcomes.

At the local level, the drive to integrate information across the medical condition will in turn drive new analysis tools for these integrated data sets that can be managed with the local level of expertise. This will include specific integration of data sets that move these local analyses beyond the national clinical or administrative analyses toward being more patient-specific. For example, Figure 5-5 illustrates an analysis of the integration of 3 years of National Adult Cardiac Surgery Database clinical data with National Death Index data for long-term outcome and with ZIP code data defining social economic status for those patients operated upon at the East Carolina Heart Institute between 2005 and 2007. Likewise, regional data-sharing arrangements, such as the Virginia Cardiac Surgery Quality Improvement project sharing adult cardiac surgery clinical and financial data, can be highly productive (Grover, 2008).

One of the most important developments in data analysis and improvement will be new patient-centric comparative effectiveness analyses. Current outcome risk models are procedure specific and generated from national data. While these are important metrics with which to evaluate risk and effectiveness, they have limited direct applicability to any individual patient in a particular healthcare setting affected by site-specific care practices and local provider influences. At the national level, the challenge will be to develop comparative effectiveness models of risk that account for multiple procedural options. The integration of clinical data resources, such as the NC2D initiative, is a critical step in this analysis development, because these comparative effectiveness risk models cannot be developed and tested based on single center or local site data. An additional challenge at the national level will be to develop models for assessing risk over the duration of the medical condition beyond the specific intervention-based episodes of care. Contributions from progress at the local level in these developments may prove extremely useful.

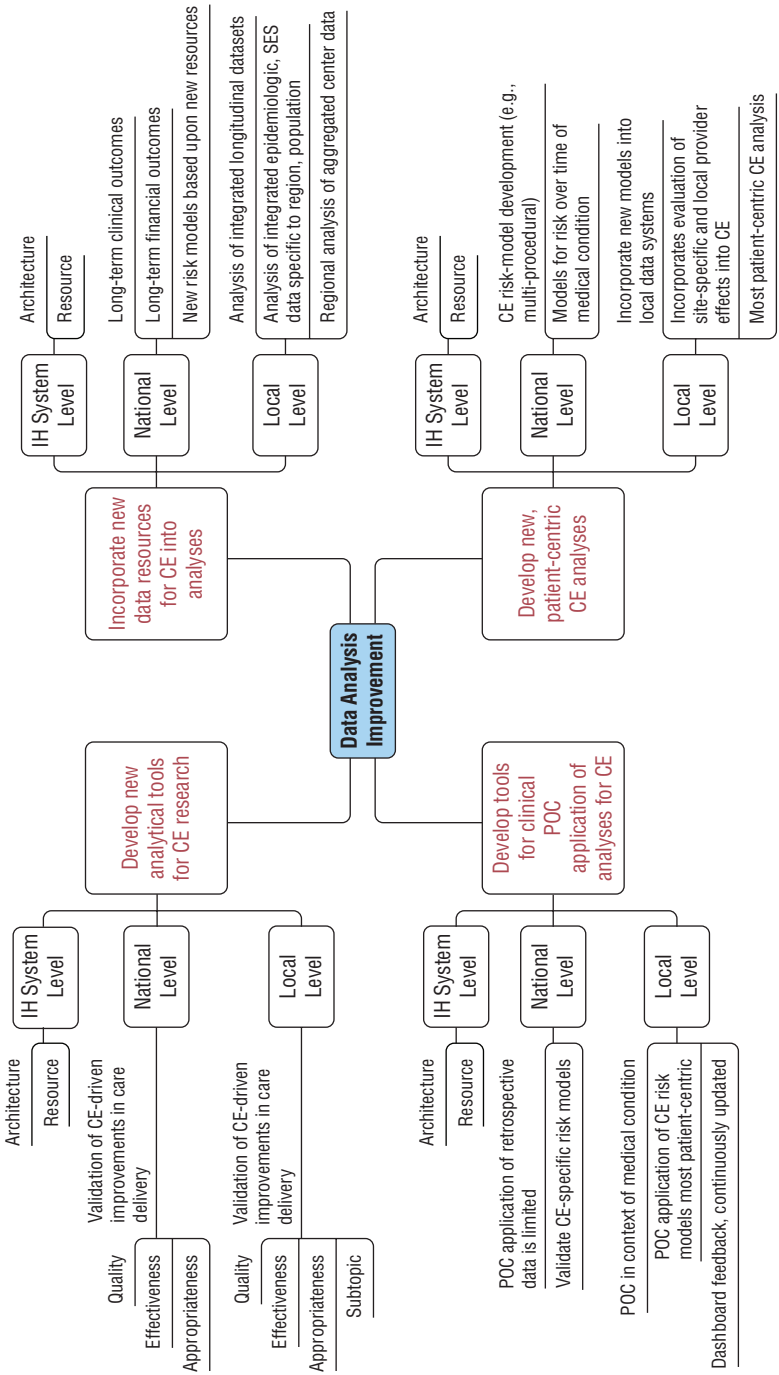
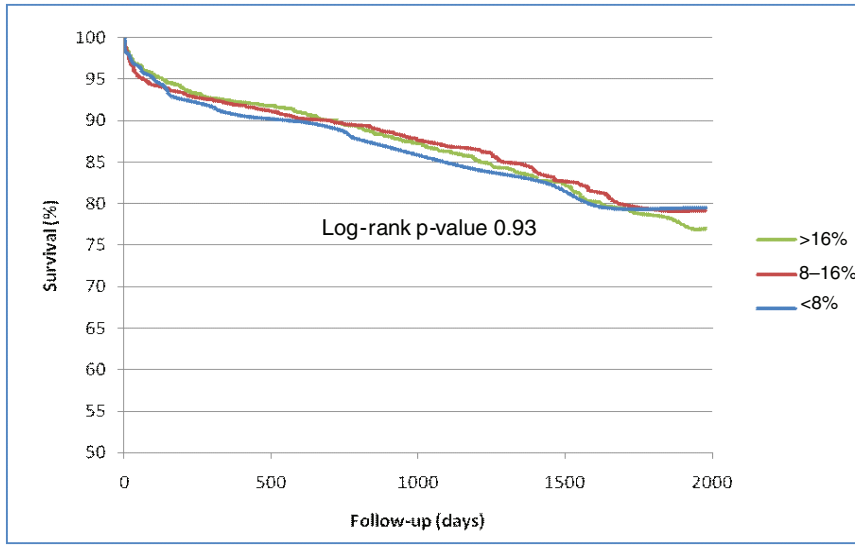


FIGURE 5-4 A similar relational map suggesting major areas of data analysis improvement for comparative effectiveness.
NOTE: CE = comparative effectiveness; IH = International Health; POC = point of care; SES = socioeconomic status.



Covariate	HR (95% CI)
>16%	0.85 (0.57, 1.28)
8–16%	0.90 (0.60, 1.37)
<8%	1.00

FIGURE 5-5 An analysis of survival following adult cardiac surgery from a single institution in a patient population operated on between July 2002 and July 2007 is shown. Periprocedural clinical data from the Society of Thoracic Surgeons National Cardiac Database, Social Security Administration Death Index data for long-term mortality outcomes, and U.S. census data for socioeconomic status based on the ZIP code in which the patients resided were linked at the patient level. The covariates were separated by the percent of population with a ZIP code at or below the poverty line from all ZIP codes within eastern North Carolina. This relatively simple analysis highlights the ability to integrate data at the local level.
NOTE: CI = confidence interval, HR = hazard ratio.

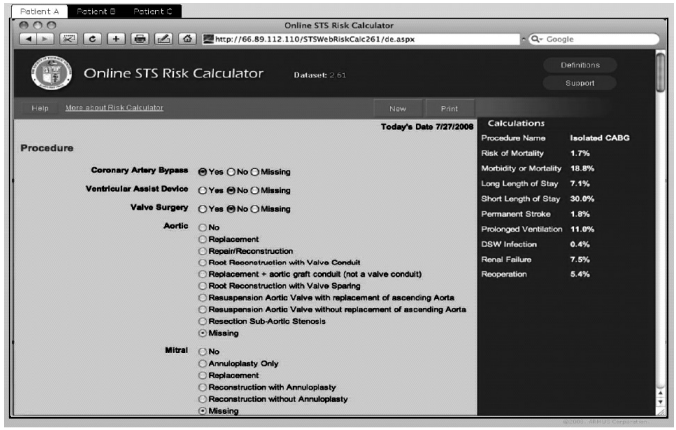
At the local level, the challenge will be to incorporate these new comparative effectiveness risk models into local data systems. Importantly, this local analysis capability incorporates site-specific and local provider effects into the comparative effectiveness dialogue between the patient and his or her providers, a key component to informed patient choice (Weinstein et al., 2007) and shared decision making (King and Moulton, 2006).

The next step in this analysis improvement, then, is to develop tools for clinical point-of-care application of comparative analysis. The argument for this can be distilled as follows: for comparative effectiveness analyses to substantively affect the quality of care, they (1) must encompass preference- and supply-sensitive care practices, (2) must be available at the point of care, and (3) must be usable for multidisciplinary decision making prior to selecting the best therapeutic option for that patient. These structure (multidisciplinary approach) and process (comparative effectiveness risk models) evolutions will drive the comparative effectiveness process one step closer to true patient-centricity. This in turn creates an absolute requirement to move beyond the retrospective analysis structure used for current analyses of both clinical and administrative data sets. To accomplish this, the analysis engine needs to be embedded in the meta-layer architecture of the data repository, and a selected portfolio of straightforward but useful clinical comparative effectiveness analyses must be continuously generated and available for review in a dashboard model (Figure 5-6). By design these analyses are focused at the comparative effectiveness level and are not structured to compete with or replace the larger, robust data set analyses.

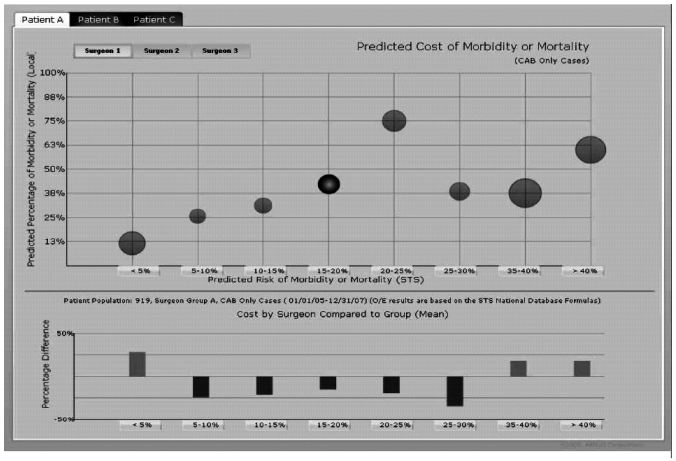
Importantly, the implementation of this approach at the local level allows for point-of-care application of these analyses in the context of the patient's medical condition. This again brings the comparative effectiveness analysis one step closer to patient-centricity. Continuously updated

FIGURE 5-6 Panel A shows a snapshot from the Society of Thoracic Surgeons (STS) Web-based online risk calculator and the data that can be generated based on this national analysis. Panels B and C show how national-level information can be brought to the local level by comparing predicted risk of morbidity or mortality at this level, influenced by site- and surgeon-specific variables, with this national risk assessment. In addition, the predicted cost of these major outcomes is illustrated. Panel B shows this analysis for surgeon #1, while Panel C shows the same analysis for surgeon #3, both of whom operate at the same institution. Panels B and C show somewhat different predicted outcomes for different categories of patient risks (gray circles), and the size of the gray circle represents cost. The dark circle highlights data for patient A. Cost per surgeon data for each of these risk categories are shown on each lower panel. All of the information on the slide is based on real clinical and financial data that have been merged together; online these three panels constitute a portion of a dynamic dashboard.
SOURCE: Analysis and presentation courtesy of G. Sziraczky, ARMUS Corporation.

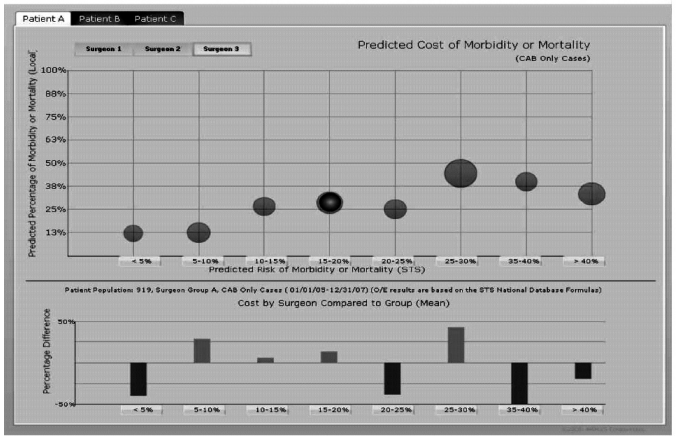
A



B



C



dashboard feedback gives providers the tools they need at the point of care to make therapeutic decisions prospectively based on comparative effectiveness (Figure 5-7).

Perhaps the last step in data analysis improvement is to develop new analytical tools for CER. This will be necessary as the new data resource developments get incorporated into clinical effectiveness studies. Health policy and outcomes research will need to validate the short- and long-term value of comparative effectiveness analyses in driving change in care delivery. At both the national and local levels, these analytical tools will affect quality, effectiveness, appropriateness, and, at the local level, efficiency of care. The research agenda will need to document and validate how the learning health system component based on comparative effectiveness affects these outcomes on an ongoing basis.

Conclusions

A broad array of infrastructure development must occur to transition to a learning health system. Critical to this effort will be the data resource development and data analysis improvement issues addressed here. Three key principles can perhaps be established based upon this analysis. The first is that this resource development and analysis improvement process must translate into infrastructure that is appropriate for dynamic, real-time availability for learning. This in turn will require the incorporation of a much broader array of data resources into the learning infrastructure and in comparative effectiveness studies than has been used in the past. Better definition of the type, source, and use of these data resources is needed, along with public–private partnership necessary to create this scope of data resources and infrastructure necessary for comparative effectiveness work. The second principle is that real-time learning will require feedback processes to be built into the research development and analysis improvement strategies. This includes data and analysis feedback to all major stakeholders, in part as a return on their investment into the infrastructure development for comparative effectiveness. The third principle is that comparative effectiveness is at its optimal usefulness when applied in a patient-centric focus at the point of care. Tools that foster real-time analysis will be an important development. These tools will be embedded in these data resources to allow real-time insights into care delivery and will be used during a shared decision-making process prior to selecting the optimal therapeutic option for that specific patient.

In aggregate, much progress has been made already. Addressing the components of data resource development and data analysis improvement outlined here will further move the agenda forward to meet the 2020 goal.

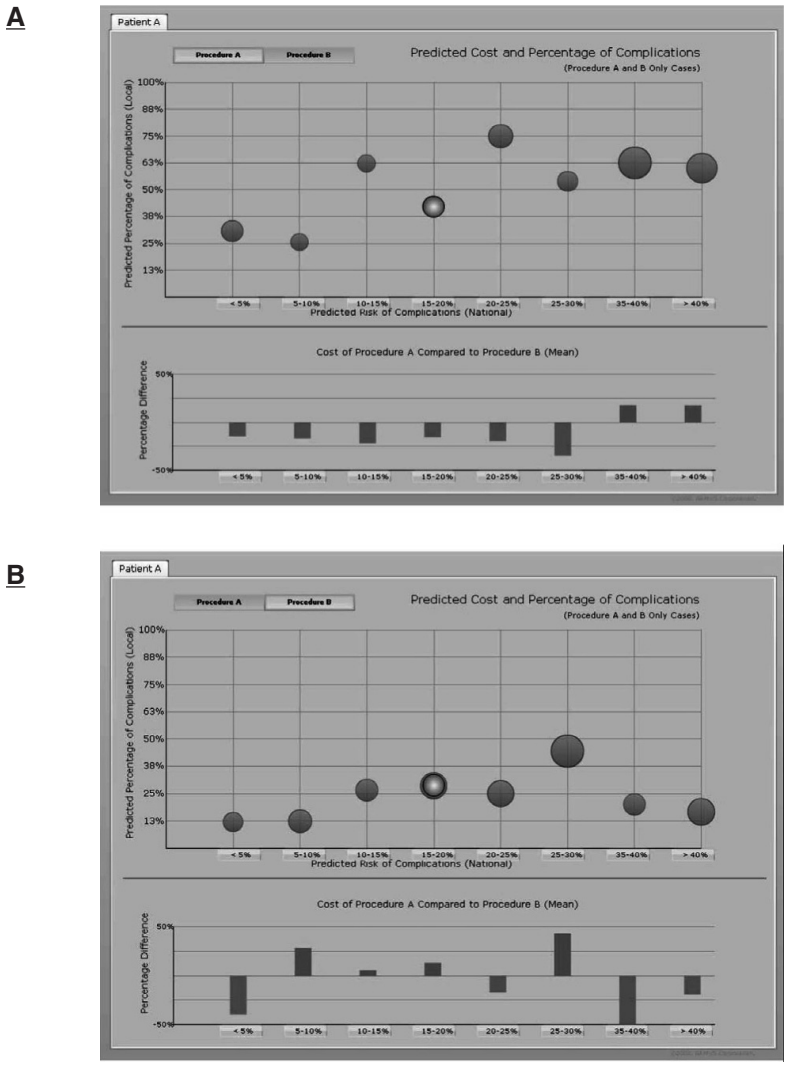


FIGURE 5-7 The dashboard concept is extended to illustrate what a potential point-of-care comparative effectiveness dashboard might look like. Based on randomized controlled trial and application data, procedures A and B are felt to be at clinical equipoise in terms of therapeutic benefit overall. This dashboard brings the comparative effectiveness analysis down to the level of patient A. In this hypothetical example, procedure B would be the option of choice for patient A based both on percentage of complications and on cost compared to procedure A.
SOURCE: Analysis and presentation courtesy of G. Sziraczky, ARMUS Corporation.

PRACTICAL CHALLENGES AND INFRASTRUCTURE PRIORITIES FOR COMPARATIVE EFFECTIVENESS RESEARCH

Daniel E. Ford, M.D., M.P.H., Vice Dean for Clinical Investigation, Johns Hopkins School of Medicine

Overview

Myriad challenges associated with the conduct of effectiveness clinical trials—particularly RCTs in practice settings—will likely reduce the immediate impact of any expanded funding for comparative effectiveness studies. This type of research is very difficult to do, and, as compared with analyzing existing data, the amount of effort it takes to complete an RCT of effectiveness in multiple practice sites represents a quantum difference in effort, time, and resources. This paper will briefly describe some of the key challenges to the efficiency of clinical research based on observations from the experience of the average investigator and of patients being asked to join a research study, it will reflect on some of the additional challenges associated with CER, and it will offer some suggestions on priorities for research infrastructure improvement that might help to improve the overall efficiency of clinical research.

Practical Challenges

Johns Hopkins University's current clinical research activities include approximately 3,500 active protocols, roughly 1,000 new protocols a year, and about 700 protocols using investigational drugs. Johns Hopkins is one of several academic medical centers with this high volume of clinical research. The National Institutes of Health (NIH) created the Clinical and Translational Science Awards (CTSA) program as one way for academic centers to consolidate existing funding and to add new funding to promote the quality and efficiency of translational research. At Johns Hopkins we have created the Institute of Clinical and Translational Research with the goal of "connecting science to people." Through our CTSA program, Johns Hopkins is collecting the data to transform the clinical research enterprise—to help define and promote what it takes with respect to manpower, efficiency, and improving the value of research.

Over the past 10 years, FDA-regulated clinical trials have precipitously moved from U.S. academic centers to community hospitals and emerging international centers. Reasons include the increasing length of time it takes to have study protocols activated and the ability to recruit participants. The length of time it takes to meet recruitment goals in many studies is also a concern. For example, the average time from the first application propos-

ing a clinical trial to being ready to enroll the first patient in the National Cancer Institute Eastern Cooperative Oncology Group network is now about 800 days for phase 2 and 3 trials (Dilts et al., 2008). These barriers appear to be greater in the United States than in emerging countries, which are in addition to the higher costs of conducting research in the United States. Efficient completion of clinical studies is especially important for CER. These studies lose their value if the practice environment changes, the costs of the interventions change (e.g., medication becomes generic), or new treatments become the de facto standard of care. Because of differences in health status, previous use of treatments, and context of healthcare delivery, comparative effectiveness studies conducted outside of the United States may not be generalizable to the United States without some careful analysis.

Why are individuals in the United States less likely to enroll in human subjects research than in some other countries? There is no one reason, and evidence is generally lacking. Most patients in the United States have health insurance and do not need to enroll in clinical studies to receive treatment for their medical conditions. The voluntary nature of joining a research study is emphasized in the United States. Consent forms are becoming longer and addressing more contingencies. One issue specific to the United States is the multiple insurance carriers that pay for care for Americans. There is no standard approach to how these insurance carriers view support for clinical research. This leads to confusion for Americans in several aspects. AHRQ has just posted a technology assessment report called "To What Extent Do Changes in Third-Party Payment Affect Clinical Trials and the Evidence Base?" (AHRQ, 2009). The review finds that there is very limited evidence available to determine if health insurance policies affect enrollment in clinical trials. Interviews with investigators found that device studies were more susceptible to problems with enrollment based on insurance status. For this discussion the statement on clinical effectiveness studies is important:

> For later-phase research—notably comparative effectiveness of existing therapies and studies of off-label uses of approved therapies—the impact of payment policy may be greater, but is not well defined. No entity assumes full responsibility for research costs, and plans to co-share expenses are in their infancy. Thus, in areas lacking sufficient evidence, especially regarding products that are already on the market, there is currently no consensus on who should pay for the evidence-generating research. (AHRQ, 2009)

Since evidence concerning insurance barriers to participating in clinical research is limited, I will discuss some examples that would be particularly

central to many comparative effectiveness studies. Let us consider a common comparative effectiveness study that compares two approved treatments that are covered by most insurers. Patients would have the following choice: They could choose one of the approved treatments that are already covered by their health insurance. If they went this direction they would know what their copayments would be and would not be concerned that any toxicity resulting from the treatment might not be covered by their health insurance. On the other hand, let us suppose they were asked to consider enrolling in a research study comparing the treatment they choose to another equally qualified treatment. They might have questions about the costs of treatment at the time of consent and generally would be told to call their insurance carriers. If they called their insurance carriers to discuss the financial implications of enrolling in a study, it is very likely they would be told their insurance does not cover research. This simple answer is technically correct, but it is misleading. In general, the study budget has been set up in such a way that the true research component of the study is paid for by the sponsor of the study. The expectation is that the health insurer will still pay for whatever treatment is provided in the trial that they would have paid for even if there was no clinical trial. However, in CER it is possible that new models of funding research studies will be created. If the study is comparing a cheap, generic medication to a new expensive medication, who will pay for the study medications? Will the medication copayment be standardized for all research participants in the protocol or allowed to vary depending on the insurance carrier? If the study team has to negotiate this with every insurance carrier a patient might have, the study will become too cumbersome and slow to be of much value. If both medications are provided free of charge this would not reflect the effect of copayments on adherence in the real world. Even if the study protocol were designed to allow standard copayments, it would be difficult to provide accurate information regarding the financial responsibilities to a potential research participant. The copayment may depend on the time of year when the patient is joining the study and whether limits for copayments have been reached at some time during the study.

Another important issue to research participants is who would pay if there are adverse events or toxicities related to a research intervention. It is important to note that when academic centers agree to conduct studies funded by commercial sponsors, the sponsors are generally required to pay for any adverse events associated with the study interventions. In contrast, studies funded by the NIH do not have any mechanism to pay for injuries related to the study intervention. Most consent forms for federally funded studies include a statement that you or your insurance company is responsible for any injuries that result from the study intervention. It is no wonder that patients in the United States would have second thoughts

about taking on the risks associated with joining studies. In the typical phase 1 study, patients may be willing to take on more risks to get randomized to a research intervention that is not currently broadly available but may represent the last chance for them. A study comparing two currently available treatments may not have the same perceived advantage to make the risks acceptable.

What can be done to simplify the financial implications of joining a comparative effectiveness study? There needs to be some way to create a model for research support that reduces the number of insurers with which a study team must interact. The CMS Clinical Trial Policy based on the 2000 National Coverage Decision has been useful in simplifying the process, and an important first step would be if all insurance companies agreed to follow this policy. The policy allows a single coverage decision so that the people who join a trial do not need to get clearance from their individual insurance carriers. Studies that have had formal peer review by a federal agency are considered to have scientific value, and CMS agrees to cover services normally delivered for that clinical condition. For example, while not paying for an investigational drug, CMS will pay for the *administration* of this investigational drug. Finally, CMS agrees to pay for evaluation of the toxicities associated with an investigational intervention.

Our experience at Johns Hopkins is that an increasing number of insurance companies do cover the associated costs in clinical trials when they are contacted by experienced staff from our insurance coverage office. However, approximately 15 percent of the patients who have already agreed to join a clinical trial are not cleared to join by their insurers. In most cases this is not the required policy of the insurer, but instead a decision by the individual's employer to not cover participation in a clinical trial. National data related to coverage of clinical trials by insurers would be valuable, but they are not easily available.

We need to make sure that patients, researchers, healthcare providers, healthcare systems, insurers, and study sponsors are all enthusiastic about participating in CER. Without the support of each group, CER is unlikely to reach its promise of informing patients and providers about best practices at the time they need to make their decisions.

Infrastructure Priorities

1. Process in Place for Getting Timely Consultation from All Stakeholders

The results of CER have implications for patients, patient families, healthcare providers, payers, and the manufacturers of healthcare treatments. While it is important to seek the perspectives of all stakeholders before designing and interpreting comparative effectiveness studies, the pro-

cess must be streamlined. The added value of each additional consultation as the study design is finalized should be measured. Ideally, representative standing panels should be available for timely consultation.

2. Streamline Initiation of Study Through Both Institutional Review Boards and Contracting Mechanisms

Institutional review boards (IRBs) and their required reviews are frequently cited as a barrier to timely starts of studies. Multicenter trials may need to get approval from multiple IRBs. In response to these delays, many have suggested that central IRBs have advantages. Central IRBs can be more efficient in creating the consent form and efficient initial review. However, protection of human subjects is much more than crafting the consent form. Local oversight is needed for the more frequent issues related to human subjects protection, including training and monitoring of competence of research teams, proper consenting of participants, timely recognition of adverse events, and supervising investigational drug services. IRBs are responsible for the conduct of the entire study, not just approving the consent form. Supervision by a central IRB that has little ability to monitor and implement local corrective action plans does not seem as desirable as a local IRB. On the other hand, IRBs have to be given appropriate resources to run in an efficient manner. At Johns Hopkins there are now five separate IRBs that all meet on a weekly basis, supported by an electronic IRB application and tracking system. Each IRB member is paid to serve on the IRB, and there are 25 additional staff to support the IRBs. With this level of support, the IRB can provide quality reviews with most approvals coming in less than 30 days.

Common and expected issues related to IRBs and contracting review should be examined and policies created. For example, institutions should have uniform policies related to when they will allow a practice to be covered by their IRB. The requirements for training and supervision should not be invented for each study. Contracts should have standard policies related to indemnification and collection of biospecimens in study protocols.

3. Standard Policy on Insurer's Coverage of Services for Individuals in Clinical Trials

As described above, patients need to be confident that there will be no financial penalty if they receive their care in the context of a clinical trial. The easiest solution would be for all insurers to accept the policies developed by CMS. This would eliminate the need for getting the insurer's approval before the participant enrolls in the study. The approval step is one more barrier for patients deciding if they want to enter a clinical trial.

4. Enhancing the Research Capability in Hospitals and Practices Outside the Academic Center

CER includes observational studies, evidence synthesis, and RCTs. A priority for supporting practice-based clinical trials is increasing the capacity of practice-based research networks. There are now over 100 practice-based research networks, but many do not have all of the components to complete the research efficiently. At Johns Hopkins we have started the Johns Hopkins Clinical Research Network. Sponsors using this network will be assured that all staff are trained to Johns Hopkins standards, that a single point for contracting is possible, that only one IRB will complete the primary review with timely communication to other IRBs, and that Johns Hopkins Web-based research IT will be accessible throughout the practices in the network.

Practice-based research networks do not need buildings or equipment to function. They do need stable funding for the people who can organize and enhance communication between the practice-based practitioners and researchers. Practice-based research networks need funds to create the contractual agreements and quality assessments that allow research to be conducted efficiently. While accountability for productivity is necessary, the costs associated with recruiting multiple practices anew for each study are considerable. Small amounts of funding to sustain practice networks would be extremely valuable as they would help minimize the costs of recruiting and training practices.

5. Stronger Partnerships Between Researchers and Healthcare Systems

For some research questions, the best approach is to randomize by provider or healthcare setting. Healthcare systems need to more carefully consider the possibility of randomization as they roll out new programs. At the same time, researchers have to realize that healthcare organizations have their own timelines and cannot wait indefinitely before they begin implementation of new programs. Delays in initiating studies caused by the need for multiple submissions for funding are particularly damaging for clinical effectiveness studies. More rapid grant cycles may be needed to increase the likelihood that healthcare organizations and researchers are able to work together on more rigorous evaluation of new healthcare interventions.

6. Need for More Research Staff

While the focus is often on the principal investigators when discussing research personnel capabilities, research now requires a team much larger

than just the principal investigator. For CER, research coordinators are needed who are experts in recruiting and retaining research participants. IT professionals with expertise in Web-based data entry and tracking systems for community practices are also in short supply. Analysts who are expert in preparing large administrative data sets and assisting with statistical analysis are in short supply. Another need is for biostatisticians who are expert in analysis of cluster RCT designs and sophisticated methods for assessing causal relationships from observational studies.

Concluding Observations

Efficient, valuable CER requires the enthusiastic support of multiple stakeholders including patients, healthcare providers, healthcare plans, and the research community. Unfortunately, if one of these stakeholders has limited participation, the study will not progress, and the value of the research will be limited. CER infrastructure needs to provide long-term support so that research becomes a common occurrence in the delivery of care in the United States.

TRANSFORMING HEALTH PROFESSIONS EDUCATION

Benjamin K. Chu, M.D., M.P.H., President,
Kaiser Foundation Health Plan and Hospitals
Southern California Region

Overview

In the health professions we enter our respective fields because we want to improve the lives of our patients. We are taught first to do no harm. We are trained to use our knowledge and our devotion to lifelong learning to relieve suffering and improve the health of those we care for. We hold sacred our duty and our responsibility to our patients.

As professionals, we try to learn from the experience of those who preceded us. We apply their knowledge and their experience about what treatments have been effective for our patients. Nonetheless, even with this ingrained dedication to the principles of professionalism and the years of devoted study of existing knowledge, much of how we practice medicine is determined by the finite, cumulative set of experiences we gather from one-on-one interactions with our patients. Furthermore, on a community level, our health system operates on the principle that the sum of all those one-on-one interactions with well-trained professionals will lead to a healthier community.

Critical reviews of the performance of our health system against many

measures of community as well as individual health outcomes point out many gaps (Schoen et al., 2006). Few informed observers would disagree that the U.S. healthcare system is experiencing a profound crisis characterized by skyrocketing costs; inconsistent, suboptimal care; and decreasing access to care (Crosson, 2005). We spend significantly more on health care per capita than other industrialized countries, yet adults in the United States receive only about half of recommended care (McGlynn et al., 2003). In its landmark *Crossing the Quality Chasm* report (IOM, 2001), the IOM reported that "the current systems cannot do the job. Trying harder will not work. Changing systems of care will." The IOM report envisions a delivery system capable of delivering care that is safe, effective, patient-centered, timely, efficient, and equitable while meeting six challenges:

1. evidence-based care processes;
2. effective uses of IT;
3. knowledge and skills management;
4. development of effective teams;
5. coordination of care across patient conditions, services, and settings over time; and
6. use of performance and outcome measurement for continuous quality improvement and accountability.

This IOM Roundtable on Value & Science-Driven Health Care is focused on identifying the infrastructure changes needed to help the nation fulfill this vision. While there are clearly gaps in key arenas critical to supporting a learning health system that is driven toward achieving the best health outcomes for patients, the participants in the workshop also pointed to steady progress. Substantial investments in IT with decision support and patient care registry capabilities, international progress on clinical systematic reviews, and developing experience in using these and other exciting new consumer-oriented Web-based tools that build on social networking capabilities illustrate a dynamic healthcare environment striving to put in place the key elements for success. It is also a system undergoing intense scrutiny of the development and reporting of objective measures that can be used to define progress and success.

Transforming Health Professions Education

Transforming health professions education is less about training "informationalists"—comparative effectiveness and health services researchers and data analysts—than it is about creating environments for training that encourage the effective use of these new tools by teams of physicians and other health professionals in order to achieve the best outcomes across the

full continuum of care and over the lifetimes of our patients. It should not be a surprise that a system of training emphasizing individual responsibility and professionalism but without systematic tools to verify effectiveness would result in inconsistent performance. Without these tools, there is simply no way to know and no way to systematically approach addressing gaps in care. Measuring outcomes with the tools to track and evaluate strategies to achieve better outcomes is now possible. In moving to address gaps in the infrastructure for evidence-based medical practice, perhaps a more important issue is to address the motivation and capability of the delivery system to use this infrastructure to achieve best performance. To facilitate diffusion of evidence-based practices it will be important to (1) establish clear expectations for high performance along clear and measurable dimensions of care, (2) encourage adoption of appropriate IT tools that provide essential information to drive performance improvement, and (3) align our payment systems to value better outcomes.

A health system driven toward achieving best outcomes for patients without the often conflicting goal to maximize income should be one that demands and supports a robust infrastructure to optimize care using the best available evidence. Organizing to meet these expectations will force health systems to address structural gaps, including reexamining the roles and responsibilities of the range of health professionals in a more common goal of performance excellence.

Setting Expectations

The Commonwealth Fund has devoted a considerable portion of its efforts to defining and advocating for a "high-performing health system" through the work of its Commission on a High-Performing Health System. Its report cards define gaps in system-level performance. Its systematic efforts to highlight the attributes and accomplishments of health systems that strive for and achieve high performance sets a clear bar for the U.S. healthcare system (The Commonwealth Fund, 2009). The Institute for Healthcare Improvement's clarion call to save 100,000 lives and reduce harm for 5 million lives combined with evidence-based tools and collaborative efforts to help achieve these goals clearly set benchmark expectations for high performance (McCannon, 2007). Collaborative efforts to improve ambulatory care outcomes (Landon et al., 2007), the CMS/Premier Hospital Quality Demonstration Project (Premier and CMS, 2007), and a variety of pay-for-performance efforts also set high expectations for high-quality outcomes in addition to helping to define successful practices to achieve high performance. Public reporting of results on larger numbers of increasingly relevant health outcomes combined with continued efforts to demonstrate successful strategies and practices should create an environment where sys-

tematic application of the evidence base in care delivery should become the norm rather than the exception. Well-informed and self-advocating patients using robust Web-based resources could accelerate this transformation.

Adopting the Right Information Systems

Reconfiguring our health system to use the evidence to optimize outcome is nearly impossible in the world of paper records. For many, the hope for better performance lies with widespread adoption of EHRs. But as others in this Roundtable have pointed out, adoption of EHRs alone is not sufficient. Without the development and use of evidence-based decision-support instruments, registries, panel management, and other tools combined with the attitudinal, cultural, and process changes that are necessary to use them effectively in our healthcare system, the challenge cannot be met.

A high-performing 21st-century healthcare system will require coordination of care among many care components. Interoperable HIT with a full suite of evidence-based decision support tools, care registries, and panel and population analytic capabilities are key drivers for better outcomes. Adoption rates for HIT have progressed slowly and unevenly among hospitals and health centers (AHA, 2007; National Health Policy Forum, 2008). Getting to a high-performing 21st-century health system will require training future health professionals in environments that use these tools. Training programs have an obligation to create a training environment that models the best care possible. Health professional schools, academic health centers, and health professions accreditation bodies should define minimum standards for HIT needed to support a high-performing health system. A timetable for obtaining this core infrastructure should be established. Consideration should be given to a suggestion raised in this symposium that funding for health professions education, particularly Medicare graduate medical education payments (or continued funding), be explicitly tied towards helping training programs gain access to these tools. Special funding might be needed to help safety net training sites gain access to these tools.

Training program certification should also be increasingly tied to health systems that can demonstrate effective use of evidence-based medicine (EBM), teamwork, and continuous learning to push for measurable outcomes of quality care. In-patient sites should demonstrate robust performance-driven programs to improve patient safety. Primary care training should clearly allow for trainees to deliver team-based, data- and evidence-informed care in a setting that fosters coordination of care, preventive care, and optimal management of chronic conditions.

Payment Reform

Fee for service as the prevailing model for payment of health care concentrates effort on single units of interactions, often centered on an episode of illness. Payments are tied to visits, procedures, tests, or some other unit of care. Payments are not directly tied to desired or reasonable outcome. Complications of care are reimbursed as additional necessary units of service while care coordination and case management are assumed but not specifically reimbursed. As a consequence, financial incentives in our health system support more units of care regardless of the firmness of the evidence base for that care. It does not necessarily support better coordinated or managed care.

To put health care on solid evidence-based footing, financial incentives need to be aligned with reasonable expectation of the best possible outcomes. Medicare has already pushed to reduce or eliminate payments for errors in management and avoidable complications. The Medicare Payment Advisory Commission has proposed experimenting with bundled payments for an "episode of care" (MedPAC, 2008). Others have proposed evidence-informed case rate methodologies to bundle payments for the management of illnesses. Such bundled payments would include built-in payments adjusted for complexity and some fixed level of complications to encourage better care (de Brantes and Rastogi, 2008). Risk-adjusted comprehensive payments have been suggested as a way to encourage comprehensive, multidisciplinary, and well-coordinated primary care (Goroll et al., 2007). And, of course, full-risk capitation has been the financial model for the fully integrated healthcare system at Kaiser Permanente.

Needed Systems Changes

The practice of medicine is complex and has become increasingly so. There is an explosion of medical knowledge, specialization, and sophisticated procedures that can yield remarkable results, an overall shift of illnesses from acute illnesses to more difficult to manage chronic illnesses, and the welcome proliferation of effective preventive care strategies. A recent *New England Journal of Medicine* report reviewed the complexity of coordinating care throughout our nation's healthcare systems and noted that "it would take a physician 7.4 hours per working day to provide all recommended preventive services to a typical patient panel, plus 10.6 hours per day to provide high-quality long-term care." Meeting these tasks as well as coordinating care between providers and ensuring seamless transitions in care from one setting to another will require increasing dependence on multidisciplinary teams, better and perhaps larger organizations of health

services depending on interoperable EHRs, and robust decision support and panel/population care management tools (Bodenheimer, 2008).

There has been a great deal of attention paid to the concept of a medical home and its ability to improve care for all (Davis and Schoenbaum, 2007). Broadly defined, a medical home is "a physician-directed practice that provides care that is 'accessible, continuous, comprehensive, and coordinated and delivered in the context of family and community'" (AAP/ACP/AOA, 2007). Other features of this model include care coordination and integration facilitated by registries and HIT, the use of evidence-based decision support, and engagement in QI activities (Rittenhouse et al., 2008). Several studies suggest that the use of medical homes leads to improvements in care and decreased resource use (Arvantes, 2007; Paulus et al., 2008).

Team-based, evidence-informed processes have been at the core of the success of the patient safety and QI efforts spurred by collaborations sponsored by the Institute for Healthcare Improvement, the Joint Commission, and the Premier Hospital Quality Demonstration Project, among others. Interdisciplinary accountability for adherence to evidence-based protocols, bundles of safety practices, and checklists to ensure reliability serve to bolster these efforts and have yielded encouraging results. Care transitions are also amenable to team efforts to help patients avoid complications and rehospitalizations (Coleman, 2006).

Evidence-Based Medicine at Kaiser Permanente

At Kaiser Permanente, we embrace the expectation of high-quality performance on behalf of our patients. We now have real-time information on quality processes and outcomes that can serve as a guide and a measure of effective improvement efforts. We also operate under a financial capitation model that encourages better outcomes. These conditions have encouraged changes in our delivery system that have implications for changing roles for healthcare personnel. Here is an example.

Like a number of large health systems across the country, Kaiser Permanente has invested heavily in IT to give the system the full capabilities of a high-performing, learning health system. The company is approaching the final stages of implementing an EHR with Web-based capabilities that tremendously facilitate communication of information with patients and among a host of health professionals both in outpatient and inpatient settings. We have developed sophisticated registries of patients suffering from a variety of conditions and have deployed evidence-based decision-support tools and panel and population management aids. Not surprisingly, these tools highlight variations in practice and outcomes even in a system that has prided itself on adherence to best-practice protocols for many years. In a world of paper charts where the unit of clinical activity

is the patient encounter, we simply did not know what we did not know. Now, tools that allow our system to track performance down to the practitioner level have given us the ability to tailor efforts to achieve better outcomes. Often, this has required remodeling care delivery.

For example, as we developed the capacity to track a portfolio of preventive interventions and chronic disease control measures for our population of almost 3.3 million people in Southern California against evidence-supported standards, we steadily and dramatically improved, but then we reached a performance plateau. As a primary care physician-based system we depended on our increasingly harried primary care physicians to achieve the desired outcomes. Computer-generated outreach reminders and a host of other systemic strategies continued to fall short of our expected goals.

Using our electronic database and population care management tools, we looked at our patients who had large gaps in over 10 key measures covering age-appropriate preventive care and chronic disease management. To our surprise, in this population of patients who were suboptimally managed according to our guidelines, three out of five received their care at Kaiser not through their primary care physicians but through a host of specialty interactions. In any given year a primary care physician could at most address the gaps in only 40 percent of these patients. To reach the other 60 percent would require a systemwide effort that would have to involve specialty areas. Further, we realized that depending solely on physicians would not solve the problem.

As a result, in Southern California we have launched an ambitious redesign of our ambulatory care setting, emphasizing a "proactive office encounter." This effort involves every member of the healthcare team in both primary care and specialty areas to address these gaps in care. Gaps in recommended care are identified for patients weeks in advance of their scheduled visits. Receptionists and professional and ancillary staff attempt to address these gaps before, during, and after the visit. In addition, physicians are encouraged to work with staff to address gaps identified through our panel management tools. Frontline staff enthusiasm and accountability for helping patients achieve better outcomes has been noticeable in the past year as performance has rapidly improved.

Our experience in the inpatient setting is similar to other organizations that have embraced the challenge to improve patient safety in hospitals. To dramatically reduce infections, to prevent falls and avoidable complications, and to ensure optimal management of patients who present with serious illnesses requires the same team focus and devotion to systematic adherence to proven protocols and bundles of care. After reviewing the results for the use of team-based simulation training in high-risk areas, Kaiser Permanente has begun to roll out a comprehensive effort using computerized human

patient simulators with some encouraging results (Draycott et al., 2006). Further, to optimize care after hospitalization and to prevent unnecessary rehospitalizations, it has become increasingly clear that managing the transitions in care is a key component. Intensive case management, home health, home monitoring, and other approaches have helped us to reduce rehospitalization rates for chronic illnesses.

Finally, with the availability of our database tools and analytic capabilities, we have expanded our research unit to capitalize on the available information for observational studies of care and to conduct health services research. The goal is to tap into the richness of this information to develop even more effective strategies.

Conclusion

In summary, health care is moving toward a patient-centered, evidence-based health management orientation. Computerization of health records, wider use of patient care registries, greater availability of tools that allow for tracking individuals as well as populations of patients, and information-savvy consumers will drive our current fragmented health system toward one that will emphasize greater accountability, transparency of information, and higher levels of performance. Computer-assisted tools with sophisticated evidenced-based decision-support protocols combined with process changes and strict adherence to demonstrated, cost-effective bundles of care can lead to safer and better care. Gaps in preventive care and chronic disease management can be easily tracked. To correct gaps in care and to ensure safe and effective interventions, physicians and other health professionals will increasingly have to work together in teams of care and share accountability for their patients' clinical outcomes. Acute episodes of illness will require coordination of handoffs, patient safety protocols and checklists, and other interventions designed to minimize harm and maximize benefit to our patients. Chronic disease management and adherence to known effective preventive measures will become systemwide accountability requirements. The complexity of care and the huge burden placed on shorter and shorter physician–patient interactions with a multitude of different clinicians will require that other health professionals and ancillary staff be used to bridge the gaps. Every touch point, enhanced with Web-based and other communication tools, will be an opportunity to maximize care.

Health professionals in a high-performing health system will rely on a new professionalism that will build on the principles of lifelong learning, duty to our patients, and devotion to finding best outcomes for them, but they will also emphasize an obligation to optimize teamwork and ensure that care is firmly grounded on the best evidence of effectiveness. Health professionals learn through experience in taking care of patients. To create

a healthcare workforce with the skills to expertly use EBM to achieve high levels of population health, it will be necessary to create the settings to allow trainees to emulate and improve on the best models. Healthcare settings that strive for high levels of performance will inevitably move toward these more effective team models of care.

BUILDING THE TRAINING CAPACITY FOR A HEALTH RESEARCH WORKFORCE OF THE FUTURE

Steven A. Wartman, M.D., Ph.D., M.A.C.P., President, Association of Academic Health Centers; and
Claire Pomeroy, M.D., M.B.A., Vice Chancellor, Human Health Sciences, School of Medicine, University of California at Davis

Overview

> "Institutions must transcend traditional boundaries to generate new ideas and technologies. . . . And link science with policy and governance to frame questions and foster social change." (Bawa et al., 2008)
>
> "If medicine is to fulfill her great task, then she must enter the political and social life." (Rudolf Virchow, 19th-century physician)

A national consensus is emerging that the U.S. healthcare system needs a fundamental retooling as to its purpose and function. This conclusion is driven mainly by the expensive costs of the current system in the context of significant variations in the basic health parameters that reflect the well-being of American society. For example, the current healthcare delivery paradigm strongly incentivizes acute care and the provision of expensive, cutting-edge drugs and technologies to the insured. Missing in this model is a strong focus on facilitating access, incentivizing quality, prioritizing preventive care, and ensuring community-wide public health. When this is coupled with the large numbers of underinsured and uninsured citizens, the result is a fragmented and dysfunctional system that creates wide disparities amongst the population in terms of health and well-being (Wilper et al., 2008).

What is not often heard in the calls for change is the urgent need to develop a new kind of research infrastructure focused on health and health care that can guide and inform decision making during this time of needed and anticipated change in the health system. This requires the development of the evidence base for clinical practice in order to ensure that the health care delivered is both effective and optimal. The importance of enhancing and supporting CER is therefore critical and fundamental to a reformed

health system. Indeed, a recent IOM Roundtable workshop has emphasized the kind of infrastructure that is necessary in order to learn which care is best (IOM, 2008).

Revisioning the Medical Research Enterprise as a Necessary Tool to Implement Health System Change

Research holds the promise of testing and finding the answers to the challenges that face health care in the United States today. But the traditional approaches to the needed research are inadequate to discover and define the innovations needed. If improved health is to be provided for all Americans, a vision for a new kind of medical research is needed. Above all, this new kind of research seeks to discover, disseminate, and optimize the adoption of practices that advance the health of individuals and the public as a whole.

Key to this new paradigm is the principle of expanding the continuum of medical research to extend from basic discovery to community-wide health innovations in order to ensure that discoveries ultimately serve the public. Already there have been calls to reevaluate the emphasis and resource allocation afforded to the various types of research (Dougherty and Conway, 2008) (Figure 5-8).

Specifically, there has been increasing recognition that medical research investment must be expanded to support more applied research. For example, the creation of the Clinical and Translational Science Center (CTSC) program by the NIH was a dramatic call for transformation of the national medical research enterprise (CTSA, 2008).

However, this new, expanded vision of medical research remains incomplete. It is increasingly recognized that health is determined only in part by the actual delivery of health care; there are other important determinants of health that are much broader, including behavioral factors, genetic variability, and, perhaps most prominently, the social determinants of health (Marmot and Wilkinson, 2006; Tarlov and St. Peter, 2000) (Figure 5-9).

The emerging field of the social determinants of health emphasizes the fact that factors such as socioeconomic status, education, job security, access to societal resources, social support, and social empowerment are powerful determinants of the health status of a community, even more so than the specifics of healthcare payment and delivery systems. Indeed, it can be argued that the emphasis in the United States on a mostly biomedical research and healthcare delivery model—as opposed to one based more on social and environmental determinants—has contributed to less than optimal health statistics, growing disparities, and spiraling healthcare costs.

The importance of the social determinants of health is highlighted in

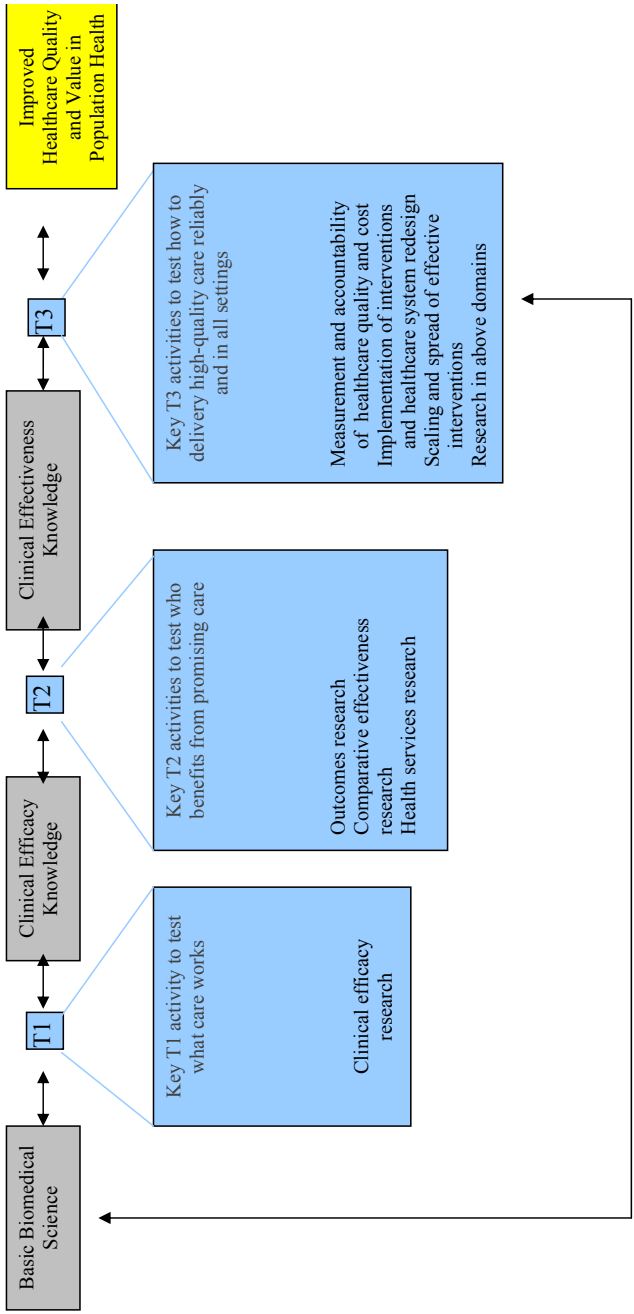


FIGURE 5-8 Expanding the research continuum.
SOURCE: Dougherty and Conway, 2008.

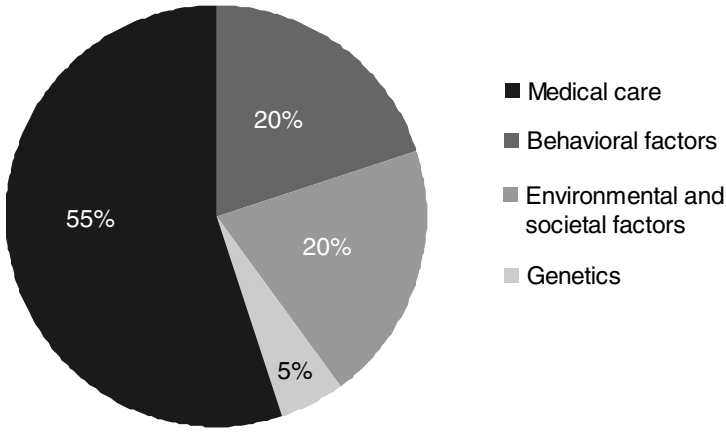


FIGURE 5-9 Determinants of health.
SOURCE: Marmot and Wilkinson, 2006.

the new "Healthiest Nation" campaign conducted by the CDC. Director Gerberding introduced the campaign by saying, "People are talking about healthcare reform, but they're not really talking about health." (Rubin, 2008). The recognition of this broader approach to health solutions is also embraced by the World Health Organization through its new Commission on the Social Determinants of Health (WHO, 2008).

Thus, although our national vision for medical research is appropriately expanding to include basic science, clinical, and T1–T3 translational research, a new, even broader view is needed. This broader view encompasses exploration of all the determinants of health, including those beyond the traditional realms of biomedical research. Let the term *health research* be used to describe studies that address all aspects of health, including biomedical research, public health research, and multidisciplinary research on the social and environmental determinants of health. This broader paradigm of *health research* is essential to ensuring that our research agenda leads to better health for all.

Approach to Achieving a New Research Vision

Responding to the urgent need for a new approach to research that explores all the facets of health requires fundamental changes in the research enterprise:

- First, this broader view of health research must be supported and facilitated. The fundamental science of translational and multidis-

ciplinary approaches must be established so as to be respected and valued in the overarching research and academic communities. That is, new disciplines, methodologies, and applications must strive to be perceived as equivalent in intellectual value and importance to basic biomedical science.

- Second, new resources must be allocated to provide funding and infrastructure for the new paradigm of health research, including expanded investment in basic science, clinical, translational, public health, and multidisciplinary social research.
- Third, new types of health research teams must be created; this workforce development will require training future researchers from a wide range of backgrounds and equipping them with new skills to conduct this research, as well as developing new partnerships and expanded venues in which to conduct health research.

The goal of this new vision is to enhance the ability of research to benefit the health of our communities more directly and more efficiently than has so far been the case. This will require a change in the way people think about and invest in research (Table 5-1).

Workforce Development for the New Research Paradigm and the Role of Academic Health Centers

This new paradigm of health research requires fundamental change. Central to this change will be the development of a new cadre of researchers, clinicians, and health leaders who have the expertise and skills to put the essential vision of this new approach into effect. To expand medical research to include translational aspects (T1, T2, T3) and to adequately address the broader definition of health (using behavioral, public health, and social determinants approaches), the health research workforce must be redefined to include individuals with diverse backgrounds and skill sets (Table 5-2). Indeed, the complexity of the challenges facing healthcare delivery today requires true multidisciplinary teams whose members bring multifaceted perspectives to finding health solutions. The new health research workforce must include not only health professionals of all types, but also engineers, sociologists, urban planners, policy experts, economists, and more. The importance of this has been pointed out in a call for a model in which "society would engage all of its facets—not just medicine and public health—in the collective act of preventing disease" (Woolf, 2008).

Academic health centers (AHCs)—consisting of schools of medicine, one or more other health professions schools, and an ownership or affiliated relationship with a teaching hospital or health system—are major players in the development of health researchers at all levels (AAHC, 2008a). Their

TABLE 5-1 An Approach to Achieving a New Research Vision

New People and Skills	• Multidisciplinary teams • Strategic faculty recruitment • Expansion and training of research support staff • New partners (e.g., industry, nongovernmental organizations, faith-based organizations, payers, government, public, diverse communities, patients, general public) • New venues (e.g., community-based research) • Training to provide new skills, including inter-professional training • Incentives within academia to support all types of health researchers (e.g., academic home, revised promotion, tenure criteria)
New Infrastructure	• Information technology investments (e.g., electronic health records, regional health information organizations, personal health records) • Biostatistics and data management support • Biorepositories • Streamlined clinical research approval processes • Efficient intellectual property policies • Links between academia, industry, and venture capitalists
New Investments and Incentives	• Expanded funding for clinical, translational, and social health research by the National Institutes of Health, National Science Foundation, foundations, others • Identification of new funding sources, especially for T2 and T3, behavioral, public health, and social health research • Increased organizational investment in translational research cores (e.g., informatics, clinical research nurses) • National coordination of research resources (e.g., informatics linkages, data sharing)

cross-cutting nature, exemplified by their multiple missions of research, education, and clinical care, suggest that they are well positioned to play leadership roles in creating this transformation. They have long served as a test bed for innovation, and they have a social mission that extends beyond purely business imperatives. They also often have university locations and affiliations that provide access to a wide range of potential collaborators from fields not traditionally located in health schools.

In addition, organizational and management trends taking place within the AHC enterprise are fostering new types of institutional integration and alignment (Wartman, 2008), which are highly supportive of new interprofessional research models. AHCs have the ability to test and disseminate new approaches to health care and also have access to the expertise of

TABLE 5-2 Workforce Development to Support Health Research

Academic Health Centers and Universities

- Commitment of leaders to train and support health researchers
- Create multidisciplinary teams of health professional researchers as mentors and role models
- Facilitate training and inclusion of broad range of other professionals on health research teams and to serve as mentors and role models
- Provide needed infrastructure to support training of and research opportunities for all types of health researchers
- Develop innovative curricula and training programs to develop health researchers
- Provide academic homes and satisfying career paths with appropriate incentives, rewards, and promotion for health researchers
- Cross-mission planning to leverage clinical investments to support health research training and research opportunities
- Streamline and update policies (e.g., intellectual property and clinical contracting) to support trainees and other health researchers

Industry, Community, and Other Nonacademic Partners

- Industry support for training, including funding, venues, and mentors
- Community partnerships for training sites, research venues, and study partnerships
- Government and NGO support for training, including funding, venues, and mentors
- Enhanced public understanding of the importance of supporting training of health researchers
- Increased philanthropic support for training of broad range of health researchers

National Policy Leaders

- Enhanced federal funding of and innovative mechanisms for training health researchers
- Policies and programs that stimulate transdisciplinary training and career opportunities for health researchers
- Removal of bureaucratic regulations that inhibit health research training and career opportunities

multidisciplinary teams. Consistent with their institutional mission of promoting societal health, they can train the next generation of researchers and clinicians in this new paradigm. They are able to look beyond their institutional walls to build multisector teams that can carry out and disseminate this new research approach with the overarching goal of improving healthcare delivery and the nation's health parameters.

To do so, AHCs will need to ensure the following:

- commitment of their own leaders to drive the expanded research approach,
- investments in new infrastructure (e.g., information technology, data, biorepositories) to support training and health research opportunities, and

- curricular and training innovation to develop multidisciplinary, multisector research teams with the skills required for this broader paradigm of health research.

Leaders of AHCs are in a strong position to contribute to the development of a workforce that can drive this new vision of health research. By lending their voices to the call for adequate and innovative funding mechanisms, these leaders can advance their institutional mission of improving health in their communities. Further, by ensuring that their organizations provide the needed culture and infrastructure to support new research approaches, they can provide the opportunity for health research. They can facilitate the partnerships with government, industry, and community groups needed for health research. Finally, academic health leaders can commit to and facilitate training the future workforce of health researchers.

Workforce development for health research requires reaching across disciplines within the medical field and beyond. Researchers (and clinicians and educators) from different subspecialties within medical research institutions must interact beyond their disciplinary silos. As encouraged by the CTSC programs, health professionals from medicine, nursing, pharmacy, dentistry, veterinary medicine, and allied health fields can come together to form research teams. And these teams should extend to colleagues in other fields ranging from engineering to policy to humanities and beyond.

These new multidisciplinary teams must also look beyond the walls of academia. Creating partnerships with a variety of nonacademic constituencies who can contribute to defining workforce needs and providing varied training venues will be critical to the success of this new vision, as discussed below.

Curriculum innovation is central to creating a workforce with the skills and experience to succeed in this new paradigm. Key to this innovation will be interprofessional training. If researchers are to work together in teams, they must first learn in teams. Innovative programs that bring various health profession educational programs together (e.g., nurses and physicians) are being developed at several AHCs but need to become more common. Programs such as the Howard Hughes Medical Institute's "Med Into Grad" program expose Ph.D. basic science trainees to clinical experiences so that they can understand the potential impact of their future work on improving health (HHMI, 2009).

This new breed of health researchers will require specialized training. Programs designed for scholars at all levels—medical, nursing, dental, pharmacy, and other health profession students; residents, fellows, and other advanced trainees; and junior (and senior) faculty—should provide the specialized skills needed for health research. While the training must be tailored to individual needs, opportunities to learn study design, research

ethics, informatics, leadership skills, business skills, and more should be made available. Clinical trainees will desire exposure to core research resources, such as gene sequencing, specialized assays, and animal models, while basic science trainees will seek out exposure to clinical settings in order to provide context for their work. Programs available through the NIH (K–30, K–12, M.D.–Ph.D.), several foundations, and others are beginning to meet these needs but need to be adequately funded and made available to more trainees.

But this curriculum innovation cannot be confined to graduate and professional education. Additional priorities include

- K–12 outreach programs that expose young people to the joys of translational research,
- staff training for all members of research teams,
- training of community partners, and
- public awareness campaigns that emphasize the range of health research impacts.

AHCs must provide clear career paths for health researchers who pursue these newer types of research, as they have for traditional biomedical researchers. AHCs must value clinical and translational researchers and actively recruit them and work to retain them by providing research and scholarly opportunities. Under this new research paradigm, clinical and translational research cannot be considered an "add-on" activity for busy clinicians who may have no special training in research. Clinical and translational researchers who have specific education and expertise in translational research must be strategically recruited to support the research agenda of the AHC. An even greater challenge will be to provide career paths and appropriate rewards to researchers from nonmedical disciplines who participate in health research.

The creation of an academic home for this new cadre of faculty has been endorsed by the CTSC program and addressed in a variety of ways at institutions across the country. Recruitment packages for translational researchers that provide the resources, environment, and protected time for faculty success are essential. The redesign of promotion and tenure criteria that will recognize and reward contributions to clinical, translational, and social research as well as to basic science investigation are critical. Finally, reasonable salaries and job security must be available to all health researchers.

AHCs also must assume responsibility for ensuring that adequate and appropriate institutional resources are available to health researchers during their training years and later, when they become faculty members. Traditional research support services and policies in most academic institutions

were generally designed to support basic science investigators. Now, it is important to design systems that also support clinical, translational, and social health researchers. In many cases, this will mean expanded administrative support for attracting clinical research grants, proposal preparation, intellectual property advice, conduct of clinical trials, data management, statistical support, financial monitoring, and more. Systems geared to support and incentivize research, rather than manage risk, need to be prioritized by institutional leadership.

AHCs must therefore invest in new research infrastructure. Ideally, this would involve "one-stop shopping," user-friendly research support offices that reduce the bureaucratic barriers too often encountered by faculty and trainees. Just as academic institutions have made major investments in basic science infrastructure over the years, AHC leaders should consider directing dollars to support expanding research IT services and informatics, data repositories, biorepositories, IT links with other institutions, clinical research units, and others. In addition, leadership endorsement and investment in programs that bring researchers together with community and business leaders can enhance the research enterprise.

AHCs are uniquely positioned to do cross-mission planning that can expand the capacity for all types of health research. By examining clinical investments as they are being made, incremental investments can be used to promote research opportunities. For example, many institutions are making large investments in EHRs. By proactively ensuring that the systems are designed to support health research, total investment can be less than if the clinical and research systems were purchased separately or the clinical system was retrofitted to accommodate the research needs. As an example, one institution has increased its capacity for community-based research by extending its telemedicine patient care network to serve as a backbone for community-based research projects (Nesbitt et al., 2006). Similarly, as institutions create community outreach programs, they should consider how clinical outreach can support and be supported by research initiatives in these same locales.

Overall, workforce development at AHCs will be most effective when the institutional values and priorities are aligned with the goals of this new research paradigm. By making it clear that the clinical mission is improved by health research, by committing to educate a broad range of health researchers, and by making sure that these academic missions are enhanced by high-quality clinical programs, the tripartite mission can be enhanced by alignment and leveraging of resources.

Workforce Development: The Role of Industry, Community, and Other Nonacademic Organizations

Academic–industry partnerships are critical to the successful translation of discoveries. While academia has traditionally focused on basic science discoveries and industry has focused on transitioning innovations to the market, the gap known as the "valley of death" must be bridged in innovative ways. Industry can play a key role in workforce development as well as in health research, especially translational studies.

Industry should serve as a strong partner in research workforce development. Joint training programs in which graduate students, postdoctoral candidates, professional trainees, and faculty spend time in both the industry and academic settings allow access to the expertise available in private companies. Other options include industry sites for internships or the conducting of research trials, didactic or "hands-on" teaching by industry leaders, and direct funding of workforce development by industry. Similar joint initiatives can be developed with government units and nongovernmental service organizations.

Successful academic–industry partnerships are central to the ability of institutions to provide opportunities and funding to both trainees and faculty throughout their careers. By funding workforce development and ongoing career opportunities for health researchers, including those at AHCs, industry can make an investment in their future workforce and potential colleagues and collaborators for health research.

An essential component of successful health research is the involvement of the community constituencies affected by the research. The growing interest in community-based participatory research recognizes the importance of the public's involvement in research (Jones and Wells, 2007). Components include community input into setting research priorities, sharing of approaches to build trust for community participation in research, and joint academic–nongovernment organization (NGO) leadership of research and dissemination of progress and findings to community. To determine the true efficacy of new medical innovations, it is essential that research not be carried out only in the ivory towers of academic research institutions but also be conducted in real-world community settings.

The inclusion of community partners, including industry, private healthcare providers, health maintenance organizations (HMOs), payers, schools, churches and other faith-based organizations, NGOs, government, and others will require ongoing trust building, communication, and openness to new research approaches.

The public has consistently reported great support for medical research, especially research that benefits patients in clearly understandable ways. Public messaging is key to sustaining the broad public support for bio-

medical research, as well as enhancing understanding of the broader goals of health research, and must be the responsibility of all. One part of this message is that declining clinical reimbursements to AHCs, especially those that assume a significant "safety net" role, represents a significant threat to the ability to develop the workforce needed for health research and to conduct the research. Messages should include appeals for appropriate societal funding for health research workforce development and also for philanthropic support of training and new types of research.

Workforce Development—The Role of National Policy Makers

This fundamental transformation in medical research to encompass all aspects of health research and the development of a trained workforce dedicated to such research can only happen with support from national policy makers. Decision makers should

- endorse the importance of health research as essential to leveraging the biomedical discoveries in which society has invested and improving the health of the country;
- provide adequate funding for the full range of health research, including appropriate resources for the needed workforce development, through a combination of public and private sources; and
- institute regulations and laws that facilitate health research and remove existing barriers to such research.

While a careful examination of allocations to the NIH for basic, clinical, and translational medical research is an essential part of this process, equally important will be a discussion of how public health and social determinants research will be funded. It is vital that the nation's basic science research infrastructure continues its impressive scientific progress. But particular attention to adequate funding of the CTSCs and of workforce training programs is urged. Chronic underfunding of health outcomes, epidemiologic, quality improvement, public health, preventive, and other clinical and translational research at the NIH and other agencies (AHRQ, CDC) must be reversed. Conceivably, a new federal entity whose mission is the fostering of all aspects of health research could be developed. Regardless, it will be important to identify funding mechanisms to support health research that encourages transdisciplinary approaches that reach out to nontraditional partners in other fields. Innovative funding mechanisms should be considered for health research, including the possibility of contributions from industry (e.g., pharmaceutical, devices, hospital, HMOs, payers). A strong business case can be made for this since the investment

will likely yield innovations that provide value by identifying cost-effective and high-quality practices.

In addition, policy makers should support regulatory and legal changes that stimulate the translation of discoveries to the market. Bureaucratic barriers that slow progress should be removed while ensuring that the safety of the public and trial participants is not compromised. For example, HIPAA procedures could be modified to facilitate identification of clinical research subjects while maintaining protections (AAHC, 2008b). A new approach that reaches beyond the traditional restraints of academia and private industry is essential to achieving the ambitious goal of health research to truly improve the health of our nations.

Finally, demonstrable linkages must be developed between the findings of health research and patient care delivery. These linkages, in our opinion, must go beyond the development of guidelines. Rather they should be integrated into clinical care through a combination of quality measures and payment reimbursement policies. In so doing, a self-perpetuating cycle of improvement would be incorporated into the nation's healthcare system.

Policy makers, academic institutions, industry, and community organizations must work together to ensure that the potential benefits of health research are fully realized and that the research workforce is optimized. Only through such cooperation can the vision of improved health for all be realized.

Conclusions

The nation needs to vigorously engage in a new health research agenda in order to achieve the goal of a healthier future for the United States and the global community. To accomplish this, a new research paradigm is needed that supports the full range of health research to ensure that new discoveries benefit patients. The adoption of this new research paradigm will require

- an expanded health research workforce with diverse perspectives and skills in which people work together in multidisciplinary teams and in venues throughout the community; and
- a health research workforce supported and incentivized by both academic institutions and other organizations with improved infrastructure and increased resources.

This ambitious but essential vision requires that health workforce issues be made a policy priority for the nation (AAHC, 2008c) and that academic and patient care delivery structures be redesigned to support the new paradigm of health research. Ultimately, the impact of this vision will depend

upon our political and social will to invest in this new type of research, and then to successfully link the new discoveries to the actual delivery of health care and design of health policy.

PUBLIC–PRIVATE PARTNERSHIPS

Overview

Consideration of how best to take advantage of existing infrastructure—e.g., data resources, expertise, technology platforms—will be important in developing and implementing priority infrastructure elements. Public–private partnerships are seen as an effective way to link some of health care's disparate component elements and to draw productively on the respective assets of participating stakeholders, and these partnerships offer an approach to developing, supporting, and nurturing productive relationships among stakeholders who come to CER from different perspectives and with diverse motivations, bridging gaps between stakeholders and removing barriers to cooperation. Absent such a mechanism for bringing relevant parties to the same table, fundamental differences in institutional cultures can impede or even preclude stakeholder-to-stakeholder communication. Public–private partnerships not only create a space in which collaboration can safely take place, but they also offer a structure and operational guidelines, typically tailored to a specific partnership by the participants, that help foment and facilitate cooperative work. The papers here report briefly on public–private partnerships from the perspectives of health plans, the federal government, and industry. Panelists were asked to describe relevant current or planned public–private partnership efforts and suggest ways that similar efforts could help build the infrastructure for CER.

Health Plans

Carmella A. Bocchino, R.N., M.B.A., Executive Vice President of Clinical Affairs and Strategic Planning, America's Health Insurance Plans

Health plans are strongly committed to working with stakeholders in both the public and private sectors to develop tools and other resources and programs to help ensure that patients and providers have the information they need on safety, effectiveness, and value to make sound health-care decisions. Toward this end, and often in partnership with federal agencies, health plans have created comprehensive databases that can be mined to identify potential safety problems as well as opportunities to improve care and care delivery. Many public–private partnerships have focused on providing information to individual agencies or health plans to

explore particular questions in depth, and emerging partnerships seek to broaden the scope and scale of these initial efforts. Briefly reviewed below are examples of learnings from such efforts that illustrate the potential of public–private partnerships as key buildings blocks for the infrastructure required for expanded CER. Key challenges, barriers, and opportunities are also discussed.

Health plan data is consistently collected and used within health plans to help provide timely information, which is important for point-of-care decision making, to clinicians and patients. Unfortunately, the United States wastes millions of dollars annually on medical treatments that may not work. In addition, when evidence exists and is not implemented, there are both human and financial consequences. Several examples from Kaiser Permanente illustrate how these data can be effectively harnessed to generate insights that improve care. A recent study of hip replacements (Meier, 2008), conducted by Kaiser, yielded important information about which devices work best for whom, information that Kaiser was able to quickly share across its network of physicians. Similarly, a case-control study of Kaiser's data on the cardiovascular effects of cyclooxygenase-2 (COX-2) inhibitors yielded results that have been used by the FDA and others to develop policy and practice guidelines about the use of COX-2 inhibitors. A logical extension of these examples is to find opportunities to draw upon data and share information more broadly, and many have called for the development of a national data system as a central part of the nation's health research infrastructure. Several current data-sharing efforts are outlined below to illustrate the potential.

Data-Sharing Efforts

The U.S. Renal Data System (USRDS) was established over 20 years ago to help Medicare, the major payer for renal dialysis and transplantation, collect data on the end-stage renal disease patient population and assess potential quality and safety problems as well as program costs. Funded by the National Institute of Diabetes and Digestive and Kidney Diseases at the NIH and CMS, the USRDS provides a means for organizations to share data sets and to collect, analyze, and distribute information on the end-stage renal disease population in the United States. In practice, this data system has served as an effective, proven national data registry that assesses trends in mortality, end-stage renal disease rates, treatment modalities, and morbidity. With the increased calls by clinicians and policy makers alike for a more comprehensive, national data registry, the USRDS merits attention as a possible model to move this effort forward.

Research and Surveillance Networks

There are several examples of health plans working together to create research and surveillance networks. These collaborations include work with the government and private sectors to address important research and public health concerns. The research and surveillance work conducted by these groups provides benefits to the health plans and to the community as a whole.

Among the notable health-plan community efforts to advance the development of needed clinical evidence is the HMO Research Network (HMORN). This umbrella organization was established in 1993 and is a consortium of 15 HMO organizations that have formal, recognized research capabilities. Distributed geographically, the participating HMOs collaborate to develop and implement common study designs, share standardized data, and provide learning for all participants. The power of combining the HMO data sets is that it gives each organization much richer, more reliable information than any one HMO could gather by itself. The work of the HMORN has impacted healthcare delivery through improved decision making and real-time learning across the different plans and has led to peer-reviewed studies on postmarketing surveillance and drug safety, population-based chronic care improvement models, and surveillance of acute diseases, such as rapid detection of immediate potential environmental or biological threats. The HMORN has partnered with several federal agencies on projects that illustrate the utility of data linkages and networks of research resources for improving the nation's health. These collaborative efforts with the government have led to the creation of collaborative projects under the umbrella of the HMORN, including the Cancer Research Network (CRN) and the Centers for Education and Research on Therapeutics, and have expanded the breadth of collaboration to include work with academic medical researchers.

The HMO CRN, an expansion of the HMORN, is a consortium of 14 research centers based in integrated healthcare delivery organizations. The CRN works with the National Cancer Institute to assist with the mining of very large data sets and focuses on the characteristics of patients, clinicians, communities, and health systems that lead to the best possible outcomes in cancer prevention and care. Multidisciplinary intervention research addresses cancer prevention, early detection, treatment, survivorship, surveillance, and end-of-life care. These models also provide some insights into how these partnerships can be effectively leveraged to improve health research.[2]

The Vaccine Safety Datalink (VSD) was established in 1990 as a col-

[2] See http://crn.cancer.gov (accessed September 8, 2010) for project list.

laborative effort between the National Immunization Program of the CDC, leading vaccine researchers, and eight large HMOs. Designed to monitor and evaluate vaccine safety, the VSD is a large, linked-data model that allows real-time analysis of vaccination and medical records of more than 6 million people for the purposes of conducting objective, population-based vaccine safety research. The VSD employs a distributed data model whereby the scope of work is decided in advance and only those data deemed necessary to support the agreed-on scope of the research are culled from available data sets and aggregated. Confidentiality of individual medical information is protected as only the data necessary to assess vaccine safety or adverse events is collected and aggregated. The VSD is the largest component in the CDC's vaccine safety surveillance efforts. It has also pioneered research methods for conducting this work, and it serves as the gold-standard model for the development of other similar surveillance networks. VSD is a valued resource that allows researchers at the CDC and health plans to conduct studies that provide information about the short- and long-term effects of specific vaccines on various populations. Rather than relying on reports from vaccine manufacturers or solely on a passive reporting system to identify possible safety issues, VSD offers a rich data resource that can be accessed quickly, by CDC and vaccine investigators employed by health plans, to continually monitor vaccine safety. These data enable both rapid analyses of specific vaccines and large-scale, retrospective studies of people who have experienced unusual or severe adverse reactions to vaccines. VSD provides a comprehensive data resource enabling researchers to examine virtually all patient health events during the time a patient is enrolled in a health plan. Moreover, if a patient moves to another health plan, an effort is made to continue the patient's enrollment in the program.

VSD research and data are often used to inform the decisions of the Advisory Committee on Immunization Practices and, as such, provide guidance and direction for public health policy in the United States and the rest of the world. The VSD infrastructure has also provided the opportunity for research on both the safety and effectiveness of many vaccines.

The VSD model also offers ample flexibility as new information needs emerge. As an example, when concern was raised about a potential association of meningococcal vaccine and Guillain-Barré syndrome (GBS), the CDC needed to conduct a rapid study to determine the reality of this association. For this particular request, the data on 6 million lives in the VSD data set were not sufficient for the analysis needed. In response, the America's Health Insurance Plans (AHIP) organized 5 additional large, national health plans to provide a data set population of 60 million people. Initial analysis of this larger data set determined that there was not a causal link between the meningococcal conjugate vaccine (MCV) and GBS. AHIP, in collaboration with six health plans and with funding provided from the

vaccine manufacturer, has continued this surveillance work on MCV and GBS using a protocol that mirrors the methodology of the VSD.

This experience demonstrates the considerable potential inherent in public–private partnerships around the analysis of large data sets to answer questions important to public health, and the particular importance of developing and using standard methodologies for these efforts.

Quality Improvement Efforts

While these partnerships have focused on safety, similar distributed data models are being developed for use in quality measurement and reporting efforts. Supported by a grant from the Robert Wood Johnson Foundation, the AHIP Foundation and the Quality Alliance Steering Committee at the Brookings Institution are currently working on the National Data Aggregation Initiative (NDAI) to develop and implement a standard methodology for aggregating data across multiple health plans for provider performance measurement and reporting. One of the goals of NDAI is to combine private-sector and Medicare data to generate physician performance measures. To achieve this goal, the AHIP Foundation and the Quality Alliance Steering Committee have been working closely with CMS and AHRQ to align the NDAI with the Generating Medicare Physician Quality Performance Measurement Results project.

Like the VSD model, the NDAI will be based on a distributed data model that involves retention of protected health information at the health plans and transmission of measure results and provider identification data to a hub at the AHIP Foundation for aggregation, provider matching, specialty assignment, and reporting. The AHIP Foundation and the Quality Alliance Steering Committee are seeking to implement a distributed data model that is flexible and can accommodate various types of measures in future iterations that rely on a variety of data sources, such as laboratory, registry, and EMR data.

One of the main advantages of the NDAI will be the ability of various stakeholders to conduct regional and national comparisons of provider quality and analyze regional variations in provider quality similar to the Dartmouth Atlas, facilitated by the implementation of a standard methodology. In addition, the distributed data model envisioned under the NDAI holds promise for other efforts, such as the FDA Sentinel Initiative.

The FDA, as discussed elsewhere in this chapter, is exploring opportunities to harness other forms of data, such as those derived from EHRs. Initial discussions of a consolidated industry approach to work with the FDA have included national, regional, and local health plans; the Udall-Reagan Foundation; and other stakeholders.

Challenges and Lessons Learned

The initiatives discussed above underscore the inherent value in developing the infrastructures and tools to aggregate and analyze these data across populations. Collectively, they demonstrate that although there are many different types of health information available within health plan data sets, approaches have been developed to standardize these data so that they can be brought together to answer specific questions relevant to decisions faced by clinicians and patients. These efforts also potentially provide initial building blocks toward a national infrastructure for long-term safety and effectiveness research.

As a first step, it will be necessary to come to an agreement on what shared methodology might facilitate comparative analyses. A great deal of clinical research currently exists, but in forms that impede comparison across studies. An additional need is to build a data system infrastructure that makes it possible not only to mine existing data, but also to identify and track potential issues in safety and effectiveness in health care. For example, when a Massachusetts plan noticed a spike in hospitalization rates across all their plans at a particular time of year, they were able to monitor this pattern over 2 years and work with the local public health agencies to confirm that the increase in hospitalizations was occurring during the influenza season. This data then contributed to a study to examine whether influenza complications were driving those admissions or whether the increase in hospitalizations could be attributed to the patients not vaccinated for influenza. The Massachusetts study points again to the power of the data available in the model of aggregated models described above, in the sense that there is both a capability to bring the data to the questions and to also bring the questions to the data. "Bringing data to the questions" helps close the evidence gap and provide valuable information to patients and clinicians.

A significant challenge facing these efforts has been the standardization of data and the need to adjust many factors when only administrative data (vs. clinical data) are available. As health plans develop and use more robust EHRs, they are learning how to compile these clinical data, but an increasingly central issue is that EHRs have not been created to produce the data needed to answer questions important to understanding quality or clinical effectiveness. For example, many EHRs do not enable researchers to cull data related to the performance of providers. As physician offices and practice sites are reengineered to support the most effective use of EHRs, at the same time it will be necessary to ensure that those records are designed with the inherent capacity to produce the information needed—not only for performance evaluation and quality improvement, but also for research. Without thoughtful and appropriate design, EHRs will not capture the

information needed to advance the challenging work of analyzing and filling gaps in evidence.

A second challenge concerns governance. Essential to the sustainability of these efforts is a governance structure with appropriate roles for government and other stakeholders. Essential questions pertaining to the governance model need to be considered, such as funding for the establishment of and long-term sustainability of the infrastructure, study priorities, use of aggregated data, and secondary use of these data beyond the initial scope of work. Health plans, working as part of collaboratives, have produced several different governance structures, both formal and informal, that could serve as models for the Sentinel Network and other developing partnerships.

A third challenge has to do with the tensions that often emerge when the insurance sector has partnered with academic institutions on specific clinical research projects. There is a diversity within the industry on plan engagement with research, but the common challenge that has emerged is that while clinical research is generally regarded as a public good and the NIH is well supported by public funds, the kinds of research, such as comparative analyses, that are important to improving health care are not generally the kinds of research funded in this manner. Discussion is needed on sustainable funding approaches for these types of research.

Finally, the conduct of clinical research or trials implies the collection of data that is intended to be combined with other data and publicly shared. That is, the data are intended to be used to help educate and to provide lessons that can be drawn from what the evidence says or does not say. Conversely, the implicit nature of clinical research is that it is not intended to be reserved for use within a specific institution for its own QI and learning, without public release. A related challenge concerns the ownership of data. Patient advocates argue that data rightfully belongs to patients. Numerous discussions are ongoing about not contributing data to research or QI efforts without patient consent. A strong case can also be made, however, that these collected data are a public good that should be used for research, scientific investigation, and filling the evidence gap. More attention is needed to articulating the case for these data as a public good in that their use can ultimately result in better patient care.

Public–Private Partnerships and Comparative Effectiveness Infrastructure Development

Lessons learned from efforts to promote and develop shared data resources suggest several immediate priorities for comparative effectiveness infrastructure development, including developing standardized methodologies and ensuring data transparency. To develop greater national-level

capacity for comparative effectiveness work, additional focus will be needed on prioritizing key issues for research, on developing sustainable funding mechanisms, and on how to best build upon existing data sources, organizations, and initiatives.

As discussed elsewhere in this publication, clinical registries are providing an increasingly rich resource for clinical data, and there may be an expanded role for registries as part of this work. Although currently a source of very good information, registry data are not usually shared broadly.

Also important to consider are the organizations engaged in the analysis of these data. Several existing programs are engaged in work vital to the CER agenda. For example, the Blue Cross/Blue Shield Association Technology Evaluation Center (TEC) program has, for some time, conducted studies to assess the effectiveness of select technologies. The TEC program has developed criteria for these assessments, and these criteria and their reports are made publicly available. The Institute for Clinical and Economic Review (ICER) is a recently established academic comparative effectiveness initiative based at the Massachusetts General Hospital's Institute for Technology Assessment. Informed by the priorities of payers, clinicians, and patients, and based at arms' length from coverage decision-making structures in government and the private sector, ICER links clinical effectiveness and comparative value in a rating system that is consistent and rigorous, yet flexible enough to be directly useful to multiple decision makers. ICER is positioned to analyze the strength of evidence and the value of technologies and to provide additional information to both consumers and employers about how technologies, drugs, and devices that show high value can most effectively be moved into the marketplace. A related component of this assessment includes the cost–benefit analysis of these new applications. Many drugs or technologies may not have sufficiently high value to be considered without greater cost sharing on the part of consumers who want them.

Federal Agencies

Rachel E. Behrman, M.D., M.P.H., Associate Commissioner for Clinical Programs and Director of the Office of Critical Path Programs, Food and Drug Administration

Providing opportunities for key stakeholders, such as patients, researchers, and other members of the public sector, to work together with regulators and other government agencies on issues of common interest has been critical to progress in many areas of health care. The development of drugs to treat human immunodeficiency virus/acquired immunodeficiency

syndrome was one of the FDA's early experiences with collaboration. The combined efforts of government, industry, and activists demonstrated that broad stakeholder engagement often accelerates progress in the development and availability of therapeutics. This experience also demonstrated that such development programs cannot be expected to address all relevant questions—a circumstance that was anticipated and accepted by all stakeholders. Although much was learned about therapies as companies pooled their data, and although the work led to development of different paradigms of drug approval, the effort did not provide all the answers about long-term outcomes.

This paper reviews key lessons learned from several existing public–private partnerships and offers suggestions concerning how these efforts might inform or contribute to expanded capacity for comparative effectiveness work.

The FDA and Public–Private Partnerships

Two FDA initiatives, the Critical Path Initiative (CPI) and the Sentinel Initiative, a collaboration being launched under the CPI, continue the tradition of collaboration with key stakeholders and seek to address some of the limitations identified in earlier projects. Each initiative has the potential to foster broader collaboration among stakeholders, and initial work suggests several key areas for future work.

The Critical Path Initiative The FDA's CPI is aimed at stimulating and facilitating a national effort to modernize the sciences through which FDA-regulated products are developed, evaluated, and manufactured. The initiative was launched in March 2004 with the release of the report *Innovation/Stagnation: Challenge and Opportunity on the Critical Path to New Medical Products* (FDA, 2004).[3] The report diagnosed the scientific reasons for the recent decrease in the number of innovative medical products submitted for approval—a decrease that was puzzling in light of recent advances in biomedical sciences and disappointing from a public health perspective. The report noted the rising complexity and unpredictability of medical product development and called for a concerted effort to modernize the scientific tools (e.g., in vitro tests, computer models, qualified biomarkers, innovative study designs) and harness the potential of bioinformatics technologies to evaluate and predict safety, effectiveness, and manufacturability of candidate medical products. The report also called for a national effort to identify specific critical path activities that, if car-

[3] See http://www.fda.gov/oc/initiatives/criticalpath/whitepaper.html (accessed September 8, 2010).

ried out, would help modernize the critical path sciences. The goal of this effort is to reap the expected public health benefits from the promises of biomedical advances of the 21st century.

In March 2006, the FDA released the *Critical Path Opportunities List* (FDA, 2006).[4] Developed with extensive public input, this list describes the areas of greatest opportunity for improvement in the product development sciences and provides 76 concrete examples of how new scientific discoveries—in such fields as genomics and proteomics, imaging, and bioinformatics—could be applied during medical product development to improve the accuracy of the tests used to predict the safety and efficacy of investigational medical products.

The CPI, designed fundamentally to help foster collaboration among stakeholders, is one of the FDA's top priorities. The agency is building on its unique position to partner with other federal agencies, patient groups, academic researchers, industry, and other stakeholders to identify areas ripe for improvement and to coordinate, develop, and disseminate solutions to scientific hurdles that are impairing the efficiency of developing, evaluating, and manufacturing of FDA-regulated products.

A number of collaborations have been formed under the CPI, including a public–private partnership co-founded by the FDA and Duke, called the Clinical Trials Transformation Initiative, which has the goal of modernizing the clinical trial enterprise. Other collaborations, involving, for example, the NCI, the NIH, the Juvenile Diabetes Research Foundation International, the Critical Path Institute, and industry partners, are working on a range of projects, including the development of an artificial pancreas, the use of imaging in cancer drug development, warfarin dosing, standards development, and bioinformatics projects.[5]

A collaboration of particular relevance to the EBM effort was announced by the FDA in May 2008. As discussed in more detail in the following section, the Sentinel Initiative has the potential to directly inform any effort to monitor product use, including benefits and risks, and ultimately might be useful in comparative analyses.

Sentinel Initiative The Sentinel Initiative has the ultimate goal of developing and implementing the Sentinel System—a national, integrated, electronic framework and approach for monitoring medical product safety. The framework is envisioned initially to be a distributed network in which data

[4] See www.fda.gov/oc/initiatives/criticalpath/reports/opp_list.pdf (accessed September 8, 2010).

[5] For a list of projects launched during 2007 either by the FDA or with FDA participation, see FDA's CPI Web page at http://www.fda.gov/oc/initiatives/criticalpath/report2007.pdf (accessed September 8, 2010).

holders will retain data and provide reports in response to queries from interested parties. Sophisticated links between involved parties are assumed as a necessary component of the broad collaboration. This national electronic framework would enable the FDA (and ultimately perhaps others) to query remote data sources (with the consent and permission of the data owners), using appropriate security and privacy safeguards, for specific information about marketed medical products. The Sentinel Initiative will respond in part to the FDA Amendments Act of 2007 (FDAAA), which calls for active postmarket safety surveillance and analysis. The congressional mandate contained in the FDAAA provides an initial focus for the Sentinel System on medical product safety; however, as envisioned, the Sentinel System could ultimately provide a basic infrastructure for all FDA-regulated products.

The Sentinel Initiative is in the very early stages of development. The FDA's first step has been to create a broad public forum for the exploration of relevant issues, and a number of short-term contracts have been awarded to research specific aspects of the initiation (e.g., who should be involved, what kinds of privacy and security concerns are of particular importance, how such a partnership should be governed). The immediate goal of the initiative is to develop a public–private partnership, hosted by a nonprofit organization that would oversee the implementation of the system. As mentioned elsewhere in this publication, there are many examples of successful efforts to build and analyze shared data resources around specific interests, and, as the Sentinel report explains, a number of collaborations are also under way that will directly inform Sentinel.[6] However, the infrastructure that will support the Sentinel System is envisioned to be a sustained and comprehensive national data resource that is broadly available to many stakeholders.

It is important to note that the Sentinel System will augment, but not replace, current FDA activities. The FDA will continue efforts to modernize and optimize its systems for spontaneous reporting. These systems and processes will remain an important part of the agency's postmarket surveillance systems. The Sentinel System is envisioned as an important new resource for efficiently detecting meaningful safety signals and investigating important questions about medical products—a key component of the FDA's work.

Key Challenges and Lessons Learned

The Sentinel Initiative drives important fundamental changes and introduces important innovations. However, significant issues remain to be

[6] FDA's Sentinel Report is available at http://www.fda.gov/oc/initiatives/advance/reports/report0508.pdf (accessed September 8, 2010).

resolved. For example, research methods and data analysis tools will need to be developed to ensure the production and validation of timely, reliable, and secure information. The distributed network approach of the Sentinel Initiative addresses some concerns about patient privacy, but other challenges remain. It is imperative to engage parties that collect, aggregate, and market data and to illustrate the critical need and business case for a sharper focus on outcomes research to improve the nation's health. For these issues, developing the appropriate governance structures and policies will be critical. Another challenge will be to ensure that the infrastructure developed considers and meets the needs of all parties while putting appropriate safeguards into place. Questions related to data access, use, and stewardship will have to be resolved.

These activities highlight many issues that will also be of central importance in the development of infrastructure for comparative effectiveness. Priority setting is critical in order to provide a common focus for all stakeholders as well as to identify key opportunities to develop smart and small pilot projects. Financing is a continual challenge, particularly given that infrastructure development is a long-term and expensive proposition. Continued attention is also needed to the governance of collaborations. A fourth and crucial area for work is data transparency. Progress in these areas is needed to ensure that analyses are conducted and reported responsibly and to avoid the development of unvetted, low-quality information. Finally, issues about how to handle proprietary data and patentable tools or processes will remain key areas of importance for all potential participants.

Public–Private Partnerships and Comparative Effectiveness Infrastructure Development

Public–private partnerships will be critical for the successful development of a national infrastructure for expanded CER as part of the IOM EBM effort. As with the Sentinel Initiative, the government alone cannot lead us to where we need and could be as a nation with respect to health. The FDA has focused on partnering with others because collaboration provides the best opportunity for substantial engagement by key stakeholders on issues of common interest and, therefore, a greater likelihood of success. The IOM effort is a large and complex project, and no one entity has the expertise, the resources, or the energy to carry it out alone. In addition, it will be important to create a nimble infrastructure to respond to dynamic and evolving research needs. Such an effort will require the engagement and participation of all sectors across the healthcare system. A government approach, possibly relying on legislation, may only slow progress.

Lessons learned from the Sentinel Initiative may be very useful for the IOM effort. Of particular benefit might be small collaborative pilots, simi-

lar to those under way as part of the Sentinel Initiative, that are making use of existing large databases to identify and test the tools and processes that will be needed to perform postmarket monitoring. Similar tools will be needed for comparative evidence analyses. Additional considerations that may be useful as the CER project evolves (or as pilots are identified that could inform the project) include the following:

- What specific tools need to be developed?
- What are the specific goals of a particular collaboration?
- How should specific projects or tasks be prioritized? And who should be tasked with setting priorities?
- Which stakeholders would be most beneficial to and interested in a particular collaborative project?
- Which organization or organizations can best take the lead on a specific project?
- How can needed short- and long-term resources be obtained?
- How can research results be made available to the community without undermining proprietary or patent interests?
- How do specific collaborations contribute to the larger effort?
- What time frames can realistically be set for short- and long-term goals?

Partnership formations will require careful vetting by all parties so that everyone involved has confidence in the successful operation of the partnership. With each partnership comes added confidence in what it will take to make a successful partnership. However, each partnership will be different, raising new questions and unique hurdles.

Health Product Developers

William Z. Potter, M.D., Ph.D., Vice President, Franchise Integrator Neuroscience, Merck Research Laboratories

Two examples of public–private partnerships that have productively linked industry, government, academia, and other stakeholders to address issues of common concern in health care are the Biomarkers Consortium (BC) and the Alzheimer's Disease Neuroimaging Initiative (ADNI). This paper briefly describes the processes of developing and sustaining these partnerships, as well as some of the key lessons learned that can inform the development of infrastructure for expanded CER. Some suggestions for priority areas for work and opportunities for greater engagement by the health product developer sector are also discussed.

Biomarkers Consortium

The BC, founded in 2006, was established to advance the discovery, development, and approval of biological markers to support new drug development, preventive medicine, and medical diagnostics. The consortium is a major public–private biomedical research partnership with broad participation from stakeholders across the health enterprise, including government, industry, academia, and patient advocacy and other nonprofit private-sector organizations. In addition to the Foundation for the NIH, founding members include the NIH, the FDA, and the Pharmaceutical Research and Manufacturers of America. Other partners in the consortium include CMS and the Biotechnology Industry Organization.

Imperative to a successful partnership is the careful delineation of specific areas of research focus that protect individual interests of consortium members, and, after some discussion, consortium organizations agreed to work together to accelerate the identification, development, and regulatory acceptance of biomarkers in four areas: cancer, inflammation and immunity, metabolic disorders, and neuroscience. Additional goals of the consortium include the conduct of joint research in "precompetitive" areas with partners that share common interest in advancing human health and improving patient care; that speed the development of medicines and therapies for detection, prevention, diagnosis, and treatment of disease; and that make project results broadly available to the entire research community.

An example from neuroscience illustrates another key to the consortium's success. As an initial focus, the group looked at the placebo response, a fundamental issue of common concern to all stakeholders. An important question for the field is the relative efficacy of antidepressants, but even the efficacy of antidepressants vs. placebo is often unclear. Consider the physician, or any other caretaker, who diagnoses and would like to treat a patient for depression. Trial results demonstrate that the placebo response is often enormously variable, ranging from some 20 percent up to as much as 60 percent in very large trials with up to 100 to 150 patients per arm. These findings raise significant questions about the validity of these data, the study design, or the diagnosis. Healthcare providers are interested in developing and using their data to clarify the quality of treatment and to determine the best possible course of care, but health product manufacturers also have an intense competitive interest in such data—particularly in improving the quality of data in this space and the analyses needed to inform critical healthcare decisions.

The BC addressed this set of issues by creating a metadata set. As outlined in the Foundation for the NIH's Consortium Placebo Data Sharing proposal, ideal characteristics for implementation include identical study design; extensive characterization of each subject (e.g., more than

FDA requirements); data elements stored in standard, easily shared data systems; and appropriate informed consent. Such ideals are of course difficult to realize regularly on a macro level, and the consortium decided to focus initially on an area in which common public and private study design were likely: antidepressant trials conducted since the introduction of selective serotonin reuptake inhibitors. Around this focus, the consortium has initiated several collaborative efforts, including a Depression Rating Scale Standardization Team (DRSST), a Placebo Response Collaborative Study Group, the National Institute of Mental Health Placebo Database Workshop, and placebo databases from Alzheimer's disease trials.

Discussions leading to the development of these projects began in 2000, and the group is beginning to put the needed infrastructure in place through the Foundation for the NIH. Many of the lessons learned from these discussion, will help to accelerate the development of infrastructure for CER work. Key barriers include the need for an internal champion within each company to work a proposal; meeting costs of full-time equivalent and data management; skepticism by industry, NIH, and academic leadership that learnings of value can be gained; and variable legal opinion as to intellectual property and medicolegal risks.

Alzheimer's Disease Neuroimaging Initiative

Another noteworthy public–private partnership is the ADNI. Started in 2004, this large research project seeks to define the rate of progress of mild cognitive impairment and Alzheimer's disease in order to develop improved methods for clinical trials in this area and also to provide a large database that will improve design of treatment trials. It is hoped that the project will provide information and methods that will help lead to effective treatments and prevention efforts for Alzheimer's disease. The project has funding from the National Institute of Aging, the National Institute of Bioimaging and Bioengineering, Pharmaceutical Research and Manufacturers of America, and several foundations.

ADNI brings together organizations from the public and private sectors, including government agencies, corporations, consumer groups, and other stakeholders, to work collaboratively to determine the right tools to understand the efficacy and effectiveness of drugs for Alzheimer's disease. Participants in the initiative collaborate via an infrastructure that, while complex, enables cross-sector communication and work and has produced promising initial results. For example, both complex clinical data and intricate brain imaging data are now readily accessible using the Web, in close to real time. Anyone who is interested in developing ways to look at complex data can mine these data, and this approach has begun to return remarkable findings. Underlying the success of this partnership is how

it addresses the important issue of data transparency. Through different portals, the data are available both to researchers and, with unprecedented access, to the general public.

Given that researchers can manage this complexity of data with existing tools in the realm of Alzheimer's disease, these results imply that similar applications are likely for other sets of data. An important lesson from this work is that real data can be made accessible, in real time, in the public domain, and yield useful results. More broadly, the success of this project underscores and justifies the benefits of a consortium approach, particularly when the scientific methodologies employed are adequately rigorous and the questions are sufficiently important.

Public–Private Partnerships and Comparative Effectiveness Infrastructure Development

As a national infrastructure for CER is being developed, leadership will be needed from the federal government to develop the incentive structures, through legislation and regulation, that are important to advance issues related to data standards and data sharing; however, despite the important "pull" provided by legislation, there are opportunities for immediate work that do not require legislation. Some possible focus areas include

- engagement of industry leadership (e.g., identifying and encouraging industry champions; fostering collaboration of industry, NIH, and academic leadership around common issues and concerns),
- making the case for broad stakeholder participation around key questions and issues,
- developing national research priorities, and
- establishing methods and collaborative agreements for data collection and use.

REFERENCES

AAHC (Association of Academic Health Centers). 2008a. *The academic health center: Evolving organizational models*. http://www.aahcdc.org/policy/reddot/AAHC_Evolving_Organizational_Models.pdf (accessed September 13, 2008).

———. 2008b. *HIPAA creating barriers to research and discovery*. http://www.aahcdc.org/policy/reddot/AAHC_HIPAA_Creating_Barriers.pdf (accessed September 15, 2008).

———. 2008c. *Out of order, out of time: The state of the nation's health workforce*. Washington, DC: AAHC. http://www.aahcdc.org/policy/AAHC_OutofTime_4WEB.pdf (accessed September 5, 2008).

AAP/ACP/AOA (American Academy of Pediatrics and American College of Physicians and American Osteopathic Association). 2007. *Joint principles of the patient-centered medical home*. http://www.medicalhomeinfo.org/Joint%20Statement.pdf (accessed August 5, 2010).

ACC (American College of Cardiology). 2008. *Clinical statements/guidelines*. http://www.acc.org/qualityandscience/clinical/statements.html (accessed September 15, 2008).

ACCF (American College of Cardiology Foundation). 2008. *NCDR: National Cardiovascular Data Registry*. http:www.ncdr.com/WebNCDR/ANNOUNCEMENTS.ASPX (accessed September 15, 2008).

AHA (American Hospital Association). 2007. *Continued progress: Hospital use of information technology.* http://www.aha.org/aha/content/2007/pdf/070227-continuedprogress.pdf (accessed September 10, 2008).

AHRQ (Agency for Healthcare Research and Quality). 2009. *Horizon scan: To what extent do changes in third-party payment affect clinical trials and the evidence base?* http://www.cms.hhs.gov/determinationprocess/downloads/id67aTA.pdf (accessed September 20, 2008).

Arvantes, T. 2007. *North Carolina seeks expansion of primary care program*. http://www.aafp.org/online/en/home/publications/news-now/government-medicine/2 (accessed September 15, 2008).

Bach, P. B. 2007. Coverage with evidence development. In *The learning healthcare system: Workshop summary*, edited by L. Olsen, D. Aisner, and J. M. McGinnis. Washington, DC: The National Academies Press. Pp. 39-45.

Bawa, K. S., G. Balachander, and P. Raven. 2008. A case for new institutions. *Science* 319(5860):136.

Bodenheimer, T. 2008. Coordinating care—A perilous journey through the health care system. *New England Journal of Medicine* 358(10):1064-1071.

Bravata, D. M., A. L. Gienger, K. M. McDonald, V. Sundaram, M. V. Perez, R. Varghese, J. R. Kapoor, R. Ardehali, D. K. Owens, and M. A. Hlatky. 2007. Systematic review: The comparative effectiveness of percutaneous coronary interventions and coronary artery bypass graft surgery. *Annals of Internal Medicine* 147(10):703-716.

Califf, R. M., R. A. Harrington, L. K. Madre, E. D. Peterson, D. Roth, and K. A. Schulman. 2007. Curbing the cardiovascular disease epidemic: Aligning industry, government, payers, and academics. *Health Affairs* 26(1):62-74.

Coleman, E. A. 2006. The care transition intervention results of a randomized controlled trial. *Archives of Internal Medicine* 166:1822-1828.

The Commonwealth Fund. 2009. *Commission on a high-performance health system*. http://www.commonwealthfund.org/About-Us/Commission-on-a-High-Performance-Health-System.aspx (accessed August 5, 2010).

Connecting for Health. 2006. *The Connecting for Health common framework*. http://www.connectingforhealth.org (accessed February 23, 2008).

Crosson, F. J. 2005. The delivery system matters. *Health Affairs* 24(6):1543-1548.

CTSA (Clinical and Translational Science Awards). 2008. *Clinical and Translational Science Awards*. http://www.ctsaweb.org/ (accessed September 13, 2008).

Daemen, J., E. Boersma, M. Flather, J. Booth, R. Stables, A. Rodriguez, G. Rodriguez-Granillo, W. A. Hueb, P. A. Lemos, and P. W. Serruys. 2008. Long-term safety and efficacy of percutaneous coronary intervention with stenting and coronary artery bypass surgery for multivessel coronary artery disease: A meta-analysis with 5-year patient-level data from the arts, ERACI–II, MASS–II, and SOS trials. *Circulation* 118(11):1146-1154.

Davis, K., and S. Schoenbaum. 2007. *Medical homes could improve care for all*. http://www.commonwealthfund.org/Content/From-the-President/2007/Medical-Homes-Could-Improve-Care-for-All.aspx (accessed August 5, 2010).

de Brantes, F., and A. Rastogi. 2008. *Evidence-informed case rates: Paying for safer, more reliable care*. New York: The Commonwealth Fund.

Dilts, D. M., A. Sandler, S. Cheng, J. Crites, L. Ferranti, A. Wu, R. Gray, J. MacDonald, D. Marinucci, and R. Comis. 2008. Development of clinical trials in a cooperative group setting: The Eastern Cooperative Oncology Group. *Clinical Cancer Research* 14(11):3427-3433.

Dougherty, D., and P. H. Conway. 2008. The "3T's" road map to transform U.S. health care: The "how" of high-quality care. *Journal of the American Medical Association* 299(19):2319-2321.

Douglas, P. S., B. Khandheria, R. F. Stainback, N. J. Weissman, E. D. Peterson, R. C. Hendel, M. Blaivas, R. D. Des Prez, L. D. Gillam, T. Golash, L. F. Hiratzka, W. G. Kussmaul, A. J. Labovitz, J. Lindenfeld, F. A. Masoudi, P. H. Mayo, D. Porembka, J. A. Spertus, L. S. Wann, S. E. Wiegers, R. G. Brindis, M. R. Patel, M. J. Wolk, and J. M. Allen. 2008. ACCF/ASE/ACEP/AHA/ASNC/SCAI/SCCT/SCMR 2008 appropriateness criteria for stress echocardiography: A report of the American College of Cardiology Foundation Appropriateness Criteria Task Force, American Society of Echocardiography, American College of Emergency Physicians, American Heart Association, American Society of Nuclear Cardiology, Society for Cardiovascular Angiography and Interventions, Society of Cardiovascular Computed Tomography, and Society for Cardiovascular Magnetic Resonance endorsed by the Heart Rhythm Society and the Society of Critical Care Medicine. *Journal of the American College of Cardiology* 51(11):1127-1147.

Draycott, T., T. Sibanda, L. Owen, V. Akande, C. Winter, S. Reading, and A. Whitelaw. 2006. Does training in obstetric emergencies improve neonatal outcome? *BJOG: An International Journal of Obstetrics and Gynaecology* 113(2):177-182.

FDA (Food and Drug Administration). 2004. *Innovation/stagnation: Challenge and opportunity on the critical path to new medical products*. Washington, DC: Department of Health and Human Services.

———. 2006. *Critical path opportunities list*. Washington, DC: Department of Health and Human Services.

Ferguson, T. B., Jr. 2008. On the evaluation of intervention outcome risks for patients with ischemic heart disease. *Circulation* 117(3):333-335.

Ferguson, T. B., Jr., B. G. Hammill, E. D. Peterson, E. R. DeLong, and F. L. Grover. 2002. A decade of change—Risk profiles and outcomes for isolated coronary artery bypass grafting procedures, 1990-1999: A report from the STS National Database Committee and the Duke Clinical Research Institute. Society of Thoracic Surgeons. *Annals of Thoracic Surgery* 73(2):480-489; discussion, 489-490.

Ferguson, T. B., Jr., E. D. Peterson, L. P. Coombs, M. C. Eiken, M. L. Carey, F. L. Grover, and E. R. DeLong. 2003. Use of continuous quality improvement to increase use of process measures in patients undergoing coronary artery bypass graft surgery: A randomized controlled trial. *Journal of the American Medical Association* 290(1):49-56.

Frisse, M. E. 2006. Comments on return on investment (ROI) as it applies to clinical systems. *Journal of the American Medical Informatics Association* 13(3):365-367.

Frisse, M. E., J. K. King, W. B. Rice, L. Tang, J. P. Porter, T. A. Coffman, M. Assink, K. Yang, M. Wesley, R. L. Holmes, C. Gadd, K. B. Johnson, and V. Y. Estrin. 2008. A regional health information exchange: Architecture and implementation. *American Medical Informatics Association Annual Symposium Proceedings* 212-216.

Giugliano, R. P., and E. Braunwald. 2007. The year in non-ST-segment elevation acute coronary syndrome. *Journal of the American College of Cardiology* 50(14):1386-1395.

Goroll, A. H., R. A. Berenson, S. C. Schoenbaum, and L. B. Gardner. 2007. Fundamental reform of payment for adult primary care: Comprehensive payment for comprehensive care. *Journal of General Internal Medicine* 22(3):410-415.

Grover, F. L. 2008. The bright future of cardiothoracic surgery in the era of changing health care delivery: An update. *Annals of Thoracic Surgery* 85(1):8-24.

Hall, B. L., M. Hirbe, B. Waterman, S. Boslaugh, and W. C. Dunagan. 2007. Comparison of mortality risk adjustment using a clinical data algorithm (American College of Surgeons National Surgical Quality Improvement Program) and an administrative data algorithm (Solucient) at the case level within a single institution. *Journal of the American College of Surgeons* 205(6):767-777.

HHMI (Howard Hughes Medical Institute). 2009. *Med into Grad Initiative: Integrating Medical Knowledge into Graduate Education.* http://www.hhmi.org/grants/institutions/medintograd.html (accessed September 13, 2008).

Hlatky, M. A., D. B. Boothroyd, K. A. Melsop, M. M. Brooks, D. B. Mark, B. Pitt, G. S. Reeder, W. J. Rogers, T. J. Ryan, P. L. Whitlow, and R. D. Wiens. 2004. Medical costs and quality of life 10 to 12 years after randomization to angioplasty or bypass surgery for multivessel coronary artery disease. *Circulation* 110(14):1960-1966.

IOM (Institute of Medicine). 2001. *Crossing the quality chasm: A new health system for the 21st century.* Washington, DC: National Academy Press.

———. 2006. *Performance measurement: Accelerating improvement.* Washington, DC: The National Academies Press.

———. 2007. *The learning healthcare system: Workshop summary.* Washington, DC: The National Academies Press.

———. 2008. *Learning healthcare system concepts v.2008: Annual report of the Roundtable on Evidence-Based Medicine.* Washington, DC: The National Academies Press.

James, B. 2007. Feedback loops to expedite study timeliness and relevance. In *The learning healthcare system: Workshop summary*, edited by L. Olsen, D. Aisner, and J. M. McGinnis. Washington, DC: The National Academies Press. Pp. 152-162.

Johnson, K. B., C. S. Gadd, D. Aronsky, K. Yang, L. Tang, V. Estrin, J. K. King, and M. Frisse. 2008. The Midsouth eHealth Alliance: Use and impact in the first year. *American Medical Informatics Association Annual Symposium Proceedings* 333-337.

Jones, L., and K. Wells. 2007. Strategies for academic and clinician engagement in community-participatory partnered research. *Journal of the American Medical Association* 297(4):407-410.

King, J. S., and B. W. Moulton. 2006. Rethinking informed consent: The case for shared medical decision-making. *American Journal of Law and Medicine* 32:429-501.

Krumholz, H. M., Y. Wang, J. A. Mattera, L. F. Han, M. J. Ingber, S. Roman, and S. L. Normand. 2006. An administrative claims model suitable for profiling hospital performance based on 30-day mortality rates among patients with an acute myocardial infarction. *Circulation* 113(13):1683-1692.

Landon, B. E., L. S. Hicks, A. J. O'Malley, T. A. Lieu, T. Keegan, B. J. McNeil, and E. Guadagnoli. 2007. Improving the management of chronic disease at community health centers. *New England Journal of Medicine* 356(9):921-934.

Malenka, D. J., B. J. Leavitt, M. J. Hearne, J. F. Robb, Y. R. Baribeau, T. J. Ryan, R. E. Helm, M. A. Kellett, H. L. Dauerman, L. J. Dacey, M. T. Silver, P. N. VerLee, P. W. Weldner, B. D. Hettleman, E. M. Olmstead, W. D. Piper, and G. T. O'Connor. 2005. Comparing long-term survival of patients with multivessel coronary disease after CABG or PCI: Analysis of BARI-like patients in northern New England. *Circulation* 112(9 Suppl.): I371-I376.

Marmot, M., and R. G. Wilkinson. 2006. *Social determinants of health,* 2nd ed. Oxford, UK: Oxford University Press.

McCannon, C. J. 2007. Miles to go: An introduction to the 5 million lives campaign. *Joint Commission Journal on Quality and Patient Safety* 33(8):447-484.

McGlynn, E. A., S. M. Asch, J. Adams, J. Keesey, J. Hicks, A. DeCristofaro, and E. A. Kerr. 2003. The quality of health care delivered to adults in the United States. *New England Journal of Medicine* 348(26):2635-2645.

MedPAC (Medicare Payment Advisory Commission). 2008. *Report to the Congress: Reforming the delivery system*. http://www.medpac.gov/documents/Jun08_EntireReport.pdf (accessed September 14, 2008).

Meier, B. 2008. A call for a warning system on artificial joints. *The New York Times*, July 29. http://www.nytimes.com/2008/07/29/health/29iht-29hip.14863139.html (accessed September 15, 2008).

Miller, R. 2008. *The information networks required: Information technology requirements*. Presented at the Institute of Medicine Workshop Learning What Works: Infrastructure Required for Comparative Effectiveness Research. Washington, DC, July 30, 2008.

National Health Policy Forum. 2008. *Health information technology adoption among health centers: A digital divide in the making*. http://209.85.173.104/search?q=cache:obn8UTmh9Y4J:www.nhpf.org/pdfs_bp/BPHealth (accessed September 13, 2008).

Nesbitt, T. S., S. L. Cole, L. Pellegrino, and P. Keast. 2006. Rural outreach in home telehealth: Assessing challenges and reviewing successes. *Telemedicine Journal and eHealth* 12(2):107-113.

NRC (National Research Council). 2009. *Computational technology for effective health care: Intermediate steps and strategic directions*. Washington, DC: The National Academies Press.

Paulus, R. A., K. Davis, and G. D. Steele. 2008. Continuous innovation in health care: Implications of the Geisinger experience. *Health Affairs* 27(5):1235-1245.

Porter, J. P., J. Starmer, J. King, and M. E. Frisse. 2007. Mapping laboratory test codes to LOINC for a regional health information exchange. *American Medical Informatics Association Annual Symposium Proceedings* 1081.

Porter, M. E., and E. O. Teisberg. 2006. *Redefining health care: Creating value-based competition on results*. Boston, MA: Harvard University Press.

———. 2007. How physicians can change the future of health care. *Journal of the American Medical Association* 297(10):1103-1111.

Premier and CMS (Centers for Medicare & Medicaid). 2007. *Centers for Medicare & Medicaid Services (CMS)/Premier Hospital Quality Incentive Demonstration project: Project findings from year two*. http://www.premierinc.com/quality-safety/tools-services/p4p/hqi/resources/hqi-whitepaper-year2.pdf (accessed September 13, 2008).

Rittenhouse, D. R., L. P. Casalino, R. R. Gillies, S. M. Shortell, and B. Lau. 2008. Measuring the medical home infrastructure in large medical groups. *Health Affairs* 27(5):1246-1258.

Roden, D. M., J. M. Pulley, M. A. Basford, G. R. Bernard, E. W. Clayton, J. R. Balser, and D. R. Masys. 2008. Development of a large-scale de-identified DNA biobank to enable personalized medicine. *Clinical Pharmacology & Therapeutics* 84(3):362-369.

Rubin, R. 2008. CDC launches campaign to make the USA a healthier nation. *USA Today* 4D.

Schoen, C., K. Davis, S. K. How, and S. C. Schoenbaum. 2006. U.S. health system performance: A national scorecard. *Health Affairs* 25(6):w457-w475.

Singh, M., B. J. Gersh, S. Li, J. S. Rumsfeld, J. A. Spertus, S. M. O'Brien, R. M. Suri, and E. D. Peterson. 2008. Mayo clinic risk score for percutaneous coronary intervention predicts in-hospital mortality in patients undergoing coronary artery bypass graft surgery. *Circulation* 117(3):356-362.

Smith, P. K., R. M. Califf, R. H. Tuttle, L. K. Shaw, K. L. Lee, E. R. Delong, R. E. Lilly, M. H. Sketch, Jr., E. D. Peterson, and R. H. Jones. 2006. Selection of surgical or percutaneous coronary intervention provides differential longevity benefit. *Annals of Thoracic Surgery* 82(4):1420-1428; discussion, 1428-1429.

Stead, W. 2006. Providers and EHR as a learning tool. In *The learning healthcare system: Workshop summary*, edited by L. Olsen, D. Aisner, and J. M. McGinnis. Washington, DC: The National Academies Press. Pp. 268-275.

Stead, W., and J. M. Starmer. 2007. Beyond expert based practices. In *Evidence-based medicine and the changing nature of health care*, edited by M. B. McClellan, J. M. McGinnis, E. G. Nabel, and L. M. Olsen. Washington, DC: The National Academies Press. Pp. 95-105.

Taggart, D. P. 2006. Thomas B. Ferguson lecture: Coronary artery bypass grafting is still the best treatment for multivessel and left main disease, but patients need to know. *Annals of Thoracic Surgery* 82(6):1966-1975.

Tarlov, A. R., and R. F. St. Peter. 2000. *The society and population reader: A state and community perspective* (Vol. 2). New York: New Press.

Tung, R., S. Kaul, G. A. Diamond, and P. K. Shah. 2006. Narrative review: Drug-eluting stents for the management of restenosis: A critical appraisal of the evidence. *Annals of Internal Medicine* 144(12):913-919.

Wartman, S. A. 2008. Toward a virtuous cycle: The changing face of academic health centers. *Academic Medicine* 83(9):797-799.

Weinstein, J. N., K. Clay, and T. S. Morgan. 2007. Informed patient choice: Patient-centered valuing of surgical risks and benefits. *Health Affairs* 26(3):726-730.

Wennberg, J. E., E. S. Fisher, and J. S. Skinner. 2002. Geography and the debate over Medicare reform. *Health Affairs* (Suppl. Web Exclusive):w96-w114.

Wennberg, J. E., E. S. Fisher, J. S. Skinner, and K. K. Bronner. 2007. Extending the p4p agenda, part 2: How Medicare can reduce waste and improve the care of the chronically ill. *Health Affairs* 26(6):1575-1585.

WHO (World Health Organization). 2008. *Social determinants of health*. http://www.who.int/social_determinants/en/ (accessed September 13, 2008).

Wilper, A. P., S. Woolhandler, K. E. Lasser, D. McCormick, D. H. Bor, and D. U. Himmelstein. 2008. A national study of chronic disease prevalence and access to care in uninsured U.S. adults. *Annals of Internal Medicine* 149(3):170-176.

Woolf, S. H. 2008. The power of prevention and what it requires. *Journal of the American Medical Association* 299(20):2437-2439.

6

Moving Forward

INTRODUCTION

Expanded capacity to develop and apply insights from comparative effectiveness research (CER) is central to providing better information for healthcare decisions of patients and their caregivers, understanding how to appropriately tailor care of individual patients, and ensuring greater overall value of health care delivered. The wide range of challenges and needs—with respect to the work, information and data networks, and workforce required for CER—discussed at the workshop helped identify some initial priorities for immediate action. The pace of progress depends on aligning incentives, encouraging needed culture change, and developing tools to better support both the generation of insights from care through CER and the seamless application of findings in clinical decision making.

Given the large scope and scale of transformation needed, moving forward will require a long-term strategy that prioritizes and sequences needs, engages all stakeholders, and builds sustained, cross-sector support. The final session of the workshop featured discussion of key elements of a roadmap, quick hits, and opportunities to build support. This chapter includes a synthesis of this session's discussion, as well as an overview of the workshop's common themes and possible follow-up actions to be considered for ongoing multistakeholder involvement through the Institute of Medicine (IOM) Roundtable on Value & Science-Driven Health Care.

THE ROADMAP—POLICIES, PRIORITIES, STRATEGIES, AND SEQUENCING

Maximizing the returns on the considerable investment of time, human resources, and money inherent to expanding CER capacity will require an understanding of the broad range of policy needs, how existing resources might be best deployed, and a sense for how quickly various elements can be taken to the scale needed. Stuart Guterman, senior program director for the Commonwealth Fund's Program on Medicare's Future, outlines six broad policy areas in need of consideration: data, methods, workforce, organization, translation, and financing. Guterman noted that a roadmap for progress would identify clear end goals for each area, priority needs within and between categories, and key actors or existing infrastructure that could help initiate the activities needed. The following summarizes key points from workshop discussions to provide a starting point.

- *Data.* Capacity is needed to produce data relevant to healthcare decision making by providers, patients, payers, and policy makers; to ensure that the value of data is maximized by integrating these data and establishing sustainable system linkages; and to develop systems that make data and information available to appropriate users when, where, and how they are needed. While these data exist to some extent, attention needs to be focused on current gaps and how emerging health information technologies (HITs) might be applied to meet new data needs. As electronic health record (EHR) capacity is adopted across the nation, the focus should be on ensuring their usefulness—beyond billing and administration—for research and decision support. For existing resources, work to integrate and link disparate data sources and develop standards will provide an important foundation for progress, but enhanced data sharing among stakeholders will require attention to privacy and data stewardship issues.
- *Methods.* Research approaches are needed that meet the needs of CER end users by providing timely information that is relevant to real-world patients and decisions faced at the point of care. Priority needs include a better understanding of current methods most useful for particular questions, the development of methodological standards for these approaches, and investment to accelerate new innovative methods that enable real-time analysis and learning. Infrastructure to support methods development and to streamline the conduct of clinical research will help accelerate this work.
- *Personnel development.* A cadre of professionals is needed—from across healthcare sectors—dedicated to and trained in the use of

tools and techniques for developing and applying comparative effectiveness information. Current training and education programs need to be reevaluated on how they might be best aligned with the workforce needs of CER, and investment could be targeted to organizations and investigators that foster the type of multidisciplinary and translational activities characteristic of CER. Clarification of the scope and anticipated scale of research needs will help guide resource use.

- *Organization.* Prioritization and coordination across the many organizations engaged in various aspects of evidence development—primary research, synthesis, and translation—will enable more efficient production of needed information. Assessment of capacities that would most benefit from centralized coordination is paramount, as is the exploration of the governance and funding structures that might facilitate greater public–private collaboration around issues of common concern.
- *Translation.* Additional focus is needed to ensure evidence is applied in healthcare decision making. A focus on processes, tools, and related capacities that ensures the broader application of evidence in clinical decision making might begin with an immediate focus on guidelines development processes. Also suggested are efforts to bring research closer to the practice environment and enable learning in the postmarket environment.
- *Financing*. Sufficient and sustained funding will be needed to establish and support CER and its application as an integral part of the U.S. healthcare system. While funding mechanisms will depend in part upon the nature of any entity charged with this work, immediate attention can be focused on other funding mechanisms, such as incentives and grants, that might be better structured to support advances in the development and application of comparative effectiveness information.

Comments by three respondents highlighted specific opportunities. Elaine Collier, assistant director for clinical research at the National Center for Research Resources, suggested that the National Institutes of Health (NIH) Clinical and Translational Science Awards (CTSA) program could play a role in expanded CER by virtue of its fundamental focus on infrastructure development targeting a variety of areas: using informatics to advance research; engaging physicians, patients, and other stakeholders in developing a better understanding of how advances in health care are made; training the next generation of investigators and supporting professional development; transforming processes and making them more effective across institutions; and emphasizing rigor and quality of studies in

study design, biostatistics, and clinical research ethics. The CTSA program could help foster the development of transformational strategies locally at individual institutions as well as help disseminate these successful strategies at a national level.

Jane Horvath, senior director for global public policy at Merck, suggested giving attention to ensuring adequate support for the many infrastructure needs identified. The development of a national electronic data system for health care, for example, requires the development of a sound business model and funding through private channels or consensus on such a resource as a social good to be supported largely by public funds. Device and pharmaceutical manufacturers can contribute to CER and help to fill evidence gaps, but emphasis is also needed on supporting the innovation these companies bring to the system. Changes in public policies, such as patent law, licensure, and Stark regulations, might help to support CER by stimulating premarket research and the generation of more data prior to a product being licensed.

Bruce H. Hamory, executive vice president and chief medical officer emeritus at Geisinger Health System, suggested that guidelines are needed that better ensure that evidence development results in lowered costs and better health care in the real world. Existing registries for cardiovascular disease were cited as resources that truly inform patient decision making, and Hamory emphasized the central importance of resolving data governance and privacy regulation issues to facilitate the development of similar resources. Other priorities from his perspective included the development of new models for information sharing, infrastructure for CER-related training, accelerated translation of studies into evidence, and models of financing tied to patient benefits and cost reductions.

COMMON THEMES IN WORKSHOP DISCUSSIONS

Common themes reoccurring in the 2 days of discussion are summarized in Box 6-1, and elaborated in the text that follows:

- ***Care that is effective and efficient stems from the integrity of the infrastructure for learning.*** The number of medical diagnostics and treatments available to patients and caregivers is increasing, but the knowledge about their effectiveness—in particular, their comparative effectiveness—is not keeping pace. This is in part a function of the rate of change, but it is also a product of capacity that is both underdeveloped and, as several participants noted, substantially fragmented, which leads to gaps, inefficiencies, and inconsistencies in the work. The accelerating rate of change in the interventions requiring effectiveness assessment compels a substantial shoring up

in the level of effort, the nature of the effort, and the coordination of the effort in order to produce the necessary insights into the right care for different people under different circumstances.

- ***Coordinating work and ensuring standards are key components of the evidence infrastructure.*** Several presentations highlighted the point that substantial activity is currently under way in effectiveness research, including work on comparative effectiveness, but the activities are fragmented and often redundant in both structure and function. The fact that the application of evidence lags behind its production is in part a function of the disparate and "siloed" approaches between and within organizations seeking and developing information. The notions of infrastructure for evidence development therefore also include the capacity for greater coordination in the setting of study priorities; the development of

BOX 6-1
Infrastructure Required for Comparative Effectiveness Research: *Common Themes*

- Care that is effective and efficient stems from the integrity of the infrastructure for learning.
- Coordinating work and ensuring standards are key components of the evidence infrastructure.
- Learning about effectiveness must continue beyond the transition from testing to practice.
- Timely and dynamic evidence of clinical effectiveness requires bridging research and practice.
- Current infrastructure planning must build to future needs and opportunities.
- Keeping pace with technological innovation compels more than a head-to-head and time-to-time focus.
- Real-time learning depends on health information technology investment.
- Developing and applying tools that foster real-time data analysis is an important element.
- A trained workforce is a vital link in the chain of evidence stewardship.
- Approaches are needed that draw effectively on both public and private capacities.
- Efficiency and effectiveness compel globalizing evidence and localizing decisions.

systematic decisions for the conduct of CER, systematic reviews, and guideline development; and ensuring the consistent translation of developed information. The identification of priority conditions, evaluation, and evidence gaps is needed in order to target limited resources, especially for high-cost or high-volume procedures and interventions.

- ***Learning about effectiveness must continue beyond the transition from testing to practice.*** "The learning process cannot stop when the label is approved," one meeting participant pointed out. Premarket testing for the safety and effectiveness of various interventions cannot assess the results for all populations or the circumstances of use and differences in practice patterns, so gathering information as interventions are applied in practice settings should represent a key focus in designing the infrastructure to learn which care is best. Local coverage decisions and private insurer use of coverage with evidence development approaches were cited as opportunities to learn as a part of the care process.
- ***Timely and dynamic evidence of clinical effectiveness requires bridging research and practice.*** The historical insulation of clinical research from the regular delivery of healthcare services evolved to facilitate data capture and control for confounding factors. With the prospect of electronically enhanced data capture, and statistical approaches to improve analysis, as well as increasing demand to keep pace with technologic innovation, this divide increasingly limits the usefulness of research results. Efforts are under way to better engage health delivery organization, practitioners, patients, and the community in research prioritization, conduct and results dissemination.
- ***Current infrastructure planning must build to future needs and opportunities.*** Research is often driven more by the methods than the questions. In fact, both are important, and infrastructure planning must account for both the key emerging healthcare questions and the key emerging CER opportunities. Emerging questions include those related to the management of multiple co-occurring chronic diseases of increasing prevalence in an aging population, the improved insights into individual variation relevant to both treatments and diagnostics, and the impact of innovation in shortening the lifecycle of any particular intervention. Emerging tools include innovations in trial design, the development of new statistical approaches to data analysis, and the development of electronic medical and personal health records.
- ***Keeping pace with technological innovation compels more than a head-to-head and time-to-time focus.*** Much of the current discus-

sion about CER has emphasized the need for more clinical trials and more head-to-head studies. Although there are numerous examples of diagnostic and treatment interventions for which such studies are needed, the notion of a research process that essentially offers periodic and static determinations is inherently limited. Especially with the rapid pace of change in the nature of interventions and the difficulty, expense, and time required to develop studies—and the challenges of ensuring the generalizability of results in the face of limitations of traditional approach to randomized controlled trials (RCTs)—a first-order priority for effectiveness research is the establishment of infrastructure for a more dynamic, real-time approach to learning. Leveraging new tools, such as HIT, should allow for a more networked and distributed approach to information sharing and evidence creation.

- ***Real-time learning depends on HIT investment.*** It was noted that collecting data is the most time-intensive part of trials and studies, and information technology (IT) is critical to streamlining this work. Moreover, it is the key to accelerated learning from broader-based clinical experience. We heard that "[t]he type of learning needed cannot be paper based." The increasing complexity of the factors involved in understanding the effectiveness of clinical options under different circumstances requires a blend of database access and computing power that can only be provided from broadly applied HIT. Although not in itself sufficient to produce the information required for better medical care management, it is a necessity for the continuous improvement expected of a learning health system. A policy framework for privacy and security will be necessary to build and maintain public trust that information will be protected as it is shared.
- ***Developing and applying tools that foster real-time data analysis is an important element.*** The scope and scale of evidence needs suggests that innovation is needed across the range of research methods, from making clinical trials faster and less expensive to moving beyond RCTs to better address practical circumstances, using registries, observational databases, and other emerging data resources. If full advantage is to be taken of HIT, statistical tools and analytic algorithms that can be embedded in databases to allow real-time insights will be important. Similarly, tools are needed that will allow findings to be drawn from databases built on different vendor platforms, using semantic technology to integrate currently disparate medical data, in order to develop the next generation of statistical tools for the analysis of clinical data, including the

building of models that allow insights to be generated by virtual studies.

- ***A trained workforce is a vital link in the chain of evidence stewardship.*** As in many other domains, progress in CER will relate to the capacity to develop and maintain the broad and diversely skilled workforce needed. Mention was often made of that factor as a determining element as well for progress in development of the learning health system. Given the pace of change in the number and variety of clinical interventions as well as in the tools and approaches to assessing them, there is a need to ensure that these developing opportunities are matched by the skills of the workforce. This includes training and education in the methodologies of research design, translating research, guideline development, and maintaining and mining clinical records. But it also includes attention to reorienting the education of frontline caregivers around their emerging responsibilities for access, interpretation, and discussion with patients of a dynamic evidence base, as well as helping to ensure the availability and integrity of the clinical data that shape conclusions on evidence.
- ***Approaches are needed that draw effectively on both public and private capacities.*** Several times in the course of the meeting it was pointed out that although the total investment in CER in the United States is substantial, it is inefficient because of the absence of a vehicle for common priority setting and coordination of efforts and because the work on effectiveness done by private companies in product development and testing is usually not accessible to the broader community. In aggregate, private investment often far exceeds public investment in assessing a given intervention, but even when available, studies associated with an enterprise with a commercial stake may be viewed suspiciously. Several models are in development to establish public–private collaborative efforts to improve the efficiency and effectiveness of the work.
- ***Efficiency and effectiveness compel globalizing evidence and localizing decisions.*** Two presentations featured international work, including the Cochrane Collaboration on evidence synthesis, and efforts in Ontario, Canada, to develop and apply insights about the comparative effectiveness of clinical interventions. Reference was made throughout the meeting to work going on elsewhere in the world and, in particular, to work at the National Institute for Health and Clinical Excellence in the United Kingdom. This brought clearly into play the need to ensure that, where possible, common work to assess an intervention's clinical effectiveness—or collective work to assess the body of evidence—be collaborative

and well coordinated across boundaries, while also being mindful that different cultural and policy environments may lead to different decisions at the local level.

KEY FACTORS AND NEEDS

Workshop speakers described a number of implications of the current state of play for the development of an infrastructure for CER (Box 6-2). These included the following:

- ***Several elements are involved in infrastructure development.*** Developing the infrastructure for CER has at least five dimensions: putting in place the physical capacity (i.e., the hardware); developing the analytic tools and methods; training the workforce; establishing processes for efficient and effective operation; and shaping the strategy for attention and phasing. Presentations at the meeting described and discussed in qualitative terms the needs and challenges in each of these dimensions and offered "opening bid" quantitative estimates on the needs for the IT infrastructure, as well as for investments in human capital. Refinements of these first approximations will be needed, as will additional clarity on the

BOX 6-2
Key Factors and Needs for Expanded Comparative Effectiveness Research Capacity

- Several elements are involved in infrastructure development:
  - putting in place the physical capacity (i.e., the hardware),
  - developing the analytic tools and methods,
  - training the workforce needed,
  - establishing processes for efficient and effective operation, and
  - shaping the strategy for attention and phasing.
- Strategies and priorities for infrastructure application include the following:
  - conduct of systematic reviews,
  - conduct of primary research,
  - clinical registry resources,
  - introduction of health information technology throughout practice,
  - fostering public and private collaboration, and
  - keeping focus on the utility and impact of a networked world.

analytic tools, processes, and strategies for a stronger infrastructure for research into effective health care.

- ***Strategies and priorities for infrastructure application.*** The dimensions noted above represent in certain ways the functional dimensions of relevance to the infrastructure that is needed for effectiveness research. There are phasing considerations as well, in part driven by the ability and need to take actions even without additional resources and in part driven by the time required to set in motion the necessary activities. Suggestions for key strategies and priorities for progress included the following:
  - ***Conduct of systematic reviews.*** There is an immediate need to improve the conduct, coordination, and consistency of systematic reviews—a point that, in effect, echoed the recommendations of the 2008 IOM report *Knowing What Works* in *Health Care: A Roadmap for the Nation*, presented by a member of that committee.
  - ***Conduct of primary research.*** Similarly, the approach to primary research on effectiveness needs a more systematic means of determining priorities, better tools and more streamlined designs, and a deeper bench workforce to do the work.
  - ***Clinical registry resources.*** In making the transition to a pattern of real-time, continuous learning in health care—in effect, creating a beta approach to clinical data systems that generate learning—the technologies for clinical registries and in the field of registry development, maintenance, and improvement will need to be strengthened.
  - ***Introduction of HIT throughout practice.*** In the area of IT development, the issues include the need to install appropriate hardware in virtually every clinical setting, the incorporation into operating software of design elements that are pegged to research activities and embedded analytic tools, the incorporation of design elements used in decision assistance, and training of the required workforce to work with this technology.
  - ***Fostering public and private collaboration.*** A longer-term development needed to sustain the growth and improvement of the infrastructure will include the design of approaches that foster meaningful public and private collaboration in support of the research activities.
  - ***Keeping focused on the usefulness and impact of a networked world.*** Also important to guide strategy development for the long term are approaches designed to take advantage of the resources emerging in our increasingly networked world—the opportunities for which hints are provided by recent develop-

ments, such as the Patients Like Me Web site, the health maintenance organization research group, the registries of the Society of Thoracic Surgeons, and even information made available by such resources as Google and Wikipedia.

QUICK HITS—THINGS THAT CAN BE DONE NOW

Long-term support for CER will benefit from quick hits or actions that can be taken now with near-term results that might help demonstrate the benefit of CER. To open the session, W. David Helms, president and CEO of AcademyHealth, noted several opportunities for collaborative efforts by stakeholders to lay the groundwork for a national capacity for CER. These efforts should target accelerating congressional action to establish a platform for CER and increasing federal funding for CER; articulating the case for CER (e.g., developing a database that better characterizes the volume and costs of current CER work, drawing upon Health Services Research Projects in Progress database, clinicaltrials.gov, the Cochrane Collaborative, and others); examining models for national capacity (e.g., the Canadian Health Services Research Foundation's $100 million support for enhanced knowledge transfer); and educating state policy representatives and Medicaid officials about the potential and needs for CER. Work can also begin immediately to build up the needed workforce—through a reexamination of funding and support streams, alignment of education and training programs around a broad array of research methods, and development of the means for improved communication across disciplines (e.g., clarification of terminology and methods). [Note: Subsequent to this meeting, Congress increased the national capacity for CER with the establishment, in the ACA of 2010, of Patient-Centered Outcomes Research Institute, previously described.]

Two respondents also offered recommendations, and a summary of these comments. Lynn Etheredge, a consultant with the rapid learning project at George Washington University, suggested an immediate focus on new technologies. While potentially offering new benefits to medical care, collectively these products are also a primary driver of healthcare cost increases. A national system to collect and analyze information about these products in the postmarket environment would enhance our understanding of safety and effectiveness is real-world populations. Existing infrastructure that might be used to accelerate the development of such a system includes the authority conferred to the Food and Drug Administration for the Sentinel System, and the Centers for Medicare & Medicaid Services (CMS) Coverage with Evidence Development policies. A national research plan could guide requirements for reporting to a new national data registry that would collect from the point at which a prod-

uct or drug is first introduced into the market or is covered by Medicare. Second, Etheredge noted the need to develop a national biobank initiative to begin work to relate genetic information with clinical data collected via EHRs. This will potentially improve diagnoses as well as our understanding of the heterogeneity of treatment responses. A national biobank could be based in large part on data resources currently in place, including the NIH compilation of genomewide association studies and genetic data available from major healthcare providers. Similar resources could be constructed for specific diseases, drawing on the considerable progress already evident in the development of such resources as the Alzheimer's Disease Neuroimaging Network and the cancer Biomedical Informatics Grid. Progress could also be accomplished in linking Medicaid data on disabled and chronically ill patients nationally. Another potential model is the Oregon Community Health Information Network, which is engaged in bringing Electronic Privacy Information Center applications to its safety-net clinics. Third, Etherege cited the huge potential in developing programs focused on Medicaid populations. Nationally, databases and EHRs will cover most populations except the Medicaid, disabled, and high-needs populations. Given the billions of dollars spent each year on dual eligibles (Medicare and Medicaid users), a small initial investment in better understanding the care of this population is needed. He suggested that starting up research registries and databases using just some of the states that could look in detail at the seriously disabled and chronically ill populations would be a feasible project with immediate returns.

David Shulke, executive vice president of the American Health Quality Association, suggested that the existing national network of Medicare Quality Improvement Organizations (QIOs) could be used to ensure the translation and application of evidence into practice. Currently providers and practitioners find it difficult to find, integrate, and use new data in clinical practice. The QIOs are a national network of organizations or private contractors that facilitate the adoption of evidence-based medicine and could easily also be used to facilitate the use of comparative effectiveness information. He urged that this be factored into the next Medicare contract cycle, which starts in 2011.

Other suggestions during the open discussion included a demonstration project on the use of personal health records to promote patient involvement in the management of their own chronic disease and an exploration of evidence available in other countries. The Independent Drug Information System—a stakeholder partnership that engages providers, insurers, and patients—and the Pragmatic Approach to Comparative Effectiveness group, which is exploring the use of Bayesian techniques to streamline the efficient generation of new knowledge, were both suggested as important new initiatives that might serve as models for various aspects of CER capacity.

Issues for Possible Roundtable Follow-Up

Throughout the course of discussions, a number of items were identified as candidates for follow-up attention by the Roundtable on Value & Science-Driven Health Care:

- ***Better characterization of the elements of the infrastructure.*** Building on the work sponsored by the Roundtable on workforce needs and IT infrastructure, continue to improve the initial estimates and pursue similar assessments related to requirements for new analytic tools and methods, establish processes for efficient and effective operation of the fields of work, and shape the strategy for attention and phasing. Include examples of effective work at the institutional level.
- ***Clarification of the nature of the "prework" needed for a more systematic approach to the necessary RCTs.*** Even though a more practical portfolio of research approaches is essential, the RCT offers the key standard for the rigor required for certain circumstances. Their most effective deployment requires attention to issues of the criteria indicating the need for an RCT, the issues and priorities to be assessed, the best structure of the research questions, and improved approaches to trial design, conduct, and data collection.
- ***More focus on the infrastructure needed for guideline development, implementation, and evaluation.*** Several issues could be productively engaged, including transparency and collaboration across professional groups on improving consistency in the methods, standards, rules, and participants in guideline development and approaches to implementation.
- ***Share meeting discussions with organizational stakeholders in elements of the infrastructure.*** Examples given included the National Quality Forum, the Association of American Medical Colleges, the Association of Academic Health Centers, the Quality Improvement Program, and CMS/Department of Health and Human Services in the context of development of the 10th QIO statement of work, the American Hospital Association Quality Forum, International Society for Pharmacoeconomics and Outcomes Research, and provider groups.
- ***Devote additional attention to data stewardship issues.*** Because the basic resource for effectiveness research is the clinical data system, the Roundtable needs to catalyze more discussion on the integrity of this resource, including issues of maintenance, privacy, and data ownership.

- ***Identify possible incentives.*** Look at how subsidies and reimbursement regulations can stimulate increased use of HIT in medical care, increased use of HIT for application of evidence, and increased use of HIT for the development of evidence.
- ***Expand engagement of the business case and demand function for infrastructure investment.*** Give additional attention to the economic or business case for employers to appreciate the investment and its necessity in improving value from health care, the case for more attention by states, the case for the personal health record deployment to drive more patient–provider interaction, and work on the consequences of not investing.
- ***More focus on the issues of strategies and infrastructure for implementing findings on effectiveness.*** Since evidence is virtually useless if not applied, the Roundtable could give more attention to understanding the infrastructure needs for effective guideline implementation.
- ***Sponsor discussions on training and health professions education reorientation.*** With greater appreciation for team-based, networked information stewardship roles by caregivers, the health professions groups should be recruited for collaborative consideration of the training implications.
- ***Provide information on the Roundtable's Web site.*** The resources of the workshop presentations and discussions should be posted on the Web site—slides, links, and speaker contact information.

Building Support

Although an enhanced focus on CER will build upon the existing infrastructure and activities, it still marks a significant shift in the nation's approach to clinical research and practice. Healthcare stakeholders generally view CER as an important tool for ensuring that healthcare decisions are based on the best science; but additional work is needed to effectively communicate between stakeholders and with the public and policy makers about needed investments and potential returns from CER. Mary Woolley from Research!America led an open discussion session on opportunities to build support for the wide-ranging investments and developments articulated throughout the workshop. Her comments along with suggestions offered by three respondents are summarized below.

Four fundamental requirements for building support are important to consider: (1) clarity about the ultimate goal, (2) understanding the target audience, (3) ensuring all stakeholders are involved, and (4) understanding the context. Using this framework, Woolley offered several suggestions on key opportunities to build support for expanded development and use of

CER. Much of the discussion about needs for infrastructure requires an inside knowledge of the specific aspects of health care. However, a simply stated goal, such as "By 2020 at least 90 percent of all clinical decisions will be based on the best available evidence," is fairly easy to understand and might therefore serve as a better way to frame the many needs articulated during this workshop to all stakeholders—including the public and policy makers. Making the case will face challenges as CER is in some sense a fundamental change in current and long-held practices in research and health care. Developing support will therefore require broad consensus and clarity about the fundamental CER value proposition and agreement about core goals. Strategies for change must therefore factor in the need for value acceptance and culture change.

Also necessary is clarity about the target audiences, as effective communications need to be tailored to the interests and concerns of different stakeholders. For each audience, there are various keys to building support: anticipating questions, but also listening carefully to questions posed in order to better understand the needs of a particular audience; engaging in clear communication and crisp, well-conceived messaging; and keeping in mind that those not engaged in CER-related work, specifically including the public and policy makers, are not as well versed as stakeholders in the terminology and concepts of CER. Woolley offered an example from personal experience: that *health outcomes*, while a commonly used term within healthcare policy, was simply not understood well by the media and other audiences. A simple language shift to "better health" was more readily understood and perceived as speaking directly to the public's interests. Personal stories, effective metaphors, sound bites, and strong, crisp messages need to replace large reports laden with jargon.

Immediate wins, or quick hits, that help to illustrate the potential of CER will also be essential to building support. Understanding public opinion (which needs to be constantly gauged) and the context for communications is also important. As of July 2008, research has demonstrated that the American public continues to be concerned about the cost of health care, and ranks it as the top long-term challenge facing the nation. How might communications about healthcare reform and CER better reflect or illustrate key opportunities to address these concerns?

Communication should not be unidirectional, but rather it should be structured to fully engage all stakeholders involved in infrastructure building. Such broad engagement will require the identification and strong support of champions who can lead the effort as well as the commitment of all interested stakeholders in taking action. Educating legislators about the importance and relevance CER is one aspect of this work, but speaking to other audiences—including those in our own social networks—will be instrumental in developing the broad-based support needed. Fine-tuning

communications strategies as lessons are learned about their effectiveness will strengthen the overall effort.

Respondent Stephen Gorshow, from the Aetna coverage policy unit, observed that individual participation in health plans inherently creates a unique opportunity for the plan to follow members over the continuum of health care. He suggested, however, that the fact that individuals obtain care from multiple providers creates gaps in the quality of care, an issue that calls for research to determine the extent and impact of this problem, albeit with adherence to principles of patient privacy. Particularly in an era of consumer-directed health plans, and toward a goal of encouraging healthy practices among individuals, research is also needed to determine the effects that changes in benefit design have on health-related behaviors. Similarly, benefit issues need to be researched in ways that inform intelligent decisions about what should and should not be covered. In addition, research should be conducted to determine the optimal balance between cost-sharing by health plan members and protection for the plan against the high cost of the high-tech care available today for more serious illnesses. Also advocated is a broader buy-in to available guidelines for wise use of high-tech radiology as well as study of the driving forces behind the need for private insurers to conduct programs to stem the overuse of such technologies. These areas of needed research are directly relevant to how patients experience health care, and they may therefore also be important in improving communications with patients about the benefits of comparative effectiveness.

A second respondent, David Longnecker, director of healthcare affairs at the Association of American Medical Colleges, emphasized the importance of having clear, distinct goals as an important tool for building support for CER. Large, big-picture goals are necessary, but smaller interim goals can also help drive rapid progress. In this respect, while it is important to develop a research enterprise that can begin to narrow the many gaps in evidence, streamlining and supporting the translation of research into practice might be an area where progress can be made quickly. An important interim goal might therefore be to significantly improve the application of known evidence; to that end, research is needed on how to best motivate and support physicians and health professionals to implement evidence-based care. In addition, Longnecker suggested some "quick hits" that might be helpful in building support. The Department of Veterans Affairs EHRs, National Consortium of Clinical Databases, and the Dartmouth Atlas databases are existing data resources that could be mined now to expand comparative effectiveness knowledge. Also, citing how the work of the Leapfrog group in improving quality was greatly accelerated by CMS payment policies, Longnecker suggested that strategies are needed, particularly financial rewards for practitioners, to help move CER forward.

A third respondent, Eva Powell, director of the Health Information Technology Project at the National Partnership for Women and Families, observed that while consumer advocacy organizations are in favor of moving CER forward, the issue of preserving and protecting patient privacy rights remains of paramount importance. Indeed, she argued, the fundamental validity of data is dependent on privacy being protected. Powell advocated for reframing the conversation around patient privacy, shifting it from a focus on barriers to participation in studies to a mindset that underscores the essential importance, for the greater good, of study participation. To ensure transparency in this realm, an agreed-on set of standards and a policy framework that covers all participating entities is required, as are changes in Health Insurancc Portability and Accountability Act legislation.

Consumers do not have all the information or the simple tools needed to become fully engaged in CER. Some work is already under way to educate and communicate with the public and consumers, but additional efforts are needed to ensure that the media is also educated. Because of their work with local constituencies, QIOs are possible conduits to the public and the media on health and health quality issues. In developing messages about the benefits of CER in terms of its value to consumers, a value case from the consumer perspective needs to be articulated and disseminated. Important to understand in this regard is that making such a case on the basis of improving efficiency or cost does not resonate with consumers, as efficiency is often perceived by consumers as an attempt to deny them care and, while recognizing that cost is important, consumers tend to be wary of decision making based solely on cost. Value in health care, which includes consideration of the benefits received by patients, is more likely to resonate with consumers.

Appendix A

Learning What Works Best: The Nation's Need for Evidence on Comparative Effectiveness in Health Care

AN ISSUE OVERVIEW

IOM ROUNDTABLE ON EVIDENCE-BASED MEDICINE

September 2007 version. This Issue Overview was prepared at the request of the IOM Roundtable Working Group on Sustainable Capacity by J. Michael McGinnis, LeighAnne Olsen, Katharine Bothner, Daniel O'Neill, and Dara Aisner.

MARCH 2009 UPDATE

The American Recovery and Reinvestment Act of 2009

In the time since the preparation of this white paper, $1.1 billion of federal funds have been provided by Congress, through the American Recovery and Reinvestment Act of 2009 (ARRA), to increase national capacity for clinical effectiveness research. AHRQ (Agency for Healthcare Research and Quality) has received $700 million of these funds, of which $400 million will be transferred to the Office of the Director of NIH (National Institutes of Health) to conduct or support comparative effectiveness research (CER) activities.

An additional $400 million will be allocated at the discretion of the Secretary of HHS (Department of Health and Human Services) to:

> "...accelerate the development and dissemination of research assessing the comparative effectiveness of health care treatments and strategies, through efforts that: (1) conduct, support, or synthesize research that compares the clinical outcomes, effectiveness, and appropriateness of items, services, and procedures that are used to prevent, diagnose, or treat diseases, disorders, and other health conditions; and (2) encourage the development and use of clinical registries, clinical data networks, and other forms of electronic health data that can be used to generate or obtain outcomes data."

The recommendations from an Institute of Medicine consensus committee report and from a newly established Federal Coordinating Council on CER within HHS will be considered by the secretary's office in designating activities to receive funds. Members of the 15-member council will be federal employees or officers appointed by the President, at least half of which will have clinical expertise.

LEARNING WHAT WORKS BEST
THE NATION'S NEED FOR EVIDENCE ON COMPARATIVE EFFECTIVENESS IN HEALTH CARE

Contents

Acknowledgments

This Issue Overview on current and needed capacity for comparative effectiveness research was developed by staff of the IOM Roundtable on Evidence-Based Medicine, initially as background material for the activities of the Roundtable's Sustainable Capacity Working Group, and to inform discussion at the March 19, 2007 Roundtable meeting.

Guidance on structure and content were provided by members of an advisory group including: Carmella Bocchino (AHIP), Queta Bond (Burroughs Wellcome Fund), Kathy Buto (Johnson & Johnson), Steve Galson (FDA), Mark McClellan (AEI-Brookings Joint Center for Regulatory Studies), Lisa Payne-Simon (Blue Shield of California Foundation), Diana Petitti (University of Southern California), Jean Slutsky (AHRQ), Sean Tunis (Center for Medical Technology Policy), and Gail Wilensky (Project Hope).

We also extend special thanks to those who took the time to review and comment on various sections or draft versions of this paper, including: Wade Aubry (HealthTech), Tanisha Carino (Avalere Health), Nancy Featherstone (AstraZeneca), Mark Gibson (OHSU), Carmen Hooker Odom (North Carolina HHS), Michael Johns (Emory), Peter Juhn (Johnson & Johnson), Doug Owens (Stanford), Steve Pearson (AHIP), Steve Phurrough (CMS), Eugene Rich (House Committee on Ways and Means, Health Subcommittee), Wayne Rosenkrans, Jr. (AstraZeneca), and Jeff Shuren (FDA).

IOM staff who contributed include: Dara Aisner, Katharine Bothner, Michael McGinnis, LeighAnne Olsen, and Daniel O'Neill.

Roundtable sponsors: The work of the Roundtable is supported by the

Agency for Healthcare Research and Quality, AHIP (America's Health Insurance Plans), AstraZeneca, Blue Shield of California Foundation, Burroughs Wellcome Fund, California Healthcare Foundation, Centers for Medicare and Medicaid Services, Department of Veterans Affairs, Food and Drug Administration, the HWG Fund, Johnson & Johnson, sanofi-aventis, and Stryker.

Suggested citation Institute of Medicine. 2007. ***Learning what works best: The nation's need for evidence on comparative effectiveness in health care***. http://www.iom.edu/ebm-effectiveness.

IOM ROUNDTABLE ON EVIDENCE-BASED MEDICINE
Working Group on Sustainable Capacity*

LEARNING WHAT WORKS BEST
The Nation's Need for Evidence on Comparative Effectiveness in Health Care

INTRODUCTION

A core objective for the nation is achieving the best health outcome for every patient. This objective simply cannot be accomplished until we have better evidence on which to base healthcare decisions, as well as more effective application of the knowledge we have. Each is vitally important. We know, for example, that failure to deliver proven interventions is a substantial challenge to the quality of health care for Americans—and is a key concern of the IOM Roundtable on Evidence-Based Medicine. Yet, with the current pace of change, the most rapidly growing problem is our inability to produce the needed evidence in a timely fashion. This paper provides background for discussion about the evidence gap—the fact that the nation's capacity has fallen far short of the need to produce reliable and practical information about the care that works best. Medical care decision-making is now strained, at both the level of the individual patient and the level of the population as a whole, by the growing number of diagnostic and therapeutic options for which evidence is insufficient to make a clear

* Jack Rowe, Columbia University (Chair); Adam Bosworth, Google; Helen Darling, National Business Group on Health; Michael Johns, Emory University; Steve MacMillan, Stryker Corporation; Mark McClellan, AEI-Brookings; Richard Platt, Harvard University; Steve Udvarhelyi, Independence Blue Cross; Bill Weldon, Johnson & Johnson; Janet Woodcock, FDA. The material here is a staff paper prepared at the request of the working group. Information on the IOM Roundtable on Evidence-Based Medicine may be obtained at www.iom.edu/ebm.

choice. Biomedical insights and medical innovation continue to advance opportunities to increase the health and life-span of the American public, yet to capitalize fully on this potential requires enhanced capacity to ensure that decisions, in the face of increasing complexity, can be supported and guided by the best available scientific information.

Health care in the United States underperforms on many dimensions. At the macro level, with per capita expenditures more than 20 percent higher than any other country in the world and more than twice the average expenditure for European countries[1], the nation ranks well below others on key health indices—28th in overall life expectancy at birth and 23rd in infant survival [2, 3]. In part this is because people often do not receive the care they need. One study found that, where evidence exists, only about 55 percent of recommended services were actually delivered [4]. In part it is also because the services people receive are not always necessary or the right ones for them. The intensity of services for similar conditions with similar results—in particular, for procedures such as lumbar surgery, hysterectomy, and bypass surgery, where discretion plays a stronger role—can vary by as much as a factor of 20 depending simply on where one lives. In Idaho Falls, Idaho, 4.6 lumbar fusions were reported per 1,000 Medicare enrollees annually, as compared with 0.2 in Bangor, Maine, with no difference in the outcomes [5]. Similarly, wide geographic variations have been reported for conditions such as hip fracture, colorectal cancer, and acute myocardial infarction as well as end-of-life care [6], with a nearly fourfold difference in cardiac bypass surgery rates, a phenomenon primarily related to the region's number of cardiac catheterization labs per capita rather than illness rates [7]. One estimate suggested that only 27 percent of the weighted discrepancy in Medicare spending across regions could be explained by population illness levels [7], and if all regional spending levels matched those of the lowest decile, Medicare could see savings of up to $40 billion (1996 dollars) without compromising health status [8]. Clearly, more does not by itself equate to better—and the variation is greater for conditions in which the evidence is more limited.

Ultimately, the central challenge is not primarily a matter of overuse or underuse of services but instead is related to the lack of available evidence to achieve the right care for any given patient. Information on which to compare the results from drugs with the same purpose is often not available. For example, both Lucentis and Avastin are promising new drugs for treatment of macular degeneration, but head-to-head information on the relative outcomes is not available—and one costs about 20 times as much as the other. Similarly, different approaches to radiation therapy—intensity-modulated radiotherapy and conformal radiotherapy—have very different costs but currently inadequate information on which to base clinical judgments. And the pace of introduction of new genetic prognostic tests is on

an exponential course without the necessary evidence about the results of clinical decisions and outcomes.

Estimates range widely concerning the proportion of medical care in the United States that is based on, or supported by, adequate evidence [9-14]. However, given concerns about the extent to which this information may be generalized and the quality of the evidence that is used, some place this figure at well below half. Regardless of the precise level, there is no question about the need for improvement. Part of the challenge is the appropriate delivery of what has already been proven effective. Medical care is becoming more complex with the increase in multifaceted chronic diseases, the development of new interventions, and the pressures to reduce the time of patient-provider interaction in the face of greater administrative burdens. New care management approaches, decision support systems, and incentives will be required to help providers and patients work together to ensure that the care delivered is the care that is known to be most effective.

The most rapidly growing problem may soon relate not so much to shortfalls in applying what is known—a clearly significant problem—as to the inability for evidence of comparative clinical advantage to keep pace with innovation. It is both a capacity investment and a resource allocation problem. Of the nation's more than $2 trillion annual health expenditure—nearly half of it borne by government—currently less than 0.1 percent [15, 16] is invested in assessing the comparative effectiveness of available interventions. Although about 5 percent of overall health expenditures is devoted to research, most is devoted either to basic research or to product development [17], as opposed to assessing how well medical treatment options perform. If trend data were kept, it would likely reveal that the proportion of expenditures devoted to this assessment "budget" was actually shrinking every year, yet the complexity of clinical decisions continues to compound.

A testament to the power of innovation is the fact that new pharmaceuticals, medical devices, biologics, and procedures are introduced constantly, and the pace is quickening. From 1991 to 2003 the number of medical device patents per year doubled, from 4,500 to nearly 9,100 [18]. From 1992 to 2001 the total biotechnology patents granted per year tripled, from more than 2,500 to nearly 7,800, and the number of biopharmaceutical patents granted in the United States increased nearly four-fold, from about 1400 to nearly 5,200 [19, 20]. In the same period, annual sales of biologics and pharmaceuticals more than doubled [21]. Between 1993 and 2004 there was a more than 80 percent increase in the number of prescriptions received by Americans [22]. Data for the growth in procedures are more difficult to obtain; however, as one example, between 1989 and 1995 specialized procedures in major teaching hospitals nearly tripled [23]. A recent review by the Kaiser Family Foundation suggests that half or more of the

growth in medical spending in recent years is attributable to technological change [24].

Much, but certainly not all, of this change has resulted in better care. Diagnostic imaging services, for example, grew more rapidly than any other type of physician service under Medicare. Between 2000 and 2005 spending on radionuclide imaging (RNI) doubled from $6.6 billion to $13.7 billion [25]. Yet an American College of Cardiology Foundation technical panel convened in 2005 to assess the appropriateness of cardiovascular RNI imaging for 52 indications [26] found that the lack of clear evidence on the best and most effective uses yielded strong disagreement on the appropriate circumstances.

In addition to the growth in use of drugs, devices, biologics, and procedures, the world of health care is about to experience dramatic new insights concerning the variation in individual response to different diagnostic and treatment interventions. The 3 billion base pairs of the human genome have now been sequenced, revealing the 99.9 percent of the genetic code that is common to all humans. Additional differences, such as the gain or loss of regions of the genetic code, increase the variation between two random individuals by five- to ten-fold. Cataloging and characterizing these differences by haplotype mapping and other initiatives is currently in progress and will begin to reveal how people vary in their susceptibilities to diseases and their responses to diagnosis and treatment.

The age of personalized medicine will soon be a reality, if the capacity can be developed to contend with its insights. Today the average clinical encounter already requires a health provider to manage more variables than would be considered reasonable given what is known about the capabilities of the human mind, and over the next decade that same encounter will require contending with perhaps an order of magnitude more [27]. The traditions of developing evidence through one-at-a-time studies and relying for quality assurance on the recall capacity of an individual provider are no longer practical.

Over the long term, substantial changes will emerge in the way the nation goes about generating and applying evidence for clinical decision making. A learning healthcare system is one in which the clinical research paradigm depends more judiciously on the serial conduct of randomized controlled trials—important, but often too expensive, untimely, and of limited applicability—and draws more heavily on electronic health records (EHRs) to generate evidence as a natural by-product of the clinical experience. But while these longer-term capacities emerge, substantial near-term improvement will be necessary in our capacity to assess the relative effectiveness of different interventions—to understand what works best for whom under what circumstances. We need better understanding, agreement, and focus on the value we get from our health care—including what

constitutes value and how to measure it. Without this capability, it is likely that the inefficiencies that currently characterize the U.S. healthcare system will be compounded, perhaps considerably. Conversely, a more systematic and sustained effort to develop evidence on comparative and "real-world" effectiveness should stimulate more investment in research on innovation that will deliver better outcomes and greater value.

Engaging the immediate need for a much stronger and sustained capacity to meet the need for guidance on the clinical effectiveness of medical interventions is the subject of this paper. Discussion follows on the perspectives of the various stakeholders, the current capacity and activities on clinical effectiveness research, the key functional needs to be met, and, finally, some possible approaches to addressing the issues, including consideration of decision principles, governance, funding, and public support.

IMPLICATIONS FOR STAKEHOLDERS

Better evidence is essential to securing trust in our healthcare system. In the face of uncertainty borne of insufficient evidence, patients, providers, insurers, and health product companies frequently find themselves at odds and distrustful of each others' motives. Concern about the shortfall in the national capacity to determine what medical care is actually best for whom is shared among many stakeholders. Most important, of course, are the patients who receive medical care and the health providers who deliver it, but large stakes are also held by healthcare delivery organizations, insurers, manufacturers, and others engaged in various aspects of health care, with the shared goals of improving patient health and delivering the best value. Increasing the level of investment in clinical effectiveness research, and doing so in a comparative fashion, is key to facilitating significant gains toward these shared goals. Roundtable teams are currently reviewing the perspectives and action prospects as the various sectoral stakeholders work to improve the prospects for the development of a learning healthcare system. Following are some of the more important implications of accruing substantially better information on clinical effectiveness.

Consumers-Patients

Each patient should be able to feel confident that there is solid evidence that the care received is the appropriate care for his or her circumstances. Yet, increasingly, this notion is strained. The American public has traditionally expressed strong support for medical care, research, and technology development, while also expressing a strong interest in both individual patient prerogative and better information to aid decision making. But with the increasing complexity of care and an increasing awareness of its

shortfalls, confidence in healthcare delivery is beginning to wane. Some of the concern may relate to increased costs borne directly by patients as a result of increased prices, related coverage reductions, and system inefficiencies—an increase of about 50 percent in out-of-pocket expenditures from 1994 to 2004 [28], on top of an 87 percent increase in premium costs for family coverage alone since 2000 [29]—but much is based in perceptions about quality. In a 2005 Research!America survey, 60 percent of Americans said they didn't believe that the United States has the best healthcare system in the world, 41 percent said they knew of a time when they or a family member had received the wrong care, and 56 percent said there should be more investment in clinical and health services research [30]. A variety of recent initiatives, some by patient groups and some by medical and scientific groups, including the Institute of Medicine's 2001 report *Crossing the Quality Chasm* [31], have emphasized the need for health care to focus more on the delivery of individualized, patient-centered care—including an active role for patients in making informed choices about their health care, with careful consideration of known risks and benefits, transparency in the quality of care, and continual assessment of the performance. A central precondition for each of these is a substantial enhancement of the evidence on the interventions that are most effective for any given circumstance. Most of the challenge in this respect is in ramping up the capacity to generate and apply the necessary evidence, but other issues related to patient perceptions must also be addressed. Because some patients may view the results of comparative effectiveness studies as potentially limiting to their choices, care must be taken in the application and interpretive processes to ensure understanding and appropriateness. Similarly, in order to reduce patient confusion as recommendations change and to improve the support for use of clinical data for new insights on effectiveness, it is important to help patients better understand the dynamic nature of evidence and the need, in a learning healthcare system, to draw upon the healthcare delivery experience to continually refine scientific understanding of the safety and effectiveness of healthcare interventions in a practical setting.

Health Professionals

Health providers remain the central decision makers in health care, although increasingly they are making those decisions in close consultation with their patients. Ultimately, expanded efforts to make better information available on the comparative effectiveness of interventions are undertaken to better equip physicians and other health providers with the information needed to deliver the right care and to foster a more informed and supportive practice context. No health professional should be put in the position of uncertainty about the evidence in support of the care provided at his or

her behest. Yet, with the current pace of advances in medical procedures, pharmaceuticals, devices, and biotechnology, a sometimes confusing array of choices is presented for patients and their healthcare providers, making decisions about the best care for the circumstances increasingly complex. The lack of critical information and studies necessary to inform these decisions already place in jeopardy both the effectiveness and the efficiency of the medical care delivered to Americans. The pace at which this problem will grow with new innovations and insights is uncertain, but routine availability of comparative effectiveness information is a basic need for health professionals seeking to deliver high-value care to patients.

Healthcare Delivery Organizations

The integrity and reputation of healthcare delivery organizations is dependent on their ability to ensure the quality and appropriateness of the care delivered within their walls. In part, this is a function of the health professionals they employ and the culture they foster. Increasingly, it is also a function of the soundness of the systems they build to ensure that the necessary care is delivered and the delivered care is necessary. Any decision support system is only as good as the information built into the model and should include the comparative advantages or disadvantages of different diagnostic and therapeutic options. Especially as healthcare delivery organizations become more skilled in team management and as they build systems to improve and measure the consistency of the care with established performance standards, the rate-limiting factor will be the baseline information on the comparative effectiveness of available options.

Healthcare Manufacturers

The sector with the largest economic stake in better capacity to generate and apply information about the effectiveness of clinical interventions is the healthcare manufacturing sector—the companies that make the pharmaceuticals, devices, and biological products that provide the backbone for much of health care and its progress. It is a given that healthcare manufacturers, focused as they are on returns on investment, inherently understand the importance of improving the value proposition in patient care. But their stake goes deeper. Manufacturers directly bear the economic burden of delays and inefficiencies related to market entry requirements that are poorly linked to coverage approval processes; the absence of a clinical safety and effectiveness monitoring process that functions beyond the one-at-a-time tracking of new interventions, and the consequences of public and shareholder backlash when problems are identified too late. Given the aggregate size of the healthcare investment, improvements should

be achieved with economies of scale, better coordination of the efforts to produce the studies and systems necessary, and more transparency in the use of clinical data to draw conclusions. A stronger capacity to identify earlier when new interventions yield superior results should yield advances in both value and the pace of progress. Without a sizable improvement in our evaluative capacity, the slower pace of understanding how and when interventions work best will retard the application of potential innovations.

Employees-Employers

Over half of the nation's health expenditures are borne by the private sector, including a sizable share by employers [32]. For the fourth consecutive year, chief executive officers of U.S. companies have cited healthcare costs as their number one economic concern [33]—a concern compounded by questions about the quality of the return on investment, when per capita expenditures rank 50 percent higher than any other country in the world. Employers now pay 78 percent more for health care than 5 years ago, and it has been suggested by some that this increased financial burden makes it more difficult for American companies and workers to compete in the global marketplace [34]. Increasingly, healthcare costs are associated with reductions in the depth and breadth of employer-based health insurance coverage for U.S. workers and are often cited as a factor reducing the ability of companies to remain competitive. Without better information on which to improve the focus on innovations that work, rising costs and growing utilization will continue to contribute to the upward trajectory of healthcare spending, eroding the efficiency and value of our health expenditures. Those expenditures are projected to reach 20 percent of the gross domestic product by 2015—raising the burden for households, businesses, and government, yet not yielding concomitant gains in value [35].

Insurers

Insurers represent the front line of the economic choices that have to be made on payment for healthcare services, with a fiduciary responsibility to purchasers to ensure that payments are devoted to care that is most appropriate and that returns the most value to their clients. This means drawing conclusions about the comparative advantages or disadvantages of proposed diagnostic or treatment interventions, in the face of a paucity of such information [36]. The larger insurers maintain analytic staffs to assess the existing literature, and most also contract with organizations that conduct formal systematic reviews, but such reviews are limited by the shallowness of primary research studying the effectiveness of interventions in a practical setting—either compared to a placebo or to other alternative

choices. Insurers perhaps most acutely feel the need for much more reliable, rigorous, transparent, and impartial comparative effectiveness information to make decisions in the growing marketplace of medical interventions.

Government Agencies

Government currently accounts for about 45 percent of health expenditures in the United States, although if foregone tax revenues for employer-based health benefits are factored in, the number may be closer to 58 percent [37-40]. Most of the direct government expenditures are for reimbursement for Medicare and Medicaid clients, but the federal government also provides services directly to 9.2 million members of the uniformed services and their dependents, to 5.3 million patients who are served by the Veterans Health Administration, and to 1.5 million clients of Indian Health Service facilities. The federal government pays for the health coverage of approximately 3.7 million employees through the Federal Employees Health Benefit Plan (FEHBP), with enrollees selecting health plans from a number of competing insurance carriers. Whether acting as a payer or a provider, the government also has a basic responsibility to ensure that its clients receive the care that is most appropriate and of the greatest value. Those responsible for the relevant decisions encounter the same disadvantages as private sector decision makers with respect to the lack of information on the comparative effectiveness of candidate interventions. In addition, the Food and Drug Administration, which is required by law to make its judgments based on the safety and efficacy of a given intervention, is increasingly under pressure to provide perspectives on the relative advantages of proposed new approaches but generally lacks the studies on which to base such counsel.

CURRENT ACTIVITIES IN CLINICAL EFFECTIVENESS RESEARCH

Activities currently under way to assess the effectiveness of healthcare interventions may be generally characterized as broad, based in multiple loci, largely uncoordinated, far short of the need, and under-resourced. Presented here are definitions of the basic terms used to describe the various categories of activity, followed by a summary of the major institutions engaged in the work. A more detailed description of the activities is presented in Appendix One.

Terms

Clinical effectiveness research can be described as either primary or secondary. For our purposes, primary refers to the direct generation of

evidence through the use of a specific experimental methodology. Secondary refers to the systematic gathering and evaluation of primary research information to further the understanding of common conclusions or disparate results.

Primary Clinical Effectiveness Research

In this respect, primary clinical effectiveness research refers to the specific design and implementation of structured research protocols to produce data on the results of one or more diagnostic or therapeutic interventions of interest. Examples include certain randomized controlled trials, practical clinical trials, cluster randomized trials, observational studies, and cohort studies, including registries. Some of these studies focus only on the *efficacy* of an intervention—the extent to which an intervention produces a beneficial result under ideal circumstances. But many also examine the *effectiveness* of an intervention when used under ordinary circumstances—including evaluations in broader patient populations and healthcare delivery settings or analyses of the relative risks and benefits of competing therapies. Both types of evaluation are important to an understanding of which interventions work best, for whom, and under what circumstances.

Evidence Synthesis

Evidence synthesis or *secondary clinical effectiveness research* refers to the structured assessment of evidence from multiple primary studies to derive conclusions which are considered to have greater weight than an individual study alone. The types of evidence synthesis include *systematic review* and *technology assessment*, both of which describe a systematic method of identifying, assembling, and interpreting a body of data to validate or extend the interpretation of single trials, lend context to individual trials, and, where possible, arrive at common conclusions. Systematic reviews are frequently published through the peer-reviewed literature, while many assessments are tailored more narrowly to assist in policy or practice decision making.

Comparative Effectiveness

Within the overall umbrella of clinical effectiveness research, the most practical need is for studies of comparative effectiveness, the comparison of one diagnostic or treatment option with one or more others. In this respect, primary comparative effectiveness research involves the direct generation of clinical information on the relative merits or outcomes of one intervention in comparison with one or more others, and secondary comparative

effectiveness research involves the synthesis of primary studies to allow conclusions to be drawn. Secondary comparisons of the relative merits of different diagnostic or treatment interventions can be done through a collective analysis of the results of multiple head-to-head studies, or else indirectly, in which case the treatment options have not been directly compared to each other in a clinical evaluation, and inferences must be drawn based on the effect of each intervention relative to a specific comparison, often a placebo.

Other Related Terms

Other relevant terms used in the context of clinical effectiveness research discussions include cost-effectiveness analysis and cost-utility analysis. In *cost-effectiveness analysis*, the economic cost per specified unit of health gain—e.g., reduced mortality or morbidity—is determined for a given intervention or family of interventions. This allows for one measure of the relative value of an intervention to be estimated in comparison to alternatives. *Cost-utility analysis* is a form of cost-effectiveness analysis that estimates the cost of a specific utility gain, usually to the patients targeted—e.g., quality-adjusted life-years—for an individual intervention. Finally, the term *health services research* refers broadly to the multidisciplinary field of scientific investigation that studies how the effectiveness of health care for different populations is shaped by various systemic factors such as social factors, financing systems, organizational structures and processes, health technologies, and personal behaviors affecting access to health care, the quality and cost of health care, and measures of population health and well-being [41].

Clinical Effectiveness Research in the United States

It is difficult to characterize precisely the national expenditures on clinical effectiveness research, but the investment is clearly far short of the need. If only 1 percent of the nation's $2 trillion healthcare bill were devoted to understanding the effectiveness of the care purchased, the total would come approximately to $20 billion annually. In contrast, only $15 million annually has been specifically appropriated by Congress to the Agency for Healthcare Research and Quality (AHRQ), under section 1013 of the Medicare Prescription Drug, Improvement, and Modernization Act (MMA), for comparative effectiveness research, the issue for which the shortfall is greatest. The total of all appropriations to all federal agencies—the National Institutes of Health (NIH), the Veterans Health Administration (VHA), the Department of Defense (DOD), the Centers for Medicare and Medicaid Services (CMS), the Food and Drug Administration (FDA),

AHRQ, and the Centers for Disease Control and Prevention (CDC)—for all health services research amounts to about $1.5 billion, only a small portion of which is devoted to clinical effectiveness research [41].

In addition to these federal appropriations, the insurance industry currently invests substantially in assessment of clinical interventions, and healthcare manufacturers account for the largest contribution to clinical effectiveness research. It is reported, for example, that Pharmaceutical Research and Manufacturers of America (PhRMA) member companies spend in the range of $15 billion annually for Phase III and IV clinical trials of pharmaceuticals, including examination of clinical effectiveness [21]. Data are not currently available on the amount of direct expenditures by professional societies on primary and secondary clinical effectiveness research. Even accounting for the support from all involved public and private institutions, the aggregate national commitment to assessing the effectiveness of clinical interventions is likely well under 1 percent—far below the standard that any company would expect to invest in work to evaluate and improve its products.

Primary Clinical Effectiveness Research

At the federal level, several agencies of the Department of Health and Human Services (HHS) sponsor research and maintain databases that provide insights on the clinical effectiveness of healthcare interventions. NIH is the largest federal sponsor of clinical research, and its component institutes have supported some work on the comparative effectiveness of health interventions, although these types of assessments are not a major focus of attention—an estimated $660 million total since 1982, in comparison with the aggregate NIH budget of $250 billion in that period [40]. Through the National Library of Medicine the NIH also maintains, in collaboration with FDA, the national inventory of clinical trials at the ClinicalTrials.gov Web site. CMS does not conduct clinical research directly, but its data systems, demonstration and evaluation activities, and coverage policies offer powerful resources for assessing and monitoring clinical effectiveness. Claims data are maintained on the more than 42 million Americans served under Medicare and on the 47 million low-income people covered under Medicaid. With the passage of the Medicare Modernization Act Part D benefit, which makes Medicare enrollees eligible for prescription drug coverage, extensive new opportunities were presented to assess clinical effectiveness in a post market environment, by linking Part D data to data from Parts A, B, and C in the conduct of public health research. Recently, CMS has launched the Coverage with Evidence Development initiative, beginning with the development of a registry that will track the experience of Medical patients receiving implantable cardioverter defibrillators (ICDs). This

allows coverage of services for certain populations on which existing effectiveness information is limited, contingent upon clinical effectiveness data collection via a registry or other mechanism.

FDA, CDC, and AHRQ are HHS agencies that also contribute to clinical effectiveness research. FDA contributions come through requirements that manufacturers establish basic safety and efficacy information as part of the drug, biologic, and device approval processes; it also collects data as part of its adverse event reporting process and related post-marketing surveillance work. Recently FDA has proposed the development of an integrated national network, a Sentinel Network, formed through a series of public-private partnerships using new electronic information technology systems to collect and analyze medical product safety information and then disseminate it to healthcare practitioners and patients at the point of care. In general, the FDA does not require comparative effectiveness information to approve the marketing of an agent or innovation. CDC funds some health services research to guide decisions on public health services and systems, and some of this research may include an examination of the effectiveness of some therapies in the area of infectious disease or vaccines [41]. In addition, CDC maintains a number of national data systems—vital statistics, health examinations, and health interview surveys—that are important resources for certain types of clinical studies.

AHRQ, the lead federal agency for health services research, has a mandate from section 1013 of the Medicare Prescription Drug, Improvement, and Modernization Act along with a related $15 million appropriation to perform research with a focus on the outcomes, comparative clinical effectiveness, and appropriateness of pharmaceuticals, devices, and healthcare services. The provision has led to the creation of AHRQ's Effective Health Care Program, which has three components: synthesizing existing studies into Comparative Effectiveness Reports by Evidence-based Practice Centers; developing evidence, including research aimed at filling knowledge gaps about treatment effectiveness (DEcIDE centers); and improving communication of complex scientific findings to a variety of audiences (Eisenberg Center).

At the Department of Veterans Affairs, the VHA has an active clinical research program in its 157 medical centers and more than 1,300 ambulatory, residential, and home and community-based sites of care, facilitated by a state-of-the-art computerized clinical data system, with 8 million patient records, which allows the rendering of large quantities of patient data for analyses on a wide variety of clinical research topics. With more than 3,000 researchers throughout the system and specially designated "centers of excellence," the VHA devoted some $400 million in 2005 to clinical research in addition to the resources devoted to the related medical care.

A substantial share of this work focuses on clinical effectiveness, including some on comparative effectiveness.

Outside the federal government, primary and secondary (evidence synthesis) research is both conducted and sponsored by health manufacturers, insurers, healthcare delivery organizations, and professional societies. Industry-sponsored trials represent a significant proportion of health manufacturer investments in research and development and about 40 percent of pharmaceutical R&D investments goes to the phase 3 and 4 trials that have particular relevance to clinical effectiveness [21]. Many of these studies are conducted with academic investigators and others are managed by contract research organizations. Relatively few of the studies are comparative, or head-to-head studies [37, 38].

Large healthcare systems, such as large health maintenance organizations, also have the capacity to use both their clinical and administrative data systems for primary clinical effectiveness research—some sponsored out of their own resources, others sponsored by NIH or commercial concerns. For example, Kaiser Permanente, Geisinger Health, and Group Health all have efforts to use their large databases to learn what works best for which patients, extending insights from primary studies of efficacy or effectiveness. The Health Maintenance Research Network (HMORN) is composed of 15 HMOs covering over 15 million individuals and working cooperatively on effectiveness research. Data from their large record systems and registries are useful in defining appropriate use of interventions in subgroups of patients, including the post-introduction monitoring of results from new diagnostics and treatments. Similarly, payer-related data consortia probe the use of linked health insurance claims information to assess the clinical effectiveness of various interventions. In July 2006, the AQA Alliance and the Hospital Quality Alliance (HQA) announced the formation of a joint effort to combine resources to identify, collect, and report data across the variety of care settings they represent; a variety of data are being aggregated, including information on the quality of physician performance, cost-of-care measures, and the quality of care for specific conditions, such as heart attack or pneumonia, as well as other measures.

Recently, as a bridge of sorts between primary research and the use of existing data, more emphasis has been given to the potential for decision models in estimating the relative benefits and harms for different interventions [42]. Most cost-effectiveness analyses depend on the construction of decision models, but such models can also be useful in estimating purely clinical outcomes that might have important implications for decisions to use or not to use a particular intervention. Modeling may be particularly helpful in decisions on the application of diagnostic testing and screening tests, by revealing the likely yield from such tests when applied on a population basis. For example, such models proved important to the develop-

ment of recommendations on the ages and periodicity for cancer screening [43]. Similarly, they can be useful in determining the optimal sequencing of radiologic tests and of other tests used in sequence for diagnosis, particularly when supplemental information is available at large databases [44]. With the growth of information from registries and other large databases, models will likely be deployed much more frequently as evidence sources and decision tools.

Evidence Synthesis

Much of the work to marshal evidence for conclusions about clinical effectiveness takes the form not of primary data generation and analysis but of systematic reviews and meta-analyses of existing studies. Such secondary clinical effectiveness research is sponsored and conducted by a variety of organizations with overlapping and intersecting activities and interests, including federal agencies, state agencies, insurer and insurer-related organizations, independent assessment centers, professional groups and societies, university centers, and consortia. At the federal level perhaps the best known is the work of the U.S. Preventive Services Task Force (USPSTF), which has since 1984 conducted systematic reviews of the evidence in support of clinical preventive services, applied rigorous criteria to classify and rate the level of the evidence, and, based on the overall strength of the evidence for a given condition, offered conclusions and recommendations. The USPSTF is now sponsored by AHRQ, and its approach has set a standard reference point for much of the subsequent work on synthesizing evidence and making clinical recommendations. Building on this work, AHRQ has established a network of 13 AHRQ-sponsored evidence-based practice centers (EPCs) which review literature, perform technology assessments, and produce evidence reports including comparative effectiveness reviews. One of the AHRQ EPCs, at Oregon Health & Sciences, operates the Drug Effectiveness Review Project (DERP) on behalf of 13 state Medicaid programs to assess the pharmaceuticals to be provided under Medicaid. Consumers Union uses DERP findings as the basis for its *Best Buy Drugs Program*. In a related AHRQ- and FDA-sponsored effort, the 11 Centers for Education and Research on Therapeutics (CERTS) each focus on a specific patient population or therapeutic area in conducting research on ways to advance the optimal use of drugs, biologics, and medical devices by identifying best practices. CMS draws directly on AHRQ-sponsored technology assessments for use by its Medicare Coverage Advisory Committee to inform CMS coverage decision making. Other federal evidence synthesis efforts include the NIH consensus development conferences and state-of-the-science conferences, and the VA Technology Assessment Program for devices, drugs, procedures, and organizational and supportive systems used in health care.

In addition to the federally sponsored efforts, the largest insurers, such as United Healthcare, Wellpoint, Aetna, Humana, and CIGNA, have substantial in-house capabilities for the conduct of evidence reviews, and they also commission more formal assessments from technology assessment entities, such as the Blue Cross Blue Shield Association's Technology Evaluation Center (TEC), ECRI, Hayes, and Cochrane (see below). Another insurer-related activity is the Academy of Managed Care Pharmacy (AMCP), through which health plans utilize comparative analysis when developing formularies. Currently, most clinical practice guidelines are produced in association with physician specialty societies and published through multiple modalities. There are more than 150 medical specialty societies in the United States, and many are engaged in some form of evidence review and guideline development. The work is also performed beyond the medical societies. Altogether, more than 300 organizations have published at least one guideline on the AHRQ-supported Web site, guidelines.gov. The national guideline clearinghouse currently contains more than 1,900 individual summaries. There is substantial variability in the approaches of the various societies, with some of them conducting the reviews by informal staff-generated activities and others engaged in large, multifaceted, and structured consultations involving the multiple related organizations with similar interests in an issue. Because of the overlapping nature of the interests, issues, and approaches, a number of collaborative efforts have emerged. The best known of these is the Cochrane Collaboration, an international effort (see Appendix Two), including a U.S. center, which sponsors systematic reviews using carefully developed common standards on a wide variety of issues in health care. Groups such as the Grading of Recommendations Assessment, Development and Evaluation Working Group (GRADE) and Appraisal of Guidelines Research and Evaluation (AGREE) have formed to develop standards for evidence syntheses and clinical practice guidelines. Similarly, the Evidence-Based Medicine Roadmap Group, an effort led by AHIP including participation from several interested organizations, is working to develop a more consistent and transparent system for assessing comparative clinical effectiveness.

International Context

Clinical effectiveness assessment activities have a utility that crosses national borders, and a growing involvement in technology assessment and clinical effectiveness evaluation has developed internationally. The explicit incorporation of cost-effectiveness evaluation tends to be a common feature of the work abroad, prompted by the need for better information on which to make decisions about the use of public monies. Implicit in the systematic evaluation of cost for an intervention is the comparison of the intervention

with alternatives for care, as a means to arrive at an assignment of value for investment. Thus, the majority of comparative effectiveness analysis performed in an international context is based, to some degree, on the cost of the intervention. Programs in Australia, Canada, France, Germany, Sweden, and the United Kingdom are described in Appendix Two, as are the multinational efforts under the European Union, EUnetHTA, and the Cochrane Collaboration.

ACTIVITIES AND NEEDS RELATED TO COMPARATIVE EFFECTIVENESS RESEARCH

As broad and variegated as the interest and activity around clinical effectiveness are, the aggregate capacity is very thin and substantially short of the need. Because there are scant resources for the support of primary comparative effectiveness research (CER)—head-to-head studies—much of the work done is secondary evidence synthesis. Even the systematic reviews and technology assessments are often uncoordinated and draw on inconsistent standards for effectiveness determinations, underscoring the need for a substantially greater and more systematic approach to the work. The areas of activity and changes needed have been characterized in various ways [45-51], and can be grouped as indicated in Table A-1 below, closely reflecting those discussed at a recent Health Industry Forum meeting [51].

TABLE A-1 Prominent CER Activities and Needs

1.	Studies of comparative effectiveness ("head to head")
2.	Systematic reviews of comparative effectiveness
3.	Assessment of comparative value/cost effectiveness
4.	Coordinated priority setting and execution
5.	Improved study designs and research methods
6.	Better linkage of studies of efficacy, safety, and effectiveness
7.	Appropriate evidence standards consistently applied
8.	Consistent recommendations for clinical practice
9.	Guidance for coverage and funding
10.	Dissemination, application, and public communication

1. Studies of comparative effectiveness

The most rapidly growing problem in health care today is the need for better information on the head-to-head effectiveness of diagnostic and therapeutic interventions. Care related to breast cancer offers a good example. Not long ago breast cancer was treated as a single entity with one therapeutic approach—radical mastectomy—but now the condition is seen as a complex cluster of conditions whose many different forms require different diagnostic approaches, ranging from characterization of tumor cell

mass and estrogen sensitivity to genetic predisposition to metastases, and also require different surgical, radiologic, and chemotherapeutic treatments. Despite the fact that breast cancer is one of the most extensively studied conditions, the rapid and encouraging development of different diagnostic approaches and treatment interventions is substantially outstripping the ability to understand what is best for a given individual. Indeed, bone marrow transplantation for breast cancer represents one of the most prominent examples of an aggressive intervention applied prematurely because adequate studies had not been done—and it was ultimately proved ineffective. The challenge of too little information on the relative effectiveness of different tools for diagnosis and treatment is a growing problem across the spectrum of diseases and conditions. When people become ill, the question that rivets the attention is, "What's best for me?" Yet, as options increase, providers increasingly do not have the answer to that question. Most biomedical research investments are devoted to understanding basic biological processes and mechanisms, which ironically both feed the development of new approaches, but, by virtue of the tacit displacement of research on clinical effectiveness, impede the ability to learn what actually works best for individuals and for populations. Even within an overall pool of research on clinical effectiveness that is inadequate, the share devoted to comparative assessments of interventions may be fairly characterized as very small. With a need this compelling, the consequences are also compelling, ranging from the prospect of lost opportunities and harm for patients to the wholesale dissipation of resources and to public and provider confusion. Put simply, there is an acute need for: (1) a substantial increase in support for primary research into the comparative effectiveness of diagnostic and treatment interventions; (2) a means of determining the priority for the studies that are most compelling; and (3) entirely different ways of conducting the research to accelerate the generation of clinical effectiveness insights, including improved approaches stemming from the use of electronic health records to generate evident from routine clinical experiences (see study designs and research methods below).

2. *Systematic reviews of comparative effectiveness*

There are currently multiple means by which secondary research on the evidence of effectiveness is performed in the United States, as described earlier. The primary limitation on systematic reviews of the comparative effectiveness of various interventions is the paucity of primary research results on which to draw. There are also several general categories of issues in need of additional systematic evaluation both for determinations related to the evidence, where it exists, and recommendations relating to additional studies needed. These include issues related to the comparative evaluation of different drugs within a single class, evaluation of different

drugs in different classes for the same condition, evaluation of medical procedures, comparison of the effectiveness of procedural versus pharmaceutical approaches to individual diseases, and the comparative effectiveness of different approaches to diagnosis given a suspected disease. Other challenges exist as well, including those related to coordinating studies of interest, ensuring consistency of approaches, and improving the compatibility of findings. For example, systematic reviews by different institutions on the use of epidural steroid injections for the treatment of sciatica and on stress ulcer prophylaxis [52] in critically ill patients have yielded discordant results [48, 53]. In part, these sorts of inconsistencies arise from the possibility of legal impediments to payer organizations collaborating on analyses that relate to their coverage determinations, but they also relate to the absence of a consensus on the conceptual framework of key decision categories and how standards for necessary levels of evidence might vary by category. So the need exists both for additional support for systematic reviews of comparative effectiveness and for application of consensus approaches and standards.

3. *Assessment of comparative value/cost effectiveness*

Understanding the value returned for investment is a basic obligation in the stewardship of resources. The essential elements of value include knowing whether and on what dimensions an intervention works (outcome), knowing how well it works in comparison to alternatives (relative outcome), and knowing the economic cost (cost per outcome unit). There are multiple ways in which cost can be considered as part of medical decision making. Methodologies such as cost-effectiveness and cost-utility analysis use cost data in specific ways to place an economic metric on individual interventions. In the United States, the use of economic considerations as a determining factor in medical decision making remains controversial. CMS, the largest single payer for healthcare services, does not have the authority to use cost as a factor in making national coverage determinations, although it does use costs to the Medicare system in choosing which technologies to evaluate, and local CMS contractors are empowered to use cost in certain determinations (e.g., "lowest cost alternative"). The notion of knowing the relative cost-effectiveness of interventions is becoming more important as the costs of health care continue to rise, consume a greater part of the national economy, and have a significant impact on individual, business, and government financial health. This is particularly true for considering interventions of comparable clinical effectiveness but very different costs. As economic implications become ever more acute, so does the need for consistent, validated, and transparent means of assessing value.

4. *Coordinated priority setting and execution*

With resources as scarce as they are for comparative effectiveness assessments, and with the need for directing them rapidly to the most pressing issues, it will be of key importance to develop a means of coordinating the establishment and execution of needed studies and assessment priorities. While there are many organizations involved in activities related to clinical effectiveness research and comparative effectiveness evaluation, these efforts are fragmentary, frequently occur in isolation, and, as a result, can be duplicative or even contradictory. A systematic approach to linking the research agendas and outputs of these various organizations would facilitate an effort to increase the national capacity for research in regards to both primary and secondary research. A focus on high-volume or high-cost interventions is natural, but many considerations pertain. Various factors enter into the consideration of priorities, such as the impact of the condition in play (morbidity and mortality), the number of intervention options, their expense, the potential for significant impact on outcome, cost, or quality of life, and more. A first step in this respect would be the development of a mechanism whereby the perspectives of the key stakeholders can be reflected and criteria developed for making determinations. In addition, some consideration should be given to the appropriate methodology for analysis. Certain issues warrant higher priority for head–to-head trials, but others might be more appropriate for observational, registry, or database studies. This issue is further examined in the section on expanded methodologies.

5. *Improved study designs and research methods*

Clearly central to progress is the development of improved tools. Chief among the needs in this respect is the development of new study designs and statistical tools that will allow an expanded use of electronic health records in order to generate insights on clinical effectiveness. As our ability to manage large quantities of data electronically continues to increase, the use of electronic records to drive observational databases and registries will increase. Thus, another key element to consider, in terms of methodology, is a more informed and practical approach to the use of observational studies, clinical registries, and data aggregation. The randomized controlled trial (RCT) has been so successful in bringing order and validity to insights from the clinical experience that it has come to be known as the gold standard for studies of clinical effectiveness. But the RCT has many designs and applications that vary depending on the circumstances. Similarly, there are many approaches to gathering and assessing observational data, and the nature of the datasets themselves is changing rapidly. The pace of EHR development has the potential to expand the data pool dramatically, as does the focus on improving and standardizing the data inputs and the development of new statistical approaches to evaluating the data. When underlying

data are adequate, modeling can be used to make determinations of relative effectiveness within populations, and it has the potential to contend with the large populations needed to substratify patients and substantially change the approach to conducting comparative studies [54]. In many ways the notion of a "gold standard" is a misleading characterization. The basic challenge is to select the approach that is most informative from those practical to accomplish, identifying and acknowledging the potential for flaws. Certainly this is a more feasible approach, given the impossibility of relying on serial RCTs to provide all, or even a sizable fraction, of the comparative insights needed. So, both improved study methods and a capacity for the provision of technical assistance in the design of clinical studies are vital to progress.

6. *Better linkage of studies of efficacy, safety, and effectiveness*

With the need being so acute for a larger study base on which to base determinations of clinical effectiveness, the merits of better linkage of existing resources and activities are obvious. Viewed in vertical and horizontal dimensions, this might be termed "working to enhance the efficiencies of scope and the efficiencies of scale." With respect to scope, the parameters of any study design will affect the parameters of its applicability. Noted earlier were some of the problems encountered, for example, by manufacturers when studies designed to test safety and efficacy came up short with respect to the proof of effectiveness needed for coverage decisions by payers. Better communication at the outset among manufacturers, product approval authorities, and payers might expedite the evaluation of an intervention's performance in comparison with existing and alternative approaches, thus expediting the decision-making process. Some have termed this planning for the full life cycle of clinical research, from the early stages of testing through post-introduction monitoring. There is a need for vehicles to foster this type of communication. Similarly, with respect to scale, there is potential benefit in enhancing mutual awareness of similar clinical studies in order to foster the potential for collaborative or synergistic work. One existing resource of this sort is clinicaltrials.gov, described earlier as a mechanism for registering all clinical trials, both to increase patient recruitment as well as to facilitate information gathering once the trials are in process or have been recently completed. Perhaps even more important than the prospects for benefits from linking clinical trial work are the potential benefits from linking clinical data. Several efforts are currently under way in this regard, including the work of AHIP to foster data sharing through the ambulatory and hospital quality alliances and also the work of the RHIOs (Regional Health Information Organization) to advance platform and language compatibility. The VHA, with its vanguard work to implement electronic health records, has dramatically improved the linkage of records throughout its

system, setting the stage for broader use in clinical effectiveness research. The pace of progress in such linkage activities is dependent on a number of factors, including progress in standardizing terminology, improving platform compatibility, developing novel search algorithms, and patient privacy through mechanisms to de-identify data. Accordingly, there is a need for a well-positioned entity to steward and coordinate the activities of the various participants dealing with the different components of these interlocking issues.

Post-approval monitoring Particular mention is warranted of the implications of enhanced data system linkage—the scale dimension—for improved post-approval monitoring of new interventions and for fostering the pace of innovation and learning. Phase IV studies on the performance of new products are now usually independently designed and funded and are generally implemented on a product-by-product basis. As the prospect expands for linked clinical information systems, many of which are already well into application (e.g., VHA, DOD, Kaiser, EPIC users), so does the prospect for the availability of information of sufficient scale to provide much earlier insights into sub-group safety and effectiveness issues not fully resolved before an intervention's introduction into practice. An intermediate variation on this theme is the development of patient registries to monitor the performance of new interventions as they are introduced into new populations for which pre-introduction trial evidence is suggestive of effectiveness, but not conclusive—as in the CMS initiative on Coverage with Evidence Development. Once a registry is established, in principle its use could be expanded to monitor certain other interventions as well. In each of these circumstances, there is a need for better capacity to coordinate, monitor, and validate the work.

7. *Appropriate evidence standards consistently applied*

The impact of reliable research results is ultimately determined by the routes to their application, and the first stop along such a route is the set of standards used to judge the reliability and implications of the findings. Currently, several issues slow progress in this respect, including the limited flexibility of established concepts in adapting to different clinical circumstances and forms of evidence; the lack of a vehicle to foster agreement on the approach to application of standards of evidence; and a certain amount of resulting inconsistency in the determinations. Concerning the first issue, in part the need is to adapt approaches to the different nature of devices, diagnostics, pharmaceuticals, and procedures. For example, evaluation of many devices needs to take account of provider training and facility experience as well as of their sometimes short life cycle. Similarly, "blinding" in an RCT on a surgical procedure is largely impractical. Indeed, concerning what has come to be known as the evidence pyramid—with RCTs at the

top and professional opinion at the base—while the pyramid offers a key conceptual touchstone, it cannot be applied without adaptation to circumstance, and the rules or guidance or adaptive principles do not yet exist. Some collaborative efforts are under way to harmonize the approach to the grading of evidence and recommendations, such as those of the international GRADE working group, and to find more appropriate approaches to matching standards with needs, such the AHIP Evidence-Based Medicine Roadmap Group project to develop a matrix or mosaic approach to characterizing evidence. But work of this sort is still early in its development and limited in its acceptance. Consequences to clinical effectiveness determinations resulting from inconsistencies related to standards of evidence are relatively few, but they do occur. Perhaps the most substantial problem in this respect relates to the uncertainty those developing new interventions have concerning the standards they will have to meet in providing proof of effectiveness.

8. *Consistent recommendations for clinical practice*

There exists substantial heterogeneity in the work of various professional societies and other groups to develop guidelines for clinical practice. Some organizations and activities have well-established and formalized protocols for their guideline development activities. Examples include the U.S. Preventive Services Task Force (USPSTF), operated with AHRQ sponsorship, and the guidelines activities jointly sponsored by the American College of Cardiology and the American Heart Association. But other activities are often more ad hoc in nature. Guidelines.gov has served as a clearinghouse of clinical practice guidelines and, in association with the USPSTF, has fostered greater consistency in formatting. But the quality and specificity of individual guidelines remain variable, and although the format seems more standardized, the approach is often not. Some prominent examples of inconsistency in guidelines include those related to mammography for average-risk women over 40, prostate-specific antigen (PSA) testing for men, and Human immunodeficiency virus (HIV) screening for the general population [55]. Fostering greater consistency in the application of standards to produce clinical guidelines is an obvious priority.

9. *Guidance for coverage and funding*

The goal of increasing the capacity for research on comparative effectiveness is to improve the basis for medical decision making at many levels, not the least of which is for those who must make determinations on funding and coverage for interventions. Each payer has developed its own approach to obtaining the best available analyses on which to base coverage determinations, but they also generally report that the majority of the decisions they make have to be made with what they consider insufficient

evidence in hand. One of the problems in the current fragmented and under-resourced activities to establish clinical effectiveness is that studies focused on the issue of efficacy and safety for product approval purposes often lack the information needed for coverage decisions down the line—e.g., how well an intervention works in different sub-populations, or how results apply to populations with multiple, interacting conditions and treatments, or how it compares to alternatives. Examples of manufacturers who reported having received FDA approval for use but having delayed the because of coverage uncertainties include those who manufacture certain artificial spinal discs and ankles, Regranex for wound healing, and the Mammotome Breast Biopsy System [56]. Approaches are therefore needed both to improve the extent to which the information requisite for coverage decision making is incorporated early in the research planning process and to foster the capacity for using the post-approval monitoring period to further elucidate issues of effectiveness and safety.

10. Dissemination, application, and public communication

Ultimately, the extent to which better evidence drives healthcare improvement in this country will be determined by the public's clear and strong demand for better evidence that is more fully applied according to what is most appropriate for an individual patient's circumstances. As a result, a major key to progress will be effective communication—both on the specifics of what is learned about effective treatment as results are gleaned and on the dynamic nature of evidence, its evolution, and how it is gauged. Currently, the frequency with which recommendations are in conflict, previously accepted practices are disproved, and messages—commercial and other—are broadcast advocating one intervention or another, often leaves the public, and sometimes providers, uncertain, skeptical, and even fearful. Some of these issues will be ameliorated if greater consistency can be achieved in the standards used to interpret evidence and develop guidelines, perhaps through the development of a trusted and reliable source that might validate determinations or through which determinations and recommendations might be filtered. But more will be necessary in using communication and marketing principles more effectively both to make validated determinations more accessible to the consumer public and to help inform and educate the public about how evidence continually evolves and how to judge its state of play for a given issue at a given point.

Noted below in Table A-2 is a summary of these key challenges in redressing the pressing shortfalls in the development and application of clinical evidence.

TABLE A-2 Prominent CER Activities and Needs—Key Challenges

Issue	Key Challenges
Head-to-head studies	Scant resources; rapidly increasing need; comparison choice
Systematic reviews	Few primary studies; inconsistent methods; uncoordinated
Comparative value insights	Little agreement on metrics or role of costs; cost fluctuation
Priority setting	Fragmentation; inefficiency; no mechanism for coordination
Study designs and tools	Clinical trial time/cost/limits; large dataset mining methods
Research lifecycle links	Efficacy-effectiveness disjuncture; post-approval surveillance
Evidence standards	Standards not adapted to needs; inconsistency in application
Practice guidance	Disparate approaches; conflicting recommendations
Coverage guidance	Narrow evidence base; limited means for provisional coverage
Application tools	Public misperceptions; incentive structures; decision support

MODELS FOR A STRONGER APPROACH TO COMPARATIVE EFFECTIVENESS RESEARCH

The compelling potential health and economic consequences for Americans call for a substantial response to narrow the rapidly growing gap between the available evidence on clinical effectiveness and the evidence necessary for sound clinical decision making. From a practical perspective, over the long run the rate of progress in both the development and the application of evidence on effective care will depend on advances in the use, standards, and interoperability of health information technology throughout all practice settings. In the meantime, and even when that capacity is in place, much stronger investment in dedicated comparative studies is essential. Various organizations and recent public articles have called for a quantum increase—several billion dollars—in the level of investment in comparative effectiveness research and the creation of a sustainable new capacity to ensure that the highest priorities are addressed most effectively and efficiently and that new approaches are developed to accelerate the pace, reliability, and consistency of the results [41, 46, 47]. Publicly conducted reviews of the concerns related to the problems with the COX-2 inhibitors have underscored the fact that the need is not only for improvements in the framework for pre-market testing of interventions,

but for improved assessment of interventions in active use and an integrated process that brings the two into closer alignment. Because of inconsistencies and overlaps even at its currently limited level of work, the need is also not only for substantially increasing the level of assessment activity but also for restructuring its administration to ensure the efficiency of implementation—in effect to ensure the presence of a trusted, independent capacity to ensure the successful execution of a program of work in a transparent fashion, with the active engagement of key stakeholders as part of the process. The notion of independence is central because of past instances of political will intervening in scientific processes, as is scientific credibility to the patients and health professionals who will ultimately determine the success of the work.

Presented below are several approaches to developing the sustainable capacity needed for studies on the comparative effectiveness of healthcare diagnosis and treatment. They are grouped into four categories according to the nature and source of the funds for their support—incremental funding, public funding, private funding, and public-private funding (Table A-3). Each of the approaches is based on an existing or recent model of some kind—ranging from government agencies such as the Agency for Healthcare Research and Quality to the user fee model of the Food and Drug Administration to the operation of the highway trust-funded Transportation Research program of the National Research Council in The National Academies. Although presented as discrete models for discussion purposes, they are clearly not mutually exclusive. Many variations on the themes are possible, and several have been the focus of recent discussion and recommendations [46-48, 51]. All are succinctly described below, with elaboration on some presented in Appendices Three through Six. A brief

TABLE A-3 Models for Enhancing Capacity

Incremental funding augmentations
- Incremental model

Public funded entity
- Executive branch agency model
- Independent government commission model
- Legislative branch office model

Private funded entity
- Operating foundation model
- Investment tax credit cooperative model

Public-private funded entity
- User fee public model
- FFRDC public model
- Independent cooperative model
- Independent quasi-governmental authority model

summary of some of the pros and cons raised for the three most prominent proposals is presented in the following section on Decision and Implementation Considerations.

Incremental Funding Augmentations

Incremental Model

The approach most resembling the status quo is to incrementally grow existing activities in both the public and private sectors. In this scenario, the expectation would be for AHRQ to receive gradually increased appropriations for its comparative effectiveness research program, NIH to steadily increase its priority on investing in clinical effectiveness research as part of its Roadmap Initiative, industry working earlier with FDA and payers in conversations to anticipate post-market needs from the outset of pre-market testing, CMS to expand its Coverage with Evidence Development initiative, and insurers and manufacturers to increase their investment in comparative effectiveness research. The formation of an interagency collaborative is one that might seek to utilize the strengths of the individual agencies while reducing the burden on any single agency. A partnership between AHRQ, NIH, and FDA, for example, could potentially carry out a number of the functional elements described in the previous section. While there are ample precedents for interagency collaboration, most are on a scale too limited for true coordination, collaboration, and harmony in determinations, in particular given the potential scope of the activities required.

Publicly Funded Entity

An alternative to the incremental approach is to establish a dedicated entity with substantial funding (see Implementation Considerations below) as a largely public-funded executive branch agency or independent government agency, or as an adjunct to Congress. Three of the major reviews to date of this issue have expressed the view that the work is substantially in the public interest and is therefore best funded directly out of public funds [41, 46, 47].

Executive Branch Agency Model

The most straightforward public funded model is an expanded and appropriated mandate to an existing or newly created federal agency.

AHRQ-based The single agency whose mandate most closely parallels these priorities is AHRQ, which undertakes many of the functional components

discussed in the previous section, including prioritization of conditions or interventions to be evaluated and conducting systematic reviews. As AHRQ has an existing framework for many components through its Effective Health Care program, it may have the greatest flexibility to accommodate a mandate with increased functionality. A concern about this approach is that the current mandate of AHRQ goes far beyond the Effective Health Care program, and an expansion of functionality, such as taking on significant components of an increased capacity for comparative effectiveness research, could become the dominant force within AHRQ and supersede its other activities.

NIH-based There are other agencies that address some of the issues related to comparative effectiveness, such as NIH. NIH has strong credibility and experience with both primary and secondary clinical effectiveness research. However, the structure of NIH with institute divisions that each have oversight on their specific research agendas, combined with the mandate of the NIH to conduct basic and translational research, is less in line with the outcomes-oriented approach of comparative effectiveness research.

Other HHS-based Another option is the formation of an entirely new agency with HHS to undertake the elements of comparative effectiveness information. This new agency would rely on collaboration with other agencies, but as a lead organization would have the capacity to take on many of the functional components described. A clear advantage to this approach is that in creating a new agency, a new approach could be used to establish the governance of the agency to maximally insulate it from political influence.

Other executive-based Many other examples exist of independent, free-standing operational federal government agencies with presidentially appointed chief executives, including the National Science Foundation, the National Aeronautics and Space Administration, the Export-Import Bank, and the Small Business Administration. An expressed concern relevant to all Executive Branch models is that a Cabinet-level agency using funds appropriated by Congress, and subject to political oversight, may be especially vulnerable to political influence in its work.

Independent Government Commission Model

Several models exist of federally funded independent government commissions or agencies with operational programs and responsibilities. Although not operating major research programs of the sort considered here, their governance seeks shared non-partisan oversight of the program

of policies and activities. Two such agencies are the Consumer Product Safety Commission and the Federal Trade Commission.

Consumer Product Safety Commission (CPSC) The CPSC is a three-commissioner independent agency charged with ensuring the safety of consumer products. Although it does not maintain an extensive research program, CPSC maintains a consumer hotline and reporting system for product problems and product-related injuries, and it also operates the National Electronic Injury Surveillance System which monitors the injuries treated in about 100 active hospital emergency rooms throughout the nation. CPSC courses of action include developing voluntary and mandatory standards, requiring the issuance of consumer warnings, issuing recalls of products, and banning dangerous consumer products.

Federal Trade Commission (FTC) The FTC is an independent government agency charged with consumer protection and the elimination of anti-competitive business practices. Its five commissioners are appointed to 7-year terms by the President. No more than three of the commissioners can belong to any single political party. Its three main divisions are the bureaus of consumer protection, competition, and economics, each with analytic and enforcement responsibilities. Other examples include the Consumer Product Safety Commission, the Federal Communications Commission, the Occupational Safety and Health Review Commission, the Securities and Exchange Commission, the Postal Rate Commission, and others.

Legislative Branch Office Model

Office of Technology Assessment (OTA) From 1972 to 1995 the OTA operated to provide Congress with objective and independent analysis and advice on matters of science and technology with policy implications. OTA was governed by the Technology Assessment Board (TAB), made up of six senators and six representatives with equal representation from each party. The TAB appointed the OTA Director and a member advisory council from industry, academia, and elsewhere outside the federal government. The comptroller general of the United States and the director of the Congressional Research Service also served as members. OTA was funded out of appropriations.

Medicare Payment Advisory Commission (MedPAC) MedPAC is a 17-member committee, with its member appointed to 3-year terms by the comptroller general to serve as independent federal advisors to advise Congress on issues affecting the Medicare program. The MedPac mandate is broad, including issues related to payments for Medicare services,

evaluation of access to care and quality of care. It meets publicly to discuss policy issues and seeks input through a variety of other mechanisms. Its recommendations to Congress are issued in reports and briefings. It does not conduct a primary clinical research program, nor does it address issues outside the target Medicare population, but its mandate could be expanded to do so.

Privately Funded Entity

A clinical effectiveness research entity could be established entirely with private funds. The primary advantage of a private sector approach is that it would be perceived as not being unduly influenced by the government and political considerations. Conversely, it could be seen as having undue influence from the private organizations which fund it, and the resulting recommendations could be seen as playing to the interests of those organizations. In addition, the sustainability of the endeavor would not be guaranteed, nor would the philosophical notion of the importance of the information as a public good. Examples of two possible approaches to substantially increasing private funding are noted below.

Operating Foundation Model

A variety of foundation models exist that could serve as the basis for pooled funds from private sector stakeholders interested in reducing duplicated effort, effecting economies of scale, and accelerating the conduct of clinical research. Foundations dedicated to activities that serve the common good have preferential tax status and can be established either as non-profit corporations or as charitable trusts. They can be established as grant-making entities, for charitable activities, or they can conduct their own programs of activities, as long as they meet the provisions that determine their tax status. The several types include private, corporate, and public foundations. The corporate model is the one likely most familiar to the stakeholders cooperating in a shared program of clinical effectiveness research. In 2004 there were more than 2,500 corporate foundations operating in the United States, and the total giving of the top 50 of these foundations by total giving was nearly $2 billion dollars. Of these 50 top foundations by giving, seven were from pharmaceutical or biotechnology foundations [57].

Investment Tax Credit Model

Another approach to expanding private support, which could be used in conjunction with the establishment of a cooperation foundation enter-

prise, would be to expand the provisions governing corporate investment tax credits to encourage investments in comparative effectiveness research. The federal research and development tax credit was established to allow private organizations to take a credit against their tax liability equal to 20 percent of certain qualified expenses—specifically those expenses in excess of a predetermined base amount in order to encourage additional R&D above and beyond what a company might otherwise undertake. It has been successful in encouraging R&D investments and, if paired with the development of an independent mechanism for shared governance, pooling, priority setting, and quality control, could provide a basis for expansion of the work needed.

Public-Private Funded Entity

User Fee Public Model

One of the most prominent examples of shared public-private funding in the health arena is the user fee model to expedite the review of drug and biological products. In the 1992 Prescription Drug User Fee Act (PDUFA), the FDA was authorized to collect fees from companies, regardless of the review outcome, for each proposed new drug or biologic. In addition, companies pay annual fees for each manufacturing establishment and for each prescription drug product marketed. Industry provides the funding in exchange for FDA agreement to meet drug-review performance goals, which emphasize timeliness [22]. User fees currently fund about half of new drug review costs—which totaled about $219,841,000 in 2006, with a total budget for drug review of $515,557,000 [58]. By providing needed funds, PDUFA began to address what industry viewed as slow and unpredictable review and approval of new drug applications, while keeping FDA's high standards. The additional resources allowed FDA to devote more time for guidance to help minimize unnecessary trials and generally improve drug development, with the result that now 50 percent of new drugs are being launched first in the United States, compared to only 8 percent in the years prior to the establishment of PDUFA. This type of user fee model could be employed to support clinical effectiveness studies.

FFRDC Model

One type of quasi-governmental organization is the Federally Funded Research and Development Center (FFRDC), which is funded at least in part by the federal government, functions as a nonprofit private organization, and is managed by nongovernmental organizations, usually universities or other nonprofit institutions. FFRDCs were first established during

World War II to assist government agencies in meeting specific well-defined technical needs that could not be met by existing government agencies or normal contractor relationships. Currently there are 37 FFRDCs, and, while their missions are quite varied, there are general rules applied to their administration. The sponsoring federal agency is responsible for overall policy and oversight, following guidelines set forth by Office of Federal Procurement Policy (OFPP) Policy Letter 84-1. FFRDC funds come from the sponsoring agency requesting the work or from a line item on congressionally appropriated budgets of the sponsoring agency. In addition, an FFRDC can receive up to of 30 percent of its funding from private sources. Limits have been placed on competition for other government or commercial business with the intent of fostering a strategic relationship between an FFRDC and its sponsor as well as limiting the potential for conflict of interest that this special access may create. Existing FFRDCs fall into four general categories: policy-focused study and analysis centers (e.g., the National Defense Research Institute, administered by RAND, and the Homeland Security Institute, administered by Analytic Services, Inc.), research and development laboratories and research laboratories (e.g., Lawrence Livermore National Laboratory, administered by the University of California; the National Cancer Institute lab at Frederick, administered by Science Applications International Corp.; and others), and systems engineering and integration centers (e.g., the Aerospace Federally Funded Research and Development Center, administered by the Aerospace Corporation).

Independent Cooperative Model

Formal public-private partnerships have been established to facilitate research, and there are two good examples in the health area: the Health Effects Institute and the NIH Public-Private Partnership Program.

Health Effects Institute Chartered in 1980, the Health Effects Institute (HEI) is a nonprofit corporation that provides unbiased, independent research on the health effects of air pollution. Since its inception, HEI has published over 200 reports on over 250 funded studies that inform policy making on various pollutants, such as carbon monoxide, nitrogen oxides, and diesel exhaust. HEI is funded jointly, with roughly half of its funds from the U.S. Environmental Protection Agency (EPA) and half from the 27 companies of the worldwide motor vehicle industry. To accomplish its mission, HEI prioritizes research needs, funds and oversees projects, provides independent review of the projects, and disseminates findings. It is governed by an independent board of directors, a group of science and policy leaders not tied to HEI sponsors that sets goals and priorities, makes final funding decisions, and oversees research. The Health Research Committee, appointed

by the board of directors, selects and oversees projects, and an independent review board helps HEI staff evaluate research. In 2005 EPA grants for the Health Effects of Air Pollution Program totaled $2.2 million, and the total HEI research investments were $7.4 million. Special projects of specific programs can also be funded by other public and private institutions.

NIH public-private partnerships The NIH has for some time maintained an active program of collaborative and co-funded activities with the private sector to address issues that require outside resources. As part of the 2002 Roadmap Initiative, NIH formally established the Program on Public-Private Partnerships (PPPs) as a central resource to provide guidance and advice in their development. Partnerships are established either directly with the NIH or through the Foundation for the NIH (FNIH), an independent public foundation chartered by Congress in 1990 to support the mission of the NIH. As a public foundation, FNIH can solicit donations from nongovernmental sources and can also conduct grant or contract solicitations, reviews, awards, and management. Partners can include foundations, patient advocacy groups, and industry (pharmaceutical, biotechnology, devices, diagnostics, informatics, and other). The governance and specific roles of the partners are also tailored to the particular aims of the partnership. For example, contributions from NIH might include development of tests, assays, and diagnostics; development of databases, biobanks, and repositories; or organization of clinical trials. Industry contributions range from intellectual and fiscal support to samples and data from clinical trials. In May 2005 FNIH was involved in approximately 50 PPPs totaling $280 million. Examples are detailed in Appendix 4 and include the Biomarkers Consortium, a partnership managed by FNIH and governed by an executive committee involving NIH, FDA, CMS, industry, voluntary organizations, and advocacy groups, with the purpose of accelerating the identification, validation, and application of new biomarkers; the Osteoarthritis Initiative, a $60 million initiative partnering NIH institutes and centers with four industry partners ($22 million from Pfizer, Merck, Novartis, and GlaxoSmithKline) in a 7-year study collecting clinical, radiological, and biological data from 5,000 patients with osteoarthritis; and the Grand Challenges in Global Health, a $450 million project administered by the FNIH and prompted by a $200 million commitment from the Bill & Melinda Gates Foundation, with contributions from the U.K. Wellcome Trust and the Canadian Institutes of Health Research, supporting 43 separate research projects aimed at solving the 14 most important challenges in global health.

Independent Quasi-Governmental Authority Model

There are two prominent existing examples of approaches to the establishment of a quasi-governmental entity on clinical effectiveness: The National Academies model and the Federal Reserve model. Although the Federal Reserve does not run a research program, both The National Academies and the Federal Reserve have government mandates, are funded independently, and share certain governance features relevant to the operation of a clinical effectiveness research entity, including: public-private character, non-partisan, independent governance, shared stakeholder priority setting, and central policy authority.

National Academies One approach that takes advantage of existing independence, capacity, and reputation would be to work through The National Academies, which are governmentally chartered and have an established model for managing a cooperative national program of research in the work of the Transportation Research Board (TRB). The TRB has operated since 1920 as a division of the National Research Council and began regularly managing research projects with the establishment in 1962 of the National Cooperative Highway Research Program. In the 1990s Congress, the U.S. Department of Transportation, and the state departments of transportation asked TRB to undertake additional tasks, including management responsibilities for the Transit Cooperative Research Program, guidance of ongoing research programs such as the Long-Term Pavement Performance studies, and management of the Innovations Deserving Exploratory Analysis (IDEA) programs. The policies and activities of the Transportation Research Board are determined and directed by its executive committee acting within the overall policies of The National Academies. Executive committee members are appointed with the approval of the chairman of the National Research Council. The executive committee is composed of representatives from government, industry, and academia who are active in the areas of interest of TRB. Representatives from states, transit organizations, and universities are appointed to liaise between TRB and the organization or institution. The program is funded through resources from state transportation departments, industry associations, the U.S. Department of Transportation, and other organizations interested in the effectiveness of transportation tools and strategies. Applying this model for a program of comparative effectiveness research would place responsibility with the Institute of Medicine (IOM), perhaps in cooperation with the National Research Council. The membership of the IOM includes national and international leaders in health care, and its studies are highly regarded. As with any organization, taking on an operational role of the scope anticipated to meet the need for comparative effectiveness research would require careful consideration

and the development of new administrative and procedural approaches. However, the IOM has indicated a willingness to explore housing such an activity, possibly along the lines of the TRB precedent.

Federal Reserve The Federal Reserve System is a blended quasi-governmental system that serves as the nation's central banking system with the charge of conducting the nation's monetary policy, regulation of banking institutions, maintaining the stability of the economy, and providing financial services to depository institutions. The Fed is comprised of the Board of Governors, 12 regional Federal Reserve Banks, numerous private banks, and the Federal Open Market Committee. The Board of Governors, appointed by the President of the United States for 14-year terms, is an independent governmental agency responsible for monetary policy and overall supervision of the integrity of the banking system. It does not receive federal appropriations, and members are not accountable to any specific official, except to the President on matters of misconduct. The regional federal reserve banks are private entities, controlled by local member banks. The Federal Open Market Committee, which determines and administers the sale of government securities, is comprised of the seven members of the Board of Governors and five representatives from the Federal Reserve Banks.

DECISION AND IMPLEMENTATION CONSIDERATIONS

Establishment of a substantially expanded capacity for clinical and comparative effectiveness research will need to address a variety of implementation considerations, including, if established as a new entity, those related to funding, governance, priority setting, research conduct, and finding validation. Some of the relevant principles and administrative issues in identifying the organization's location, support, structure, and function include those that follow.

Principles

Scientific Credibility

Given the complexity of the issues and the intensity of the public interest, concerns, and stakes, scientific credibility is the most important characteristic of the organization vested with the responsibility of managing the conduct of research aimed at determining which health care works best, for whom, and under what circumstances. The organization that functions best in this respect will be the organization that is best able to gain the

trust and confidence of the public, the scientific community, and the other stakeholders involved.

Political Independence

Similarly, the conduct of the scientific enterprise must be insulated from the political processes. In any public endeavor, virtually all interests will seek—and have sought—to use political pressure to affect policies and processes. Whether the focus is funding for different projects, the wording of policies and recommendations, or the make-up of advisory committees, political influence and action has the ability to distort the integrity of the scientific process. Insulation from that influence is key.

Stakeholder Neutrality

Inherent in the notion of scientific credibility is that of stakeholder neutrality. Patients, providers, employers, manufacturers, and insurers all have driving perspectives whose interests are understandable, but whose impact must not intrude on the integrity of the scientific process. Both the organization and its management structure must be well insulated from the disproportionate sway of any particular stakeholder.

Participatory Governance

Political independence and stakeholder neutrality do not equate to the absence of their engagement in the agenda- and priority-setting process. The determination of the priorities to pursue is a policy exercise in which all relevant stakeholders have a right to engage and to which they can add value. This amounts to an appropriate involvement in the governance process, while safeguarding against interference with the structure or conduct of scientific reviews.

Investigator Integrity

While agenda and priority setting may be shared and brokered responsibilities, management and conduct of the research processes and the determination and validation of research results must be completely insulated from stakeholder influence. Vesting of responsibility for this work ought therefore to be placed, if not in an organization with a history of protecting against such influence, then in one that can construct the necessary procedural firewalls between the processes of policy and priority setting and the design, conduct, and reporting of the scientific studies.

Infrastructure Efficiency

Current work is far short of the need, but described in this paper are a number of organizations now operating with related mandates and activities. Advantage must be taken, where possible, of existing capacity for the establishment of scientific standards, the oversight and conduct of comparative effectiveness research, and the development of the requisite findings and recommendations.

Agenda Flexibility

Although a certain measure of process is a precondition for ensuring the integrity of scientific determinations, the organizational decision making, resource allocation, and program conduct must have enough flexibility to be able to respond quickly to emerging issues and changing circumstances.

Transparency of Processes and Results

The credibility, integrity, and usefulness of the studies carried out under the auspices of the responsible organization, and the determinations made, will depend upon the transparency of the work. An assumed character of any determinations, and the studies on which they are based, should be specification and availability of the data on which determinations are based, and clarity as to the processes and tools used in their evaluation.

Support and Management Issues

Funding

According to various estimates, funding adequate to meet the information needs of clinical decision makers must increase substantially, in the several billions above current levels [16, 46]. Estimates of this sort are consistent with the fact that simply investing a relatively modest 1 percent of healthcare expenditures to discern which care works best would amount to some $20 billion annually. This is a figure 10 to 20 times more than current federal expenditures for this work, but, if realized, it would clearly improve our ability to get better value from health care. Under the publicly funded approach, the two primary funding options are direct annual appropriations or a set-aside from the Medicare Trust Fund. With a public-private funded approach, proportionately matching contributions could be structured in a variety of ways, including blending small set-asides from Medicare fund expenditures, from private health insurance premiums, or

from manufacturer R&D expenditures. There can be many variations on these themes, but the key concepts are less related to the structure of the funding support than to the value of the yield for better health—the return for the outcomes and efficiency of the nation's health care.

Governance and Priority Setting

With multiple stakeholders involved, ranging from individuals like patients and providers to institutions like manufacturers and insurers, governance and priority setting should be a shared responsibility. Accomplishing this will require a governance structure that allows for broad stakeholder involvement in setting the agenda, determining priorities, and shaping and executing the decision-making principles.

Research Conduct, Findings and Recommendations

Several important considerations pertain to the conduct of the research supported, including insulating the research process from the governance and priority-setting responsibilities to ensure that the conduct of the science is fully independent of the political dynamics of priority setting; building on the research infrastructure that currently exists for the conduct of the studies; and full transparency in the conduct of the research and availability of data. Administratively, this suggests several elements to the task: identification of study priorities; triage of the research responsibilities to capable institutions; ensuring the quality and consistency of approaches among those institutions; validating the results; and fashioning guidelines and recommendations based on the results.

Summary of Common Approaches to the Independent Model

Because of the challenges to the notion of increasing comparative effectiveness research primarily through a straightforward federal appropriation to the Agency for Healthcare Research and Quality—difficulty of marshaling an appropriation at a level commensurate with the need, lack of political independence, limited ability to draw on other agencies—much of the recent discussion has focused on independent models, often with blended public and private funding. Table A-4 reviews some possible features of the three such models most commonly discussed: an agency (including an agency-linked FFRDC) model, an independent board model, and a hybrid model.

As independent entities, each of these approaches assumes the establishment of a governing board charged with priority setting, broad budget allocation, and fiduciary responsibility for execution of the program of

TABLE A-4 Comparative Effectiveness Research Enterprise Models

Activity	Federal	Independent Board	Hybrid
Reference model	NIH or AHRQ	Federal Reserve Board	NAS (IOM/TRB)
Priority setting	Agency[a]/HHS Board	Governing board/staff	Governing board/ISO[b]
Budget allocation	Agency/HHS Board	Governing board/staff	Governing board/ISO
Study selection	Agency	Governing board/staff	ISO
Design/methods	Agency	Agencies/ IOM-NAS	Agencies/ISO
Agency designation	Agency	Governing board/staff	Governing board/ISO
Study management	Agency	Agencies	Agencies
Study conduct	Agency/field	Agencies/field	Agencies/field
Study certification	Agency	Governing board/staff	ISO
Study conclusions	Agency	Governing board/staff	ISO
Dissemination	Agency	Governing board/staff	ISO
Advantages	Builds on current	Independent	Independent Builds on current
Disadvantages	Politically vulnerable	No established credibility	Other missions of ISO
	Linked to one agency	Duplicate capacity	

[a] Some proposals suggest creating an agency-associated but privately operated FFRDC (Federally Funded Research and Development Center) to give the work quasi-insulated status.

[b] ISO, independent scientific organization (e.g., Institute of Medicine, on the model of the NAS Transportation Research Board).

activities. Depending on the nature of support base for the work, this board would presumably be comprised of some mix of key stakeholders, who would be appointed in a fashion that would assure their political independence. Not explicitly discussed here are a number of important issues that also will need to be addressed, such as the loci for fostering new and improved study designs and research methods, the relationship of the findings from the proposed center to those of existing professional organizations and societies, and how the determinations might be used in payment decisions.

Agency or Agency-Linked FFRDC Model

Because an FFRDC is linked to a sponsoring federal agency that assumes responsibility for policy and operational oversight, the most logical agency sponsor of a comparative effectiveness research entity is the Agency for Healthcare Research and Quality, which has both a formal mandate and an existing, albeit small, program of such research [37, 42]. The components of AHRQ's Effective Health Care Program could be drawn closely and naturally into the work, such as the Evidence-Based Practice Centers, the Centers for Education & Research on Therapeutics, and the Developing Evidence to Inform Decisions about Effectiveness network. On the other hand, even an FFRDC operating outside the federal structure is still substantially associated with its sponsoring agency and not immune from political vulnerability. Although many public-funded entities have demonstrated the capacity for long-term stability and relative insulation from political influence, experiences indicate the difficulties faced when topics sensitive to certain constituencies are engaged. The Congressional Office of Technology Assessment was eliminated, in part on the grounds that it intruded on market forces, and the Agency for Health Care Policy and Research (now AHRQ) was nearly eliminated in 1995 when spine surgeons hostile to its recommendation on management of low back pain brought political pressure—which led to the removal of policy from the name. In addition, the primary association of the FFRDC with one agency might limit the engagement of other federal research agencies with important clinical effectiveness activities, such as NIH, FDA, CDC, and the VA. Furthermore, if the research is to be supported through combined public and private funds, the current FFRDC maximum of 30 percent private funds would have to be altered.

Independent Board Model

An alternative to the federal agency-based approach is to create an entirely new, free-standing entity along the lines of the Federal Reserve Board which governs the nation's banking system [47]. The philosophy behind the Federal Reserve model is that an enterprise that will affect the direction of 16 percent of the national economy and the health care of every citizen warrants a status with sufficient independence, size, and power to control the landscape. Although most discussions of this approach do not propose the same legislated mandates related to oversight, accountability, and reporting that are vested in the Federal Reserve Board, there are clear advantages to the creation of a capacity with similar independence and reach. In addition to a governing board for priority setting and budget allocation, some versions of this scenario would vest this board with operating

capacity for selecting specific studies, designating the locus of their conduct, and certifying the results and conclusions for dissemination. The possible disadvantages to this model include the need, as a newly established entity, to build credibility from the ground up, and the duplication of existing capacity for independent scientific review that exists in places such as the National Institutes of Health and the National Academies' Institute of Medicine and that is growing in AHRQ.

Hybrid Model

A hybrid model, blending reliance on existing independent scientific capacity with the creation of a new free-standing priority-setting entity, would establish a governing board to set priorities and allocate the necessary funds but draw on the IOM to execute the scientific activities. This would have the advantages of ensuring stakeholder involvement in the policy-determining elements of the governing board work, while constructing a firewall between the determination of policy by the governance apparatus and the conduct and oversight of the science. Based on the priorities established by the governing board, the IOM (or AHRQ/NIH, if preferred) would select the necessary studies, identify the study design and approach, designate the appropriate agency (e.g., NIH, AHRQ, CDC, FDA, VA) to identify the research organization and manage the conduct of the work, then the IOM would certify the validity of the results, fashion the conclusions, and initiate the dissemination. Additional advantages to this approach include that, with IOM, it uses an organization that has the established and acknowledged credibility and independence, and draws the capacity of all existing federal agencies into play, as appropriate. The principal disadvantage is that, since one of its basic principles is the notion of the desirability of a firewall to ensure the integrity of the science, the relationship between the governing board and the National Academy of Sciences would have to be solidly established.

CONCLUDING OBSERVATIONS

As ever-increasing options evolve in health care, current gaps in knowledge and practice about which care works best will persist or worsen without the appropriate information on which to base healthcare decisions. Innovations that are beginning to find clinical application, such as genetic profiling, are on the front end of a wave of technology yet to come. As the boom in pharmaceuticals, devices, and biologics has left us with a need for information on clinical utility, and as innovations in the pipeline come to fruition, the information gap will widen. The rate with which new interventions are introduced into the medical marketplace is currently outpacing

the rate at which information is generated on their effectiveness and the circumstances of best use. If trends continue, the ability to deliver appropriate care will be strained and may be overwhelmed.

A substantially increased capacity to conduct and evaluate research on clinical effectiveness of interventions brings many potential opportunities for significant improvement across a wide spectrum of healthcare needs. An integrated system to examine clinical effectiveness can significantly improve the speed with which evidence is developed on medical interventions through coordination of research activities, and further function to align evident clinical information gaps with research priorities. Clinical effectiveness research, through establishing the best practices in health care, serves as the cornerstone of quality improvement, and an increased ability to determine what has the most clinical utility will drive faster and more rigorous quality improvements. As electronic health records move from concept to reality, the need for coordinated efforts for clinical effectiveness research will become even more pressing but also more feasible. Interoperability among data systems will provide a wealth of data, the power of which will only be realized by a harmonized infrastructure that allows for the scale of clinical effectiveness studies required. The enhanced capability to know what works best will generate substantial and valuable information to support innovation by identifying the key areas where it is needed the most. The options reviewed here offer a sense of the possibilities and opportunities, but the need for swift action is pressing.

REFERENCES

1. OECD Health Data. 2006. *How Does the United States Compare?*
2. Mathers, C., et al. 2001. Estimates of healthy life expectancy for 191 countries in the year 2000: Methods and results. *Global Programme on Evidence for Health Policy Discussion Paper No. 38*. World Health Organization.
3. Anderson, G. F. 1997. In search of value: An international comparison of cost, access, and outcomes. *Health Affairs* 16(6):163-171.
4. McGlynn, E. A., et al. 2003. The quality of health care delivered to adults in the United States. *New England Journal of Medicine* 348(26):2635-2645.
5. Sung, N. S., et al. 2003. Central challenges facing the national clinical research enterprise. *JAMA* 289(10):1278-1287.
6. Fisher, E. S., and J. E. Wennberg. 2003. Health care quality, geographic variations, and the challenge of supply-sensitive care. *Perspectives in Biology and Medicine* 46(1):69-79.
7. Wennberg, J. E., E. S. Fisher, and J. S. Skinner. 2002. Geography and the debate over Medicare reform. *Health Affairs*. Suppl Web Exclusives: W96-114.
8. Wennberg, J. E., and M. M. Cooper, eds. 1999. *The Dartmouth Atlas of Health Care*. Chicago: American Hospital Association.
9. Gill, P., et al. 1996. Evidence based general practice: A retrospective study of interventions in one training practice. *BMJ* 312(7034):819-821.

10. Lee, H. Y., et al. 2005. Comparison of evidence-based therapeutic intervention between community- and hospital-based primary care clinics. *International Journal of Clinical Practice* 59(8):975-980.
11. Lee, J. S., D. M. Urschel, and J. D. Urschel. 2000. Is general thoracic surgical practice evidence based? *Annals of Thoracic Surgery* 70(2):429-431.
12. Moyer, V. A., A. K. Gist, and E. J. Elliott. 2002. Is the practice of paediatric inpatient medicine evidence-based? *Journal of Paediatrics and Child Health* 38(4):347-351.
13. Myles, P. S., et al. 1999. Is anaesthesia evidence-based? A survey of anaesthetic practice. *British Journal of Anaesthesia* 82(4):591-595.
14. Smith, R. 1991. Where is the wisdom...? *BMJ* 303(6806):798-799.
15. Moses, H., 3rd, et al. 2005. Financial anatomy of biomedical research. *JAMA* 294(11): 1333-1342.
16. AcademyHealth. 2005. Placement, coordination, and funding of health services research within the federal government. *AcademyHealth Report*. September.
17. Research!America. 2006. *2005 Investment in U.S. Health Research*. Online. Available at http://www.researchamerica.org/publications/appropriations/healthdollar2005.pdf.
18. Advanced Medical Technology Association. 2004. *The Medical Technology Industry at a Glance*. p. 16, Chart 3.6.
19. Biotechnology Industry Organization. 2006. *Chart: Total Biotechnology Patents Granted per Year*. Online. Available at http://www.bio.org/speeches/pubs/er/statistics.asp.
20. Foster, L., D. Nanu, and S. Alcock. 2002. Patenting in the biopharmaceutical industry—Comparing the US with Europe. *Drug Plus International* 1(1).
21. PhRMA (Pharmaceutical Research and Manufacturers of America). 2006. *Pharmaceutical Industry Profile*. p. 49, Table 4.
22. FDA (Food and Drug Administration). 2005. *Prescription Drug User Fee Act (PDUFA): Adding Resources and Improving Performance in FDA Review of New Drug Applications (Executive Summary)*.
23. Levin, R., E. Moy, and P. F. Griner. 2000. Trends in specialized surgical procedures at teaching and nonteaching hospitals. *Health Affairs* 19(1):230-238.
24. Kaiser Family Foundation. 2007. *How Changes in Medical Technology Affect Health Care Costs*.
25. Kuhn, H. 2006. *Testimony on Payment for Imaging Services under the Medicare Physician Fee Schedule, July 18, 2006*. Online. Available at http://www.hhs.gov/asl/testify/t060718.html. Accessed January 22, 2007.
26. Brindis, R. G., et al. 2005. ACCF/ASNC appropriateness criteria for single-photon emission computed tomography myocardial perfusion imaging (SPECT MPI): A report of the American College of Cardiology Foundation Quality Strategic Directions Committee Appropriateness Criteria Working Group and the American Society of Nuclear Cardiology endorsed by the American Heart Association. *Journal of the American College of Cardiology* 46(8):1587-1605.
27. Stead, W. W. 2007. *The Learning Health Care System: Workshop Summary*. Washington, DC: The National Academies Press.
28. Kaiser Family Foundation. 2006. *Trends and Indicators in the Changing Health Care Marketplace: Section One*. p. 8, Exhibit 1.8.
29. Kaiser Family Foundation and Health Research and Educational Trust. 2006. *Summary of Findings: National Survey of Private and Public Employers of Three or More Workers*.
30. Research!America. 2005. *National Survey*. Charlton Research Company.
31. IOM (Institute of Medicine). 2001. *Crossing the Quality Chasm*. Washington, DC: National Academy Press: Washington, DC.

32. Robinson, J. C. 2006. The commercial health insurance industry in an era of eroding employer coverage. *Health Affairs* 25(6):1475-1486.
33. Business Roundtable. 2006. *New CEO Survey Predicts Slow, Steady Growth for First Half of 2007; More Than Half of CEOs Say Health Care is Greatest Cost Pressure on Their Companies*. Washington, DC.
34. National Business Group on Health. 2007. Online. Available at www.businessgrouphealth.org. Accessed February 26, 2007.
35. Borger, C., et al. 2006. Health spending projections through 2015: Changes on the horizon. *Health Affairs* 25(2):w61-w73.
36. IOM (Institute of Medicine). 2006. *The Role of Purchases and Payers in the Clinical Research Enterprise: Workshop Summary*. Washington, DC: The National Academies Press.
37. Heres, S., et al. 2006. Why olanzapine beats risperidone, risperidone beats quetiapine, and quetiapine beats olanzapine: an exploratory analysis of head-to-head comparison studies of second-generation antipsychotics. *American Journal of Psychiatry* 163(2):185-194.
38. Lieberman, D. A., et al. 2000. Use of colonoscopy to screen asymptomatic adults for colorectal cancer. Veterans Affairs Cooperative Study Group 380. *New England Journal of Medicine* 343(3):162-168.
39. McDonnell, K., and P. Fronstin. 1999. *EBRI Health Benefits Data Book*. Washington, DC: Employee Benefit Research Institute.
40. NIH (National Institutes of Health). 2007. *Fact Sheet: Research Into What Works Best*.
41. Academy Health. 2005. *Placement, Coordination, and Funding of Health Services Research within the Federal Government*. Washington, DC.
42. Eddy, D. M. 2007. Linking electronic medical records to large-scale simulation models: Can we put rapid learning on turbo? *Health Affairs* 26(2):w125-136.
43. Eddy, D. 1989. Screening for breast cancer. *Annals of Internal Medicine* 1;111(5): 389-399.
44. Petitti, D. 2007. Personal communication. April 2.
45. Clancy, C. M. 2006. Getting to 'smart' health care. *Health Affairs* 25(6):w589-w592.
46. Wilensky, G.R. 2006. Developing a center for comparative effectiveness information. *Health Affairs* 25(6):w572-w585.
47. Kupersmith, J., et al. 2005. Creating a new structure for research on health care effectiveness. *Journal of Investigative Medicine* 53(2):67-72.
48. Hopayian, K. 2001. The need for caution in interpreting high quality systematic reviews. *BMJ* 323(7314):681-684.
49. Buto, K., and P. Juhn. 2006. Can a center for comparative effectiveness information succeed? Perspectives from a health care company. *Health Affairs* 25(6):w586-w588.
50. Rowe, J. W., D. A. Cortese, and J. M. McGinnis. 2006. The emerging context for advances in comparative effectiveness assessment. *Health Affairs* 25(6):w593-w595.
51. The Health Industry Forum. 2006. *Comparative Effectiveness Forum Executive Summary*. November 30. Washington, DC.
52. Cook, D. J., et al. 1996. **Stress ulcer prophylaxis in critically ill patients. Resolving discordant meta-analyses.** *JAMA* 275(4):308-314.
53. Jadad, A. R., D. J. Cook, and G. P. Browman. 1997. A guide to interpreting discordant systematic reviews. *CMAJ* 156(10):1411-1416.
54. Eddy, D., and L. Schlessinger. 2003. Validation of the archimedes diabetes model. *Diabetes Care* 26(11):3102-3110.
55. Owens, D. 2007. Personal communication. February 13.
56. Juhn, P. 2007. Personal communication. February 25.

57. Foundation Center. 2007. Online. Available at www.foundationcenter.org. Accessed February 16, 2007.
58. IOM (Institute of Medicine). 2006. *The Future of Drug Safety: Promoting and Protecting the Health of the Public*. Washington, DC: The National Academies Press.

APPENDIX ONE

CURRENT NATIONAL CAPACITY FOR CLINICAL EFFECTIVENESS RESEARCH

Information is not currently available to characterize precisely the national expenditure on clinical effectiveness research, but the investment is clearly far short of the need. If only 1 percent of the nation's $2 trillion health care bill was devoted to understanding the effectiveness of the care purchased, the total would come approximately to $20 billion annually. In contrast, only $15 million annually has been specifically appropriated by Congress to the Agency for Healthcare Research and Quality (AHRQ), under section 1013 of the Medicare Prescription Drug, Improvement, and Modernization Act (MMA), for comparative effectiveness research, the issue for which the shortfall is greatest. The total of all appropriations to all federal agencies—the National Institutes of Health (NIH), the Veteran's Health Administration (VHA), the Department of Defense (DOD), the Centers for Medicare and Medicaid Services (CMS), the Food and Drug Administration (FDA), AHRQ, and the Centers for Disease Control and Prevention (CDC)—for all health services research amounts to about $1.5 billion, only a small portion of which is devoted to clinical effectiveness research [1]. Apart from federal appropriations, the insurance industry currently undertakes significant effort and investments annually in assessment of clinical interventions, and healthcare manufacturers devote more than $15 billion dollars annually to their sponsorship of Phase III and IV clinical trials that provide insights on clinical effectiveness [2]. Data are not currently available on the amount of direct expenditures by professional societies on primary and secondary clinical effectiveness research. Even accounting for all the support from all of the sources involved, the aggregate national commitment to assessing the effectiveness of clinical interventions is likely well under 1 percent—far below the standard that any company would expect to invest in work to evaluate and improve its products. Presented below is a summary of the institutions and activities currently engaged in trying to meet the need for better information on the

effectiveness of healthcare interventions, beginning with a short description of relevant terms and definitions.

Terms

As used in this paper, clinical effectiveness research can be described as either primary or secondary. For our purposes, primary refers to the direct generation of evidence through the use of a specific experimental methodology. Secondary refers to the systematic gathering and evaluation of primary research information to further the understanding of common conclusions or disparate results.

Primary Clinical Effectiveness Research

In this respect, primary clinical effectiveness research refers to the specific design and implementation of structured research protocols to produce data on the results of one or more diagnostic or therapeutic interventions of interest. Examples include certain randomized controlled trials, practical clinical trials, cluster randomized trials, observational studies, and cohort studies, including registries. Some of these studies focus only the *efficacy* of an intervention—the extent to which an intervention produces a beneficial result under ideal circumstances. But many also examine the *effectiveness* of an intervention when used under ordinary circumstances—including evaluation in broader patient populations and healthcare delivery settings, or the relative risks and benefits of competing therapies. Both types of evaluation are important to gain an understanding of which interventions work best, for whom, and under what circumstances.

Evidence Synthesis

Evidence synthesis, or *secondary clinical effectiveness research,* refers to the structured assessment of evidence from multiple primary studies to derive conclusions, which are considered to have greater weight than an individual study alone. Terms sometimes used to describe activities of secondary clinical effectiveness research include *systematic review* and *technology assessment*, which both describe a systematic method of identifying, assembling, and interpreting a body of data to validate or extend the interpretation of single trials, lend context to individual trials, and, where possible, arrive at common conclusions. For the purpose of this discussion the two are seen as equivalent processes, although in general practice the audiences for these products are different. Systematic reviews are frequently published through the peer-reviewed literature or other mechanisms to inform policy and practice, while technology assessment is often carried out

to assist in decision making on a policy level regarding the effectiveness of a new intervention.

Comparative Effectiveness

Within the overall umbrella of clinical effectiveness research, the most practical need is for studies of comparative effectiveness, the comparison of one diagnostic or treatment option with one or more others. In this respect, 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, and secondary comparative effectiveness research involves the synthesis of primary studies to allow conclusions to be drawn. Secondary comparisons of the relative merits of different diagnostic or treatment interventions can be done either directly, through collective analysis of the results of multiple head-to-head studies, or indirectly, in which case the treatment options have not been directly compared to each other in a clinical evaluation, and inferences must be drawn based on the relative effect of each intervention to a specific comparison, often a placebo.

Other Related Terms

Other relevant terms often used in the context of clinical effectiveness research discussions include cost-effectiveness analysis and cost-utility analysis. In *cost-effectiveness analysis* the economic cost per specified unit of health gain—e.g., reduced mortality or morbidity—is determined for a given intervention or family of interventions. This allows for one measure of the relative value of an intervention to be estimated in comparison to alternatives. *Cost-utility analysis* is a form of cost-effectiveness analysis estimating the cost of a specific utility gain, usually to the patients targeted—e.g., quality-adjusted life-years—for an individual intervention. Finally, the term health services research refers broadly to the multidisciplinary field of scientific investigation that studies how systemic factors shape the effectiveness of health care for different populations, including how social factors, financing systems, organizational structures and processes, health technologies, and personal behaviors affect access to health care, the quality and cost of health care, and measures of population health and well-being [1].

Primary Clinical Effectiveness Research

Conduct of primary research that produces data on the clinical effectiveness of candidate diagnostic and therapeutic interventions is sponsored and managed by a variety of federal agencies, including NIH, CMS, VHA,

FDA, CDC, and AHRQ. Similarly, a number of private initiatives provide information important to primary clinical effectiveness research, including the sponsorship by the healthcare industry of study protocols for their products, the coordinated work of health maintenance organizations to use their clinical databases to study clinical effectiveness, and work involving various payers in data consortia to improve the availability of information for tracking intervention safety and effectiveness. Efforts are under way through the network of regional health information organizations (RHIOs) to advance the progress and interoperability of electronic health data systems, which will enhance the capacity for primary clinical effectiveness research [3-5]. Finally, the use of mathematical modeling to construct clinical scenarios is a form of primary research in that it creates new information by building on established work and uses biological science and probability assumptions to project and test outcomes.

National Institutes of Health

NIH is the largest federal sponsor of clinical research. No direct mandate or prioritization process exists at the NIH for performing comparative effectiveness trials, but many trials of this nature have been conducted, sometimes in collaboration with CMS or AHRQ. Over the past quarter century, several institutes have sponsored or co-sponsored large multicenter comparative effectiveness studies, including studies of balloon angioplasty plus carotid stenting vs. carotid endarterctomy for stroke prevention; behavioral therapy vs. pharmacotherapy for alcoholism; various antipsychotics for schizophrenia; newer vs. older antihypertensive and lipid-lowering agents to prevent heart attacks; lung volume reduction surgery vs. standard medical care for emphysema patients; use of intensive therapy vs. routine glucose monitoring and control in the prevention of diabetes complications; addition of Nevirapine vs. Azidothymidine (AZT) alone in reducing mother-to-child transmission of HIV; combination therapy vs. single drug therapy (doxaziosin, finasteride) to treat benign prostatic hyperplasia; and use of implantable cardiac defibrillators (ICD) vs. antiarrythmic drugs in preventing sudden cardiac death. These comparative effectiveness trials represent an investment of about $660 million since 1982—representing a small share of its total budget of approximately $300 billion over that period [6, 7]. Through the National Library of Medicine, the NIH, in collaboration with the FDA, also maintains a national computer searchable database, ClinicalTrials.gov, to track all federal and privately sponsored clinical trials. This resource was mandated as part of the FDA Modernization Act and requires that any clinical trial being conducted to support an investigational new drug application be included in the database. A related effort, GlobalTrialBank.org, is under way under the auspices of the Ameri-

can Medical Informatics Association with the goal of forming a worldwide repository of clinical trial protocols and data, with the aim of facilitating evidence generation, as well as comparison and evaluation of trials.

Centers for Medicare and Medicaid Services

CMS does not conduct clinical research directly, but its data systems, demonstration and evaluation activities, and coverage policies offer powerful resources for assessing and monitoring clinical effectiveness. Claims data are maintained on the more than 42 million Americans served under Medicare and on the 47 million low-income people covered under Medicaid. Medicare and Medicaid each maintain databases of administrative data which can be accessed by authorized researchers. These data have been a rich resource for the study of practice variation patterns, the effects of volume on surgical outcomes, disparities in medical care, and other issues related to profiles of health services delivery and quality. In addition, they have been used to study the relative effectiveness of different interventions, such as bariatric surgery and treatments for myocardial infarction and asthma [8-10]. The VHA databases have been recently linked with Medicare data to create a merged database, the Diabetes Epidemiology Cohort (DepIC), which allowed an increased accuracy in the determination of diabetes prevalence in the VA population and which has the potential for significant contributions to insights on the comparative effectiveness of interventions [11]. With the passage of the Medicare Modernization Act (MMA) Part D benefit, which makes Medicare enrollees eligible for prescription drug coverage, extensive new opportunities are presented for assessing clinical effectiveness in a post-market environment by linking Part D data to data from Parts A, B, and C in the conduct of public health research. CMS has thus far been largely precluded from using its demonstration and evaluation authority to provide clinical effectiveness insights because this authority is required to be used only for delivery-related demonstrations that are budget neutral. With the recent launching of the CMS initiative on Coverage with Evidence Development, the prospect is emerging for a new capacity to generate important information on the most effective application of clinical interventions by requiring that coverage, for certain populations on which existing effectiveness information is limited, be contingent on subsequent data collection via a registry or other mechanism. Under the CED initiative, for example, CMS has entered into partnership with the American College of Cardiology to maintain a registry that will track the experience of medical patients receiving implantable cardioverter defibrillators (ICDs). Other CED possibilities include FDG-PET for suspected dementia, off-label uses of colorectal cancer drugs (co-sponsored with NCI), and use of home oxy-

gen therapy in moderate COPD (co-sponsored with National Heart, Lung, and Blood Institute [NHLBI]).

Food and Drug Administration

Although FDA supports selective, relatively small-scale intramural research projects, it does not conduct primary clinical effectiveness research. It requires such work by manufacturers to establish basic safety and efficacy information as part of the drug, biologic, and device approval processes (see *Healthcare Industry* below). FDA also collects and maintains data directly as part of its adverse event reporting process and related post-marketing surveillance work, and it frequently seeks commitments from sponsors to conduct post-market surveillance as a condition of approval. One key issue, which has been the subject of recent scrutiny, is the absence of a practical mechanism to ensure the conduct of the post-market study commitments, with the result that many are never completed. An IOM report on drug safety notes that of over 1,200 drug and biologic post-market commitments, about two-thirds had yet to be initiated [12]. Recently, FDA has proposed the development of an integrated national network, a sentinel network, formed through a series of public-private partnerships using developing electronic information technology systems to collect, analyze, and disseminate medical product safety information to healthcare practitioners and patients at the point of care. In general, the FDA does not require comparative effectiveness information to grant market approval of an agent or innovation.

Veterans Health Administration

The VHA has a strong clinical research program as a result of its 157 medical centers and more than 1,300 ambulatory, residential, and home- and community-based sites of care, and a state-of-the-art computerized clinical data system with 8 million patient records that allows the rendering of large quantities of patient data for analyses on a wide variety of clinical research topics. With more than 3,000 researchers throughout the system and specially designated "centers of excellence," the VHA devoted over $400 million in 2005 to clinical research, with a nearly equivalent amount coming from its medical care account to support research efforts. In addition to VHA-appropriated funds, research activities are supported through NIH grants and through the conduct of research sponsored by pharmaceutical companies. A substantial share of this research focuses on clinical effectiveness. Multicenter clinical trials investigate the best therapy for various conditions, including AIDS, alcoholism, schizophrenia, stroke, and Parkinson's disease, as well as health and rehabilitation services research.

Some examples of this research that have yielded valuable information on comparative clinical effectiveness include coronary artery bypass grafts vs. percutaneous coronary intervention for revascularization of high-risk patients [13], intravenous administration vs. subcutaneous administration of erythropoietin for severe anemia in hemodialysis patients [14], colonoscopy vs. sigmoidoscopy for primary screening of colon cancer [15], arthroscopic lavage vs. arthroscopic débridement vs. placebo [16], and new vs. old drugs for schizophrenia [17].

Centers for Disease Control and Prevention

CDC funds some health services research to guide decisions on public health services and systems, and some of this research may also examine effectiveness of some therapies in the area of infectious disease or vaccines [1]. In addition, CDC maintains a number of national data systems that are important resources for certain types of studies. National Vital Statistics are collected by CDC through a cooperative arrangement between states and the National Center for Health Statistics (NCHS). NCHS also operates the National Health Interview Survey and the National Health and Nutrition Examination Survey, which use personal interviews and health examinations to track important trends in health status and health-related behavior. Along with other information collected from health records, these data are used to evaluate the impact of health policies and programs and support other research activities related to health trends, services, and status. Also operated out of CDC, the Vaccine Safety Datalink (VSD) is a collaborative project with America's Health Insurance Plans (AHIP), eight large managed care organizations, and the Immunization Safety Office of CDC. Using administrative data sources, VSD's linked database has been used to provide comprehensive medical and immunization histories, provide guidance on immunization policy decisions, and provide a platform for research on the safety of vaccines [18].

Agency for Healthcare Research and Quality

AHRQ uses its more than $300 million annual budget to support health services research, improve the quality of health care, and promote evidence-based decision making. Among federal agencies, AHRQ carries a mandate through section 1013 of the Medicare Prescription Drug, Improvement, and Modernization Act (MMA), and a related $15 million appropriation, to perform research with a focus on the outcomes, comparative clinical effectiveness, and appropriateness of pharmaceuticals, devices, and healthcare services. The provision has led to the creation of AHRQ's Effective Health Care Program, which has three components: synthesizing

existing studies into comparative effectiveness reports (CERs) by evidence-based practice centers, developing evidence including research aimed at filling knowledge gaps about treatment effectiveness (DEcIDE centers), and improving communication of complex scientific findings to a variety of audiences (Eisenberg Center). Of these, the Developing Evidence to Inform Decisions about Effectiveness (DEcIDE) network supports work in 13 centers to perform primary research on clinical effectiveness.[1] The DEcIDE centers have, collectively, access to de-identified medical data for over 50 million patients, including Medicare's 42 million beneficiaries, which allows accelerated practical studies about the outcomes, comparative clinical effectiveness, safety, and appropriateness of healthcare items and services. *Initial research* focuses on the outcomes of prescription drug use and other interventions for which randomized controlled trials would not be feasible or timely or else would raise ethical concerns that are difficult to address. Comparative effectiveness studies initiated through DEcIDE include atypical vs. conventional antipsychotic medications, beta-adrenergic antagonists, new therapies for glucose control in diabetes mellitus, and extended use of clopidogrel in drug-eluting stents. AHRQ also maintains the Medical Expenditure Panel Survey (MEPS)—including the Household Component Survey, the Medical Provider Component Survey, and the Insurance Component Survey—which are used in policy-related and behavioral research on healthcare use, spending, and insurance coverage [19].

Healthcare Industry

Industry-sponsored trials represent a significant proportion of the billions invested by health manufacturers in research and development. For example, well over half of the $37 billion that pharmaceutical companies spent in 2004 on R&D was devoted to the conduct of Phase I to IV clinical trials. Of this total, about 40 percent went to the Phase III and IV trials that have particular relevance to clinical effectiveness [2]. Total R&D spending (including engineering research and pre- and post-marketing activities) by the top 100 device manufacturers was estimated to be $47.8 billion in 2003. Many of these studies are conducted with academic investigators and others are managed by contract research organizations (CROs). The data from these studies are considered proprietary and reside with the companies sponsoring the trials. While some industry-sponsored work is devoted to

[1] University of Illinois at Chicago; Acumen, LLC—Data Development for Part D; Johns Hopkins University; University of Maryland at Baltimore; Brigham and Women's Hospital—Antipsychotic Treatment in Older Adults, Outcome Sciences; University of North Carolina; Duke University; RTI International; University of Pennsylvania School of Medicine; Harvard Pilgrim Health Care, Inc.; University of Colorado at Denver; and Health Sciences Center, Vanderbilt University Medical Center.

comparative, or head-to-head, studies with other products, they have been a small percentage of the studies that industry sponsors, perhaps because of the sense of potential bias from a commercially sponsored comparative study [15, 20].

Integrated Delivery Systems

Large healthcare systems, such as large health maintenance organizations (HMOs), have the capacity to use both their clinical and administrative data systems for primary clinical effectiveness research—some sponsored out of their own resources, others sponsored by NIH or commercial concerns. Kaiser Permanente, Geisinger Health, and Group Health, for example, have efforts to utilize their large databases for improved understanding of what works best for which patients. Kaiser Permanente has integrated participation in multiple networks for information gathering into its routine approach to cancer care, including its own detailed database as well as participation in several national databases, such as the NIH-sponsored Cancer Research Network (CRN) [21]. At Intermountain healthcare, evaluation of large patient databases on diabetes management allowed for a rapid series of process changes which dramatically improved diabetes control [22]. Leadership has grown out of these systems for collaborative work on a number of dimensions. The Health Maintenance Organization Research Network (HMORN) is composed of 15 HMOs covering over 15 million individuals. There are formal research programs that utilize the electronic medical records within each individual organization, and the HMORN is also a designated center for effectiveness. Through this linkage, projects are examining asthma drug use, safety during pregnancy, survival of COPD patients exposed to specific therapies, antibiotic use in children, and many other areas. Similarly, the National Cancer Institute–funded HMO Cancer Research Network is a collaborative effort whose aim is to improve the effectiveness of preventive, therapeutic, and supportive interventions for cancer care (crn.cancer.gov).

Patient Registries

Patient registries provide an ever-expanding primary source of information about the performance of different interventions and are sponsored and maintained by many groups—e.g., companies, professional societies, academic researchers, and government agencies. AHRQ defines a patient registry as an organized system with a stated predetermined purpose that uses observational study methods to collect uniform data in order to evaluate specified outcomes for populations defined by a particular disease, condition, or exposure. There are multiple types of registries, including

product, health services, and disease and condition. Many of the hundreds of patient registries in the United States are sponsored at individual academic medical centers or by organizations centered around specific diseases, while others are product registries maintained for post-marketing surveillance. For example, the IVAX Clozapine registry, mandated by the FDA, is designed to detect adverse events associated with use of the drug. Another example is the Cystic Fibrosis Foundation patient registry, which has been in existence for nearly 40 years and is used to help improve care for patients with cystic fibrosis. The quality of patient registries—and their utility for evaluating clinical effectiveness—varies. To begin improving the consistency, quality, and utility of registries, AHRQ has recently developed a guide for the creation and utilization of registries.

Payer-Related Data Consortia

If they are appropriately linked, health insurance claims can also provide a potentially strong resource for assessing the clinical effectiveness of various interventions. In July 2006 the AQA Alliance and the Hospital Quality Alliance (HQA) announced the formation of a joint effort to combine resources to identify, collect, and report data across the variety of care settings they represent through data aggregation, including information on quality of physician performance; cost-of-care measures; quality of care for specific conditions, such as heart attack or pneumonia; and other measures. A related variation on this theme, typical of the potential for such consortia, is the emerging Health Plan Consortium for Public Health, composed of health plans covering an estimated 100 million individuals, which is aimed initially at the use of participating organizations' databases to track the long-term safety of vaccines. The scale of this project will greatly improve the timeliness, generalizability, and statistical power of the record linkage studies, and, if successful, will expand to similarly track the experience for pharmaceuticals.

Mathematical Modeling

One example of emerging work to generate evidence in a novel fashion is the use of mathematical models. At Kaiser Permanente, for example, the Archimedes project is a mathematical modeling system that creates a virtual healthcare universe in which hospitals, doctors, patients, organs, and cells are modeled using mathematical equations and other formalisms. "Trials" may then be performed against this virtual medical universe in order to make predictions about the consequences. The model spans all levels, from physiology of disease to clinical treatment and organizational operations. Data were obtained from a number of different data sources, including

epidemiological studies, clinical trials, medical center records, and hospital accounting departments. Archimedes uses this complex model to offer consulting services regarding healthcare questions such as how to design a clinical trial, which drug shows better effectiveness, or how much money an organization might save through a specific treatment plan. Because modeling builds upon existing trial data, systems like Archimedes might offer an efficient way to develop comparative information on competing therapies.

Evidence Synthesis (Secondary Clinical Effectiveness Research)

Much of the work to marshal evidence for conclusions about clinical effectiveness takes the form not of primary data generation and analysis, but of systematic reviews and meta-analyses of existing studies. Such secondary clinical effectiveness research is sponsored and conducted by a variety of organizations with overlapping and intersecting activities and interests, including federal agencies, state agencies, insurer and insurer-related organizations, independent assessment centers, professional groups and societies, university centers, and consortia. Despite the effort and expense that goes into these analyses, often the results of the assessments are inconclusive with respect to a given intervention. Furthermore, those technology assessments that include a comparative component generally compare a new or evolving technology to an established technology, as opposed to head-to-head comparisons with other emerging technologies [23].

Federal Agencies

Most federal health agencies are involved at some level in the sponsorship of evidence synthesis for the purpose of drawing conclusions, making recommendations, or crafting policy related to the use of clinical diagnostics and treatment interventions. Perhaps the best known is the work of the U.S. Preventive Services Task Force (USPSTF), which has since 1984 conducted systematic reviews of the evidence in support of clinical preventive services, applied rigorous criteria to classify and rate the level of the evidence, and, based on the overall strength of the evidence for a given condition, offer conclusions and recommendations. The USPSTF is now sponsored by AHRQ, and its approach has set a standard reference point for much of the subsequent work to synthesize evidence and make clinical recommendations. Building on this work, AHRQ has established a network of 13 AHRQ-sponsored *evidence-based practice centers*[2] (EPCs) that review

[2] Blue Cross and Blue Shield Association Technology Evaluation Center; Duke University; ECRI; Johns Hopkins University; McMaster University; Oregon Health & Science University; RTI International—University of North Carolina; Southern California; Stanford University—

literature, perform technology assessments, and produce evidence reports including comparative effectiveness reviews. The EPCs are located at the Technology Evaluation Center of the Blue Cross and Blue Shield Association, ECRI, and a number of universities: Duke, Johns Hopkins, Oregon Health & Sciences, North Carolina, Southern California, Stanford-UCSF, Tufts-New England Medical Center, and Minnesota, as well as McMaster, Alberta, and Ottawa in Canada. Now part of the AHRQ Effective Healthcare Program, EPCs produce studies that span broad issues ranging from clinical to social and behavioral, economic, and other healthcare organization and delivery issues. The EPCs often compare the relative effectiveness of different treatments, including drugs, and also identify gaps in the research. A related AHRQ- and FDA-sponsored effort, the 11 Centers for Education and Research on Therapeutics (CERTS) each focus on a specific patient population or therapeutic area in the conduct of research on ways to advance the optimal use of drugs, biologics, and medical devices by identifying best practices. Although the research typically does not compare specific treatments in a head-to-head fashion, implicit conclusions are drawn through the definitions of best practices. To keep track of the various clinical guideline development efforts, AHRQ also sponsors the National Guideline Clearinghouse (www.guideline.gov), which was originally developed in conjunction with the AMA and AHIP and is managed by ECRI (see below).

CMS draws directly on AHRQ-sponsored technology assessments for the use of its Medicare Coverage Advisory Committee (MEDCAC) to inform CMS coverage decision making. MCAC is charged with determining whether scientific evidence is adequate to support the routine clinical use of certain interventions in the population of Medicare beneficiaries as well as with assessing how the effectiveness of a candidate intervention compares with the effectiveness of established practice. Occasionally the coverage group in CMS will conduct its own evidence syntheses or build off related work done through the Cochrane Collaboration and related technology assessment organizations.

Various NIH institutes sponsor evidence syntheses as part of their initiatives to translate research results for provider and public use. The National Heart Lung and Blood Institute has, for example, supported a variety of committees charged with developing guidance on issues such as high blood pressure treatment, blood cholesterol management, tobacco cessation programs, and asthma management. Other institutes have sponsored guideline development in a variety of areas, including management of hepatitis C, celiac disease, and total knee replacement. In addition, since

---

University of California, San Francisco; Tufts University—New England Medical Center; University of Alberta, Edmonton; University of Minnesota, Minneapolis; University of Ottawa.

1977 the NIH has sponsored consensus development conferences and state-of-the-science conferences to offer public and professional guidance on matters of important clinical concern. Recently AHRQ's EPCs have begun providing the evidence reviews for the use of the panels convened to offer the recommendations.

The Department of Veterans Affairs (VA) also sponsors evidence reviews through the VA Technology Assessment Program (VATAP). VATAP is a national program within the Office of Patient Care Services dedicated to advancing evidence-based decision making in the VA. It provides systematic reviews on key healthcare technology issues to help in the service profile decisions of senior VA policy makers. VATAP evaluations encompass devices, drugs, procedures, and organizational and supportive systems used in health care.

The Department of Defense's TRICARE Management Activity also has a process in place to evaluate medical interventions to make coverage determinations. Often, these evaluations arise out of the appeals process, through physician or beneficiary inquiry, or through TMA's monitoring of the scientific literature. The group that formulates these evaluations reviews the scientific evidence; utilizes technology assessments, which are either accessed through subscription to a private service (currently Hayes) or obtained as custom reports from a private service (currently ECRI); and reviews published policy statements. A wide variety of interventions are assessed, including pharmaceuticals, devices, and procedures.

State Agencies

Certain states have established programs to conduct technology assessments and systematic reviews to assist in their decisions about Medicaid coverage approaches and the policies of state health insurance agencies. California, for example, has established the California Health Benefits Review Program with a legislative mandate to analyze the medical, financial, and public health impacts of proposed changes in mandated health benefits. The analyses have separate components for medical effectiveness, utilization and cost analysis, and coverage impact on public health. In addition, 13 state Medicaid programs have joined in support of the Drug Effectiveness Review Project (DERP) at the Center for Evidence-Based Policy at Oregon Health & Sciences University (OHSU) to provide assistance in assessing the pharmaceuticals to be provided under Medicaid. OHSU has also initiated a parallel program, the Medicaid Evidence-Based Decisions Project (MED), designed to create a research pool of evidence-based tools, such as technology assessments (see *University Centers* below).

Insurers and Insurer-Related Organizations

The largest insurers, such as United Healthcare, Wellpoint, Aetna, Humana, and CIGNA, have substantial in-house capabilities for the conduct of evidence reviews. Although the volume of reviews conducted internally by these organizations tends to be much larger than the number from outside organizations, they also license more formal assessments from technology assessment entities such as the Blue Cross Blue Shield Association's Technology Evaluation Center (TEC), ECRI, Hayes, and Cochrane (see below), as well as the health technology assessment work of other countries, such as NICE in the United Kingdom. In addition, these larger groups will consult with appropriate specialty societies and academic centers for guidance and help on clinical effectiveness research. Smaller plans and state and regional Blue Cross/Blue Shield organizations generally do not maintain this analytic capacity and instead look elsewhere for guidance. As a result, they depend more on organizations like TEC, which conducts evidence syntheses for clients in both the public and private sectors (including Kaiser Permanente and CMS). TEC was one of the first private sector agencies devoted to assessing evidence of clinical effectiveness of new technologies in direct response to requests from healthcare decision makers. TEC performs about 30 clinical effectiveness and appropriateness reviews per year for various procedures, devices, or drugs, using five criteria, also adopted by a number of other organizations: regulatory approval; effect on net health outcome; applicability of benefits in a real-world setting; impact relative to established alternatives (in health outcomes, not cost); and availability of the intervention. The resulting reports are reviewed and ratified by TEC's Medical Advisory Panel (MAP) and then distributed to subscribing health plans and provider groups. Another insurer-related activity is the Academy of Managed Care Pharmacy (AMCP), through which health plans utilize comparative analysis when developing formularies. AMCP guidelines provide detailed information on a drug's economic value relative to alternative therapies. Particular challenges faced by payers include the determination of patient selection criteria, technologies not identifiable in the claim payment process, and the evolution of devices (once a technology has received medical necessity approval, the next generation device is generally not reviewed, although there may be significant cost issues).

Independent Assessment Centers

In addition to TEC, several other independent assessment centers provide evidence syntheses for payers, healthcare delivery organizations, providers, employers, and manufacturers. Two of the most active in this respect are ECRI and Hayes. ECRI (formerly Emergency Care Research Institute)

is an AHRQ EPC that also provides fee-based content, such as commercial information services and technical assistance. ECRI performs some comparative technical assessments, mostly on devices and technologies, but also on some drugs and procedures. Subscribing member organizations also fund ECRI's Health Technology Assessment Information Service (HTAIS), which develops and disseminates evidence-based health technology assessment information. Similarly, Hayes, Inc. is a health technology assessment company targeted primarily at healthcare professionals and, to a lesser extent, to healthcare consumers. About 275 reports, including horizon scanning (80) and technology assessments (190), are issued annually by Hayes; they evaluate medical and surgical procedures, drugs, biologics, diagnostic and screening tests, medical devices and equipment, and complementary and alternative therapies in order to facilitate decision making in coverage policy, contracting, acquisition, and health technology deployment. Reports focus on new and emerging technologies likely to have a significant impact on the cost, utilization, and quality of health care. Costs to produce a single report range from $20,000 to $500,000, depending on the complexity of the analysis and scope of the literature [23].

Two other independent activities provide services intended to set the stage for identifying priority future evidence review efforts. HealthTech, a non-profit California-based research and education organization, conducts horizon scanning to help participating healthcare and health industry organizations plan for emerging developments. HealthTech's research covers new and emerging pharmaceuticals, biotechnology, medical devices, and information technology, focusing in particular on ways to engage emerging technologies for which, by definition, good evidence is lacking for decisions about purchase, deployment, and use. Through literature review, expert interviews, stakeholder analysis, expert panels, and product/developer reviews, HealthTech develops research forecasts as decision-making tools projecting possible technologies and impacts. A related activity, currently positioned as a HealthTech affiliate, is the Center for Medical Technology Policy (CMTP). CMTP is convening groups with interests in identifying critical clinical effectiveness knowledge gaps and plans to facilitate the design and implementation of related studies.

Professional Groups and Societies

Many clinicians rely on clinical practice guidelines for advice on best practices for specific clinical situations, in some cases providing advice on how to choose between options. Currently, most clinical practice guidelines are produced in association with physician specialty societies and published through multiple modalities. There are more than 150 medical specialty societies in the United States, and many are engaged in some form

of evidence review and guideline development, and the work spans beyond medical societies. Altogether, more than 300 organizations have published at least one guideline on the AHRQ-supported website, guidelines.gov. The national guideline clearinghouse currently contains more than 1,900 individual summaries. There is substantial variability in the approaches of the various societies, some conducting the reviews by informal staff-generated activities, and others engaged in large, multifaceted and structured consultations involving the multiple related organizations with similar interests in an issue. For example, the American College of Cardiology and the American Heart Association have a carefully specified cooperative protocol for conducting their reviews and developing their guidelines.

University Centers

Much of the work of various organizations sponsoring evidence syntheses, from AHRQ and NIH to states, professional societies, and manufacturers, is conducted by university researchers, some of whom reside in university centers substantially devoted to such efforts. These primarily include the 12 universities involved as EPCs, although individual academic researchers do conduct various forms of evidence syntheses within their areas of expertise. For example, the EPC at the University of Oregon does many of the systematic reviews for the work of the AHRQ-supported U.S. Preventive Services Task Force (noted above), and it also does the reviews for the Drug Effectiveness Review Program (DERP) (noted above), which compares evidence about effectiveness and safety of drugs of the same class, and uses information to inform health policy. Between 2003 and 2006, 26 reports were commissioned, and 12 were completed. Reports are also periodically updated. Final reports are delivered to subscribers and are also publicly available. Consumers Union also uses DERP findings as the basis for its *Best Buy Drugs Program.*

Consortia

Because of the overlapping nature of the interests, issues, and approaches, a number of collaborative efforts have emerged. The best known of these is the Cochrane Collaboration, an international effort (see next section), including a U.S. center, which sponsors systematic reviews using carefully developed common standards on a wide variety of issues in health care. Groups such as the Grading of Recommendations Assessment, Development and Evaluation Working Group (GRADE), and Appraisal of Guidelines Research and Evaluation (AGREE) have formed to develop standards for evidence syntheses and clinical practice guidelines. Other consortia involved in coordinating their work on clinical effectiveness research

include international technology assessment groups, including EUnetHTA and INAHTA (see Appendix Two below), as well as those mentioned earlier, including the HMO Research Network, the Cancer Research Network, and the Vaccine Safety Datalink at CDC.

REFERENCES

1. Academy Health. 2005. *Placement, Coordination, and Funding of Health Services Research within the Federal Government*. Washington, DC.
2. PhRMA (Pharmaceutical Research and Manufacturers of America). 2006. *Pharmaceutical Industry Profile*. p. 49, Table 4.
3. The White House. *Transforming Health Care: The President's Health Information Technology Plan, Promoting Innovation and Competitiveness: A New Generation of Innovation*. 2004 Online. Available at http://www.whitehouse.gov/infocus/technology/economic_policy200404/chap3.html. Accessed February 14, 2007.
4. HHS (U.S. Department of Health and Human Services). 2007. *Office of the National Coordinator for Health Information Technology (ONC): Goals of Strategic Framework*. Online. Available at www.hhs.gov/healthit/goals.html. Accessed February 13, 2007.
5. Bartschat, W., et al. 2006. Surveying the RHIO landscape. A description of current RHIO models, with a focus on patient identification. *Journal of AHIMA* 77(1):64A-64D.
6. NIH (National Institutes of Health). 2007. *The NIH Almanac—Appropriations*—Section 2. Online. Available at www.nih.gov/about/almanac/appropriations/part2.htm. Accessed February 13, 2007.
7. NIH (National Institutes of Health). 2007. *Fact Sheet: Research Into What Works Best*.
8. Flum, D. R., et al. 2005. Early mortality among Medicare beneficiaries undergoing bariatric surgical procedures. *JAMA* 294(15):1903-1908.
9. Udvarhelyi, I. S., et al., Acute myocardial infarction in the Medicare population. Process of care and clinical outcomes. *JAMA* **268**(18):2530-2536.
10. Heaton, P. C., et al. 2006. Analysis of the effectiveness and cost benefit of leukotriene modifiers in adults with asthma in the Ohio Medicaid population. *Journal of Managed Care Pharmacy* 12(1):33-42.
11. Kupersmith, J., et al. 2007. Advancing evidence-based care for diabetes: Lessons from the Veterans Health Administration. *Health Affairs*.
12. IOM (Institute of Medicine). 2006. *The Future of Drug Safety: Promoting and Protecting the Health of the Public*. Washington, DC: The National Academies Press.
13. Stroupe, K., et al. 2006. Investigators of Veterans Affairs Cooperative Studies Program #385 (AWESOME: Angina With Extremely Serious Operative Mortality Evaluation). *Circulation* 114(12):1251-1257.
14. Kaufman, J. S., et al. 1998. Subcutaneous compared with intravenous epoetin in patients receiving hemodialysis. Department of Veterans Affairs Cooperative Study Group on Erythropoietin in Hemodialysis Patients. *New England Journal of Medicine* 339(9):578-583.
15. Lieberman, D. A., et al. 200. Use of colonoscopy to screen asymptomatic adults for colorectal cancer. Veterans Affairs Cooperative Study Group 380. *New England Journal of Medicine* 343(3):162-168.
16. Moseley, J. B., et al. 2002. A controlled trial of arthroscopic surgery for osteoarthritis of the knee. *New England Journal of Medicine* 347(2):81-88.
17. Rosenheck, R., et al. 2003. Effectiveness and cost of olanzapine and haloperidol in the treatment of schizophrenia: A randomized controlled trial. *JAMA* 290(20):2693-2702.

18. CDC (Centers for Disease Control and Prevention). 2007. *Vaccine Safety Datalink Project (VSD)*. Online. Available at www.cdc.gov/od/science/iso/research_activities/vsdp.htm. Accessed January 31, 2007.
19. AHRQ (Agency for Healthcare Research and Quality). 2007. *Medical Expenditure Panel Survey (MEPS): Overview of the MEPS Web Site*. Online. Available at www.ahrq.ov/data/mepsweb.htm. Accessed January 31, 2007.
20. Heres, S., et al. 2006. Why olanzapine beats risperidone, risperidone beats quetiapine, and quetiapine beats olanzapine: An exploratory analysis of head-to-head comparison studies of second-generation antipsychotics. *American Journal of Psychiatry* 163(2):185-94.
21. Wallace, P. J. 2007. Reshaping cancer learning through the use of health information technology. *Health Affairs* 26(2):w169-w177.
22. IOM (Institute of Medicine). 2007. *Learning Health Care System: Workshop Summary*. Washington, DC: The National Academies Press.
23. Hayes, W. 2007. Personal communication.

APPENDIX TWO

INTERNATIONAL ACTIVITIES IN CLINICAL EFFECTIVENESS RESEARCH

Clinical effectiveness assessment activities have utility that crosses national borders. A growing involvement in technology assessment and clinical effectiveness evaluation has developed internationally. The explicit incorporation of cost-effectiveness evaluation tends to be a more common feature of the work abroad, prompted by the need for better information with which to make decisions about the use of public monies. Implicit in the systematic evaluation of cost for an intervention is the comparison of the intervention with alternatives for care as a means to arrive at an assignment of value for investment. Thus, the majority of comparative effectiveness analysis that is performed in an international context is based, to some degree, on the cost of the intervention. Noted below are examples of some of the work under way in individual countries and through multicountry collaborations.

Country Specific Programs

Australia [1-3]

The Australian government's Department of Health and Ageing maintains several parallel pathways for the evaluation of various interventions: the Therapeutic Goods Administration (TGA), the Medical Services Advisory Committee (MSAC), the Pharmaceutical Benefits Advisory Com-

mittee (PBAC), and the Prostheses and Devices Committee (PDC). The Therapeutic Goods Agency is responsible for the regulation of multiple classes of therapeutic products, including pharmaceuticals (prescription and over-the-counter), devices, cellular therapies, and blood and tissue products, and it serves as the primary gateway for introduction of these interventions. Once a product is approved for use, there is a separate process for determination of coverage of regulated products. The PBAC reviews applications from pharmaceutical companies for pharmaceuticals to be subsidized under the Pharmaceutical Benefits Scheme (PBS). Each application undergoes an economic cost and utilization review. A similar process is undertaken for the evaluation of medical technologies and procedures, in which the MSAC provides advice to the Australian government as to whether new technologies and procedures should be considered for reimbursement under the Medicare Benefits Schedule (MBS). The primary aim of the MSAC is to ensure that new and existing medical procedures are supported by evidence of their safety, clinical effectiveness, and cost effectiveness. The option of interim funding can be linked to a requirement for additional data acquisition.

Canada [4]

The Canadian Agency for Drugs and Technologies in Health (CADTH) is charged with providing Canada's federal, provincial, and territorial healthcare decision makers with counsel on the effectiveness and efficiency of healthcare interventions. CADTH operates as an independent, not-for-profit organization with three main programs: Health Technology Assessment (HTA), Common Drug Review (CDR), and the Canadian Optimal Medication Prescribing and Utilization Service (COMPUS). The HTA program conducts assessments of new and existing technologies for the technology's impact on health, comparisons with alternatives, value for investment, and other health services implications. HTA work is performed both through internal capacity and through a program of extramural research contracts and grants. HTA also operates a horizon scanning program. Since 2003 the CDR has provided a centralized process for evaluation of pharmaceuticals in order to reduce the duplication inherent in each locality independently preparing an evaluation. The recommendations of the CDR advisory committee are advisory and non-binding on local decision making. A technical assistance institution, COMPUS, was initiated in 2004 to help identify and promote evidence-based best practices in drug prescribing and utilization, including maintenance of a virtual library of stakeholder-submitted best-practices guidelines.

France [5]

In 2004 the French government established the Haute Authorie de Sante (HAS) to bring together a number of activities related to patient care improvement: evaluation of drugs, procedures, and devices; publication of clinical guidelines; accreditation of healthcare providers and doctors; and dissemination of medical information. It carries out these activities through seven special committees: the Transparency Committee, which evaluates medicines; the Committee for Assessment of Devices and Health Technologies (CEPP); the Committee for Assessment of Diagnostic and Therapeutic Procedures (CEAP); Committee for Healthcare Cover for Chronic Conditions; the Committee for Practice Guidelines and Practice Improvement; the Committee for Medical Information Quality and Dissemination; and the Committee for Accreditation of Healthcare Organizations. The Transparency Committee reviews approved pharmaceuticals for reimbursement and use under national health insurance, based on both the baseline and marginal clinical benefit. It also organizes post-market surveillance for drugs. While the functions of the Transparency Committee and CEPP have existed for some time, the introduction of CEAP added the capacity to study diagnostic and therapeutic procedures with respect to their benefit, comparison with other strategies, temporal placement of the procedure in the timeline of care, necessary skills for operators, and economic assessment. HAS is supported by taxes on promotional spending by pharmaceutical companies, health insurance funds, state funding, accreditation fees, and application fees.

Germany [6]

The Institute for Quality and Efficiency (IQWiG) has been operating in Germany since October 2004 as an independent, non-profit organization responsible for the evaluation of diagnostic and therapeutic services, pharmaceuticals, clinical practice guidelines, and healthcare services. It also makes recommendations to disease management programs, and disseminates information to patients and consumers. IQWiG was established by the Federal Joint Committee (G-BA), a self-governing organization made up of physicians and health insurance funds that is responsible for the administration of health services in Germany and which is funded by contributions from statutory health insurance funds. Under the new framework, IQWiG is responsible for conducting assessments on topics requested by G-BA or the Federal Ministry of Health, and the conclusions of these reports are used for decision making regarding coverage. Most interventions require a randomized controlled trial that demonstrates efficacy in order to receive a positive recommendation from IQWiG. For the purpose of reimbursement,

medical interventions are grouped into clusters, and reference pricing is used for all treatments which are determined to fall within a specific cluster.

United Kingdom [7, 8]

The National Institute for Health and Clinical Excellence (NICE) was established in 1999 to provide patients, health professionals, and the public with authoritative guidance on public health, health technologies, and clinical practices. Three centers guide the programs at NICE: the Centre for Public Health Excellence, which develops public health guidance; the Centre for Clinical Practice, which develops clinical practice guidelines; and the Centre for Health Technology Evaluation, which assesses new and existing medicines, treatments, and procedures. The assessment process involves a technical committee (the Technology Appraisal Committee) as well as economic modeling of cost effectiveness. Technology approvals by NICE mandate coverage for those who meet the approved indications for a treatment. The work of NICE is facilitated by the availability of multiple large databases such as the General Practice Research Database, which contains longitudinal anonymized primary care data for 35 million patient-years and 3 million active patients.

Sweden [9]

The Ministry of Health and Social Affairs and the National Board of Health and Welfare are two centralized agencies that oversee operation of the Swedish healthcare system. However, the system is decentralized into localities which are primarily responsible for decision making. The Swedish Council on Technology Assessment in Health Care (SBU), formed in 1987, is an agency that undertakes health technology assessments to aid in the decision-making process for the localities; it works in cooperation with the National Board of Heath and Welfare as well as other governmental agencies such as the Medical Products Agency and the Pharmaceutical Benefits Board. Economic evaluations are included in the assessments undertaken by SBU.

Multinational Collaborations

European Union [10]

The European Medicines Agency (EMEA) was established in 1995 as a centralized means of establishing marketing authority for pharmaceuticals in the European Union. A single application process conducted through the EMEA allows for marketing approval for pharmaceuticals and biophar-

maceuticals based upon an evaluation of the quality, safety, and efficacy of the medicinal products submitted. The EMEA coordinates the evaluation of medicinal products through its network of roughly 3,500 experts. Guidance from the EMEA is utilized by member states in varying ways, and member states are free to adopt additional measures, such as those related to cost, with regard to approved agents. The EMEA has adopted a standard emphasizing the use of comparator evaluation when an established pharmacological alternative is available. Processes used by member states to arrive at relative clinical effectiveness and cost-effectiveness have not been fully clarified.

EUnetHTA [11]

A collaborative project that began in January 2006, EUnetHTA seeks to develop an organizational framework for a sustainable European Network for Health Technology Assessment. EUnetHTA involves 59 partners from 31 countries (including the United States, Canada, Australia, and Israel), and is coordinated and led by the Danish Centre for Evaluation and Health Technology Assessment. It will operate until at least 2008 in the work to create a permanent capacity. It builds on a number of predecessor efforts: EUR-ASSESS (1994 to 1997), HTA-Europe (1997 to 1999), and the European Collaboration for Assessment of Health Interventions-Health Technology Assessment (1999 to 2001). Functions under exploration include facilitation of networking, identification of opportunities for joint assessments and priority setting, development of methodology, gathering and dissemination of information, and educational activities.

INAHTA [12]

The International Network of Agencies for Health Technology Assessment (INAHTA) was established in 1993 and has 45 member organizations from 23 countries. Member organizations must be nonprofit, have a relationship with a local or national government, and be at least 50 percent publicly funded. Through a Web-based service, INAHTA seeks to facilitate the extent to which various HTA organizations can benefit from the work of other centers, through networking, communication, and process improvement.

Cochrane Collaboration [13]

Perhaps the best known multinational organization working to evaluate evidence of healthcare interventions is the Cochrane Collaboration. Founded in 1993, its major product is the Cochrane Database of System-

atic Reviews. Review groups clustered around medical specialties follow a standardized procedure to produce the systematic reviews, including the preparation, formulation of the problem, locating and selecting studies, assessment of study quality, analysis, and interpretation. The reviews are published electronically and are periodically updated as new information becomes available. A methods advisory group provides advice and assistance to the collaboration in the methodology of systematic review. Organized around disease topics, the Cochrane Collaboration now has more than 4,500 systematic reviews in its database. Among the 12 Cochrane centers worldwide, one is based in the United States, at Johns Hopkins Bloomberg School of Public Health. The central organizing functions of the collaboration are funded through subscription revenue, while individual entities are funded through a wide variety of other sources. including governmental, institutional, and private funding.

REFERENCES

1. Australian Government Department of Health and Ageing. 2006. *Public Support for Science and Innovation*. Online. Available at biblio.ind.com.gov/au/study/science/subs/sub105.rtf. Accessed 2006.
2. MSAC (Medical Services Advisory Committee). 2006. *Strenthening Evidenced-Based Health Care in Australia*. Online. Available at www.msac.gov.au. Accessed December 5, 2007.
3. O'Malley, S. P. 2006. The Australian experiment: The use of evidence based medicine for the reimbursement of surgical and diagnostic procedures (1998-2004). *Australian and New Zealand Health Policy* 3(3).
4. Canadian Agency for Drugs and Technologies in Health. 2007. Online. Available at www.cadth.ca/index.php/en/home. Accessed 2007.
5. Haute Autorite de Sante (HAS). 2007. Online. Available at www.has-sante.fr. Accessed February 14, 2007.
6. Institute for Quality and Efficiency in Health Care (IQWiG). 2007. Online. Available at www.iqwig.de/about-us.21.en.html. Accessed February 14, 2007.
7. National Institute for Health and Clinical Excellence. 2007. Online. Available at www.nice.org/uk. Accessed February 14, 2007.
8. Pearson, S. D., and M. D. Rawlins. 2005. Quality, innovation, and value for money: NICE and the British National Health Service. *JAMA* 294(20):2618-2622.
9. www.sbu.se/www.html. 2007. Online. Accessed February 14, 2007.
10. European Medicines Agency. 2007. Online. Available at www.emea.eu.int. Accessed February 14, 2007.
11. EUnetHTA. 2007. Online. Available at www.eunethta.net. Accessed February 14, 2007.
12. *INAHTA: Global Networking for Effective Healthcare*. 2007. Online. Available at www.inahta.org. Accessed February 14, 2007.
13. The Cochrane Collaboration. 2007. Online. Available at www.cochrane.org. Accessed February 14, 2007.

APPENDIX THREE

POTENTIAL MODEL: FEDERALLY FUNDED RESEARCH AND DEVELOPMENT CENTERS

Federally Funded Research and Development Center (FFRDC)

Organizations spanning both the public and private domains exist in many forms. One such quasi-governmental entity is the Federally Funded Research and Development Center (FFRDC). FFRDCs are congressionally chartered to meet the specific needs of a government agency, and strict federal oversight maintains the course of FFRDC research. An FFRDC is funded by the federal government and functions as a non-profit private organization. FFRDCs were first established during World War II to assist government agencies in meeting specific, well-defined technical needs that could not be met by existing government agencies or by the private sector. The mission and core competencies of an FFRDC are developed to support the needs of the sponsoring governmental agency. Currently there are 37 FFRDCs, and, while their missions are quite varied, there are general rules applied to their administration. The sponsoring agency is responsible for overall policy and oversight following guidelines set forth by Office of Federal Procurement Policy (OFPP) Policy Letter 84-1. FFRDCs are managed by non-governmental organizations, usually universities or other nonprofit institutions. These operating parent organizations may also operate multiple FFRDCs or non-FFRDC organizations. According to the OFPP, federal agencies should only establish an FFRDC when the agency determines that "[e]xisting alternative sources for satisfying agency requirements cannot effectively meet the sponsor's special research and development needs."

FFRDC funding comes from the organization within the sponsoring agency requesting the work or from a line item on congressionally appropriated budgets of the sponsoring agency. An FFRDC can receive up to 30 percent of its funding from private sources. Competition for other government or commercial business is limited with the intention of fostering a strategic relationship between an FFRDC and its sponsor as well as reducing the potential for conflict of interest that this special access may create. In FY 1993 the federal government spent $67.3 billion on R&D, of which $5.9 billion, or 8.7 percent, was obligated for FFRDCs.

Existing FFRDCs fall into four general categories: policy-focused studies and analyses centers, research and development (R&D) centers, research laboratories, and systems engineering and integration centers. Research laboratories differ from R&D centers in that they are usually limited to basic and applied research that seeks solutions to specific problems but

do not engage in major developmental activities. At this time the government agencies sponsoring FFRDCs include the Department of Defense, the Department of Energy, the National Science Foundation, the Federal Aviation Administration, the National Aeronautics and Space Administration, the National Institutes of Health, the Nuclear Regulatory Commission, and the Internal Revenue Service.

Six FFRDCs are briefly described below, with details offered about their governance structure, funding mechanism, and output.

National Defense Research Institute (NDRI)[3]
Sponsored by Department of Defense (DoD)
Administered by RAND Corporation

Purpose Established as an FFRDC in 1984, NDRI is the primary source of analytic expertise regarding national defense policy and strategy problems for the Office of the Secretary of Defense (OSD), the Joint Staff, the Unified Combatant Commands, the defense agencies, the United States Marine Corps, and the United States Navy. NDRI identifies and evaluates new policy options and develops alternative approaches to implement current policy. Corresponding with the responsibilities of four OSD under secretaries, NDRI research is conducted within RAND's National Security Research Division in four interconnected centers: International Security and Defense Policy Center, Acquisition and Technology Policy Center, Forces and Resources Policy Center, and Intelligence Policy Center.

Governance An advisory board made up of 13 senior DoD civilians oversees the NDRI research program as well as the DoD budget request.

Funding NDRI is funded primarily by its sponsor, DoD, with a renewable 5-year contract. Additional funding comes from the departments of its Advisory Board members [1]. Non-DoD agencies that fund NDRI include agencies from the intelligence community, the U.S. Coast Guard, the Department of State, allied governments and their ministries of defense, and foundations and private contributors. RAND supports some NSRD research with its discretionary funds. The fiscal year 2001 research and development (R&D) expenditure was $25.5 million, all of which was federal [2].

Output NDRI research findings are documented as reports, both classified and unclassified. Examples of recently published reports include *Review of Military Health Benefit Design, China's Military Modernization and the*

[3] Information summarized from NDRI Web site http://www.rand.org/nsrd/ndri.html and links therein.

Cross-Strait Balance, *Monitoring the Progress of Shipbuilding Programs*, and *How Can the Defense Procurement Agency More Accurately Monitor Progress?*

Homeland Security Institute (HSI)[4]
Sponsored by Department of Homeland Security (DHS)
Administered by Analytic Services, Inc.

Purpose Established in 2004, HSI evaluates homeland security systems and technologies during development, deployment, and use for DHS and its operating elements. HSI has four core competencies: systems evaluations; technology assessments; operational assessments; and resource and support analyses. Within Title III, Science and Technology in Support of Homeland Security, of the Homeland Security Act (P.L. 107-296) there are two provisions that call for the establishment of FFRDCs. Section 312 directs the secretary of DHS to establish an FFRDC known as the Homeland Security Institute. According to the legislation, the institute should, among other things, conduct systems analysis, risk analysis, and simulation and modeling in order to assess the vulnerability of the nation's critical infrastructure.

Governance Acting as the primary sponsor, the DHS under secretary for science and technology designates the executive agent who provides DHS oversight of HSI. The executive agent oversees HSI activities, chairs and designates the membership of the HSI Advisory Group, and designates the program manager. The advisory group oversees the research program. The program manager works to implement the decisions of the advisory group.

Funding HSI receives $10 million in annual core funding [3]. HSI core tasks are funded by the Science and Technology (S&T) Directorate. Analytic tasks are funded by various sponsors throughout DHS, including the S&T directorate.

Output HSI supplies research, studies, analyses, analytic and computational models, simulations, and other technical and analytical support. HSI also distributes weekly newsletters and publishes the *Journal of Homeland Security*.

[4] Information summarized from HSI Web site http://www.homelandsecurity.org/about.asp and links therein.

Science and Technology Policy Institute (STPI)[5]
Sponsored by National Science Foundation (NSF)
Administered by Institute for Defense Analyses (IDA)

Purpose Established in 1991, STPI (named the Critical Technologies Institute until 1998) provides technical analytical support to inform the Executive Branch as it formulates federal science and technology policy. STPI conducts research for the Office of Science and Technology Policy (OSTP) and other government users. In addition to providing information regarding science and technology developments and trends, STPI identifies roles for the federal government and other sectors to ensure long-term U.S. competitiveness. Science and technology issues addressed by STPI include national security, homeland security, critical infrastructure protection, health and the environment, space and transportation, information/telecommunication infrastructure and technology, education and training, physical sciences, engineering and technology, research and development portfolio, and development of new analytic methods and tools.

Governance The director of STPI reports to the president of IDA. IDA leadership includes a board of trustees and corporate officers. OSTP approves all projects requested by other agencies. Other agencies requesting projects transfer funds to NSF, which then contracts with STPI.

Funding STPI is funded on a project basis. The fiscal year 2001 R&D expenditure was $7.3 million, all of which was federal [2]. The NSF budget request for fiscal year 2008 for projects requested by OSTP provides $3.04 million for STPI.

Output STPI produces informal briefings as well as formal reports for the sponsoring agency. Decisions concerning public availability are strictly up to the sponsoring agency.

Lawrence Livermore National Laboratory (LLNL)[6]
Sponsored by Department of Energy (DOE)
Administered by University of California
(contract expires September 2007)

Purpose Established in 1952 as a research and development FFRDC, LLNL is one of three national laboratories that are part of National Nuclear

[5] Information summarized from STPI Web site http://www.ida.org/stpi/pages/about.html and links therein.

[6] Information summarized from LLNL Web site http://www.llnl.gov/llnl/about/ and links therein.

Security Administration (NNSA); the other two are Los Alamos National Laboratory and Sandia National Laboratories. LLNL is an applied science laboratory that provides scientific and technical expertise to ensure the safety and reliability of the nation's nuclear weapons. LLNL also supports national security needs by developing capabilities for nonproliferation. LLNL principle activities include stockpile surveillance, stockpile refurbishment, and integrated program management. LLNL's eight core competencies are physics, computing, biology, engineering, national security, lasers and optics, chemistry and materials science, and energy and environment.

Governance DOE determines the mission of LLNL. LLNL determines how best to carry out the research. University of California (UC) management provides the appropriate long-term research environment. The UC board of regents appoints the laboratory director and senior executive team (the director, the deputy director for operations, the deputy director for science and technology, the laboratory executive officer, the associate director at large, and the director's chief of staff). Historically, the amount of DOE oversight of LLNL has varied, with increases in DOE oversight in the 1980s due to environmental issues at nuclear weapons sites. The subsequent development of performance-based management by UC and DOE resulted in an increase in UC oversight.

Funding LLNL support comes largely from the NNSA Office of Defense Programs for nuclear weapons stockpile stewardship activities. National security and homeland security work is also funded by the NNSA Office of Defense Nuclear Nonproliferation, the Department of Homeland Security, various Department of Defense sponsors, and other federal agencies. The National Aeronautics and Space Administration, Nuclear Regulatory Commission, National Institutes of Health, Environmental Protection Agency, California state agencies, and private industry provide additional LLNL funding. The fiscal year 2001 R&D expenditure was $1.1 billion, of which 96 percent was federal [2]. The FY 2005 LLNL annual budget was $1.6 billion.

Output Recent examples of LLNL-developed advanced technologies include a device to detect highly enriched uranium inside cargo containers, a radiation detector, and a bioagent monitor. LLNL also provides an annual assessment report for nuclear weapons systems.

National Cancer Institute at Frederick (NCI-Frederick)[7]
Sponsored by Department of Health and Human Services (HHS)
Administered by Science Applications International Corp. (SAIC); Charles River Laboratories, Inc.; Data Management Services, Inc.; and Wilson Information Services, Inc.

Purpose A part of the National Institutes of Health (NIH), NCI-Frederick is one of two NCI campuses. The NCI's clinical researchers and the NIH Clinical Center are located on the NIH campus in Bethesda, Maryland. The NCI's Frederick campus is located within Fort Detrick, a U.S. Army base in Frederick, Maryland. NCI-Frederick is a GOCO (government-owned, contractor-operated) facility operated by four separate contractors providing services and support for (1) operations and technical support, (2) animal production, (3) computer and statistical services, and (4) scientific library services. NCI-Frederick employs more than 100 scientists to provide scientific and technical support to various NCI and NIH programs involved in the advancement of knowledge and tools for diagnosis, treatment, and prevention of human cancer and AIDS. Researchers work toward understanding the underlying genetic, molecular, environmental, and behavioral factors that contribute to human cancers, as well as developing novel tools for cancer diagnosis, treatment, and prevention. Additionally, NCI-Frederick provides scientific expertise and technology development to the NIH with its Research Technology Program (RTP) and other programs.

Governance The associate director at NCI-Frederick provides leadership and strategic planning, working with the NCI director to ensure that resources are being used to fulfill the FFRDC's mission. The Office of the Director at NCI through its Office of Scientific Operations develops and coordinates contractor requirements, providing overall scientific administrative management and program planning. The Office of the Director for the Center for Cancer Research (CCR) is also involved in strategic planning and determines CCR's scientific needs for NCI-Frederick resource allocation.

Funding The fiscal year 2001 R&D expenditure was $187 million, of which $183 million was federal [3]. The fiscal year 2008 proposed budget for the NCI provides up to $8 million for facilities repairs and improvements at NCI-Frederick [4].

[7] Information was summarized from the NCI-Frederick Web page http://web.ncifcrf.gov/about/ and pages therein.

Output Scientific research is published and disseminated to the research community in peer-reviewed publications. The following research support services are available at NCI-Frederick: animal resources; bioinformatics and biostatistics; chemistry and structural biology; mass spectrometry center; clinical research support; Economy Act/Work For Others; flow cytometry; genetics, genomics, DNA sequencing; imaging/microscopy; mass spectrometry; microarray and quantification of gene expression; protein chemistry and biophysics; proteomics; repository services; and scientific graphics and media.

Aerospace Federally Funded Research and Development Center (Aerospace)[8]
Sponsored by Department of Defense
Administered by the Aerospace Corporation

Purpose Established in 1960, Aerospace is the systems engineering and integration center type of FFRDC. Aerospace provides scientific and engineering support for both the long-term programs and the immediate needs of U.S. military and reconnaissance space programs for National Security Space. Aerospace mainly works for the Space and Missile Systems Center of Air Force Space Command and the National Reconnaissance Office. Aerospace is involved in the concept, design, acquisition, development, deployment, and operation of missions. Aerospace's core competencies, as defined by DoD, are launch certification, system-of-systems engineering, systems development and acquisition, process implementation, and technology application.

Governance The board of trustees of the Aerospace Corporation elects corporate officers and sets policy for its FFRDC. The board also supervises and directs the general management of the corporation. The 20 board members are elected from business, scientific, academic and public-service sectors to 3-year terms. The chief executive officer and president are ex officio members and are elected annually by the board.

Funding The total funding for Aerospace was $381 million for FY 1994, which includes $365.5 million from DoD obligations.

Output Aerospace has been involved in almost all national-security launch and satellite programs. For example, in fiscal year 2005 Aerospace pro-

[8] Information summarized from the Aerospace Corporation Web site http://www.aero.org/corporation/ffrdc.html and links therein.

vided oversight to 43 operational launches for the Space Missile Systems Center and the Air Force. Aerospace also provides technical analyses and assessments.

REFERENCES

1. GAO (United States General Accounting Office). 1996. *Federally Funded R&D Centers: Information on the Size and Scope of DoD-Sponsored Centers*.
2. NSF (National Science Foundation). 2002. *Federally Funded Research and Development Centers' R&D Expenditures: Fiscal Year 2001. Table 9*. Online. Available at www.nsf.gov/statistics/nsf03313/ pdf/tab9.pdf. Accessed 2007.
3. House of Representatives Committee on Science. 2007. *An Overview of the Federal R&D Budget for Fiscal Year 2007*. Online. Available at http://commdocs.house.gov/committees/science/hsy25938.000/hsy25938_0.HTM. Accessed 2007.
4. OMB (Office of Management and Budget). 2007. *FY 2008 Budget Appendix*. Online. Available at http://www.whitehouse.gov/omb/budget/fy2008/pdf/appendix/hhs.pdf. Accessed 2007.

APPENDIX FOUR

POTENTIAL MODEL: NIH PUBLIC-PRIVATE PARTNERSHIP PROGRAM

The National Institutes of Health (NIH) has long realized the importance of collaborations to tackle those issues that require outside resources, and public-private partnerships (PPPs) were set up ad hoc to fulfill such needs. One outcome of the 2002 NIH Roadmap, which recognized that partnerships between PPPs could help push scientific discoveries from the bench to the bedside, was the creation of the Program on PPPs as a central resource to provide guidance and advice to NIH and prospective partners on the development of a PPP.

Partnerships are established either directly with the NIH, often facilitated by the Program for PPP, or through the Foundation for the NIH (FNIH), an independent public foundation. The FNIH was chartered by Congress in 1990 to support the mission of the NIH. As a public foundation, FNIH can solicit donations from non-governmental sources. FNIH has many roles in PPPs beyond fund raising and donor interactions, including the management of scientific projects and the solicitation, review, award, and management of grants or contracts. In May 2005 FNIH was involved in approximately 50 PPPs totaling $280 million.

NIH PPPs are formed to address specific scientific issues, with the larger

goal of improving public health. Thus PPPs range widely in size and scope to accommodate the goals and mandates of the partners. Partners from the public/non-profit sector may include academic institutions, foundations, patient advocacy organizations, U.S. government agencies, and foreign governments. Private-sector partners may include foundations, patient advocacy groups, and members of industry (pharmaceutical, biotechnology, devices, diagnostics, informatics, and other). Governance of PPPs is individualized and includes all sectors of the partnership. All NIH PPPs comply with federal laws and NIH policies and maintain transparency, fairness, inclusiveness, and scientific rigor. Partner contributions to the scientific and fiscal aspects of the projects also vary greatly, depending on project needs.

Brief descriptions of three PPPs are provided below, with information on their governance structures, funding mechanisms, and output. The three PPPs are the Biomarkers Consortium, the Osteoarthritis Initiative, and Grand Challenges in Global Health.

The Biomarkers Consortium[9]

The Biomarkers Consortium (BC) is a public-private partnership launched in October 2006 which aims to identify and validate new biological markers, or biomarkers. The goal is to use these novel biomarkers to promote the "delivery of successful new technologies, medicines, and therapies for prevention, early detection, diagnosis, and treatment of disease." In practice, the BC will fund clinical research centers to generate data pertinent to biomarkers that will be publicly available to encourage the use of biomarkers to improve public health.

Partners BC is a partnership between public entities (NIH, FDA, and Centers for Medicare and Medicaid Services), industry (pharmaceuticals, biotechnology, diagnostics, and medical devices), non-profit organizations and associations, and advocacy groups. The partnership is managed by the FNIH.

Contributions NIH contributes to the partnership its PPP Program staff effort and Office of General Council (OGC) and Office of Technology Transfer (OTT) involvement where necessary. Individual NIH institutes and centers (ICs) provide intellectual and scientific contributions to steering committees and project teams in the design or conduct of experiments. FDA contributes intellectual and scientific input to the steering committees and project teams. Additionally, FDA's involvement helps attract partici-

[9] Information summarized from the FNIH BC Web site http://www.fnih.org/Biomarkers%20Consortium/Biomarkers_home.shtml.

pation from companies that hope to gain exposure or insight into FDA. Pharmaceutical Research and Manufactures of America (PhRMA) and Biotechnology Industry Organization (BIO) participated in establishing the BC structure. They have contributed funding for infrastructure development and support, and they are coordinating the effort and participation from all industry partners. Participating industry organizations, including biotechnology, diagnostics, and imaging industries, contribute to the steering committees and project teams. They also contribute data, assays, equipment, drugs, and other reagents. Advocacy groups representing the public are critical to fundraising, patient recruitment, public education, and public relations. FNIH provides oversight and coordination of BC activities among the partners. FNIH issues solicitations and reviews the administration and awarding of BC/FNIH grants and contracts.

Governance BC is governed by a 15-member executive committee (EC) representing founding partners and additional stakeholders: three members from NIH ICs, three from FDA, three from the pharmaceutical industry, one from CMS, one from the biotech/diagnostics/imaging industry, one from advocacy groups representing the public, and three non-voting FNIH members. The EC is involved in decision making and advising the FNIH board on financial commitments. Decision making is in the hands of the founding partners (EC), with any partner having veto ability, and all projects needing assent from all three founding partners. EC has approval checkpoints related to project concept, project plan, intellectual property, data sharing, data access, conflict of interest, and antitrust. Steering committees, which are content area focused, identify and oversee specific project activities. Project teams determine priorities and implement projects. BC policies are negotiated with principals and legal counsels concerning issues pertaining to intellectual property, antitrust issues, grantee/contractor selection, confidentiality, and conflict of interest.

Funding FNIH is responsible for fund raising and financial management. At its launching in October 2006, $1.2 million was committed by BC funding members: the Alzheimer's Association, AstraZeneca, the Biotechnology Industry Organization, Bristol-Myers Squibb, GlaxoSmithKline, the Leukemia & Lymphoma Society, Johnson & Johnson, Eli Lilly & Company, Pfizer Inc., the Pharmaceutical Research and Manufacturers of America, and F. Hoffmann-La Roche. NIH ICs also contribute funding if it is appropriate agency-wide. Specific projects may have funding from additional sources.

Output BC seeks to identify and validate biomarkers useful for disease diagnosis, including early detection, and for predicting response in order to better guide treatment decisions.

The Osteoarthritis Initiative[10]

The Osteoarthritis Initiative (OAI) is a partnership established in 2001 to bring together public and private scientific expertise and funding to generate a research resource of osteoarthritis patient data, radiologic information, and biospecimens. The OAI goal is to "create a public resource to validate prospective biomarkers, obtain early-stage input from the FDA as to acceptability of biomarkers as clinical endpoints, and ensure that validated biomarkers are as widely available as possible to further product development and the public health." OAI awarded NIH peer-reviewed contracts to four clinical centers and one coordinating center to perform a comprehensive longitudinal study of approximately 5,000 patients over 5 to 7 years.

Partners OAI is a partnership with multiple NIH ICs (National Institute of Arthritis and Musculoskeletal and Skin Diseases, National Institute on Aging, National Center for Complimentary and Alternative Medicine, National Center on Minority Health and Health Disparities, National Institute of Biomedical Imaging and Bioengineering, National Institute of Dental and Craniofacial Research, Office of Research on Women's Health, and Office of the Director). FNIH is coordinating the private sector participation of GlaxoSmithKline, Merck, Novartis, and Pfizer. The four clinical centers are at the University of Maryland/Johns Hopkins University, Memorial Hospital Rhode Island/Brown University, Ohio State University, and University of Pittsburgh. The Data Coordinating Center is University of California San Francisco.

Contributions Academic centers that receive NIH funding carry out the research and provide access to public databases and biospecimens. Private sector industry partners provide study planning, scientific input, financial support, data, and biospecimens. The public partners provide financial support, contract funding, and a biospecimen repository as well as housing the public database. FNIH is coordinating the private sector participation.

Governance An OAI Steering Committee provides scientific oversight. There is also an Observational Study Monitoring Board (OSMB).

Funding NIH and private sector participants have contributed to the funding of this $60 million PPP, with approximately $22 million support from private industry [1]. According to the FNIH, private-sector funding is $3.2

[10] Information summarized from the FNIH Web site http://www.fnih.org/programs/research_environment/osteoarthritis.shtml and the OAI Web site http://www.oai.ucsf.edu/datarelease/StudyOverview.asp.

million annually, with equal contributions from GlaxoSmithKline, Merck, Novartis, and Pfizer, and $5 million per year from the NIH.

Output OAI set out to create a public-domain resource for researchers to be used to "identify and compare the effectiveness of new treatments and to define the state of the disease." The first release of such data was in August 2006.

Grand Challenges in Global Health[11]

Launched in 2003 by the Bill & Melinda Gates Foundation, the Grand Challenges in Global Health (GCGH) initiative provides grants for a broad range of research aimed at solving the 14 grand challenges identified by scientists and health experts worldwide. As of 2005, there were 43 research projects under way in 33 countries, each working toward scientific or technical breakthroughs in one of seven project areas to improve vaccines, create new vaccines, control insect vectors, improve nutrition, limit drug resistance, cure infection, and measure health status.

Partners GCGH partners include the Bill & Melinda Gates Foundation, the Canadian Institutes of Health Research (CIHR), the Wellcome Trust, NIH, and FNIH.

Contributions The Bill & Melinda Gates Foundation, the CIHR, and the Wellcome Trust provide funding and management. FNIH also participates in managing the GCGH initiative. NIH provides scientific advice and expertise to the FNIH. Additionally, NIH may participate in identifying and funding related projects or follow-up activities resulting from GCGH studies.

Governance The GCGH scientific board includes 22 scientists and public health experts from 13 countries. Twenty of the 43 grants were managed by the FNIH as of 2005.

Funding The Bill & Melinda Gates Foundation has committed $450 million to the GCGH initiative. Additional funding was contributed from the Wellcome Trust and CIHR ($27.1 million and $4.5 million, respectively).

Output The GCGH research projects are aimed at developing and improv-

[11] Information summarized from he GCGH Web site http://www.gcgh.org/AboutUs/, the FNIH Web site http://www.fnih.org/programs/programs.shtml, the NIH PPP Web site http://ppp.od.nih.gov/pppinfo/examples.asp, and links therein.

ing methods and techniques to bring effective, affordable health care to developing countries.

REFERENCE

1. Foundation for the National Institutes of Health. 2005. *Annual Report*.

APPENDIX FIVE

POTENTIAL MODEL: NATIONAL ACADEMIES' TRANSPORTATION RESEARCH BOARD

The National Academies is a private, non-profit organization that provides independent advice to the federal government and others on scientific and technological issues. It includes the National Academy of Sciences (NAS), the National Academy of Engineering (NAE), the Institute of Medicine (IOM), and the National Research Council (NRC). NAS was congressionally chartered in 1863, and, under this charter, the NRC was created in 1916, the NAE in 1964, and the IOM in 1970. NRC is the working arm of the NAS and NAE, and it is administered jointly by NAS, NAE, and IOM. A division of the NRC, the Transportation Research Board (TRB) is considered to be the nation's primary transportation research body. Formerly named the National Advisory Board on Highway Research, TRB was established in 1920 as a source of information and research results about highway technology. Starting in the 1950s, TRB moved beyond information exchange to the management of research projects.

The stated mission of the TRB is to promote innovation and progress in transportation through research. TRB is involved in a wide range of activities to fulfill this mission. TRB activities are organized into six divisions: technical activities; studies and special programs; administration and finance; cooperative research programs; SHRPII (Strategic Highway Research Programs II); and activities of the Marine Board. The Cooperative Research Programs Division houses the transportation research activities that parallel in funding and approach several proposals for support and conduct of comparative effectiveness research under the auspices of the IOM/NRC. It includes five contract research programs that aim to identify near-term, practical solutions to problems facing transportation agencies. The five programs are the National Cooperative Highway Research Program (NCHRP), the Transit Cooperative Research Program (TCRP), the Airport Cooperative Research Program (ACRP), the National Cooperative

Freight Research Program (NCFRP), and the Hazardous Materials Cooperative Research Program (HMCRP). The programs are administered by TRB and sponsored by federal or state agencies. Of these the NCHRP is the largest and oldest program, with expenditures of $35 million in 2007. All Cooperative Research Programs expenditures totaled $57 million in fiscal year 2007 [1].

The National Cooperative Highway Research Program, started in 1962, was the first continuing research management activity of the NRC. TRB research management capacity has continued to grow significantly at the request of Congress, the U.S. Department of Transportation, and the state departments of transportations through the American Association of State Highway and Transportation Officials (AASHTO). The state departments of transportation are the sponsors of the NCHRP through voluntary set-asides—at an allocation that amounts to 5.5 percent of the apportionment from the states' Federal-Aid Highway State Planning and Research (SPR) funds.

Every year the AASHTO Standing Committee on Research solicits potential research topics from the member state transportation departments, the various AASTHO committees, and the Federal Highway Administration. The NCHRP uses evaluation panels to assess the problem statements, which are fed back for revision, and the problems are given priority rankings based on the feedback. The Standing Committee on Research then selects specific research projects based on the level of funding available. After the projects are selected, an announcement of research projects is issued, technical oversight panels are formed for each project, scopes of work are finalized, RFPs are issued, research proposals are received and evaluated, and awards are made. The other cooperative research programs noted above are run in a similar fashion, in association with their respective public agency constituencies.

A related branch, the Studies and Special Programs Division produces policy studies, syntheses of current transportation practices, information services, and the Long-Term Pavement Performance Studies. This division supports the IDEA program (Innovations Deserving Exploratory Analysis), which provides grants to investigator-initiated proposals.

SHRPII is a new program that will focus on applied research in four areas of highway transportation—safety, renewal, reliability, and capacity—with the goal of advancing highway performance and safety. The program will be funded at approximately $150 million over 7 years. SHRPII will make recommendations about procedures, practices, and applications, which may be adopted as standards or guides to recommended practices at the local, state, or federal levels.

Apart from research-related support, TRB is funded through fees provided by the U.S. Department of Transportation and other federal agen-

cies, the state transportation departments, industry associations, and other interested organizations and individuals. Sixty-seven sponsors provide the bulk of funding for TRB's core technical activities, each with a minimum annual fee of $60,000. Five sustaining affiliates provide minimum annual fees of $15,000. Additionally, 110 organizational affiliates also contribute to TRB financial support via annual fees.

REFERENCE

1. Transportation Research Board. 2006. *Annual Report*. Washington, DC: The National Academies Press.

APPENDIX SIX

POTENTIAL MODEL: THE FEDERAL RESERVE SYSTEM

The Federal Reserve System was created by Congress through the Federal Reserve Act in 1913 to serve as the central bank for the United States. Established as an entity independent from the federal government, the Federal Reserve promotes economic stability through formulating monetary policy, overseeing and regulating banks, and providing financial services to depository institutions, all with minimal political pressure. The three major components of the Federal Reserve—the board of governors, 12 regional federal reserve banks, and various member banks located throughout the country—work in concert to carry out its charge.

Consistent with its independence from the federal government, the Federal Reserve receives funding not through congressional appropriations but through interest on government securities acquired through open market operations as well as fees charged for banking services. With this income, the reserve banks pay their expenses, and any remaining earnings are turned over to the U.S. Treasury. In 2005 payments to the treasury amounted to $21.5 billion.[12]

To achieve a balance between the Federal Reserve System's independence from the federal government and the need to keep the system accountable, the Federal Banking Agency Audit Act gives the Government Accountability Office authority to audit the board of governors, the reserve banks, and the branches. Congress also requires the Federal Reserve to produce reports

[12] U.S. Federal Reserve System. 2006 (June). *2005 Annual Report of the Board of Governors of the Federal Reserve System.*

and conduct semi-annual hearings. Federal Reserve officials often testify on a wide range of issues affecting the economy and banking industry.

In addition to the reporting requirements, Congress exercises its oversight role during the appointment process for the seven members of the board of governors, who are appointed by the President with the guidance and approval of the Senate. When selecting a governor, the President carefully considers each candidate to ensure fair representation of the various economic interests and geographical divisions of the country. For instance, only one governor from any one of the Federal Reserve regions may be appointed. Furthermore, candidates cannot have held a position in a member bank for 2 years prior or at any point during their term, if appointed. The term of service for a board seat is 14 years, and the terms are staggered with one term ending every 2 years. Two governors hold the positions of chairman and vice-chairman of the board. These terms are only 4 years long, and they do not necessarily coincide with the President's term or that of each other. Each of the seven governors, including the chairman and vice-chairman, has one vote on the board.

The board of governors formulates monetary policy through open market operations carried out by the Federal Open Market Committee (FOMC). The FOMC is made up of the seven board members as well as 5 of the 12 federal reserve bank presidents. This body increases and decreases the supply of money by altering interest rates, with the long-term goal of promoting non-inflationary economic expansion. Several open market operations exist, such as the purchase and sale of government securities, the shifting of legal reserve requirements, and the resizing of the discount window. For example, when securities are purchased by the Federal Reserve, the money supply is increased, relaxing credit conditions, and when securities are sold, the money supply is decreased and credit conditions are tightened. Similarly, legal reserve requirements, the amount that depository financial institutions are required to set aside in proportion to deposits, can be increased or decreased, causing banks to have less and more money to lend, respectively. Also, rates charged by the discount window, the Federal Reserve facility for lending to eligible depository institutions, can be augmented to encourage or discourage borrowing and, consequently, lending.

In addition to formulating monetary policy, through the Truth in Lending Act, the Electronic Funds Transfer Act, and the Fair Housing Act, the board of governors oversees member banks as well as the 12 federal reserve banks.

The 12 federal reserve banks reside in Boston, New York, Philadelphia, Cleveland, Richmond, Atlanta, Chicago, St. Louis, Minneapolis, Kansas City, Dallas, and San Francisco. Each of these banks is managed by a nine-member board of directors, who serve 3-year terms. Each board of directors is divided into three classes to ensure that the interests of member banks

and the general public are represented. The federal reserve banks act as fiscal agents to the federal government, issuing, transferring, exchanging, and redeeming government securities and savings bonds. They also provide transaction accounts for the treasury and collect and disburse funds for the federal government. Furthermore, the banks carry out such routine duties as moving coin and currency out of circulation, collecting and processing checks, and tracking electronically originated credits and debits.

The stock of these 12 Federal Reserve Banks is owned entirely by the member banks of their respective districts, purchased at 6 percent of the member banks' own capital and interest, with the understanding that 3 percent must be paid and 3 percent is subject to call by the board of governors.[13] All national banks are required to be members of the Federal Reserve System. This membership is optional for state-chartered banks. As of June 30, 2005, there were 1,861 national banks and 903 state-chartered Federal Reserve member banks.[14]

APPENDIX SEVEN

THE BUSINESS CASE FOR COMPARATIVE EFFECTIVENESS RESEARCH

A Commissioned Analysis

Michael Chernew
Professor, Department of Health Care Policy
Harvard Medical School

Mark Fendrick
Professor, Department of Internal Medicine
University of Michigan Medical School

March 2007

Acknowledgment: We thank David Helms, Sean Tunis, Peter Juhn, and reviewers from the Institute of Medicine for their helpful comments.

[13] U.S. Federal Reserve System. 2006 (June). *2005 Annual Report of the Board of Governors of the Federal Reserve System.*

[14] U.S. Federal Reserve System. 2006 (June). *2005 Annual Report of the Board of Governors of the Federal Reserve System.*

Although the United States spends a far greater amount on healthcare services than other countries, it is well documented that Americans underachieve in terms of health measures such as life expectancy and infant mortality, as well as other vital health statistics. For instance, in a 2002 ranking of 23 industrialized countries, the United States tied for last with 4 other countries on healthy life expectance at age 60. Similarly, the United States' infant mortality rate ranked last, with rates more than double the average of the three leading countries (France, Japan, and Spain) [1]. While some of these discrepancies may be explained by genetic and social factors, it is well accepted that inefficiencies in clinical care—which medical services are used and to whom they are provided—contribute substantially to the discrepancy, particularly the extra spending. Systemic factors in the organization, financing, and delivery of care are part of the problem. The mere fact that Americans consistently under-use medical interventions rigorously proven to provide important health benefits and, at the same time, use substantially more services of unproven value, warrants a careful reassessment of the evidence base that informs medical decision making as well as the payment system that likely rewards inefficient behavior.

Given the complexity of clinical medicine and the financing system, the demands on the evidence base are large. When assessing the value of a specific clinical intervention, the natural tension between quality improvement and cost containment compels a transparent discussion of not only the clinical advantages a certain service may provide, but also the economic ramifications of its use. Moreover, it must be acknowledged that the value of a specific medical intervention will vary across patient groups (e.g., colonoscopy for a 55-year-old woman with a first-degree relative with colon cancer as compared to a low-risk 35-year-old woman). Thus effectiveness is not an attribute purely of the intervention, but of both the intervention and the patient. While acquiring the scientific evidence to improve medical decision making may seem a daunting task, an investment in a clinical research agenda with the specific aim of determining the right medical intervention for the right person at the right time is critical to ensure a more efficient system.

Comparative effectiveness research (CER) refers to the clinical and/or economic evaluations of specific medical interventions (including pharmaceuticals, medical devices, and medical procedures) relative to other available alternatives for a selected clinical indication. CER is very broad. In some cases, it may compare two similar approaches to a particular clinical issue (drug A to drug B, or procedure A to procedure B). In other cases, it may compare available—and dissimilar—clinical approaches (e.g., watchful waiting, radiotherapy, and surgery for prostate cancer) to determine the optimal management of a specific clinical problem. CER typically will focus on realistic decisions confronting patients and their clinicians in actual prac-

tice, and thus seldom compares interventions to placebo controls. Because of this focus on effectiveness as opposed to efficacy, these investigations will likely rely on both prospective trials and observational data to determine relative value in real-world settings, in contrast to the carefully manipulated environment of a controlled clinical trial.

Many countries other than the United States already support a centrally coordinated CER structure, typically financed with public funds. For example, the United Kingdom's National Institute for Health and Clinical Excellence (NICE) produces national guidance on the appropriate use of health technologies and the treatment and prevention of ill health [2]. In the United States, federally funded agencies such as NIH and AHRQ support CER as a part of their established practices, and several for-profit and not-for-profit entities currently engage in CER to varying degrees. However, these efforts are modest when compared to the $2 trillion annually spent on health care. Although estimates are difficult to obtain, the Coalition for Health Service Research reports that of the $31.3 billion spent by the federal government in 2005 for health research, only $1.5 billion (4.8 percent) was devoted to health services research, of which only a small fraction is likely directed to CER. AHRQ received only $15 million for CER as part of the Medicare Modernization Act. Moreover, there is little to no coordination among CER efforts.

A greater investment in CER may yield clinical and/or economic benefits. The clinical rewards would stem from CER's ability to support efforts to direct clinical practice toward interventions that yield superior outcomes and away from services that provide little or no clinical benefit. In some cases, evidence suggests that greater benefit can be achieved at an incremental cost. In these cases, policy makers and private sector managers must explicitly acknowledge the tradeoff between better health and medical spending when debating potential solutions to our system's inefficiency. Determining what level of spending should be established for services that lead to improvements in health is one of the most challenging tasks for U.S. decision makers. Several other countries, such as Australia, Great Britain, and Canada, use measures of incremental cost per health outcome gained in their decisions regarding the allocation of healthcare resources. For several political and operational reasons, such an approach is unlikely to be adopted in the United States in the near future.

Since the over-utilization of unnecessary or unproven medical services is rampant in the United States, economic benefits may accrue if CER is able to reduce utilization of these services. CER may also produce cost savings by increasing the use of those services where the clinical benefits are sufficient to reduce spending on adverse outcomes that may have otherwise occurred. There is also potential to enhance efficiency through the identification of clinical approaches that achieve the same clinical outcomes with

the utilization of less costly interventions (e.g., laparoscopic versus open cholecystectomy).

This paper examines the potential for CER to enhance the efficiency of the healthcare system. Because CER is merely a tool, the impact of CER ultimately depends on how it is used—whether it is used to inform decision makers as part of efforts to improve the way markets function or else is implemented in specific clinical areas such as guideline development or in a broader application such as health benefit design. No matter how CER is utilized, it is imperative that outputs from CER address both the clinical and financial effects, since in some instances improvements in health will require incremental expenditures.

Since there is little argument that the potential clinical and economic impact of any healthcare reform proposal will depend on the knowledge base that guides practice, such an effort is worthwhile regardless of the approach taken to reform our inefficient system. For example, consumer-driven healthcare initiatives require patients to know about the clinical and economic ramifications of their medical care choices. Centralized delivery systems require administrators to understand the effects of different clinical options to allocate resources optimally. Other initiatives, such as payment reform or pay-for-performance programs, also require a better understanding of the merits of alternative clinical strategies. Even investments in infrastructure, such as information technology, cannot achieve their full potential without the detailed clinical and economic knowledge provided by CER.

1. Inefficiency in the System

The potential for CER to enhance efficiency depends on the underlying amount of inefficiency in the current healthcare system. Inefficiency arises when clinical goals could be more efficiently achieved by use of other services or by no services at all. CER inherently addresses the inappropriate use of medical services. Three types of inappropriate use are typically discussed. First, overuse is the delivery of services that provide no (or minimal) clinical value. In certain cases of overuse (e.g., the use of antibiotic therapy for a viral infection), delivery of these services may result in adverse clinical events or a further increase in spending above the costs of the service itself (e.g., patient-level side effects due to allergic reactions as well as societal costs of future antibiotic resistance). Elimination of this type of inappropriate overuse represents the most direct way for CER to achieve cost savings.

A second form of inefficiency, misuse, is the use of certain effective services in situations where they are not clinically indicated. This pertains to the aforementioned point that the value of specific medical services differs

substantially depending to whom and when they are delivered, and we will consider this a subset of overuse.

Third, underuse represents the failure to deliver appropriate services that attain clinical benefits. Elimination of underuse may not save money in most instances, but still may be perceived as an efficient use of expenditures. However, if the provision of appropriate services can prevent clinical deterioration that would result in greater spending over an extended period (e.g., use of certain medications after a heart attack), savings may be possible.

a. Overuse Overuse in the American healthcare system takes many forms. In some cases, it is the use of an expensive treatment choice within any given treatment strategy. This may be the use of an expensive medication when a cheaper medication is available, or an expensive imaging procedure when a less expensive alternative could be used. A notable example of this behavior was the widespread use of the more expensive cyclooxygenase-2 selective non-steroidal anti-inflammatory drug (COX-2 NSAID) for pain relief in patients without risk for gastrointestinal complications. Despite several studies demonstrating no advantages in terms of pain relief when compared to older, less expensive agents, this novel class of agents became the treatment of choice. Similarly, the use of more expensive angiotensin receptor blockers (ARBs) to treat hypertension might not be indicated in patients for whom angiotensin converting enzyme (ACE) inhibitors are effective.

In other cases, overuse occurs when a more expensive treatment strategy is chosen when a less expensive strategy exists or when a procedure or test is not necessary. For example, in some cases a surgical procedure is more expensive than an alternative pharmacologic intervention (e.g., fundoplication versus proton pump inhibitor therapy for gastroesophageal reflux disease). In other cases, the reverse may be true. In some cases, the use of a diagnostic test could reduce overall costs (e.g., T-wave alternans testing prior to implantable coronary defibrillator insertion), and in other cases diagnostic tests may be unnecessary (e.g., upper gastrointestinal endoscopy in suspected ulcer disease).

In analyzing overuse, it is important to recognize that in some cases waste occurs because the same outcome can be achieved with less expense. In other cases, the more expensive approach provides marginally better outcomes at a much higher cost, in which case the more expensive choice is deemed not cost effective.

Assessing the appropriate strategy is complex. The effects of a treatment strategy may accrue over a long period of time, and proper analysis will incorporate the costs of complications at a later date. Strategies that seem less expensive at first assessment may in fact be more expensive if they entail more complications over a lifetime. Moreover, costs beyond those of

the medical care system, such as effects on worker productivity, should be included.

A second complexity arises because patients are heterogeneous. The efficient choice for one patient may not be efficient for another. For example, the use of COX-2 selective NSAIDs may be appropriate for those patients in need of pain relief who are also at high risk of gastrointestinal complications of traditional NSAIDs (despite the possibility of increased risk of coronary adverse events), whereas this same class of medication is likely to provide no better (and perhaps worse) health outcomes at a higher cost for patients at low gastrointestinal risk but high risk for cardiovascular complications. Often, procedures appropriate in some patient groups diffuse into others, generating overuse due largely to this heterogeneity of risk and benefit.

One major challenge for CER is to assess how to get the right service to the right patient at the right time. A similar critical challenge for the health services research community is to devise systems to limit use to the set of patients for whom the service provides value, thereby eliminating overuse. The fundamental tenet of clinical medicine is *primum non nocere,* "first do no harm." In today's complex environment this principle should extend beyond the clinician-patient relationship to include health systems operations.

There are many estimates of waste in the American healthcare system associated with the excessive use of medical interventions. These are documented by several broad categories of literature.

i. Geographic variation The literature documents variations in use across geographic regions, even after controlling for clinical differences. For example, small area analysis, presented in the Dartmouth Atlas of Health Care, compares the use of resources, practice patterns, and spending levels across 306 hospital referral regions to draw inferences about the quality and cost of care provided [3]. Such studies have revealed discrepancies in care for conditions such as hip fracture, colorectal cancer, and acute myocardial infarction.

Fisher et al. looked at end-of-life care spending to examine costs and outcomes. They found that residents of high-spending regions received 60 percent more care but did not have better quality of care or outcomes [4]. Each 10 percent increase in regional end-of-life spending was associated with a 0.3-0.12 percent increase in risk for death, depending on the original condition [5]. Similarly, cardiac bypass surgery rates exhibited about a fourfold range of variation. These rates were strongly correlated with the numbers of per capita cardiac catheterization labs in the regions but not with illness rates as measured by the incidence of heart attacks in the region [6].

In another study, mortality rates and quality-of-life measures were compared for patients undergoing coronary angiography in Texas, where the utilization of the procedure is high (45 percent), and for similar patients in New York, where utilization is low (30 percent). After adjusting for case mix differences, the researchers found no health advantages associated with greater utilization, suggesting that savings associated with reduced utilization of the procedure in Texas could be achieved with no deleterious clinical consequences [7].

One estimate suggests that, in aggregate, only 27 percent of the weighted variation in Medicare spending across regions can be explained by population illness levels [6]. If spending levels in all regions were made to match those in the lowest decile (age-, sex-, and race-adjusted), then Medicare could see savings of up to $40 billion in 1996 dollars [8].

ii. Inappropriate use The second body of research that addresses waste in the system attempts to directly measure how frequently certain medical services are delivered for medically inappropriate indications. Results from this literature often demonstrate high levels of inappropriate use. For example, a 1993 study of members of seven managed care organizations found that about 16 percent of hysterectomies performed were judged to have been clinically inappropriate, and 25 percent of the patients underwent hysterectomy for uncertain indications [9]. A more recent study (in 2000) on hysterectomies found more dramatic results. Among hysterectomies performed in a capitated medical group in Southern California, 70 percent of cases were judged to have been inappropriate, according to RAND appropriateness criteria. Of the 497 women studied, 71 had hysterectomies for conditions covered by three recent ACOG criteria sets. The recommendation for hysterectomy was judged inappropriate for 53 percent of that subset by the RAND criteria and for 76 percent according to the ACOG criteria [10].

In other cases, the rates of inappropriate use are relatively low, but there is a wide range of situations in which appropriateness is uncertain, which demonstrates the need for a stronger evidence base. For example, in one study, 4 percent of coronary angiographies performed at 15 hospitals in New York State were rated inappropriate; another 20 percent were rated uncertain. The rate of inappropriate use varied from 0 percent to 9 percent among hospitals, but the difference was not significant [11]. In another study, 4 percent of percutaneous transluminal coronary angioplasty (PTCA) performed at 15 hospitals in New York State were rated inappropriate; another 38 percent were rated uncertain. The inappropriate rate varied from 1 percent to 9 percent by hospital, the uncertain rate from 26 percent to 50 percent [12].

Trends toward inappropriate and uncertain use appear in other clini-

cal areas as well. Reviewing cases of new-onset chest pain not due to myocardial infarction at one of five Los Angeles-area hospital emergency departments revealed that 7 percent of those who received some form of diagnostic cardiac testing had tests that were judged to be inappropriate. A literature review on cases of metastatic renal cell cancer (MRCC) rated 46.9 percent of treatments as inappropriate and 25.8 percent as uncertain [13]. A review of Medicare patients in three geographic areas revealed that 32 percent of the sample had carotid endarterectomy for inappropriate reasons, and 32 percent for uncertain reasons [14]. Seventeen percent of diagnostic upper gastrointestinal endoscopy procedures for Medicare patients were performed for inappropriate indications, and 11 percent were performed for uncertain indications [15]. In cases of hospital use, 23 percent of admissions were judged to be inappropriate and an additional 17 percent could have been avoided by the use of ambulatory surgery [16].

These studies often examine a specific intervention (e.g., upper gastrointestinal endoscopy or percutaneous coronary angioplasty) and evaluate the usefulness in a number of clinical indications. Most of the appropriateness research focuses on high unit cost services. However, significant expenditures associated with overuse may accrue from inappropriate utilization of low unit cost services if they are used in sufficient volume (e.g., routine blood testing, imaging procedures). Moreover, many of the studies cited above are based on data from the 1980s. The more recent small area variations literature suggests that substantial inappropriateness likely still exists, but much more work is needed in the area if we are to better understand, and address, the inefficiencies in the system.

These findings of substantial variation in practice patterns and often large rates of inappropriate use highlight the fact that the merit of a specific medical intervention depends on the precise reason for use. Thus, in most situations, detailed patient-specific information is required before reporting whether the use of a drug, test, or device is worthwhile.

It is important to recognize that one cannot say that a particular medical service is always appropriate or always inappropriate. Consider an example in the area of diagnostic imaging: radionuclide cardiovascular imaging (RNI). This is but one type of diagnostic imaging, but understanding the appropriateness of imaging as a whole is crucially important. Diagnostic imaging services reimbursed under Medicare's physician fee schedule have grown more rapidly than any other type of physician service. Between 2000 and 2005, spending doubled from $6.6 billion to $13.7 billion [17]. In 2005 the American College of Cardiology Foundation convened a technical panel to assess the appropriateness of RNI for 52 indications [18]. Of the 52 indications, 13 were deemed inappropriate, 27 appropriate, and for 12 the appropriateness was uncertain. Moreover, there was not even consensus on all of the indications for which RNI was deemed appropriate.

For example, for 6 of the 27 indications deemed appropriate, there was strong disagreement among the panelists about that designation. Much more research is needed to reduce the level of clinical uncertainty and move the system toward efficient practice patterns.

However, CER will not be sufficient to eliminate overuse. Even when identified, system factors and the complexities of care limit the ability of the system to eliminate the waste. Research on these system factors, including patient- and system-oriented interventions such as benefit design and clinician/hospital reimbursement, will be needed to complement CER and to allow development of the systems needed to realize the potential offered by CER.

b. Underuse Paradoxically, while overuse in the healthcare system is common, underuse of medical services rigorously determined to provide substantial clinical benefit is also widespread. While the small area variation discussion commonly focuses on overuse, similar aggregate-level outcomes in high-expenditure areas and low-expenditure areas imply that some of the small area variation may be due to underuse. For example, among patients with heart attacks who were considered "ideal candidates" for beta-blockers, those who actually received the needed drug ranged from 5 percent to 92 percent of patients among the 306 Dartmouth Atlas Hospital Referral Regions (HRRs) [6].

A substantial portion of underuse reflects the failure of individuals or their physicians to use preventive services or to manage their chronic illnesses as the scientific evidence would recommend. CER is needed to improve our ability to identify when variation represents underuse and when it represents overuse so that the system can respond appropriately. However, as with overuse, CER will not be sufficient to eliminate underuse. While the clinician-patient relationship plays a critical role in this shortcoming, systemic effects such as access to care, benefit design, and ability to pay are also likely contributors, and more research examining these factors will be needed to improve the ability of the system to integrate CER findings into practice.

2. Effects of Inefficiency on Key Stakeholders

Inefficiency in the healthcare system, particularly that which leads to unnecessary expenditures, affects all stakeholders. Both overuse and underuse reduce the value of the resources devoted to the healthcare system. The enormous incremental costs associated with this inefficiency are borne throughout the system.

a. Individuals Individuals, whether they use the system or not, pay for these inefficiencies in several ways and are unmistakably worse off. First, in some

cases, individuals pay out of pocket for services (e.g., total body imaging scans) that provide little value in terms of clinical outcomes. Second, the financial costs associated with waste are reflected in higher healthcare premiums. These are paid by workers either directly or, because higher healthcare costs lead employers to pay lower wages, indirectly [19]. Third, higher costs for public programs are financed by taxpayers. The costs of the largest public program, Medicare, rose 8.9 percent to $342.0 billion between 2004 and 2005 [20]. Furthermore, projections suggest that Medicare will grow at an annual rate of over 9 percent between 2005 and 2015 [21]. The growth of Medicare spending will represent a serious burden for taxpayers and a significant challenge for policy makers. It is well established that the tradeoff between access to medical care and how to pay for it is a complex and extremely political issue.

Fourth, high healthcare costs are also associated with declining rates of health insurance coverage [22]. To the extent that greater waste leads to fewer covered individuals, those that are un- and under-insured must bear greater financial risk and suffer the consequences of diminished access to valuable care in the event that such care is needed.

Finally, inefficiency generates additional adverse consequences for patients already engaged in the system. Specifically, the over-consumption of care often entails clinical risk as well as financial costs. Over the past decade, the "patient safety" movement has brought to light the extent of the clinical and economic ramifications of avoidable medical errors. For example, hospital-acquired infections are estimated to be responsible for between $3.5 billion and $5.7 billion in excess healthcare costs each year [23, 24]. Under-utilization also generates suboptimal clinical outcomes as patients forego utilization of important services.

b. Employers The clinical and financial effects of inefficient care delivery on other stakeholders are more complex. To the extent that employers bear a large fraction of the costs associated with inefficiency, they are adversely affected. As mentioned above, standard economic models supported by empirical evidence suggest that, over time, employers shift the costs of higher healthcare spending to workers in the form of lower wages. However, in the short run, employers (or the shareholders) may bear some of the costs of inefficiency. Moreover, the ability to shift cost to workers is limited for retiree expenses, suggesting that shareholders will bear the costs of inefficiency for this population of workers. Employers may also bear some of the administrative costs associated with managing healthcare benefits in an environment of rising costs and considerable inefficiency.

c. Health insurers The fiscal implications of inefficiency for insurers are also complicated. To the extent that cost increases can be anticipated, they may

be included in premiums. However, as premiums escalate, the demand for coverage may be dampened, suggesting that, on balance, insurers will find it challenging to remain profitable in a rising cost environment over the long run. Yet with the challenge comes opportunity. If insurers can develop ways to address the problems of inefficiency in the healthcare system, substantial profit opportunities may arise.

d. Providers of healthcare services Providers of healthcare services—especially those whose income is related to productivity, not quality of care—may be one stakeholder group that benefits from inefficiency. Since one group's expense is another's revenue, the payments for unnecessary interventions are income for the providers of those services. Thus, while no physician or hospital may intentionally, or even knowingly, provide unnecessary services, they likely reap some financial gain from the services delivered, necessary or not. The magnitude of this effect for particular providers depends on the extent to which they deliver unnecessary care. Providers of necessary care would not be adversely affected by reductions in the use of unnecessary services.

Reductions in the use of unnecessary care may offer indirect benefit to providers in the long term. Specifically, higher costs lead to fewer people with coverage. This may place a burden on providers who are increasingly called on to provide uncompensated care. Providers may also benefit from any reductions in inefficient care because they may find this type of cost containment preferable to other approaches (such as fee reductions).

e. Manufacturers The impact of inefficiency (and efforts to reduce inefficiency) on manufacturers is much the same as on the providers of those services. Any reduction in utilization may be a reduction in revenue, but the effects will target low-value or unnecessary services. Manufacturers that have the potential to make important clinical advances can thrive in a low-waste environment. Moreover, relative to other cost containment efforts that may impact manufacturers, efforts to reduce unnecessary use of certain medical products may be preferable.

3. Uses of CER

Discussions about CER frequently focus on the use of these evaluations to assist in development of practice guidelines or in coverage and payment decisions. While CER could be used in these specific endeavors, CER is needed for more far-reaching efforts to improve the efficiency of the healthcare system. The critical nature of a comprehensive CER agenda arises because of the lack of controlled assessments of available therapeutic options and the substantial amount of patient heterogeneity that exists. Waste generally arises when services that are valuable in some clinical situ-

ations are applied to other indications. CER is an essential tool to determine which intervention should be delivered to which person and in what clinical circumstance.

By facilitating improved targeting of both the clinical intervention and the specific patient population, the information provided by CER can benefit key stakeholders, particularly patients and payers. Specifically, by reducing the uncertainty about which treatment course is most appropriate, CER can decrease the frequency that patients receive inappropriate care, reducing costs and the potential for harmful medical errors. Similarly, CER can facilitate efforts to develop coverage policy and design value-based insurance packages, which should enhance the return on healthcare expenditures made by payers—private or public [25]. Taking the perspective of the provider, the effects of CER on utilization will depend on both the nature of the product and the incentives in place to use the service. If coverage and reimbursement levels reflect the findings of CER (i.e., payment based on clinical effects, not exclusively on production costs), providers and manufactures of high-value services should find that the CER increases their market share. However, the demand for low-value services will likely (appropriately) decline. Given that the burden of proof necessary to demonstrate value in the marketplace may intensify, so might the costs to perform the requisite CER studies.

A particular concern for providers is that cost containment efforts designed to eliminate the use of unnecessary services often inadvertently lead to restrictions on the provision of needed care. In almost all of the studies that report the appropriate indications for the use of a specified intervention, the appropriateness is "uncertain" in a significant portion of situations. Recall that there are few instances where the use of a specific drug, diagnostic test, or procedure is always appropriate or inappropriate. This underscores the need for a CER agenda that is able to measure health and economic impact on a granular level that will ultimately target those specific circumstances when certain interventions should and should not be used.

While the evidence examines both under- and overuse of selected medical services, one cannot accurately predict the net effect of a more efficient system on expenditures. This is related to the tradeoffs of how a better evidence base drives the increased use of more valuable services (and likely increases expenditures) and slows the utilization of low-value interventions (and decreases spending).

Individual CER studies may not always suggest that the least expensive course of action is the appropriate course of action—recall that "the most expensive therapy is the one that doesn't work." However, medical culture tends to be driven toward the adoption of new, expensive services, and cost growth has widely been attributed to the development and diffusion of new

medical services [25-27]. Therefore, on balance, we would expect that CER would tend to dampen spending to a level below that which would otherwise occur, because the "adopt everything for everyone" mentality would be replaced with an "adopt when appropriate" paradigm. For example, a 2006 study examined whether some stable, high-risk patients with persistent total occlusion of the infarct-related coronary artery should undergo percutaneous coronary intervention (PCI) in addition to receiving optimal medical therapy [28]. Although use of this procedure in such cases was not universal, the authors reported an inclination among physicians toward its use. In this case, a randomized trial demonstrated that PCI did not improve clinical outcomes, suggesting that resources could be saved by foregoing the procedure. The trend would likely have been toward greater use, and the CER-suggested lower use was medically appropriate.

Since the literature on diffusion of medical technology clearly shows a preference among U.S. clinicians to use new interventions before definitive clinical data are available, one can safely assume that the clinical data provided by a CER agenda will improve the quality of care. However, it should not be assumed that the completion and implementation of a CER agenda will save money in the short term. The short-term financial consequences will depend on how CER is used and on whether the savings incurred to lower rates of use of low-value interventions will offset the added expenses of the increased use of higher-value services.

While enhancing the health of Americans is a noble goal, we acknowledge that cost containment is an integral and inevitable part of the future healthcare policy. Without a strong investment in CER, patients and providers are more likely to face unintended "across the board" restrictions on the provision of valuable care because of the fiscal pressures that are being imposed on public and private health care payers. Whether these are manifested by fewer insured individuals or by the underinsurance of those with some type of benefits, CER provides the knowledge base by which providers of high-value services can advocate their continued use, using accepted scientific approaches to make their case. The findings of research that directly compares the pros and cons of available treatment options from numerous perspectives will be important for clinical practice, regardless of the cost containment/benefit reform approaches being considered. Cost containment efforts that rely on an improved evidence base are likely preferable to current efforts to drive all practice toward those of the lowest cost. Findings from CER should be used to better target, not to limit, care.

The exact mechanisms by which CER will lead to enhanced efficiency will vary based on the level of detail of the data generated by the studies and the ability of the system to implement the findings in everyday practice. On the quality improvement side, similar challenges have been identified in studies examining the suboptimal uptake of evidence-based practice

guidelines. From the financial perspective, cost-sharing approaches aim to control spending by making patients pay more at the point of service. Most efforts to raise patient out-of-pocket costs have resulted in higher costs across all services (with the possible exception of some preventive health services). It has been demonstrated that financial disincentives are often placed at the patient level, making adherence with evidence-based care difficult. Yet when faced with higher costs, patients often make poor clinical decisions, which in fact could, in some cases, lead to greater overall costs. Thus, the alignment of clinical *and* financial incentives is a necessary component to ensure the attainment of an efficient delivery system. The status quo has been unable to align quality improvement and cost containment initiatives. In fact, in some instances they actually compete with one another, contributing directly to ineffiency [29].

Such an alignment of incentives is possible in the setting of improved clinical evidence—driven by CER—and health benefit reform. Value-based insurance design (VBID) represents a "clinically sensitive, fiscally responsible" approach that advocates keeping patient out-of-pocket payments low on high-value services and raising them on services of no or marginal clinical value. Similar processes can be developed for clinician payment (e.g., payment based on quality of care delivered, not productivity). Implementation of such a scheme, in any form, would require greater CER since the relative value of services would be based directly on the findings. The advantages of such an alignment of clinical and economic incentives are obvious when compared to the current approach of untargeted "across the board" cost-sharing schemes, where the rates of both non-essential and essential services are negatively affected by higher out-of-pocket rates. By using incentives to encourage the use of high-value services and discouraging low-value ones, VBID has the potential to achieve marked increases in the efficiency of the healthcare system.

Supply-side-oriented healthcare reform approaches could also benefit from added investment in and coordination of CER. Certainly, coverage policy and clinical guidelines require such knowledge. But other initiatives, such as provider education, disease management, or pay-for-performance programs, all require an understanding of which services provide value in which settings and how quality and cost metrics can be designed in a clinically meaningful way.

4. *Conclusions*

Healthcare cost growth has placed a growing strain on our healthcare financing system. Although there is no consensus about how we can address the healthcare cost issue, most stakeholders would probably agree that the resources devoted to health care must be allocated more efficiently. This will entail being able to identify situations when more resources are necessary to

overcome the problem of underuse of highly valued services that improve health, as well as when money is being wasted on interventions that do not improve health, or worse, actually produce adverse consequences.

Regardless of the reform approach considered—market-based health savings accounts or a system administered through a single source—enhanced efficiency will require more detailed knowledge about the relative effectiveness of different interventions in specific clinical indications. All vested stakeholders should encourage investment in an infrastructure that prioritizes and undertakes investigations that yield practical information on which services to provide to which patients and when. Our healthcare system is too complex and too large to be guided without an appropriate knowledge base. Moreover, because innovation in the healthcare sector is substantial, investment in an infrastructure that would allow the assessment of the clinical and economic impact of new and existing diagnostic and treatment modalities is essential.

Creating this infrastructure will require a substantial investment. For those who consider the upfront investment necessary to create such an infrastructure to be unaffordable, it is imperative to contemplate the costs of the status quo that propagate tremendous inefficiency.

REFERENCES

1. Schoen, C., et al. 2006. US health system performance: A national scorecard. *Health Affairs Web Exclusive*, W457-475.
2. Kaufman, J. S., et al. 1998, Subcutaneous compared with intravenous epoetin in patients receiving hemodialysis. Department of Veterans Affairs Cooperative Study Group on Erythropoietin in Hemodialysis Patients. *New England Journal of Medicine* 339(9):578-583.
3. Fisher, E. S., and J. E. Wennberg. 2003. Health care quality, geographic variations, and the challenge of supply-sensitive care. *Perspectives in Biology and Medicine* 46(1):69-79.
4. Fisher, E. S., et al. 2003. The implications of regional variations in Medicare spending. Part 1: The content, quality, and accessibility of care. *Annals of Internal Medicine* 138(4):273-287.
5. Fisher, E. S., et al. 2003. The implications of regional variations in Medicare spending. Part 2: health outcomes and satisfaction with care. *Annals of Internal Medicine* 138(4):288-298.
6. Wennberg, J. E., E. S. Fisher, and J. S. Skinner. 2002. Geography and the debate over Medicare reform. *Health Affairs* Suppl Web Exclusives:W96-114.
7. Guadagnoli, E., et al. 1995 Variation in the use of cardiac procedures after acute myocardial infarction. *New England Journal of Medicine* 333(9):573-578.
8. Wennberg, J. E., and M. M. Cooper, eds. 1999. *The Dartmouth Atlas of Health Care*. Chicago: American Hospital Association.
9. Bernstein, S. J., et al. 1993. The appropriateness of hysterectomy. A comparison of care in seven health plans. Health Maintenance Organization Quality of Care Consortium. *JAMA* 269(18):2398-2402.
10. Broder, M. S., et al. 2000. The appropriateness of recommendations for hysterectomy. *Obstetrics and Gynecology* 95(2):199-205.

11. Bernstein, S. J., et al. 1993. The appropriateness of use of coronary angiography in New York State. *JAMA* 269(6):766-769.
12. Hilborne, L. H., et al. 1993. The appropriateness of use of percutaneous transluminal coronary angioplasty in New York State. *JAMA* 269(6):761-765.
13. Halbert, R. J., et al. 2006. Treatment of patients with metastatic renal cell cancer: A RAND Appropriateness Panel. *Cancer* 107(10):2375-2383.
14. Winslow, C. M., et al. 1988. The appropriateness of carotid endarterectomy. *New England Journal of Medicine* 318(12):721-727.
15. Kahn, K. L., et al. 1988. The use and misuse of upper gastrointestinal endoscopy. *Annals of Internal Medicine* 109(8):664-670.
16. Siu, A. L., et al. 1986. Inappropriate use of hospitals in a randomized trial of health insurance plans. *New England Journal of Medicine* 315(20):1259-1266.
17. Kuhn, H. 2006. *Testimony on Payment for Imaging Services under the Medicare Physician Fee Schedule*. July 18. Online. Available at http://www.hhs.gov/asl/testify/t060718.html. Accessed January 22, 2007.
18. Brindis, R. G., et al. 2005. ACCF/ASNC appropriateness criteria for single-photon emission computed tomography myocardial perfusion imaging (SPECT MPI): A report of the American College of Cardiology Foundation Quality Strategic Directions Committee Appropriateness Criteria Working Group and the American Society of Nuclear Cardiology endorsed by the American Heart Association. *Journal of the American College of Cardiology* 46(8):1587-1605.
19. Gruber, J. 1994. The incidence of mandated maternity benefits. *American Economic Review* 84(3):622-641.
20. CMS Office of Public Affairs. 2007. *Press Release: CMS Releases U.S. Health Spending Estimates Through 2005*. January 9. Online. Available at http://www.cms.hhs.gov/apps/media/press/release.asp?Counter=2069&intNumPerPage=10&checkDate=&checkKey=&srchType=&numDays=3500&srchOpt=0&srchData=&keywordType=All&chkNewsType=1%2C+2%2C+3%2C+4%2C+5&intPage=&showAll=&pYear=&year=&desc=&cboOrder=date. Accessed January 22, 2007.
21. Borger, C., et al. 2006. Health spending projections through 2015: Changes on the horizon. *Health Affairs* 25(2):w61-73.
22. Chernew, M., D. M. Cutler, and P. S. Keenan. 2005. Increasing health insurance costs and the decline in insurance coverage. *Health Services Research* 40(4):1021-1039.
23. Burke, J. P. 2003. Infection control—A problem for patient safety. *New England Journal of Medicine* 348(7):651-656.
24. Hollenbeak, C. S. 006. Factors associated with risk of surgical wound infections. *American Journal of Medical Quality* 21(6):29S-34S.
25. Chernew, M. E., et al. 1998. Managed care, medical technology, and health care cost growth: A review of the evidence. *Medical Care Research and Review* 55(3):259-288; discussion 89-97.
26. Cutler, D. M. 1995. *Technology, Health Costs, and the NIH*. Cambridge, MA: Harvard University and the National Bureau of Economic Research.
27. Newhouse, J. P. 1992. Medical care costs: How much welfare loss? *Journal of Economic Perspectives* 6(3):3-21.
28. Hochman, J. S., et al. 2006. Coronary intervention for persistent occlusion after myocardial infarction. *New England Journal of Medicine* 355(23):2395-2407.
29. Chernew, M. E., A. B. Rosen, and A. M. Fendrick. 2006. Rising out-of-pocket costs in disease management programs. *American Journal of Managed Care* 12(3):150-154.

Appendix B

Comparative Effectiveness Studies Inventory Project

A commissioned activity for the IOM Roundtable on Evidence-Based Medicine

In 2008, a working group was formed at the request of the IOM Roundtable on Evidence-Based Medicine to identify and consider compelling comparisons of clinical interventions in order to inform and advance discussions of criteria important to determining priorities for assessment, to identify a list of possible interventions to be considered, and to offer observations on the kinds of studies needed. Material was presented in draft form at the July 2008 IOM workshop on infrastructure needs for comparative effectiveness research and delivered in final form in March 2009. An abridged version of the text follows, drawn from material produced by Douglas B. Kamerow, M.D., M.P.H., consultant and working group convener, and is based in part on meetings and conversations with working group members. This portion and the complete paper have not been reviewed and should not be considered a product of the Institute of Medicine, the roundtable, or the individual members of the working group.

STUDY TOPIC SUMMARIES

Diagnosis and prognosis of breast cancer using genetic tests: HER2 and others

Drug treatment of depression in primary care

Drug treatment of epilepsy in children

Gamma knife surgery for multiple types of intracranial lesions: comparison with surgery and/or whole brain radiation

Inguinal hernia repair: open vs. minimally invasive

Outcomes of percutaneous coronary interventions in hospitals with and without onsite surgical backup

Over-the-counter drug treatment of upper respiratory tract infections in children

Prevention and treatment of pressure ulcers

Screening hospital inpatients for methicillin-resistant *Staphylococcus aureus* infection

Tobacco cessation: nicotine replacement agents, oral medications, combinations

Treatment of acute thrombotic/embolic stroke: clot removal, reperfusion drugs

Treatment of ADHD in children: drugs, behavioral interventions, no Rx

Treatment of chronic atrial fibrillation: drugs, catheter ablation, surgery

Treatment of chronic low back pain

Treatment of localized prostate cancer: watchful waiting, surgery, radiation, or cryotherapy

Use of erythropoiesis-stimulating agents in the treatment of hematologic cancers

Diagnosis and Prognosis of Breast Cancer Using Genetic Tests: HER2 and Others

Brief description of the condition or problem

In 2008, there were an estimated 185,000 new cases of breast cancer in the United States, and more than 40,000 women died from breast cancer. Breast cancer has long been known to be a heterogeneous disease, varying by estrogen receptor expression, tumor grade, and patient age. These characteristics affect diagnosis, prognosis, and treatment. More recently, gene expression status has been added to the list of variations. Multiple studies have proposed many ways to organize gene expression differences in breast cancer. Most genomic-based studies have reported that breast cancer constitutes at least four different entities, with differing gene expression profiles, molecular markers, prognosis, and treatment sensitivities. What are the clinical and treatment implications of these profiles and the resulting tumor markers?

Available treatments or interventions

Many treatments for breast cancer exist, such as chemotherapeutics and surgery. Some breast cancers express tumor markers that provide specific targets for treatment. For example, human epidermal growth factor receptor 2 (HER2), a protein involved in normal cell growth, is found in high levels on some breast cancer cells. This so-called HER2 overexpression is seen in 15 to 20 percent of breast cancers and is a good predictor of benefit from treatment with trastuzumab (Herceptin), a monoclonal antibody against HER2. It may also correlate with the response to other treatments and to the overall prognosis. Another tumor marker, Oncotype DX, is a multi-gene assay that was developed as a prognostic tool for women with estrogen receptor-positive, lymph node-negative breast cancer.

Current evidence

Prospective studies have shown that women with HER2-positive breast cancer can decrease their risk of recurrence and death by between one-third and one-half by using trastuzumab. Many clinical practice guidelines and health system guidance documents now routinely approve HER2 testing and trastuzumab treatment for women with recurrent breast cancer. Prospective studies of Oncotype DX are less common. It is less widely used as a prognostic tool, although it is approved by some health systems and has been shown to affect the management of patients with early-stage breast cancer.

Issues needing research, and conclusions

Three important characteristics of all tumor markers are their utility (when and how they are used: risk determination, malignancy, prognosis, response to therapy, etc.); the proven magnitude of their effect (the differences in outcome seen in marker-positive and marker-negative patients); and their reliability (measurement standardization, accuracy, and repeatability). Generally, the strongest evidence is gathered from prospective studies in which markers are used to determine treatment or other allocation groups, although retrospective analysis can also produce valid results. *Depending on the marker being evaluated, either prospective studies or retrospective analyses will be needed to assess whether the marker should be used in clinical care*. Such comparative research is needed for all new genetic markers that are introduced.

References

Andre, F., J. Domont, and S. Delaloge. 2007. What can breast cancer molecular sub-classification add to conventional diagnostic tools? *Annals of Oncology* 18(Suppl 9):ix33-36.

Carney, W. P., R. Neumann, A. Lipton, et al. 2003. Potential clinical utility of serum HER-2/*neu* oncoprotein concentrations in patients with breast cancer. *Clinical Chemistry* 49:1579-1598.

Chiuri, V. E., G. Leo, and V. Lorusso. 2007. Clinical and therapeutic perspectives of gene expression profiling for breast cancer. *Annals of Oncology* 18(Suppl 6):vi5856-5862.

Dendukuri, N., K. Khetani, M. McIsaac M, et al. 2007. Testing for HER2-positive breast cancer: A systematic review and cost-effectiveness analysis. *Canadian Medical Association Journal* 176:1429-1434.

Dhesy-Thind, B., K. I. Pritchard, H. Messersmith, et al. 2008. HER2/*neu* in systemic therapy for women with breast cancer: A systematic review. *Breast Cancer Research and Treatment* 109:209-229.

Hayes, D. F., A. D. Thor, L. G. Dressler, et al. 2007. HER2 and response to paclitaxel in node-positive breast cancer. *New England Journal of Medicine* 357:1496-1506.

Henry, N. L., and D. F. Hayes. 2006. Uses and abuses of tumor markers in the diagnosis, monitoring, and treatment of primary and metastatic breast cancer. *The Oncologist* 11:541-552.

Ross, J. S., J. A. Fletcher, K. J. Bloom, et al. 2004. Targeted therapy in breast cancer. *Molecular and Cellular Proteomics* 3:379-398.

Seidenfeld, J., D. J. Samson, B. M. Rothenberg, et al. 2008. HER2 Testing to Manage Patients with Breast Cancer or Other Solid Tumors. Evidence Report/Technology Assessment No. 172. AHRQ Publ No. 09-E001. Rockville, MD: Agency for Healthcare Research and Quality.

Valabrega, G., F. Montemurro, and M. Aglietta. 2007. Trastuzumab: Mechanism of action, resistance and future perspectives in HER2-overexpressing breast cancer. *Annals of Oncology* 18:977-984.

Wolff, A. C., M. E. H. Hammond, J. N. Schwartz, et al. 2007. American Society of Clinical Oncology/College of American Pathologists guideline recommendations for Human Epidermal Growth Factor Receptor 2 testing in breast cancer. *Journal of Clinical Oncology* 25:118-145.

Drug Treatment of Depression in Primary Care

Brief description of the condition or problem

Lifetime prevalence of depression is up to 17 percent in community surveys; the prevalence of major depression in primary care settings is between 6 and 8 percent. Depression leads to much disability and high utilization of health care, and depression is ranked by the World Health Organization as the fourth leading cause of worldwide disease burden. Suicide occurs in 3.5 percent of depressed patients, an increase of up to 20 times the general population rates. Depression also contributes to increased morbidity and mortality from other medical disorders, such as cardiovascular disease and diabetes. Depression is a chronic and recurring disease in most patients, which contributes greatly to its heavy disease burden.

Available treatments or interventions

Depression is treated most commonly with psychotherapy and drugs. Because of the lack of availability and the cost of psychotherapy in primary care, most primary care patients with depression are treated with drugs. Two major classes of drugs are used: older (and less costly) tricyclic antidepressants (TCAs); and newer (and more costly) agents, including selective serotonin reuptake inhibitors (SSRIs) and serotonin and noradrenaline reuptake inhibitors (SNRIs).

Current evidence

While most patients with clinical depression are seen in primary care settings, most research on depression treatment has been conducted in patients referred from psychiatrists and psychiatric emergency departments—that is, specialty care settings. Research in primary care has found that both TCAs and SSRIs are more effective than placebos in treating depression. Research assessing the newer agents in primary care has found no consistent benefit of one drug over others in terms of efficacy or adverse effects. It is difficult to draw clinically meaningful guidelines for drug choice from the current studies. In addition, many studies are of demographically homogeneous populations and thus do not allow conclusions about depressed patients who are members of minority groups or in the adolescent or elderly age groups.

Issues needing research, and conclusions

Large, head-to-head randomized trials with heterogeneous primary care populations of depressed patients are needed to better understand the benefits and harms of the various antidepressant agents. Observational studies using large databases may provide some of the needed evidence, but without randomization, conclusions may be difficult to obtain.

References

Arroll, B., S. Macgillivray, S. Ogston, et al. 2005. Efficacy and tolerability of tricyclic antidepressants and SSRIs compared with placebo for treatment of depression in primary care: A meta-analysis. *Annals of Family Medicine* 3:449-456.

Cipriani, A., P. Brambilla, T. A. Furukawa, et al. 2005. Fluoxetine versus other types of pharmacotherapy for depression. *Cochrane Database of Systematic Reviews*, Issue 4. DOI: 10.1002/14651858.CD004185.pub2.

Culpepper, L. 2006. Chronic depression: Treatment in primary care. *Primary Care Companion to the Journal of Clinical Psychiatry* 8:104-105.

Culpepper, L. 2006. Managing depression in primary care. *Primary Care Companion to the Juornal of Clinical Psychiatry* 8:88-92.

Kroenke, K., S. L. West, R. Swindle, et al. 2001. Similar effects of paroxetine, fluoxetine, and sertraline in primary care: A randomized trial. *JAMA* 286:2947-2955.

Pignone, M., B. N. Gaynes, J. L. Rushton, et al. 2002. Screening for Depression. Systematic Evidence Review No. 6. AHRQ Publication No. 02-S002. Rockville, MD: Agency for Healthcare Research and Quality.

Simon, G. 2001. Choosing a first-line antidepressant. *JAMA* 286:3003-3004.

Williams, J. W., C. D. Mulrow, E. Chiquette, et al. 2000. A systematic review of newer pharmacotherapies for depression in adults: Evidence review summary. *Annals of Internal Medicine* 132:743-756.

Drug Treatment of Epilepsy in Children

Brief description of the condition or problem

Although individual seizures are relatively common in children, epilepsy as an ongoing challenge is less so. A heterogeneous group of disorders characterized by spontaneous, recurrent seizures, epilepsy affects somewhere between 4 and 8 per 1,000 children. Epileptic conditions in childhood vary in diagnostic criteria, management, and clinical outcomes, and many of them are distinct from adult epilepsies as well. Also, drug treatment in children is different from that in adults because their pharmacokinetics vary with age.

Available treatments or interventions

Until the 1990s, six major anticonvulsant drugs were available for children: phenobarbital, phenytoin, carbamazepine, valproic acid, ethosuximide, and the benzodiazepines. These traditional drugs were not optimal because many children with epilepsy did not have their seizures controlled by them or had significant side effects. In the last decade, 10 or more new anticonvulsant drugs have been approved for use in the United States, including felbamate, fosphenytoin, gabapentin, lamotrigine, levetiracetam, oxcarbazepine, pregabalin, tiagabine, topiramate, and zonisamide. Their indications vary by seizure type or epileptic syndrome and mode of delivery.

Current evidence

Considerations when selecting anticonvulsant drugs in children include seizure type, side effects, mechanism of action, potential drug interactions, pharmacokinetic profile, ease of initiation, need for laboratory monitoring, and cost. In general, there is no evidence that any of the newer anticonvulsant drugs are more efficacious than the traditional drugs in preventing seizures. They often do, however, offer improved pharmacokinetic profiles, fewer drug-drug interactions, and/or better tolerability than the older treatments, albeit at a much higher price. Guidelines from authorities have varied in their recommendations for use of the newer drugs. Some recommend consideration of some of the newer drugs for initial treatment of new onset partial epilepsy, for instance, while other guidelines state that they should be used only when older drugs are ineffective, result in intolerable adverse affects, or are contraindicated.

Issues needing research, and conclusions

Prospective head-to-head trials among the newer anticonvulsant drugs and with traditional drugs among children with similar epileptic disorders are needed, along with careful assessment of side effects and drug-drug interactions. It is unlikely these data can be obtained from case series or analysis of existing treatment data.

References

Connock, M., E. Frew, B.-W. Evans, et al. 2006. The clinical effectiveness and cost-effectiveness of newer drugs for children with epilepsy. *Health Technology Assessment* 10:1-118.

Donner, E. J., and C. Snead. 2005. New drugs in the treatment of epilepsy in children. *Current Problems in Pediatric and Adolescent Health Care* 35:398-419.

Marson, A. G., R. Appleton, G. A. Baker, et al. 2007. A randomised controlled trial examining the longer-term outcomes of standard versus new antiepileptic drugs. *Health Technology Assessment* 11:1-102.

NICE (National Institute for Health and Clinical Excellence). 2004. N*ewer Drugs for Epilepsy in Children*. Technology Appraisal 79. Online. Available at www.nice.org.uk/TA079guidance. Accessed December 2006.

Wheless, J. W., D. F. Clarke, A. Arzimanoglou, et al. 2007. Treatment of pediatric epilepsy: European expert opinion, 2007. *Epileptic Disorders* 9:353-412.

Gamma Knife Surgery for Multiple Types of Intracranial Lesions: Comparison with Surgery and/or Whole Brain Radiation

Brief description of the condition or problem

Gamma knife surgery (GKS) is a type of stereotactic radiation treatment developed in the 1960s. It delivers a high dose of radiation to a lesion in a single session (hence "surgery") by converging multiple beams of ionizing radiation on a carefully targeted area, with minimal radiation exposure

to surrounding areas. The resulting radiation injury to targeted cells' DNA results in the desired effect, e.g., cell death or decreased growth.

Available treatments or interventions

Traditionally, GKS has been used for small (<4 cm), isolated intracranial lesions or tumors. For example it is used as primary treatment for trigeminal neuralgia, arteriovenous malformations (AVMs), and acoustic neuromas. More recently it has been employed to treat brain metastases and as adjunctive treatment (to surgery and whole-brain radiation) for malignant tumors such as glioblastoma multiforme. GKS has been used to treat movement disorders such as Parkinson's disease.

Current evidence

GKS is intuitively appealing, given that it is noninvasive (does not require craniotomy) and is generally less costly than brain surgery for the equivalent procedure. That said, very few randomized trials have been performed comparing GKS to conventional surgical, microsurgical, or radiation treatments. For example, a systematic review of GKS in the treatment of trigeminal neuralgia found 23 case series but no randomized trials.

Whereas GKS has become an accepted treatment for certain types of small intracranial lesions, the need for randomized evidence is made clear by the experience with GKS as adjunctive treatment for malignant, infiltrative tumors such as glioblastoma multiforme. While a retrospective case series presented significant benefits of adding GKS to surgery and conventional radiation therapy, when a randomized trial was performed comparing the two strategies no differences in survival outcomes were found.

Issues needing research, and conclusions

The experience with GKS is an example of the general point that technologies require comparative evidence before they are adopted for new indications. A successful track record in treating one condition combined with case series of use for new indications does not add up to compelling evidence for the new use. *Comparative research is necessary to insure that simple "technology creep" is not occurring.* Thus, head-to-head trials—with whatever treatment is currently the standard—are necessary before new indications for GKS are routinely approved.

References

Borcsek, B., I. Boncz, C. Dozsa, et al. 2005. Systematic literature review of the gamma knife. *Italian Journal of Public Health* 2:178.

Chamberlain, M. C. 2006. Treatment options for glioblastoma. *Neurosurgical Focus* 20(4): E2.

Crowley, R. W., N. Poutatian, and J. P. Sheehan. 2006. Gamma knife surgery for glioblastoma multiforme. *Neurosurgical Focus* 20(4):E17.

Gilbert, M. R. 2006. New treatments for malignant gliomas: Careful evaluation and cautious optimism required. *Annals of Internal Medicine* 144:371-373.

Lim, J. N. W., and L. Ayiku. 2004 (January). The clinical efficacy and safety of stereotactic radiosurgery (gamma knife) in the treatment of trigeminal neuralgia.

Medicare Services Advisory Committee. 2007. Gamma knife radiosurgery. Medicare Services Advisory Committee:117.

NICE (National Institute for Health and Clinical Excellence). 2004. S*tereotactic Radiosurgery for Trigeminal Neuralgia Using the Gamma Knife. Interventional Procedure Guidance 85*. Online. Available at http://www.nice.org.uk/nicemedia/pdf/ip/IPG085guidance.pdf.

Nwokedi, E. C., S. J. DiBiase, S. Jabbour, et al. 2002. **Gamma knife stereotactic radiosurgery** for patients with glioblastoma multiforme. *Neurosurgery* 50:41-47.

Souhami, L., W. Seiferheld, D. Brachman, et al. 2004. Randomized comparison of stereotactic radiosurgery followed by conventional radiotherapy with carmustine to conventional radiotherapy with carmustine for patients with glioblastoma multiforme. *International Journal of Radiation Oncology Biology Physics* 60:853-860.

Yamakami, I., Y. Uchino, E. Kobayashi, et al. 2003. **Conservative management, gamma-knife** radiosurgery, and microsurgery for acoustic neurinomas: A systematic review of outcome and risk of three therapeutic options. *Neurological Research* 25:682-690.

Inguinal Hernia Repair: Open vs. Minimally Invasive

Brief description of the condition or problem

Inguinal hernias, protrusions of the peritoneum (with or without abdominal contents) through a defect or weakness in the abdominal wall, are very common in men, with about a 25 percent lifetime risk. Thus, surgical hernia repair is one of the most common operations in general surgery.

Available treatments or interventions

For more than 100 years, inguinal hernia surgery was an open procedure with general or (more recently) regional or local anesthesia, although the surgical approach has varied over time. Traditional methods of hernia repair used sutures to close the defect. Newer open techniques involve inserting synthetic mesh in an attempt to decrease recurrent hernias. More recently, minimally invasive surgery techniques, using laparoscopic instruments, have been introduced. The two major laparoscopic approaches are transabdominal pre-peritoneal (TAPP) surgery, in which mesh is inserted through the peritoneum to prevent herniation, and totally extraperitoneal repair (TEP), in which the peritoneum is not entered, and the mesh is used to seal the hernia from the outside.

Current evidence

Evidence and practice clearly favor the use of mesh for open and laparoscopic repairs, mainly for reduced recurrence rates. Randomized trials have found that laparoscopic repairs take longer and are more expensive to

perform than open repairs and require general rather than local or regional anesthesia. Recovery is quicker with laparoscopic surgery, and there is less persisting pain and numbness at the surgical site. There is a question of whether laparoscopic surgery has a higher serious complication rate than open surgery, and the recurrence rate may be higher with laparoscopic surgery. Few randomized trials have compared TAPP vs. TEP techniques, but the newer TEP is technically more difficult to perform than TAPP and takes longer. Both types of laparoscopic surgery are performed more rapidly as surgeons gain experience with the technique used. Because the peritoneum is not entered, the TEP technique may reduce the risk of damage to intra-abdominal organs.

Issues needing research, and conclusions

Because of the necessity for a learning curve for laparoscopic techniques, combining results from earlier trials may not reflect current outcomes. *Large-scale data analysis may answer the question of whether laparoscopic hernia repair is more effective and more cost-effective than open repair.* If it does not, then a large randomized trial needs to be undertaken. Also, large trials with long-term follow-up are needed comparing the two major laparoscopic approaches, TAPP and TEP.

References

Gokalp, A., M. Inal, G. Maralcan, et al. 2003. A prospective randomized study of Lichtenstein open tension-free versus laparoscopic totally extraperitoneal techniques for inguinal hernia repair. *Acta Chirurgica Belgica* 103:502-506.

Gould, J. 2008. Laporoscopic versus open inguinal hernia repair. *Surgical Clinics of North America* 88:1073-1081.

Den Hartog, D., A. H. M. Dur, W. E. Tuinebreijer, et al. 2008. Open surgical procedures for incisional hernias. *Cochrane Database of Systematic Reviews*, Issue 3. DOI: 10.1002/14651858.CD006438.pub2.

McCormack, K., B. L. Wake, C. Fraser, et al. 2005. Transabdominal pre-peritoneal (TAPP) versus totally extraperitoneal (TEP) laparoscopic techniques for inguinal hernia repair: A systematic review. *Hernia* 9:109-114.

McCormack, K, N. Scott, P. M. Go, et al. 2003. Laparoscopic techniques versus open techniques for inguinal hernia repair. *Cochrane Database of Systematic Reviews*, Issue 1. DOI: 10.1002/14651858.CD001785.

NICE (National Institute for Health and Clinical Excellence). 2004. *Laparoscopic Surgery for Inguinal Hernia Repair*. Technology Appraisal Guidance 83. Online. Available at www.nice.org.uk/TA083guidance. Accessed September 2007.

Scott, N., P. M. Go, P. Graham, et al. 2001. Open mesh versus non-mesh for groin hernia repair. *Cochrane Database of Systematic Reviews*, Issue 3. DOI: 10.1002/14651858.CD002197.

Wake, B. I., K. McCormack, and C. Fraser. 2005. Transabdominal pre-peritoneal (TAPP) vs totally extraperitoneal (TEP) laparoscopic techniques for inguinal hernia repair. *Cochrane Database of Systematic Reviews,* Issue 1. DOI: 10.1002/14651858.CD004703.

Outcomes of Percutaneous Coronary Interventions Performed in Hospitals with and without Onsite Surgical Backup

Brief description of the condition or problem

More than 1 million Americans have a myocardial infarction—heart attack—each year, with about a 40 percent mortality rate. Urgent or primary percutaneous coronary intervention (PCI) by cardiac catheterization is recommended in the setting of an ST-segment elevation myocardial infarction [STEMI]), in order to dilate occluded coronary arteries and improve blood flow. More than 1 million PCIs are performed annually in the United States, both in the treatment of STEMI and on an elective basis.

Available treatments or interventions

In the setting of an acute myocardial infarction, establishing cardiac reperfusion as soon as possible is critically important. With a STEMI, PCI has been shown to be superior to pharmacologic reperfusion (fibrinolytic therapy) if undertaken within 90 minutes of first medical contact. Over 90 percent of primary PCIs are successful. In a small number of unsuccessful cases, however, urgent cardiac surgery is required. PCI can be performed in hospitals with and without cardiac surgery capabilities, and many hospitals (especially in suburban and rural areas) have cardiac catheterization facilities but do not have cardiac surgery onsite. Patients prefer care in their local hospitals, and in many cases hospitals with cardiac surgery capabilities are too distant for immediate access. The question is thus whether it is necessary to have cardiac surgery immediately available onsite for primary PCI and/or for elective PCI.

Current evidence

Numerous case series and case-control studies have shown that PCI can be done safely in selected individual hospitals without onsite cardiac surgery. In these studies, PCI success rates and mortality rates are similar to those at hospitals with onsite surgery. An analysis of national Medicare data for patients aged 65 and older, however, found greater mortality for elective PCIs at hospitals without onsite surgery, and elective PCIs are much more common than primary PCIs. As with many procedures, outcomes are usually better in centers with higher volume, once corrected for case mix. Evidence-based guidelines currently recommend that primary PCI may be allowed in hospitals without onsite cardiac surgery if there is a proven process for emergency transport and if the operator and facility meet standards of performance and volume. These guidelines recommend against performance of elective PCI at hospitals without cardiac surgery onsite.

Issues needing research, and conclusions

Studies are needed to assess the safety and outcomes of PCI performed in settings with and without onsite cardiac surgery, both for primary and elective interventions. What facilities, training, procedure volumes, and transport capabilities are necessary for excellent outcomes? Retrospective analysis of national data, if it includes patients under 65 years of age, might provide needed answers. Given excellent results in specific centers without onsite surgery, a randomized, controlled trial would be ethical and would provide more definitive guidance.

References

Canadian Cardiovascular Society, American Academy of Family Physicians, American College of Cardiology, et al. 2008. 2007 focused update of the ACC/AHA 2004 guidelines for the management of patients with ST-elevation myocardial infarction. *Journal of the American College of Cardiology* 51:210-247.

Frutkin, A. D., S. K. Mehta, T. Patel, et al. 2008. Outcomes of 1,090 consecutive, elective, nonselected percutaneous coronary interventions at a community hospital without onsite cardiac surgery. *American Journal of Cardiology* 101:53-57.

King, S. B. 3rd, S. C. Smith Jr., J. W. Hirshfeld Jr., et al. 2008. 2007 focused update of the ACC/AHA/SCAI 2005 guideline update for percutaneous coronary intervention. *Journal of the American College of Cardiology* 51:172-209.

Roussanov, O., G. Sestacio, M. Capuno, et al. 2006. Nonemergent percutaneous coronary interventions in a Veterans Affairs medical center without onsite cardiac surgery. *American Heart Journal* 152:909-913.

Ting, H. H., G. Raveendran, R. J. Lennon, et al. 2006. A total of 1,007 percutaneous coronary interventions without onsite cardiac surgery. *Journal of the American College of Cardiology* 47:1713-1721.

Weaver, W. D. 2004. Is onsite surgery backup necessary for percutaneous coronary interventions? *JAMA* 292:2014-2016.

Wennberg, D. E., F. L. Lucas, and A. E. Siewers. 2004. Outcomes of percutaneous coronary interventions performed in a center without and with onsite coronary artery bypass graft surgery. *JAMA* 292:1961-1968.

Over-the-counter Drug Treatment of Upper Respiratory Tract Infections in Children

Brief description of the condition or problem

The average child in the United States has 6 to 8 upper respiratory infections (URIs) or colds annually, more if enrolled in day care. URIs are overwhelmingly self-limiting, viral infections, lasting up to 10 days and producing symptoms that include sneezing, cough, rhinorrhea, and fever. Nonprescription, over-the-counter (OTC) cough and cold medications are widely used for children in the United States, with about 4 million children younger than 12 years being treated each week. U.S. sales of OTC cough and cold medicines have been estimated to be at least $3.5 billion annually.

About 95 million packages of such medication are purchased each year for use by children.

Available treatments or interventions

More than 100 different OTC cough and cold preparations are available. They include individual agents and combinations of the following types of drugs: antipyretics, to reduce fever; decongestants, to relieve nasal congestion; antitussives, to reduce cough; and antihistamines and expectorants, to reduce or thin mucous production. Antihistamines are also employed in "nighttime" formulations to induce sleep.

Current evidence

Antipyretics are safe and effective in reducing fever in children when used as directed. Multiple randomized trials of variable quality have evaluated the other OTC medications for cough and cold and have not found clear evidence of efficacy. There is no good evidence for or against the effectiveness of OTC medicines for acute cough in children. Furthermore, there is clear documentation that OTC cough and cold medications can be toxic, even when used as directed. Also, older children have misused the OTC antitussive dextromethorphan in attempts to induce hallucinations.

Issues needing research, and conclusions

Large-scale, simple head-to-head trials of OTC cough and cold medications are needed to establish clearly if they have effectiveness in reducing the severity and/or duration of URI symptoms in children.

References

American Academy of Pediatrics. 1997. Use of codeine- and dextromethorphan-containing cough remedies in children. *Pediatrics* 99:918-920.

Chang, A. B., and W. B. Glomb. 2006. Guidelines for evaluating chronic cough in pediatrics. *Chest* 129:260S-283S.

Dart, R. C., I. M. Paul, G. R. Bond, et al. 2008. Pediatric fatalities associated with over the counter (nonprescription) cough and cold medications. *Annals of Emergency Medicine*. DOI: 10.1016.annemergmed.2008.09.015.

FDA (Food and Drug Administration). 2008. *FDA Statement following CHPA's announcement on nonprescription over-the-counter cough and cold medicines in children*. Online. Available at http://www.fda.gov/bbs/topics/NEWS/2008/NEW01899.html.

Ford, J. A. 2009. Misuse of over-the-counter cough or cold medications among adolescents: Prevalence and correlates in a national sample. *Journal of Adolescent Health*. DOI: 10.1016/j.jadohealth.2008.10.140.

Gunn, V. L., S. H. Taha, E. L. Liebelt, et al. 2001. Tocicity of over-the-counter cough and cold medications. *Pediatrics* 108(3):e52.

Irwin, K. A. 2007. Use of over-the-counter cough and cold medications in children younger than 2 years. *Journal of Pediatric Health Care* 21:272-275.

Merenstein, D., M. Diener-West, A. C. Halbower, et al. 2006. The trial of infant response to diphenhydramine. *Archives of Pediatric and Adolescent Medicine* 160:707-712.

Sharfstein, J. M., M. North M, and J. R. Serwint. Over the counter but no longer under the radar—Pediatric cough and cold medications. *New England Journal of Medicine* 357:2321-2324.

Smith, S. M., K. Schroeder, and T. Fahey. 2008. Over-the-counter medications for acute cough in children and adults in ambulatory settings. *Cochrane Database of Systematic Reviews,* Issue 1. DOI: 0.1002/14651858.CD001831.pub3.

Substance Abuse and Mental Health Services Administration. Office of Applied Studies. 2008. *NSDUH Report: Misuse of Over-the-Counter Cough and Cold Medications among Persons Aged 12 to 25*. January 10. Rockville, MD: Substance Abuse and Mental Health Services Administration.

Prevention and Treatment of Pressure Ulcers

Brief description of the condition or problem

A pressure ulcer is an area of localized damage to the skin and underlying tissue caused by pressure, shear, friction, and/or moisture. Pressure ulcers are usually caused by skin compression between a bony prominence and an external surface. Most pressure ulcers arise in hospital and nursing home settings, with an incidence rate of up to 5 percent in hospitals. Incidence rises with patient age, length of stay, and physical incapacity. The presence of pressure ulcers in the elderly is associated with a fivefold increase in mortality rate, with inpatient mortality rates as high as 33 percent.

Available treatments or interventions

Preventive interventions include good nutrition, frequent turning of bedbound patients, prevention of contact with excess moisture, pressure relieving beds, bed pads, and chair pads, and frequent repositioning of wheel-chair bound patients. Once pressure ulcers have occurred they can be treated with all of the above plus dressings of various types, multiple types of topical treatments, and antibiotics when appropriate.

Current evidence

There is good evidence supporting most preventive/nursing measures, such as frequent turning of bedbound patients and prevention of excess moisture. Less evidence is available comparing specific types of dressings, cleansing, and treatments for ulcers themselves. Some studies have found no evidence to support ultrasound treatment of pressure ulcers.

Issues needing research, and conclusions

Good randomized trials are needed comparing alternating pressure beds with continuous low-pressure supports. Good trials with adequate sample sizes are needed for wound dressings and cleansing substances.

Newer topical healing agents, such as nerve growth factor, should be compared head to head with other proven treatment methods.

References

American Medical Directors Association. 2008. *Pressure Ulcers in the Long-Term Care Setting*. Columbia, MD: American Medical Directors Association.

Consortium for Spinal Cord Medicine. 2000. *Pressure Ulcer Prevention and Treatment Following Spinal Cord Injury*. Washington, DC: Paralyzed Veterans of American.

Grey, J. E., S. Enoch, and K. G. Harding. 2006. Pressure ulcers. *BMJ* 332:472-475.

Hollisaz, M. T., H. Khedmat, and F. Yari. 2004. A randomized clinical trail comparing hydrocolloid, pheytoin and simple dressings for the treatment of pressure ulcers. *BMC Dermatology* 4:1-9. Online. Available at www.biomedcentral.com/1471-5945/4/18.

Landi, F., L. Aloe, A. Russo, et al. 2003. Topical treatment of pressure ulcers with nerve growth factor. *Annals of Internal Medicine* 139:635-641.

Langer, G., A. Knerr, and O. Kuss. 2003. Nutritional interventions for preventing and treating pressure ulcers. *Cochrane Database of Systematic Reviews,* Issue 4. DOI: 10.1002/14651858.CD003216.

McInnes, E., S. E. M. Bell-Syer, J. A. C. Dumville, et al. 2008. Support surfaces for pressure ulcer prevention. *Cochrane Systematic Reviews,* Issue 4. DOI:10.1002/14651858.CD001735.pub3.

Moore, Z. E. H., and S. Cowman. 2005. Wound cleansing for pressure ulcers. *Cochrane Database of Systematic Reviews,* Issue 4. DOI:10.1002/14651858.CD004983.pub2.

NICE (National Institute for Health and Clinical Excellence). 2005. *The Prevention and Treatment of Pressure Ulcers*. NICE Clinical Guideline 29. Online. Available at www.nice.org.uk/CG029.

Registered Nurses Association of Ontario. 2002. *Assessment and Management of Stage I to IV Pressure Ulcers*. Nursing Best Practice Guideline. Toronto, Ontario: Registered Nurses Association.

Royal College of Nursing. 2005. *The Management of Pressure Ulcers in Primary and Secondary Care*. London: Royal College of Nursing.

Sari, A. B.-A., K. Flemming, N. A. Cullum, et al. 2006. Therapeutic ultrasound for pressure ulcers. *Cochrane Systematic Reviews,* Issue 3. DOI:10.1002/14651858.CD001275.pub2.

Thomas, D. R. 2003. The promise of topical growth factors in healing pressure ulcers. *Annals of Internal Medicine* 239:694-695.

Screening Hospital Inpatients for Methicillin-resistant *Staphylococcus aureus* Infection

Brief description of the condition or problem

Hospital infections take an enormous toll in the United States, leading to nearly 100,000 deaths and added costs of $6.5 billion annually. Methicillin-resistant *Staphylococcus aureus* (MRSA) is a leading pathogen in hospital infections, thought to be responsible for 18,000 deaths per year in the United States. MRSA is easily transmitted in hospitals and is a frequent cause of hospital outbreaks. Up to 25 percent of hospital inpatients colonized with MRSA will develop infection. MRSA infection is particularly a problem in surgical patients, now accounting for the majority of *Staphylo-*

coccus aureus surgical site infections. MRSA infections are associated with higher morbidity and mortality rates, longer hospital stays, and greater costs than methicillin-sensitive *Staphylococcus aureus* infections.

Available treatments or interventions

Standard hospital infection control procedures (increased hand washing, room cleaning and disinfection, and isolation of MRSA-infected patients) have the goal of preventing cross-contamination.

Since MRSA-colonized patients are a risk to themselves and others, their timely identification has become a priority in hospitals around the world. One method introduced to decrease the spread of MRSA is termed "active surveillance culturing" (ASC). In ASC, patients' nares are typically cultured on or before admission to detect MRSA carriers. They can then be treated and/or isolated to prevent infection and spread of the organism.

Until recently, conventional bacteriological culture techniques were used to detect MRSA carrier status, requiring at least 24 hours and usually 2 to 3 days for results. Recently, rapid detection assays for MRSA using polymerase chain reaction (PCR) technology have become commercially available, shortening detection times from days to hours in optimal settings.

Some infection control experts have argued that ASC should become widespread or universal. Some hospital systems require MRSA screening for all hospital patients on admission; others screen only high-risk or surgical patients. State legislatures have also become vocal on this issue, introducing and occasionally passing legislation requiring hospital MRSA screening.

Current evidence

Standard hospital infection control procedures can and have led to decreased MRSA infection rates.

The use of ASC is both expensive and complex and, like all interventions, has adverse effects. While the PCR tests have good sensitivity (about 80 to 90 percent) and excellent specificity (up to 98 percent) they are expensive, costing about five times as much as conventional cultures. When employed on a large scale, the cost implications are enormous, not just for testing but also for downstream costs related to treatment, isolation, and subsequent care. With an increase in identification of MRSA carriers, many more patients will be isolated by being placed on contact precautions, which has negative implications both for the patients' care and their mental status.

MRSA screening can clearly identify MRSA carriers, and PCR screening can identify carriers more quickly than conventional culture techniques. Although it costs more, PCR screening thus can reduce the number of inap-

propriate pre-emptive isolation days while awaiting conventional culture results.

Much of the data supporting ASC comes from single hospital intervention descriptions, in which screening is part of an overall infection control program and there are no control patients for comparison. Recently published randomized trials of rapid screening tests have reported mixed results, confirming the ability of PCR to identify MRSA carriers quickly but failing to find decreases in MRSA acquisition and/or surgical infection rates resulting from the interventions.

Issues needing research, and conclusions

Despite much research, discussion, and attention, this topic could benefit from further research. *Carefully designed controlled trials of various approaches are needed to compare the effects of conventional and rapid MRSA screening with each other and with other infection control measures, such as contact precautions, environmental decontamination, and colonization eradication*. Which patient groups will benefit most from screening in MRSA-endemic settings? Will screening decrease overall surgical MRSA infection rates in addition to preventing progression from carrier to infected status in screened individuals? Is there a role for screening in low-prevalence settings? In addition, costs should be evaluated along with effectiveness, given the increased expenses associated with ASC, especially using new rapid assays.

References

Buhlmann, M., K. Bogli-Stuber, S. Droz, and K. Muhlmann. 2008. Rapid screening for carriage of methicillin-resistant *Staphylococcus aureus* by PCR and associated costs. *Journal of Clinical Microbiology* 46:2151-2154.

Dancer, S. J. 2008. Considering the introduction of universal MRSA screening. *Journal of Hospital Infections* 69:315-320.

Diekema, D. J., and M. Climo. 2008. Preventing MRSA infections; finding it is not enough. *JAMA* 299:1190-1192.

Harbath, S., C. Fankhauser, J. Schrenzel, et al. 2008. Universal screening for methicillin-resistant *Staphylococcus aureus* at hospital admission and nosocomial infection in surgical patients. *JAMA* 299:1149-1157.

Jeyaratnam, D., C. J. M. Whitty, K. Phillips, et al. 2008. Impact of rapid screening tests on acquisition of methicillin resistant *Staphylococcus aureus*: Cluster randomised crossover trial. *BMJ* 336:927-930.

Wilcox, M. H. 2008. Rapid screening for MRSA is no more effective at reducing acquisition than conventional screening. *BMJ* 336:899-900.

Tobacco Cessation: Nicotine Replacement Agents, Oral Medications, Combinations

Brief description of the condition or problem

The use of tobacco products is the leading cause of death in America, resulting in an estimated 440,000 deaths and more than 5 million years of productive life lost annually. Although smoking prevalence has decreased in men from almost 50 percent in the 1950s to 24 percent today, now 19 percent of women smoke, and smoking prevalence is higher in many minority groups and among persons with serious mental illness. Seventy percent of smokers would like to quit, but the combination of nicotine addiction and social habituation makes it difficult for many people to stop smoking.

Available treatments or interventions

In addition to behavioral counseling, which has been shown to enhance quit rates when used with medication, two major categories of pharmaceutical agents have been proven effective: nicotine replacement therapy (NRT) and oral medications. Available nicotine replacement preparations include transdermal patches, gum, inhalers, nasal spray, and lozenges. Oral medications include varenicline and buproprion SR (slow release formulation).

Current evidence

All of the above therapies have been proven superior to placebo in helping smokers quit. Some meta-analyses have shown that certain combinations of therapies (such as nicotine patch combined with gum) are more effective than the patch alone. Based on the pharmacology of the treatments, patient preferences, and clinical experience, guidelines now recommend combining various NRT agents with each other and with oral medications for greatest possible effect.

Issues needing research, and conclusions

Most research has examined single cessation medications with or without behavioral counseling in research settings. *Research is needed that compares combinations of drugs, both combinations of NRT agents and the use of NRT agents and oral drugs together.* Such head-to-head research is needed in general populations of smokers and for specific subpopulations—women, pregnant women, adolescents, older smokers, individuals with psychiatric disorders, and minority populations. Research should also examine prescription vs. over-the-counter medications and the use of medication combinations with and without counseling from clinicians and in inpatient as well as outpatient settings. Since most evidence is from research settings, effectiveness research from more "real-world" settings is needed.

References

Bohadana, A., F. Nilsson, T. Rasmussen, et al. 2000. Nicotine inhaler and nicotine patch as a combination therapy for smoking cessation. *Archives of Internal Medicine* 160:31288-31234.

Cummings, K. M. and A. Hyland. 2005. Impact of nicotine replacement therapy on smoking behavior. *Annual Reviews of Public Health* 26:583-599.

Fiore, M. C., C. R. Jaèn, T. B. Baker, et al. 2008. *Treating Tobacco Use and Dependence: 2008 Update*. Clinical Practice Guideline. Rockville, MD: U.S. Department of Health and Human Services, Public Health Service.

Hughes, J. R., S. Shiffman, P. Callas, et al. 2003. A meta-analysis of the efficacy of over-the-counter nicotine replacement. *Tobacco Control* 12:21-27.

Treatment of Acute Thrombotic/embolic Stroke: Clot Removal, Reperfusion Drugs

Brief description of the condition or problem

Almost 800,000 strokes occur each year in the United States, leading to almost $70 billion in direct and indirect costs. Cerebrovascular disease, including stroke, is the third leading cause of death in the United States, resulting in 145,000 deaths in 2005. Of all strokes, about 87 percent are ischemic in origin, meaning that blood flow is obstructed in an artery in the brain. This cuts off the blood supply to the part of the brain supplied by that artery. Most ischemic strokes are caused by a blood clot that forms at the site of obstruction (thrombus) or travels there from the heart (embolus). If the clot is not dissolved or removed, that part of the brain dies and its functions are lost.

Available treatments or interventions

Intravenous tissue plasminogen activator (tPA), which can dissolve thrombi in the setting of an acute thrombotic stroke, is the best-researched and most widely available reperfusion drug. Other, more powerful, or different thrombolytic drugs, such as tenecteplase and reteplase, are available, but not all are FDA approved for use in stroke treatment. Intra-arterial delivery of fibrinolytic agents to the site of the thrombus can potentially expand the time window for reperfusion therapy and decrease the risk of systemic bleeding from the procedure, but it requires more technical capability. Multiple types of percutaneous endovascular therapies are also available for physical clot removal, including clot retrieval devices, suction thrombectomy devices, laser or ultrasound devices, mechanical clot disruption and fragmentation devices, and intracranial angioplasty with or without stents.

Current evidence

Intravenous thrombolytic therapy has been shown to significantly improve acute ischemic stroke outcomes. As a result, all major guidelines recommend that tPA be used in patients with ischemic strokes within 3 hours of the onset of symptoms, assuming lack of contraindications and appropriate facility and clinician experience and capabilities. Despite these guidelines, however, it is estimated that less than 10 percent of stroke patients receive thrombolytic therapy. Reasons include lack of clinical appropriateness, inadequate facilities or inexperienced clinicians, clinician attitudes about the medico-legal consequences of thrombolytic therapy, and delayed presentation or evaluation.

Most of the large number of clot-removal technologies have not been tested in randomized trials. They offer the potential to increase the number of patients with acute stroke who can receive early (but not necessarily restricted to the first three hours) interventions to improve their clinical outcomes.

Issues needing research, and conclusions

Head-to-head trials are needed to compare various thrombolytic agents and percutaneous clot-removal technologies, both individually and combined. The goal is to determine the best strategies for patients, depending on type of stroke, time of presentation, medical comorbidities, and available resources.

References

Adams, H. P. Jr, G. del Zoppo, M. J. Alberts, et al. 2007. Guidelines for the early management of adults with ischemic stroke. *Stroke* 38:1655-1711.

Albers, G. W., P. Amarenco, J. D. Easton, et al. 2008. Antithrombotic and thrombolytic therapy for ischemic stroke. *Chest* 133(6 Suppl):630S-669S.

American Heart Association. 2009. *Heart Disease and Stroke Statistics—2009 update*. Dallas, TX: American Heart Association.

Gupta, R., N. A. Vora, and M. B. Horowitz. 2006. Multimodal reperfusion therapy for acute ischemic stroke. *Stroke* 37:986-990.

Liang, B. A., R. Lew, and J. A. Zivin. 2008. Review of tissue plasminogen activator, ischemic stroke, and potential legal issues. *Archives of Neurology* 65:1429-1433.

Mielke, O., J. M. Wardlaw, and M. Liu. 2004. Thrombolysis (different doses, routes of administration and agents) for acute ischaemic stroke. *Cochrane Database of Systematic Reviews*, Issue 4. DOI:10.1002/14651858.CD000514.pub2.

Molina, C. A., and J. L. Saver. 2005. Extending reperfusion therapy for acute ischemic stroke. *Stroke* 36:2311-2320.

NICE (National Institute for Health and Clinical Excellence). 2008. *Stroke: Diagnosis and Initial Management of Acute Stroke and Transient Ischaemic Attack*. NICE Clinical Guideline 68. Online. Available at www.nice.org.uk/CG068.

Stead, L. G., R. M. Gilmore, F. Bellolio, et al. 2008. Percutaneous clot removal devices in acute ischemic stroke. *Archives of Neurology* 65:1024-1030.

Wardlaw, J. M., G. J. del Zoppo, T. Yamaguchi, et al. 2003. Thrombolysis for acute ischaemic stroke. *Cochrane Database of Systematic Reviews*, Issue 2. DOI:10.1002/14651858.CD000213.

Treatment of ADHD in Children: Drugs, Behavioral Interventions, no Rx

Brief description of the condition or problem

Attention deficit/hyperactivity disorder (ADHD) is a heterogeneous behavioral disorder characterized by hyperactivity, impulsivity, and inattention. It is the most common behavioral disorder in school-age children, with a prevalence of 8 to 12 percent in the United States. The disorder varies in severity, resulting in mild to severe impairment of daily activities.

Available treatments or interventions

Non-pharmacologic treatments include parental interventions, behavioral interventions, and school interventions. Drug used to treat ADHD include stimulants (both immediate- and modified-release), atomoxetine, and other medications. No treatment is an option as well, especially with milder cases.

Current evidence

Drug therapy is widely regarded as the mainstay of treatment for moderate to severe ADHD, with clear evidence of superiority of all drug types over placebo. Behavioral therapy and other non-pharmacologic treatments have also been shown to be more effective than placebo but have generally not been found to be as good as drug therapy. The combination of drug and non-drug therapy is widely recommended but has rarely been proven to be superior to drug therapy alone. Newer, more expensive extended-release preparations of stimulants and other classes of drugs have been proven effective vs. placebo, but they have not consistently been shown to be better than less expensive immediate-release stimulant formulations. The threshold for the need for drug and/or behavioral treatment is not clearly defined.

Issues needing research, and conclusions

Further research is needed to better define the effectiveness of various behavioral and educational therapies in the overall management of ADHD, in combination with drug therapy. Is behavioral therapy necessary or sufficient treatment for children with mile ADHD? Also, head-to-head trials of immediate- and modified-release stimulants and other medications could better define the relative benefits of these medications in different populations of children and adolescents, especially with respect to costs.

References

American Academy of Child and Adolescent Psychiatry. 2007. Practice parameter for the assessment of and treatment of children and adolescents with attention-deficit/hyperactivity disorder. *Journal of the American Academy of Child and Adolescent Psychiatry* 46:894-921.

King, S., S. Griffin, Z. Hodges, et al. 2006. A systematic review and economic model of the effectiveness and cost-effectiveness of methylphenidate, dexamfetamine and atomoxetine for the treatment of attention deficit hyperactivity disorder in children and adolescents. *Health Technology Assessment* 10:1-162.

Klassen, A., A. Miller, P. Raina, et al. 1999. Attention-deficit hyperactivity disorder in children and youth. *Canadian Journal of Psychiatry* 44:1007-1016.

McLennan, J. D. 2006. Deciding on stimulant use for childhood ADHD. *Canadian Family Physician* 52:940-942.

NICE (National Institute for Health and Clinical Excellence). 2008. *Attention Deficit Hyperactivity Disorder,* NICE Clinical Guideline 72. Online. Available at www.nice.org.uk/CG072.

Smoot, L. C., L. A. Boothby, and R. C. Gillett. 2007. Clinical assessment and treatment of ADHD in children. *International Journal of Clinical Practice* 61:1730-1738.

Treatment of Chronic Atrial Fibrillation: Drugs, Catheter Ablation, Surgery

Brief description of the condition or problem

Atrial fibrillation (AF), in which the upper chambers of the heart beat in a chaotic unorganized fashion, is the most common heart rhythm abnormality seen in clinical practice, affecting about 2 million Americans. It can lead to blood clots, stroke, heart failure, and increased mortality. Many patients feel fluttering in their chest from AF and are symptomatic from decreased pumping of blood to the body, leading to weakness, dizziness, and/or shortness of breath. Some patients have asymptomatic AF, but they still may be at increased risk of stroke. AF may be paroxysmal, which comes and goes, or chronic, which is persistent.

Available treatments or interventions

Drug treatment includes three types of agents. Anticoagulants such as warfarin and antiplatelet drugs such as aspirin are used to prevent blood clots and strokes. Arrhythmia drugs, including propafenone, flecainide, sotalol, and amiodarone, are intended to prevent the fibrillation rhythm. Drugs in a third group are used to control the fast heart rate that often comes with AF. They include digoxin, various beta-blocker drugs, and some calcium channel blockers. Electrical cardioversion, in which an electric shock is administered to an anesthetized patient, can also convert AF to a regular heart rhythm. Catheter ablation techniques can be used to treat AF without surgery by passing a catheter into the left atrium of the heart and using radiofrequency energy to disrupt the electrical impulses that

lead to AF. Finally, open heart surgery can be used to disrupt the electrical impulses, but this is rarely used unless there is another reason (e.g., heart valve repair) for open heart surgery.

Current evidence

Many patients with AF are elderly and have other heart disease problems that complicate their treatment. Others are younger with so-called "lone" AF. Traditionally, the treatment goal was to keep all patients out of AF by using arrhythmia drugs with or without cardioversion. Arrhythmia drugs do not always successfully control heart rhythms, however, and they have major side effects, up to and including mortality. Recent studies have found that many patients, especially elderly, relatively sedentary patients, have similar overall clinical and quality of life-related outcomes with drugs that control heart rate rather than rhythm. Many of these drugs also have fewer side effects than arrhythmia drugs. Some experts recommend rhythm control as an initial strategy for younger, symptomatic patients presenting with AF for the first time and rate control for patients over age 65 with heart disease and who may be unsuitable for electric cardioversion.

Traditionally, catheter ablation was considered only when drug and cardioversion therapy failed. Some recent research, however, mainly in younger patients, has reported high rates of conversion to regular rhythms with catheter ablation, obviating the necessity of years of medications. Long-term follow-up is not yet available for these patients and the studies were done in a small number of highly selected medical centers.

Issues needing research, and conclusions

The costs and effectiveness of catheter ablation in large groups of different populations of patients with AF needs to be determined. As more centers become skilled in catheter ablation techniques, large effectiveness rather than efficacy studies will be useful to determine the best clinical guidance for patients with diverse risk factors and comorbidities. Most likely new, randomized trials will have to be undertaken to answer these questions, as current data are probably unsuitable for systematic review.

References

Andrikopoulos, G., S. Tzeis, N. Maniadakis, et al. 2009. Cost-effectiveness of atrial fibrillation catheter ablation. *Europace* 11:147-151.

Cox, J. L. 2005. The central controversy surrounding the interventional-surgical treatment of atrial fibrillation. *Journal of Thoracic and Cardiovascular Surgery* 129:1-4.

Khargi, K., B. A. Hutten, B. Lemke, et al. 2005. Surgical treatment of atrial fibrillation: A systematic review. *European Journal of Cardiothoracic Surgery* 27:258-265.

Mead, G. E., A. Elder, A. D. Flapan, et al. 2005. Electrical cardioversion for atrial fibrillation and flutter. *Cochrane Database of Systematic Reviews*, Issue 3. DOI:10.1002/14651858.CD002903.pub2.

Morady, F., H. Oral, F. Pelosi, et al. 2007. *The Treatment of Atrial Fibrillation. University of Michigan Electrophysiology Service*. Online. Available at http://www.med.umich.edu/1libr/aha/Pt_ed_A1204.pdf.

NICE (National Institute for Health and Clinical Excellence). 2006. *Atrial Fibrillation*. NICE Clinical Guideline 36. Online. Available at www.nice.org.uk/CG036.

Oral, H., C. Pappone, A. Chugh, et al. 2006. Circumferential pulmonary-vein ablation for chronic atrial fibrillation. *New England Journal of Medicine* 354:934-941.

Wanzi, O. M., N. F. Marrouche, D. O. Martin, et al. 2005. Radiofrequency ablation vs. antiarrhythmic drugs as first-line treatment of symptomatic atrial fibrillation. *JAMA* 293(21):2634-2640.

Wood, M. A., and K. A. Ellenbogen. 2006. Catheter ablation of chronic atrial fibrillation—The gap between promise and performance. *New England Journal of Medicine* 354:967-969.

Treatment of Chronic Low Back Pain

Brief description of the condition or problem

Acute low back pain serious enough to disrupt daily routines affects about 70 percent of adults sometime during their lives, and it is the second most frequent symptomatic cause for visiting a doctor in the United States. In the vast majority of patients, low back pain resolves within a few weeks with conservative medical management. But about 5 percent of patients go on to have nonspecific chronic low back pain, defined as pain lasting longer than 12 weeks without a specific treatable cause (such as tumor or fracture). These patients experience significant pain and functional impairment and consume 90 percent of all the healthcare costs associated with low back pain.

Available treatments or interventions

The main choice in managing chronic low back pain is whether to continue medical management or undergo spinal surgery.

Medical management can include the following interventions individually or in combination: pain, anti-inflammatory, and other medications; physical and rehabilitation therapy and other types of exercise programs: acupuncture and acupressure: cognitive-behavioral therapy: spinal manipulation: massage: and other treatments.

The mainstay of surgical treatment for nonspecific low back pain is spinal fusion surgery, which can be performed several ways. The intent of the surgery is to stabilize the spine and remove the cause of the chronic low back pain. Between 150,000 and 300,000 spinal fusion surgeries are attempted each year, costing over $16 billion in annual hospital charges.

Current evidence

Most evidence on pharmaceutical treatment of chronic low back pain comes from placebo-controlled trials. Short-term trials have found small to moderate levels of effectiveness for decreasing pain and increasing function for acetaminophen, antidepressants, antiepileptic drugs, benzodiazepines, non-steroidal anti-inflammatory drugs, and opioids. Few head-to-head trials have been conducted. Few trials have included multiple medications, which are commonly prescribed by doctors.

Several non-pharmaceutical therapies show similar effectiveness versus placebo in the treatment of chronic low back pain. These treatments include acupuncture, acupressure, psychological counseling, interdisciplinary rehabilitation, exercise, massage, spinal manipulation, and yoga. Again, few head-to-head trials have been conducted, and trials including multiple interventions are also uncommon.

At least six randomized trials have compared lumbar fusion surgery with some type of intensive nonsurgical management for chronic low back pain, with conflicting results. Questions have been raised about the varying exclusion criteria and generalizability of these studies. Although some patients were clearly helped by the surgery, it is difficult to define the characteristics of patients most likely to benefit.

Issues needing research, and conclusions

For the treatment of chronic low back pain, most published research of drug and non-invasive nondrug therapy is composed of placebo-controlled trials of single treatments. There is a paucity of research of dual-agent treatment, although most patients are treated this way. Also, head-to-head trials are needed to evaluate the relative effectiveness of drug and non-drug treatments for this condition.

But the key need in this area is research on the comparative effectiveness of spinal fusion surgery and nonsurgical approaches. Which types of patients are most likely to benefit from surgery, and when should it be undertaken in the course of the illness? Because of the significant disability associated with chronic low back pain, cost-effectiveness studies are also crucial in helping to decide which approaches return patient function to normal most quickly. Comparative studies are also needed of different surgical approaches and techniques, as well as the appropriate use of procedures such as intradiscal electrothermal therapy and the appropriate role for newer artificial spinal disks.

References

Andersson, G. B. J., N. A. Mekhail, and J. E. Block. 2006. Treatment of intractable discogenic low back pain: A systematic review of spinal fusion and intradiscal electrothermal therapy. *Pain Physician* 9:237-248.

Chou, R. and L. H. Huffman. 2007. Medications for acute and chronic low back pain: A review of the evidence for an American Pain Society/American College of Physicians clinical practice guideline. *Annals of Internal Medicine* 147:505-514.

Chou, R. and L. H. Huffman. 2007. Nonpharmacologic therapies for acute and chronic low back pain: A review of the evidence for an American Pain Society/American College of Physicians clinical practice guideline. *Annals of Internal Medicine* 147:492-504.

Deyo, R. A. 2007. Back surgery: Who needs it? *New England Journal of Medicine* 356:2239-2243.

Fairbank, J., H. Frost, J. Wilson-MacDonald, et al. 2005. Randomised controlled trial to compare surgical stabilization of the lumbar spine with an intensive rehabilitation programme for patients with chronic low back pain: The MRC spine stabilisation trial. *BMJ* 330:1233-1238.

Henschke, N., C. G. Maher, K. M. Refshauge, et al. 2007. Low back pain research priorities: A survey of primary care practitioners. *BMC Family Practice* 8:40.

Hsieh, L. L., C. Kuo, L. H. Lee, et al. 2006. Treatment of low back pain by acupressure and physical therapy: Randomised controlled trial. *BMJ* 332:696-698.

Ibrahim, T., I. M. Tleyjeh, and O. Gabbar. 2008. Surgical versus non-surgical treatment of chronic low back pain: A meta-analysis of randomised trials. *International Orthopaedics* 32:107-113.

Koes, B. 2005. Surgery versus intensive rehabilitation programmes for chronic low back pain. *BMJ* 330:1220-1221.

Melloh, M., C. Roder, A. Elfering, et al. 2008. Differences across health care systems in outcome and cost-utility of surgical and conservative treatment of chronic low back pain: A study protocol. *BMJ Musculoskeletal Disorders* 9:81.

Sherman, K. J., D. C. Cherkin, J. Erro, et al. 2005. Comparing yoga, exercise, and a self-care book for chronic low back pain. *Annals of Internal Medicine* 143:849-856.

Soegaard, R., F. B. Christensen, T. Christiansen, et al. 2007. Costs and effects in lumbar spinal fusion: A follow-up study in 136 consecutive patients with chronic low back pain. *European Spine Journal* 16:657-668.

Zeller, J. I. 2006. Artificial spinal disk superior to fusion for treating degenerative disk disease. *JAMA* 296:2665-2667.

Treatment of Localized Prostate Cancer: Watchful Waiting, Surgery, Radiation, Hormone Therapy, or Cryotherapy

Brief description of the condition or problem

In the United States in 2008, there were an estimated 186,000 new cases of prostate cancer and more than 28,000 deaths from this disease. Prostate cancer is, aside from skin cancers, the leading cause of new cancers and is the second leading cause of cancer deaths in men. However, up to 70 percent of men are found to have cancer in their prostate gland at the time of death, most often in low-grade microscopic amounts that posed no threat to the man during life. Thus, prostate cancer is common, diagnosed in one of six men, but it is fatal in only 3 percent of men. This leads to the perplexing problem of trying to decide which prostate cancers to treat and what treatment to use. Most prostate cancers are now diagnosed by elevated prostate-specific antigen (PSA) levels in routine testing, although

some are still detected on clinical examination of the prostate or from patient symptoms.

Available treatments or interventions

Major options for localized prostate cancer include watchful waiting, in which the PSA level is followed and no treatment is given initially; radical prostatectomy, which is usually an option only if the cancer is confined within the capsule of the prostate; radiation therapy, with external-beam radiation or implanted radioisotopes (brachytherapy); hormonal therapy with various anti-androgen hormones; and cryotherapy, in which probes are introduced into the tumor to freeze and kill malignant cells.

Current evidence

No randomized trial data favor one treatment or combination of treatments for prostate cancer over another. Patients and their doctors need to balance limited evidence of treatment efficacy with important known side effects as well as patient preferences and the availability and quality of local services.

If the cancer is confined to the prostate, then surgery offers the chance for a cure and is often offered to younger men, especially if the biopsies show an aggressive tumor. Prostatectomy has significant side effects, however, such as impotence and incontinence. It is not usually done in patients who have cancer that has spread beyond the prostate capsule.

External radiation can be effective in extending survival in men with cancer that has spread locally and is often preceded by hormonal therapy to shrink the tumor mass. It does not require a hospital stay or recuperation from surgery, although it has significant side effects as well, including impotence and rectal irritation and injury. Brachytherapy has fewer side effects because the radiation dose is lower than external beam radiation, but the clinical effectiveness is not as well demonstrated as other types of radiation.

Cryotherapy has not been tested in many randomized trials comparing it to more conventional treatments, and its effectiveness has mainly been documented in case series. Watchful waiting is often employed in older men and those with concomitant serious diseases, or in younger men with relatively low PSA levels and lower-grade tumors. The literature describing its long-term effectiveness is mainly from the pre-PSA era, when cancers were detected by symptoms at a more advanced stage. It, of course, has no side effects and active monitoring allows later treatment when and if it is deemed appropriate.

Issues needing research, and conclusions

Better understanding is needed of the choice, sequencing, and combination of therapies that are most effective for localized and early prostate cancer, including trials comparing radical prostatectomy and radiation in its various forms, including drug therapy. The role of watchful waiting both in the extremely elderly and in young men with low PSA levels and low-grade tumors needs to be established with clinical trials.

References

Kumar, S., M. Shelley, C. Harrison, et al. 2006. Neo-adjuvant and adjuvant hormone therapy for localized and locally advanced prostate cancer. *Cochrane Database of Systematic Reviews*, Issue 4. DOI:10.1002/14651858.CD006019.pub2.

NCI (National Cancer Institute). 2008. *National Cancer Institute Prostate Cancer Treatment (PDQ), Health Professional Version*. Online. Available at http://www.cancer.gov/cancertopics/pdq/treatment/prostate/healthprofessional.

Postma, R.. 2006. Treatment of prostate cancer. *Annals of Oncology* 17(suppl 10), x207-210. DOI: 10.1093/annonc/mdl261.

Shelley, M., T. Wild, B. Coles, et al. 2007. Cryotherapy for localised prostate cancer. *Cochrane Database of Systematic Reviews*, Issue 3. DOI:10.1002/14651858.CD005010.pub2.

Taichman, R. S., R. D. Loberg, R. Mehra, et al. 2007. The evolving biology and treatment of prostate cancer. *Journal of Clinical Investigations* 117:2351-2361.

Zeliadt, S. B., S. D. Ramsey, D. F. Penson, et al. 2006. Why do men choose one treatment over another? *Cancer* 106:1865-1874.

Use of Erythropoiesis-stimulating Agents in the Treatment of Hematologic Cancers

Brief description of the condition or problem

Anemia, a decreased red blood cell count, is common in cancer, occurring in up to two-thirds of patients. It can be caused by the disease itself (due to bone marrow infiltration or hemolysis), by nutritional deficiencies, by the myelosuppressive effects of chemotherapy, or by a combination of these factors. Anemia causes well-documented adverse effects in cancer patients, including weakness, impaired concentration, and, most commonly, fatigue. All of these lead to a decreased quality of life, which has been amply demonstrated in research studies. Anemia may also contribute to decreased responsiveness to radiotherapy or chemotherapy.

Available treatments or interventions

Before the development of erythropoiesis-stimulating agents, transfusions were the treatment of choice for cancer-associated anemia. Although transfusions immediately reverse anemia-related symptoms, the effects are short-lived and do not affect the cause of the anemia. Further, frequent transfusions increase the likelihood of adverse effects such as alloimmu-

nization, allergic reactions, iron overload, and transmission of infectious diseases.

Recombinant human erythropoietin was developed to stimulate red cell formation and has been used in a wide range of disorders to treat chronic anemia. It does not increase hemoglobin levels immediately but its effects last longer than transfusions. It is available in two forms, alpha and beta, which are very similar in molecular characteristics and pharmacokinetics. A third erythropoiesis-stimulating agent, darbepoetin alfa, is also available. It is longer acting than erythropoietin and only needs to be given every three weeks.

Current evidence

Numerous randomized and non-randomized trials have established that erythropoiesis-stimulating agents increase hemoglobin levels and reduce transfusion requirements in cancer patients with anemia. They also decrease fatigue and increase quality of life. U.S. clinical practice guidelines recommend that an erythropoiesis-stimulating agent be started as hemoglobin levels reach or fall below 10 g/dL.

Many doctors start erythropoiesis-stimulating agents at higher levels, believing the quality of life is dramatically improved if the hemoglobin levels are kept from going lower than 11 or 12 g/dL. Although there is some research showing that the quality of life is improved when these agents are started earlier, it is controversial.

Recent studies have raised the possibility that erythropoiesis-stimulating agent use may increase the risk for thromboembolism. Thus, patients with standard risk factors for thromboembolic events—history of previous thrombosis, surgery, or prolonged immobilization—may be at increased risk of having a throbolembolic event while taking epoetin or darbepoetin. Questions have also been raised about the effects of erythropoiesis-stimulating agents on mortality rates, with some studies showing positive effects and others (particularly in the treatment of solid tumors, not hematologic malignancies) finding increased mortality.

Issues needing research, and conclusions

Comparative effectiveness research would be helpful in defining the exact benefits, costs, and harms of starting erythropoiesis-stimulating agents at different hemoglobin levels in patients with hematologic malignancies. Head-to-head trials could address the advantages and disadvantages of erythropoietin alpha and beta versus the newer preparation, darbepoetin alfa. Finally, large-scale data collection/registers might be helpful to better understand the risks as well as benefits to cancer outcomes associated with these agents.

References

Aapro, M., A. Scherhag, and H. U. Burger. 2008. Effect of treatment with epoetin-β on survival, tumor progression, and thromboembolic events in patients with cancer: An updated meta-analysis of 12 randomised controlled studies including 2301 patients. *British Journal of Cancer* 99:14-22.

Bohlius, J., S. Langensiepen, G. Schwarzer, et al. 2005. Recombinant human erythropoietin and overall survival in cancer patients: Results of a comprehensive meta-analysis. *Journal of the National Cancer Institute* 97:489-498.

Hardee, M. E., M. O. Arcasoy, K. L. Blackwell, et al. 2006. Erythropoietin biology in cancer. *Clinical Cancer Research* 12:332-339.

Jones, M., B. Schenkel, J. Just, et al. 2004. Epoetin Alfa improves quality of life in patients with cancer: Results of a meta-analysis. *Cancer* 101:1720-1732.

Oster, H. S., M. Hoffman, S. Pruchti-Sagiv, et al. 2006. Erythropoietin in clinical practice: Current use, effect on survival, and future directions. *Israel Medical Association Journal* 8:703-706.

Pronzato, P. 2006. Cancer-related anaemia management in the 21st century. *Cancer Treatments Reviews* 32:S1-S3.

Rizzo, J. D., M. R. Somerfield, K. L. Hagerty, et al. 2008. Use of epoetin and darbepoetin in patients with cancer: 2007 American Society of Clinical Oncology/American Society of Hematology clinical practice guideline update. *Journal of Clinical Oncology* 26:132-149.

Straus, D. J. 2005. Treatment of anemia with erythropoietic agents in patients with hematologic malignancies. *Supportive Cancer Therapies* 2:215-224.

Straus, D. J., M. A. Testa, B. J. Sarokhan, et al. 2006. Quality-of-life and health benefits of early treatment of mild anemia: A randomized trial of epoetin alfa in patients receiving chemotherapy for hematologic malignancies. *Cancer* 107:1909-1917.

Appendix C

Comparative Effectiveness Research Priorities *IOM Recommendations (2009)*[1]

[1] Passage from the Institute of Medicine Report, *Initial National Priorities for Comparative Effectiveness Research June 30, 2009*. National Academies Press.

5

Priorities for Study

Abstract: *The Institute of Medicine Committee on Comparative Effectiveness Research Prioritization was charged with developing a portfolio of priority topics that reflected balance across research areas, populations, type of interventions, and methodologies. The final list of 100 priority CER topics includes a large number addressing health care delivery systems, and a large number that consider racial and ethnic disparities. All but 3 of the 32 originally delineated research areas are represented. Similarly, the priority research topics include studies examining various special population categories, including individuals with rare diseases. This chapter presents the full list of priority CER topics.*

As explained in detail in Chapter 1, the Institute of Medicine (IOM) committee's statement of task charged the committee with developing a list of priority comparative effectiveness research (CER) topics and presenting those recommendations for the Secretary to consider. To develop the list, the committee obtained substantial public input (described in Chapter 3) and followed a multistage process of individual and collective deliberation (described in Chapter 4). The final portfolio, described in this chapter, contains 100 priority topics. The first half of the chapter is a "portfolio analysis," which shows the representation of research areas, study populations, comparators, and study methodologies within the final 100 topics. The second half of the chapter presents the specific CER topics prioritized by the committee, together with a description of their relevance.

ASSEMBLING A DIVERSE PORTFOLIO

As described in Chapter 4, the committee utilized the concept of a diverse research portfolio, meaning that the committee's priority topics reflect a balance of CER questions across research area (i.e., disorders by organ systems, specific populations, systems of care), study populations (i.e., men, women, children, minority groups), types of interventions (i.e., comparators, such as surgical or pharmaceutical treatments), and study methodologies (i.e., randomized controlled trials, registry studies, systematic reviews). The committee wanted to ensure that the final list of topics represents not only those diseases and conditions with the greatest effects on the health of the U.S. population, but also that it includes other diseases and conditions that disproportionately and seriously affect subgroups of the population (such as women, minorities, and children and adolescents). In addition, the committee wanted to ensure its priority topics examine a variety of interventions, including studies examining prevention, systems of care, pharmacological treatments, devices, surgery, and monitoring of disease. The committee also sought to achieve balance in the distribution of proposed methodologies so that some answers could be obtained within the 2-year framework specified by the American Recovery and Reinvestment Act (ARRA) of 2009, while other research questions would require a longer timeframe. For example, CER conducted from established databases and from systematic reviews of the current literature holds the potential to provide information relatively rapidly, whereas performance of randomized controlled clinical trials or prospective observational trials would extend well beyond the 2-year focus of the ARRA.

The committee strongly believes that CER should be conducted using "real-world" patients, so that results are readily generalizable across populations. Therefore, it is important that sponsors design CER studies to ensure adequate numbers of all relevant population and patient subgroups, including all genders and patients representing a wide range of races, ethnicities, levels of health literacy, and ages, as well as those with multiple chronic conditions.

The following sections conduct a "portfolio analysis"—an analysis of the distribution of the committee's final 100 priority topics across the portfolio variables, including (1) research areas, (2) study populations, (3) interventions, and (4) study methodologies. A successful portfolio is one that is widely distributed across these dimensions. It is important to recognize that the precision of the information in this section was limited by the procedures that were required to meet the committee's deadline. In the future, thorough topic nomination development requires interaction with the nominators and other stakeholders to sufficiently develop the nomina-

tion and to ensure that the supporting evidence accurately conveys the context and the main points of the nomination (Whitlock et al., 2009).

The following sections display the distribution of the committee's priority list by the portfolio criteria: research area, population, intervention, and methodology. In addition, an interactive electronic file providing search capabilities for priority topics by portfolio criteria is available at www.iom.edu/cerpriorities. This spreadsheet will allow the reader to search, for example, all cardiovascular disease topics affecting women and children, or to study the effectiveness of procedures for their treatment. The search will also indicate which quartile the committee assigned each topic.

DIVERSITY OF RESEARCH AREAS

As described earlier, one of the committee's main methods of categorizing the proposed priority topics was by research area. The committee identified 32 categories of research areas based on disease classification, other patient conditions, and systems of care.[1] However, because many of the conditions co-occur frequently (e.g., obesity and osteoarthritis), and many of the nominated priorities mentioned both a disease and a system of care (e.g., Alzheimer's disease and nursing home care), most of the priority topics could be classified according to two or more research areas.[2] For example, a topic to study alternative strategies for treating heart disease in African American patients with diabetes could have been classified as cardiovascular disease, endocrinology (which includes diabetes care), and racial and ethnic disparities. In addition, if that research question involved comparing alternative organizational approaches to care, such as coordinated disease management programs or remote monitoring of patients' symptoms, the topic could also be classified under the health care delivery system area. In fact, among the final 100 priority topics, the average number of assignable research categories was three.

To determine whether the committee's priority list was balanced across research areas, each priority was categorized by all of the possible research areas that reasonably described it. For the purposes of this exercise, one area was designated as the primary topic. Table 5-1 and Figure 5-1 show the breakdown of the 100 final priority topics categorized by research area. In Table 5-1, the topic's primary research area is shown with assigned secondary research areas, if reported. Several areas are prominently represented.

[1] Refer to Chapter 3 to see how the committee developed the list of 32 research area categories.

[2] In the classification exercise that took place at each stage of the IOM committee's deliberations, however, each nominated recommendation was placed into only one area, which was considered its primary research area.

TABLE 5-1 Recommended Research Priorities by Research Area

Category	Primary Research Area	Secondary Research Area	Total
Health Care Delivery Systems*	23	27	50
Racial and Ethnic Disparities	3	26	29
Cardiovascular and Peripheral Vascular Disease	8	13	21
Geriatrics	2	19	21
Functional Limitations and Disabilities	2	20	22
Neurologic Disorders	6	11	17
Psychiatric Disorders	7	10	17
Pediatrics	1	15	16
Endocrinology and Metabolism Disorders	2	12	14
Musculoskeletal Disorders	5	7	12
Oncology and Hematology	6	5	11
Women's Health	5	2	7
Alcoholism, Drug Dependency, and Overdose	2	4	6
Infectious Diseases	3	2	5
Skin Disorders	3	1	4
Birth and Developmental Disorders	3	1	4
Nutrition (including obesity)	3	1	4
Immune System, Connective Tissue, and Joint Disorders	1	3	4
Eyes, Ears, Nose, and Throat Disorders	2	1	3
Trauma, Emergency Medicine, and Critical Care Medicine	1	2	3
Complementary and Alternative Medicine	3	0	3
Kidney and Urinary Tract Disorders	2	1	3
Oral Health	2	1	3
Respiratory Disease	1	2	3
Genetics and Disease	0	3	3
Gastrointestinal System Disorders	1	1	2
Palliative and End-of-Life Care	2	0	2
Sexual Function and Reproductive Disorders	0	2	2
Liver and Biliary Tract Disorders	1	1	2
Total	100	193	293

*Although this category was described as "Safety and Quality of Health Care" in the web-based questionnaire, the category was relabeled by the committee as "Health Care Delivery Systems" to be more accurate.

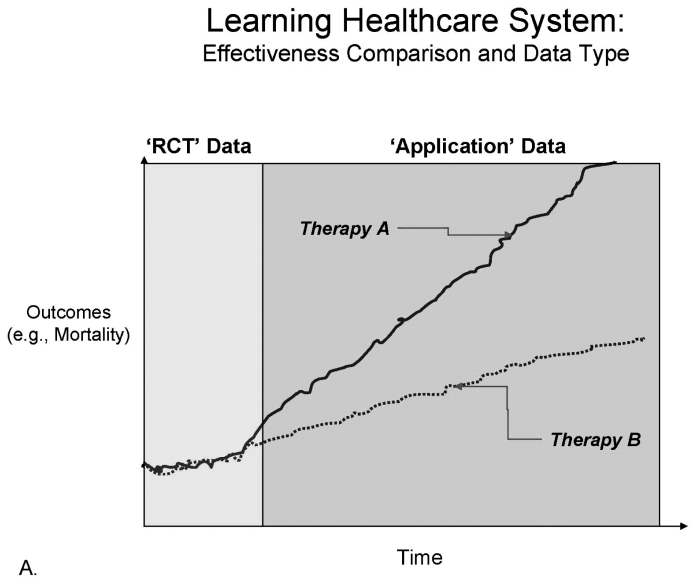


FIGURE 5-1 Distribution of the recommended research priorities by primary and secondary research areas.

Half of all topics involve a comparison to some aspect of the healthcare delivery system. Research topics categorized in this group focus on comparing *how* or *where* services are provided, rather than *which* services are provided. The prominence of health care delivery systems in the portfolio primarily reflects the interest of the public in this area, as well as the committee's belief that an early investment in CER should focus on learning how to make services more effective. Nearly one-third of the total recommended topics involve research that addresses racial and ethnic disparities and nearly one-fifth address functional limitations and disabilities. Other frequently represented areas are cardiovascular disease, geriatrics, psychiatric disorders, neurologic disorders, and pediatrics.

Twenty-nine out of the original 32 research areas are represented in the final portfolio. The missing categories include medical aspects of bioterrorism, pancreatic disorders, and regenerative medicine. The fact that there are no topics from any of these categories in the final list is less of a reflection of these categories' importance than of the fact that these categories only received 2 nominations out of the total 1,268 topics that entered the first round of voting and that the committee did not score the particular topics nominated within these categories as highly as topics in other categories. The portfolio's inclusion of 29 out of the original 32 research areas suggests that an investment in CER based on the committee's portfolio recommendations would comprehensively explore a broad spectrum of disease. It is interesting to note that, when asked for input, the public responded with recommendations that spanned a full portfolio of research areas.[3]

DIVERSITY OF POPULATIONS

A balanced portfolio should include a consideration of the demographic characteristics of the populations and subpopulations to be studied, including minority, racial, and ethnic groups; gender; and different age groups ranging from infancy to the elderly. It should also consider less obvious factors that affect health care, such as geographic location, socioeconomic status, educational achievement, and cultural differences; and it should be proportionately representative of those factors. Table 5-2 displays the 100 final priority topics categorized by study population. Many of the nominators of the priority topics selected more than one population as appropriate for the proposed research. Adults, including the elderly and the general population, are the most frequently represented study populations in the committee's portfolio. Other populations well represented in

[3] As discussed in Chapter 3, 82 percent of the committee's final priority list were nominated by the public; 18 percent were nominated by the committee during its in-depth discussion of the priority list.

TABLE 5-2 Committee's Recommended Research Priorities by Study Populations

Study Population	Number of Topics
Adults (including elderly)	36
Population at Large (general population)	28
Women	27
Special Populations (e.g., pregnant women, low income, patients with disabilities)	24
Men	22
Children and Adolescents Only	20
Elderly Only	15
Other	12
Long-Term Care	7
Ethnic Subpopulations Only	5
Adults (excluding elderly)	4
Rare Diseases	2
Total	202

NOTE: The total exceeds the total number of priority topics because respondents were allowed to select multiple populations for each topic.

the committee's portfolio are women, special populations (such as pregnant women and low-income families and individuals), men, and children and adolescents.

Based on the answers to the open-ended questions given by the questionnaire respondents, the "other" category in the table encompasses a wide variety of study populations, such as those with chronic conditions, cancer survivors, persons with psychiatric and mental disabilities, and persons at risk of developing heart disease.

DIVERSITY OF INTERVENTIONS

Another component of a balanced portfolio is that it should cover all steps in the trajectory of health care, from prevention and screening to diagnosis and treatment of acute and chronic health problems to palliative and end-of-life care. It should also reflect the full range of care modalities, from behavioral changes to pharmacological treatment to radiation to surgery. Table 5-3 displays the 100 final priority topics categorized by type of intervention or strategy proposed for the CER study. Types of comparators represented in the portfolio range from institutional and organization-based, such as management and delivery of health care, to patient-centered interventions. The patient-centered interventions range from completely

TABLE 5-3 Committee's Recommended Research Priorities by Types of Intervention

Types of Interventions	Number of Topics
Systems of Care	43
Pharmacological Treatment	36
Standard of Care	33
Behavioral Treatment	29
Prevention	24
Procedures	23
Provider-Patient Relationships	20
Treatment Pathways	19
Testing, Monitoring, and Evaluation	17
Devices	13
Alternative Treatment	9
Other	18
Total	284

NOTE: The total exceeds the total number of priority topics because respondents were allowed to select multiple interventions to be compared for each topic.

noninvasive approaches, such as ways to persuade patients to adopt healthier behavior, to major surgical procedures.

The interventions most strongly represented in the committee's portfolio are systems of care, pharmacologic treatment, and standard of care comparisons. Other frequently proposed types of interventions include behavioral treatments, disease prevention modalities, medical or surgical procedures (including radiological procedures), provider-patient forms of communication or other features of provider-patient relations, and treatment pathways (or clinical guidelines).

The list includes a broad array of diagnostic and therapeutic actions taken by primary care physicians and specialists. It also includes actions taken by other health professionals, ancillary service providers, administrators, and, importantly, health care leaders—for example, professional associations that develop treatment pathways. The "other" category includes interventions such as complementary care and economic incentives.

DIVERSITY OF STUDY METHODOLOGIES

Table 5-4 displays the division of the 100 final priority topics by study methodology. The four major methodologies identified by the committee as appropriate for CER are well represented on the committee's portfolio. Thus, the committee's portfolio provides a list of CER questions that vary

TABLE 5-4 Committee's Recommended Research Priorities by Study Methodology

Methodology	Number of Topics
Randomized Trial	49
Prospective Observational Study	46
Database Research	27
Systematic Review	23
TOTAL	145

NOTE: The total exceeds the total number of priority topics because respondents were allowed to select multiple methodologies for each topic.

widely in terms of resource requirements, timelines, and types of infrastructure necessary to conduct the research. For example, a database study using existing databases could be performed more rapidly and economically than a randomized clinical trial, but its findings and conclusions may be less definitive. The appropriate choice of method depends on the nature of the research, on whether the intervention is currently in use, on whether sufficient data are available to identify a large group of persons receiving the intervention and suitable unbiased comparator groups, and whether a range of patient outcomes is recorded.

INTRODUCTION TO FINAL LIST OF PRIORITY TOPICS

In preparing the list for presentation in this report, the committee refined the wording of each priority topic to fit a common format that indicates the research area, two or more interventions to be compared, the population, and, where appropriate and feasible, the outcomes of interest. The committee did not attempt to change the essence of the research question, or to change or add specific outcomes, nor did the committee attempt to refine the topics by specifying methodologies or comparators that the nominator did not provide. The committee fully anticipates that funding agencies, when preparing their Requests for Applications based on these priority topics, will provide details on the scope of the clinical problem, the current best practices, and the potential alternative approaches. It is ultimately the responsibility of the research teams applying for funding to propose the precise population, comparators, outcomes, and methodologies to be undertaken in the studies attempting to answer the priority questions. Moreover, a single priority topic is likely to generate alternative designs, so the committee's 100 priorities will likely provide the opportunity for many more than 100 specific research studies.

The voting process (described in detail in Chapter 4) introduced a substantial degree of subjectivity and variable weighting of topics. The com-

BOX 5-1
Round 3 Voting Procedures

One hundred fifty-five nominated research topics were considered in the committee's third round of voting. Each committee member was allocated 300 total points to distribute among the 155 topics but could not award more than 30 points to any one topic. The mean score for each topic was calculated by dividing the total points that each topic received by the number of committee members voting. The raw scores were reviewed by the committee, and the distribution of the scores provided a natural cutoff at 100 topics. The top 100 topics all received a mean of at least 1.0 points.

TABLE 5-5 Results of the IOM Committee's Final Vote for Priority Topics, by Quartile

Quartile	Mean Score	Standard Deviation	Range	
			Low	High
1	4.6	1.0	3.5	7.4
2	2.9	0.3	2.5	3.4
3	2.0	0.3	1.5	2.4
4	1.3	0.1	1.0	1.4

mittee felt that this imprecision reduced the reliability of relative rankings. Therefore, the 100 priority topics are presented grouped into quartiles, listed alphabetically by primary area of research.[4] The first quartile contains all topics with a mean score between 3.5 and 7.4 (see Box 5-1 for a brief recap of how the voting was conducted). The second quartile contains all topics with a mean score between 2.5 and 3.5. The third quartile contains all topics with a mean score between 1.5 and 2.5. The fourth quartile contains all topics with a mean score between 1 and 1.5. Refer to Table 5-5 to see the variability and ranges of the committee's votes across quartile. Table 5-6 displays the 100 priority topics by quartile. The medical terminology used in the list of priorities is defined in Appendix E.

[4] Note that 55 of the 155 nominated recommendations that appeared on the final ballot did not score high enough to be included in the final list. These 55 items are not represented in the quartiles.

TABLE 5-6 Final List of Priority Topics, by Quartile Ratings
**display within quartile does not indicate priority rank—topics are listed alphabetically by primary research area*

First Quartile (listed alphabetically by primary research area)	
CAD	Compare the effectiveness of treatment strategies for atrial fibrillation including surgery, catheter ablation, and pharmacologic treatment.
DIS	Compare the effectiveness of the different treatments (e.g., assistive listening devices, cochlear implants, electric-acoustic devices, habilitation and rehabilitation methods [auditory/oral, sign language, and total communication]) for hearing loss in children and adults, especially individuals with diverse cultural, language, medical, and developmental backgrounds.
ENDO	Compare the effectiveness of primary prevention methods, such as exercise and balance training, versus clinical treatments in preventing falls in older adults at varying degrees of risk.
GI	Compare the effectiveness of upper endoscopy utilization and frequency for patients with gastroesophageal reflux disease on morbidity, quality of life, and diagnosis of esophageal adenocarcinoma.
HCDS	Compare the effectiveness of dissemination and translation techniques to facilitate the use of CER by patients, clinicians, payers, and others.
HCDS	Compare the effectiveness of comprehensive care coordination programs, such as the medical home, and usual care in managing children and adults with severe chronic disease, especially in populations with known health disparities.
IMUN	Compare the effectiveness of different strategies of introducing biologics into the treatment algorithm for inflammatory diseases, including Crohn's disease, ulcerative colitis, rheumatoid arthritis, and psoriatic arthritis.
INFD	Compare the effectiveness of various screening, prophylaxis, and treatment interventions in eradicating methicillin-resistant *Staphylococcus aureus* (MRSA) in communities, institutions, and hospitals.
INFD	Compare the effectiveness of strategies (e.g., bio-patches, reducing central line entry, chlorhexidine for all line entries, antibiotic impregnated catheters, treating all line entries via a sterile field) for reducing healthcare-associated infections (HAI), including catheter-associated bloodstream infection, ventilator associated pneumonia, and surgical site infections in children and adults.
KUT	Compare the effectiveness of management strategies for localized prostate cancer (e.g., active surveillance, radical prostatectomy [conventional, robotic, and laparoscopic], and radiotherapy [conformal, brachytherapy, proton-beam, and intensity-modulated radiotherapy]) on survival, recurrence, side effects, quality of life, and costs.

TABLE 5-6 Continued

MS	Establish a prospective registry to compare the effectiveness of treatment strategies for low back pain without neurological deficit or spinal deformity.
NEURO	Compare the effectiveness and costs of alternative detection and management strategies (e.g., pharmacologic treatment, social/family support, combined pharmacologic and social/family support) for dementia in community-dwelling individuals and their caregivers.
NEURO	Compare the effectiveness of pharmacologic and non-pharmacologic treatments in managing behavioral disorders in people with Alzheimer's disease and other dementias in home and institutional settings.
NUTR	Compare the effectiveness of school-based interventions involving meal programs, vending machines, and physical education, at different levels of intensity, in preventing and treating overweight and obesity in children and adolescents.
NUTR	Compare the effectiveness of various strategies (e.g., clinical interventions, selected social interventions [such as improving the built environment in communities and making healthy foods more available], combined clinical and social interventions) to prevent obesity, hypertension, diabetes, and heart disease in at-risk populations such as the urban poor and American Indians.
ONC	Compare the effectiveness of management strategies for ductal carcinoma in situ (DCIS).
ONC	Compare the effectiveness of imaging technologies in diagnosing, staging, and monitoring patients with cancer including positron emission tomography (PET), magnetic resonance imaging (MRI), and computed tomography (CT).
ONC	Compare the effectiveness of genetic and biomarker testing and usual care in preventing and treating breast, colorectal, prostate, lung, and ovarian cancer, and possibly other clinical conditions for which promising biomarkers exist.
ORAL	Compare the effectiveness of the various delivery models (e.g., primary care, dental offices, schools, mobile vans) in preventing dental caries in children.
PEDS	Compare the effectiveness of various primary care treatment strategies (e.g., symptom management, cognitive behavior therapy, biofeedback, social skills, educator/teacher training, parent training, pharmacologic treatment) for attention deficit hyperactivity disorder (ADHD) in children.
PSYCH	Compare the effectiveness of wraparound home and community-based services and residential treatment in managing serious emotional disorders in children and adults.

continued

TABLE 5-6 Continued

RED	Compare the effectiveness of interventions (e.g., community-based multi-level interventions, simple health education, usual care) to reduce health disparities in cardiovascular disease, diabetes, cancer, musculoskeletal diseases, and birth outcomes.
RED	Compare the effectiveness of literacy-sensitive disease management programs and usual care in reducing disparities in children and adults with low literacy and chronic disease (e.g., heart disease).
WH	Compare the effectiveness of clinical interventions (e.g., prenatal care, nutritional counseling, smoking cessation, substance abuse treatment, combinations of these interventions) to reduce incidences of infant mortality, pre-term births, and low birth weights, especially among African American women.
WH	Compare the effectiveness of innovative strategies for preventing unintended pregnancies (e.g., over-the-counter access to oral contraceptives or other hormonal methods, expanding access to long-acting methods for young women, providing free contraceptive methods at public clinics, pharmacies, or other locations).
Second Quartile (listed alphabetically by primary research area)	
BDEV	Compare the effectiveness of therapeutic strategies (e.g., behavioral or pharmacologic interventions, the combination of the two) for different autism spectrum disorders (ASD) at different levels of severity and stages of intervention.
BDEV	Compare the effectiveness of the co-location model (psychological and primary care practitioners practicing together) and usual care (identification by primary care practitioner and referral to community-based mental health services) in identifying and treating social-emotional and developmental disorders in children ages 0-3.
BDEV	Compare the effectiveness of diverse models of comprehensive support services for infants and their families following discharge from a neonatal intensive care unit.
CAD	Compare the effectiveness of treatment strategies for vascular claudication (e.g., medical optimization, smoking cessation, exercise, catheter-based treatment, open surgical bypass).
CAM	Compare the effectiveness of mindfulness-based interventions (e.g., yoga, meditation, deep breathing training) and usual care in treating anxiety and depression, pain, cardiovascular risk factors, and chronic diseases.

TABLE 5-6 Continued

ENDO	Compare the long-term effectiveness of weight-bearing exercise and bisphosphonates in preventing hip and vertebral fractures in older women with osteopenia and/or osteoporosis.
HCDS	Compare the effectiveness of shared decision making and usual care on decision outcomes (treatment choice, knowledge, treatment-preference concordance, and decisional conflict) in children and adults with chronic disease such as stable angina and asthma.
HCDS	Compare the effectiveness of strategies for enhancing patients' adherence to medication regimens.
HCDS	Compare the effectiveness of patient decision support tools on informing diagnostic and treatment decisions (e.g., treatment choice, knowledge acquisition, treatment-preference concordance, decisional conflict) for elective surgical and nonsurgical procedures—especially in patients with limited English-language proficiency, limited education, hearing or visual impairments, or mental health problems.
HCDS	Compare the effectiveness of robotic assistance surgery and conventional surgery for common operations, such as prostatectomies.
HCDS	Compare the effectiveness (including resource utilization, workforce needs, net health care expenditures, and requirements for large-scale deployment) of new remote patient monitoring and management technologies (e.g., telemedicine, Internet, remote sensing) and usual care in managing chronic disease, especially in rural settings.
HCDS	Compare the effectiveness of diverse models of transition support services for adults with complex health care needs (e.g., the elderly, homeless, mentally challenged) after hospital discharge.
HCDS	Compare the effectiveness of accountable care systems and usual care on costs, processes of care, and outcomes for geographically defined populations of patients with one or more chronic diseases.
HCDS	Compare the effectiveness of different residential settings (e.g., home care, nursing home, group home) in caring for elderly patients with functional impairments.
KUT	Compare the effectiveness (including survival, hospitalization, quality of life, and costs) of renal replacement therapies (e.g., daily home hemodialysis, intermittent home hemodialysis, conventional in-center dialysis, continuous ambulatory peritoneal dialysis, renal transplantation) for patients of different ages, races, and ethnicities.
MS	Compare the effectiveness of treatment strategies (e.g., artificial cervical discs, spinal fusion, pharmacologic treatment with physical therapy) for cervical disc and neck pain.

continued

TABLE 5-6 Continued

ONC	Compare the effectiveness of film-screen or digital mammography alone and mammography plus magnetic resonance imaging (MRI) in community practice-based screening for breast cancer in high-risk women of different ages, risk factors, and race or ethnicity.
ONC	Compare the effectiveness of new screening technologies (such as fecal immunochemical tests and computed tomography [CT] colonography) and usual care (fecal occult blood tests and colonoscopy) in preventing colorectal cancer.
PELC	Compare the effectiveness of coordinated care (supported by reimbursement innovations) and usual care in long-term and end-of-life care of the elderly.
PSYCH	Compare the effectiveness of pharmacologic treatment and behavioral interventions in managing major depressive disorders in adolescents and adults in diverse treatment settings.
RD	Compare the effectiveness of an integrated approach (combining counseling, environmental mitigation, chronic disease management, and legal assistance) with a non-integrated episodic care model in managing asthma in children.
SKIN	Compare the effectiveness (including effects on quality of life) of treatment strategies (e.g., topical steroids, ultraviolet light, methotrexate, biologic response modifiers) for psoriasis.
TEMC	Compare the effectiveness of treatment strategies (e.g., cognitive behavioral individual therapy, generic individual therapy, comprehensive and intensive treatment) for post-traumatic stress disorder stemming from diverse sources of trauma.
WH	Compare the effectiveness and outcomes of care with obstetric ultrasound studies and care without the use of ultrasound in normal pregnancies.
WH	Compare the effectiveness of birthing care in freestanding birth centers and usual care of childbearing women at low and moderate risk.
Third Quartile (listed alphabetically by primary research area)	
ADDO	Compare the effectiveness of different opioid and non-opioid pain relievers, in different doses and durations, in avoiding unintentional overdose and substance dependence among subjects with acute and non-cancer chronic pain.
CAD	Compare the effectiveness of aggressive medical management and percutaneous coronary interventions in treating stable coronary disease for patients of different ages and with different comorbidities.

TABLE 5-6 Continued

CAD	Compare the effectiveness of innovative treatment strategies (e.g., cardiac resynchronization, remote physiologic monitoring, pharmacologic treatment, novel agents such as CRF-2 receptors) for congestive heart failure.
CAD	Compare the effectiveness of traditional risk stratification for coronary heart disease (CHD) and noninvasive imaging (using coronary artery calcium, carotid intima media thickness, and other approaches) on CHD outcomes.
CAD	Compare the effectiveness of different treatment strategies (e.g., modifying target levels for glucose, lipid, or blood pressure) in reducing cardiovascular complications in newly diagnosed adolescents and adults with type 2 diabetes.
CAM	Compare the effectiveness of acupuncture for various indications using a cluster randomized trial.
CAM	Compare the effectiveness of dietary supplements (nutriceuticals) and usual care in the treatment of selected high-prevalence conditions.
EENT	Compare the effectiveness of different treatment options (e.g., laser therapy, intravitreal steroids, anti-vascular endothelial growth factor [anti-VEGF]) for diabetic retinopathy, macular degeneration, and retinal vein occlusion.
EENT	Compare the effectiveness of treatment strategies for primary open-angle glaucoma (e.g., initial laser surgery, new surgical techniques, new medical treatments) particularly in minority populations to assess clinical and patient-reported outcomes.
ENDO	Compare the effectiveness and cost-effectiveness of conventional medical management of type 2 diabetes in adolescents and adults, versus conventional therapy plus intensive educational programs or programs incorporating support groups and educational resources.
HCDS	Compare the effectiveness of alternative redesign strategies—using decision support capabilities, electronic health records, and personal health records—for increasing health professionals' compliance with evidence-based guidelines and patients' adherence to guideline-based regimens for chronic disease care.
HCDS	Compare the effectiveness of adding information about new biomarkers (including genetic information) with standard care in motivating behavior change and improving clinical outcomes.
HCDS	Compare the effectiveness of different quality improvement strategies in disease prevention, acute care, chronic disease care, and rehabilitation services for diverse populations of children and adults.

continued

TABLE 5-6 Continued

HCDS	Compare the effectiveness of formulary management practices and usual practices in controlling hospital expenditures for products other than drugs including medical devices (surgical hemostatic products, radiocontrast, interventional cardiology devices, and others).
HCDS	Compare the effectiveness of different benefit design, utilization management, and cost-sharing strategies in improving health care access and quality in patients with chronic diseases (e.g., cancer, diabetes, heart disease).
INFD	Compare the effectiveness of HIV screening strategies based on recent Centers for Disease Control and Prevention recommendations and traditional screening in primary care settings with significant prevention counseling.
MS	Establish a prospective registry to compare the effectiveness of surgical and nonsurgical strategies for treating cervical spondylotic myelopathy (CSM) in patients with different characteristics to delineate predictors of improved outcomes.
NEURO	Compare the effectiveness of traditional and newer imaging modalities (e.g., routine imaging, magnetic resonance imaging [MRI], computed tomography [CT], positron emission tomography [PET]) when ordered for neurological and orthopedic indications by primary care practitioners, emergency department physicians, and specialists.
NEURO	Compare the effectiveness of comprehensive, coordinated care and usual care on objective measures of clinical status, patient-reported outcomes, and costs of care for people with multiple sclerosis.
NUTR	Compare the effectiveness of treatment strategies for obesity (e.g., bariatric surgery, behavioral interventions, pharmacologic treatment) on the resolution of obesity-related outcomes such as diabetes, hypertension, and musculoskeletal disorders.
ORAL	Compare the clinical and cost-effectiveness of surgical care and a medical model of prevention and care in managing periodontal disease to increase tooth longevity and reduce systemic secondary effects in other organ systems.
PSYCH	Compare the effectiveness of atypical antipsychotic drug therapy and conventional pharmacologic treatment for Food and Drug Administration-approved indications and compendia-referenced off-label indications using large datasets.
PSYCH	Compare the effectiveness of management strategies (e.g., inpatient psychiatric hospitalization, extended observation, partial hospitalization, intensive outpatient care) for adolescents and adults following a suicide attempt.

TABLE 5-6 Continued

RED	Compare the effectiveness of different strategies to engage and retain patients in care and to delineate barriers to care, especially for members of populations that experience health disparities.
SKIN	Compare the effectiveness of topical treatments (e.g., antibiotics, platelet-derived growth factor) and systemic therapies (e.g., negative pressure wound therapy, hyperbaric oxygen) in managing chronic lower extremity wounds.

Fourth Quartile
(listed alphabetically by primary research area)

ADDO	Compare the effectiveness of smoking cessation strategies (e.g., medication, individual or quitline counseling, combinations of these) in smokers from understudied populations such as minorities, individuals with mental illness, and adolescents.
CAD	Compare the effectiveness of computed tomography (CT) angiography and conventional angiography in assessing coronary stenosis in patients at moderate pretest risk of coronary artery disease.
CAD	Compare the effectiveness of anticoagulant therapies (e.g., low-intensity warfarin, aspirin, injectable anticoagulants) for patients undergoing hip or knee arthroplasty surgery.
DIS	Compare the effectiveness of focused intense periodic therapy and usual weekly therapy in managing cerebral palsy in children.
ENDO	Compare the effectiveness of different disease management strategies in improving the adherence to and value of pharmacologic treatments for the elderly.
HCDS	Compare the effectiveness of care coordination with and without clinical decision supports (e.g., electronic health records) in producing good health outcomes in chronically ill patients, including children with special healthcare needs.
HCDS	Compare the effectiveness of coordinated, physician-led, interdisciplinary care provided in the patient's residence and usual care in managing advanced chronic disease in community-dwelling patients with significant functional impairments.
HCDS	Compare the effectiveness of minimally invasive abdominal surgery and open surgical procedures on post-operative infections, pain management, and recuperative requirements.
HCDS	Compare the effectiveness of traditional behavioral interventions versus economic incentives in motivating behavior changes (e.g., weight loss, smoking cessation, avoiding alcohol and substance abuse) in children and adults.

continued

TABLE 5-6 Continued

HCDS	Compare the effectiveness of diagnostic imaging performed by non-radiologists and radiologists.
HCDS	Compare the effectiveness of different techniques (e.g., audio, visual, written) for informing patients about proposed treatments during the process of informed consent.
HCDS	Compare the effectiveness of different disease management strategies for activating patients with chronic disease.
HCDS	Compare the effectiveness of different delivery models (e.g., home blood pressure monitors, utilization of pharmacists or other allied health providers) for controlling hypertension, especially in racial minorities.
INFD	Compare the effectiveness of alternative clinical management strategies for hepatitis C, including alternative duration of therapy for patients based on viral genomic profile and patient risk factors (e.g., behavior-related risk factors).
MS	Compare the effectiveness of different treatment strategies in the prevention of progression and disability from osteoarthritis.
MS	Compare the effectiveness (e.g., pain relief, functional outcomes) of different surgical strategies for symptomatic cervical disc herniation in patients for whom appropriate nonsurgical care has failed.
NEURO	Compare the effectiveness of different treatment strategies on the frequency and lost productivity in people with chronic, frequent migraine headaches.
NEURO	Compare the effectiveness of monotherapy and polytherapy (i.e., use of two or more drugs) on seizure frequency, adverse events, quality of life, and cost in patients with intractable epilepsy.
ONC	Compare the effectiveness of surgical resection, observation, or ablative techniques on disease-free and overall survival, tumor recurrence, quality of life, and toxicity in patients with liver metastases.
PELC	Compare the effectiveness of hospital-based palliative care and usual care on patient-reported outcomes and cost.
PSYCH	Compare the effectiveness of different treatment approaches (e.g., integrating mental health care and primary care, improving consumer self-care, a combination of integration and self-care) in avoiding early mortality and comorbidity among people with serious and persistent mental illness.

TABLE 5-6 Continued

PSYCH	Compare the effectiveness of traditional training of primary care physicians in primary care mental health and co-location systems of primary care and mental health care on outcomes including depression, anxiety, physical symptoms, physical disability, prescription substance use, mental and physical function, satisfaction with the provider, and cost.
PSYCH	Compare the effectiveness of different treatment strategies (e.g., psychotherapy, antidepressants, combination treatment with case management) for depression after myocardial infarction on medication adherence, cardiovascular events, hospitalization, and death.
SKIN	Compare the effectiveness of different long-term treatments for acne.
WH	Compare the effectiveness of different strategies for promoting breastfeeding among low-income African American women.

NOTE: ADDO = Alcoholism, Drug Dependency, and Overdose; BDEV = Birth and Developmental Disorders; CAD = Cardiovascular and Peripheral Vascular Disease; CAM = Complementary and Alternative Medicine; DIS = Functional Limitations and Disabilities; EENT = Eyes, Ears, Nose, and Throat Disorders; ENDO = Endocrinology and Metabolism Disorders and Geriatrics; GI = Gastrointestinal System Disorders; HCDS = Health Care Delivery Systems; IMUN = Immune System, Connective Tissue, and Joint Disorders; INFD = Infectious Diseases Liver and Biliary Tract Disorders; KUT = Kidney and Urinary Tract Disorders; MS = Musculoskeletal Disorders; NEURO = Neurologic Disorders; NUTR = Nutrition (including obesity); ONC = Oncology and Hematology; ORAL = Oral Health; PEDS = Pediatrics; PELC = Palliative and End-of-Life Care; PSYCH = Psychiatric Disorders; RD = Respiratory Disease; RED = Racial and Ethnic Disparities; SKIN = Skin Disorders; TEMC = Trauma, Emergency Medicine, and Critical Care Medicine; WH = Women's Health.

DISCUSSION OF THE PRIORITY TOPICS BY RESEARCH AREA

The following discussion presents the items contained in the final list of 100 priority topics, grouped by primary research area. The importance of the research area is explained, with reference to the criteria used by the IOM committee members in voting.

For voting purposes, each nominated priority topic was assigned to a primary research area.[5] The remainder of this section presents the priority topics by research areas. The areas containing the most topics are presented first.

[5] As discussed in Chapter 4, the committee's subgroup reviewed all of the nominated priorities and assigned each topic to a primary research area.

Health Care Delivery Systems[6]

Almost one-fourth of the committee's recommended priority topics are classified primarily in the health care delivery system (HCDS) research area. This is a broad category that includes topics related to dissemination of CER study results; patient decision making, health behavior and care management, comparing settings of care, and utilization of surgical, radiological, and medical procedures (Table 5-7). Different dissemination techniques are proposed for study (HCDS-A) to ensure that interventions are widely adopted in practice once CER studies prove them effective. Five priority topics focus on patient decision making (HCDS-B–F) involving decision support tools and other mechanisms, such as electronic health records, to help patients make informed choices about their care. Health behaviors, such as smoking, are the subject of four topics (HCDS-G–J), which involve disease management (a comprehensive approach to caring for patients with chronic diseases), clinical guidelines (as followed by both clinicians and patients), information about genetic biomarkers and their impact on patient choice of diagnostic and therapeutic approaches, and economic incentives to adopt a healthier lifestyle. Healthcare management (HCDS-K–P) specifically addresses quality improvement, post-hospital transition support, hospital formularies for medical devices, comprehensive care coordination, population-based "accountable care," and certain health system strategies (such as revising health insurance policies). Settings of care topics (HCDS-Q–S) address remote patient monitoring, care that is not structured around office visits to physicians, including community and home-based care for elderly and chronic disease patients. Certain procedures included in the health care delivery system research area (HCDS-T–W) address robotic surgery, minimally invasive surgery, scanning and imaging performed by physicians other than radiologists, and methods of controlling hypertension.

Other groups have set a high priority on studying healthcare delivery topics. Several aspects of this expansive topic were identified as important by *Healthy People 2010*, the National Quality Forum, and the Cochrane Collaboration (Doyle et al., 2005; HHS, 2000; NPP, 2008). These aspects include access to quality health services, education and community-based programs, environmental health, food safety, health communication, medical product safety, occupational safety and health, public health infrastructure, safety and reliability of the health care system, integration and coordination of care, overuse and misuse of care, and organizational capacity.

The large number of recommended topics addressing health care and delivery reflects the dramatic variability of care from region to region, the

[6] Described in the questionnaire as "Safety and Quality of Health Care."

TABLE 5-7 Health Care Delivery Systems Priority Topics

HCDS-A	Compare the effectiveness of dissemination and translation techniques to facilitate the use of CER by patients, clinicians, payers, and others.
HCDS-B	Compare the effectiveness of shared decision making and usual care on decision outcomes (treatment choice, knowledge, treatment-preference concordance, and decisional conflict) in children and adults with chronic disease such as stable angina and asthma.
HCDS-C	Compare the effectiveness of patient decision support tools on informing diagnostic and treatment decisions (e.g., treatment choice, knowledge acquisition, treatment-preference concordance, decisional conflict) for elective surgical and nonsurgical procedures—especially in patients with limited English-language proficiency, limited education, hearing or visual impairments, or mental health problems.
HCDS-D	Compare the effectiveness of care coordination with and without clinical decision supports (e.g., electronic health records) in producing good health outcomes in chronically ill patients, including children with special health care needs.
HCDS-E	Compare the effectiveness of different techniques (e.g., audio, visual, written) for informing patients about proposed treatments during the process of informed consent.
HCDS-F	Compare the effectiveness of strategies for enhancing patients' adherence to medication regimens.
HCDS-G	Compare the effectiveness of different disease management strategies for activating patients with chronic disease.
HCDS-H	Compare the effectiveness of alternative redesign strategies—using decision support capabilities, electronic health records, and personal health records—for increasing health professionals' compliance with evidence-based guidelines and patients' adherence to guideline-based regimens for chronic disease care.
HCDS-I	Compare the effectiveness of adding information about new biomarkers (including genetic information) with standard care in motivating behavior change and improving clinical outcomes.
HCDS-J	Compare the effectiveness of traditional behavioral interventions versus economic incentives in motivating behavior changes (e.g., weight loss, smoking cessation, avoiding alcohol and substance abuse) in children and adults.
HCDS-K	Compare the effectiveness of different quality improvement strategies in disease prevention, acute care, chronic disease care, and rehabilitation services for diverse populations of children and adults.

continued

TABLE 5-7 Continued

HCDS-L	Compare the effectiveness of diverse models of transition support services for adults with complex health care needs (e.g., the elderly, homeless, mentally challenged) after hospital discharge.
HCDS-M	Compare the effectiveness of formulary management practices and usual practices in controlling hospital expenditures for products other than drugs including medical devices (surgical hemostatic products, radiocontrast, interventional cardiology devices, and others).
HCDS-N	Compare the effectiveness of comprehensive care coordination programs, such as the medical home, and usual care in managing children and adults with severe chronic disease, especially in populations with known health disparities.
HCDS-O	Compare the effectiveness of accountable care systems and usual care on costs, processes of care, and outcomes for geographically defined populations of patients with one or more chronic diseases.
HCDS-P	Compare the effectiveness of different benefit design, utilization management, and cost-sharing strategies in improving health care access and quality in patients with chronic diseases (e.g., cancer, diabetes, heart disease).
HCDS-Q	Compare the effectiveness (including resource utilization, workforce needs, net health care expenditures, and requirements for large-scale deployment) of new remote patient monitoring and management technologies (e.g., telemedicine, Internet, remote sensing) and usual care in managing chronic disease, especially in rural settings.
HCDS-R	Compare the effectiveness of different residential settings (e.g., home care, nursing home, group home) in caring for elderly patients with functional impairments.
HCDS-S	Compare the effectiveness of coordinated, physician-led, interdisciplinary care provided in the patient's residence and usual care in managing advanced chronic disease in community-dwelling patients with significant functional impairments.
HCDS-T	Compare the effectiveness of robotic assistance surgery and conventional surgery for common operations, such as prostatectomies.
HCDS-U	Compare the effectiveness of minimally invasive abdominal surgery and open surgical procedures on post-operative infections, pain management, and recuperative requirements.
HCDS-V	Compare the effectiveness of diagnostic imaging performed by non-radiologists and radiologists.
HCDS-W	Compare the effectiveness of different delivery models (e.g., home blood pressure monitors, utilization of pharmacists or other allied health providers) for controlling hypertension, especially in racial minorities.

lack of clarity of what constitutes best practice, and the desire to identify optimal systems for providing health care.

Cardiovascular and Peripheral Vascular Disease

Cardiovascular and Peripheral Vascular disease was the second-ranked topic category among the committee's top 100 priority topics. Diseases of the heart were ranked as the leading cause of death in 2005 according to the Centers for Disease Control and Prevention's (CDC's) National Vital Statistics Reports (Kung et al., 2008). Such diseases are associated with multiple comorbidities that are becoming increasingly prevalent, such as diabetes and obesity. The final priority list had eight topics (Table 5-8) dealing with ischemic heart disease (CAD-A–D) and heart failure (CAD-E), which are among the leading causes of death in all age groups (Kung et al., 2008) together with cardiac arrhythmias (CAD-F), which are among the most variably treated conditions (Wennberg, 2009). In addition, the AHRQ Effective Health Care program, *Healthy People 2010*, and the Cochrane Collaboration rank cardiovascular disease among the highest national priorities for health (Doyle et al., 2005; HHS, 2000; Whitlock et al., 2009). The committee's list also had two topics that focused on the treatment and management of peripheral vascular disorders (CAD-G–H).

Psychiatric Disorders

Across the nation, the prevalence of mental health disorders is high, and the cost of treating such disorders is substantial. The committee recommended that CER address several important psychiatric disorders (Table 5-9). AHRQ's Effective Health Care Program, *Healthy People 2010*, and the Cochrane Collaboration agree that mental health disorders are a priority research area for the nation (Doyle et al., 2005; HHS, 2000; Whitlock et al., 2009). Three topics address various strategies for managing and treating mental health disorders (ranked among the most prevalent, the most costly, and the leading causes of morbidity across all age groups) (AHRQ, 2009a,c; Kung et al., 2008) by specifically studying location of care, provider training, and various pharmacologic treatments (PSYCH-A–C). Depression contributes to suicidal ideation and suicide and is one of the leading causes of mortality across all age groups (Kung et al., 2008). The final list includes two topics addressing depression (PSYCH-D–E), and two that address early mortality (PSYCH-F) and suicide (PSYCH-G).

TABLE 5-8 Cardiovascular and Peripheral Vascular Diseases Priority Topics

CAD-A	Compare the effectiveness of aggressive medical management and percutaneous coronary interventions in treating stable coronary disease for patients of different ages and with different comorbidities.
CAD-B	Compare the effectiveness of traditional risk stratification for coronary heart disease (CHD) and noninvasive imaging (using coronary artery calcium, carotid intima media thickness, and other approaches) on CHD outcomes.
CAD-C	Compare the effectiveness of different treatment strategies (e.g., modifying target levels for glucose, lipid, or blood pressure) in reducing cardiovascular complications in newly diagnosed adolescents and adults with type 2 diabetes.
CAD-D	Compare the effectiveness of computed tomography (CT) angiography and conventional angiography in assessing coronary stenosis in patients at moderate pretest risk of coronary artery disease.
CAD-E	Compare the effectiveness of innovative treatment strategies (e.g., cardiac resynchronization, remote physiologic monitoring, pharmacologic treatment, novel agents such as CRF-2 receptors) for congestive heart failure.
CAD-F	Compare the effectiveness of treatment strategies for atrial fibrillation including surgery, catheter ablation, and pharmacologic treatment.
CAD-G	Compare the effectiveness of treatment strategies for vascular claudication (e.g., medical optimization, smoking cessation, exercise, catheter-based treatment, open surgical bypass).
CAD-H	Compare the effectiveness of anticoagulant therapies (e.g., low-intensity warfarin, aspirin, injectable anticoagulants) for patients undergoing hip or knee arthroplasty surgery.

Neurologic Disorders

The final priority list includes six topics in the area of neurologic disorders (Table 5-10). These address imaging used for diagnosing neurologic conditions (NEURO-A), treatment of headaches (NEURO-B), multiple sclerosis (NEURO-C), epilepsy (NEURO-D), and the detection, treatment, and management of dementia (NEURO-E) and Alzheimer's disease (NEURO-F). Epilepsy is one of the most costly disorders affecting adolescents (AHRQ, 2009a), while dementias disproportionately affect the elderly, and are considered national priorities by the AHRQ Effective Health Care Program (Whitlock et al., 2009).

TABLE 5-9 Psychiatric Disorders Priority Topics

PSYCH-A	Compare the effectiveness of wraparound home and community-based services and residential treatment in managing serious emotional disorders in children and adults.
PSYCH-B	Compare the effectiveness of atypical antipsychotic drug therapy and conventional pharmacologic treatment for Food and Drug Administration-approved indications and compendia-referenced off-label indications using large datasets.
PSYCH-C	Compare the effectiveness of traditional training of primary care physicians in primary care mental health and co-location systems of primary care and mental health care on outcomes including depression, anxiety, physical symptoms, physical disability, prescription substance use, mental and physical function, satisfaction with the provider, and cost.
PSYCH-D	Compare the effectiveness of pharmacologic treatment and behavioral interventions in managing major depressive disorders in adolescents and adults in diverse treatment settings.
PSYCH-E	Compare the effectiveness of different treatment strategies (e.g., psychotherapy, antidepressants, combination treatment with case management) for depression after myocardial infarction on medication adherence, cardiovascular events, hospitalization, and death.
PSYCH-F	Compare the effectiveness of different treatment approaches (e.g., integrating mental health care and primary care, improving consumer self-care, a combination of integration and self-care) in avoiding early mortality and comorbidity among people with serious and persistent mental illness.
PSYCH-G	Compare the effectiveness of management strategies (e.g., inpatient psychiatric hospitalization, extended observation, partial hospitalization, intensive outpatient care) for adolescents and adults following a suicide attempt.

Oncology and Hematology

Cancer is a leading cause of death and among the most costly conditions to treat (AHRQ, 2009a; Kung et al., 2008). Cancer is also listed as a national priority by the AHRQ Effective Health Care Program and *Healthy People 2010* (HHS, 2000; Whitlock et al., 2009). The final priority list includes six topics in this research area (Table 5-11). These include two topics involving screening technologies for colorectal and breast cancer (ONC-A–B). Breast cancer is among the most variably treated diseases, due in part to the large number of subtypes of breast cancer (Wennberg, 2009). One topic specifically addresses strategies for managing one of those

TABLE 5-10 Neurologic Disorders Priority Topics

NEURO-A	Compare the effectiveness of traditional and newer imaging modalities (e.g., routine imaging, magnetic resonance imaging [MRI], computed tomography [CT], positron emission tomography [PET]) when ordered for neurological and orthopedic indications by primary care practitioners, emergency department physicians, and specialists.
NEURO-B	Compare the effectiveness of different treatment strategies on the frequency and lost productivity in people with chronic, frequent migraine headaches.
NEURO-C	Compare the effectiveness of comprehensive, coordinated care and usual care on objective measures of clinical status, patient-reported outcomes, and costs of care for people with multiple sclerosis.
NEURO-D	Compare the effectiveness of monotherapy and polytherapy (i.e., use of two or more drugs) on seizure frequency, adverse events, quality of life, and cost in patients with intractable epilepsy.
NEURO-E	Compare the effectiveness and costs of alternative detection and management strategies (e.g., pharmacologic treatment, social/family support, combined pharmacologic and social/family support) for dementia in community-dwelling individuals and their caregivers.
NEURO-F	Compare the effectiveness of pharmacologic and non-pharmacologic treatments in managing behavioral disorders in people with Alzheimer's disease and other dementias in home and institutional settings.

subtypes, ductal carcinoma in situ (ONC-C). The topics also address the use of imaging technologies for diagnosing, staging, and monitoring all cancers (ONC-D), the use of biomarker analysis in risk assessment and treatment strategies for common cancers (ONC-E), and comparing treatment strategies for liver metastases (ONC-F).

Women's Health

Three of the five priority topics in the area of women's health emphasize conditions of particular importance among minority and underserved populations (Table 5-12). One topic addresses the prevention of unplanned pregnancies (WH-A), focusing on the effectiveness of strategies to expand access to care and systems of health care delivery. One topic focuses on alternative interventions to ensure healthy pregnancies and manage risky pregnancies in minority populations, including behavioral interventions to reduce infant mortality, preterm birth, and low birth weight (WH-B). One topic examines the optimal use of ultrasound during pregnancy (WH-C). The use of ultrasound scanning throughout gestation in both normal and

TABLE 5-11 Oncology and Hematology Priority Topics

ONC-A	Compare the effectiveness of film-screen or digital mammography alone and mammography plus magnetic resonance imaging (MRI) in community practice-based screening for breast cancer in high-risk women of different ages, risk factors, and race or ethnicity.
ONC-B	Compare the effectiveness of new screening technologies (such as fecal immunochemical tests and computed tomography [CT] colonography) and usual care (fecal occult blood tests and colonoscopy) in preventing colorectal cancer.
ONC-C	Compare the effectiveness of management strategies for ductal carcinoma in situ (DCIS).
ONC-D	Compare the effectiveness of imaging technologies in diagnosing, staging, and monitoring patients with cancer including positron emission tomography (PET), magnetic resonance imaging (MRI), and computed tomography (CT).
ONC-E	Compare the effectiveness of genetic and biomarker testing and usual care in preventing and treating breast, colorectal, prostate, lung, and ovarian cancer, and possibly other clinical conditions for which promising biomarkers exist.
ONC-F	Compare the effectiveness of surgical resection, observation, or ablative techniques on disease-free and overall survival, tumor recurrence, quality of life, and toxicity in patients with liver metastases.

TABLE 5-12 Women's Health Priority Topics

WH-A	Compare the effectiveness of innovative strategies for preventing unintended pregnancies (e.g., over-the-counter access to oral contraceptives or other hormonal methods, expanding access to long-acting methods for young women, providing free contraceptive methods at public clinics, pharmacies, or other locations).
WH-B	Compare the effectiveness of clinical interventions (e.g., prenatal care, nutritional counseling, smoking cessation, substance abuse treatment, combinations of these interventions) to reduce incidences of infant mortality, pre-term births, and low birth weights, especially among African American women.
WH-C	Compare the effectiveness and outcomes of care with obstetric ultrasound studies and care without the use of ultrasound in normal pregnancies.
WH-D	Compare the effectiveness of birthing care in freestanding birth centers and usual care of childbearing women at low and moderate risk.
WH-E	Compare the effectiveness of different strategies for promoting breastfeeding among low-income African American women.

high-risk pregnancies is highly variable, and it is not yet known whether frequency of use affects pregnancy outcomes or safety. One topic addresses the impact of birthing location on outcomes (WH-D) and, finally, the committee recommended examination of programs to promote breastfeeding in African American women (WH-E). Topics related to metabolic bone disease and cardiovascular disease as they affect women are discussed within those specific research areas.

Musculoskeletal Disorders

Although musculoskeletal disorders produce a very broad range of health problems, the committee's topics focused on two primary disorders: (1) neck and back pain, and (2) osteoarthritis, both considered to be priorities in *Healthy People 2010* (HHS, 2000). The committee recommended four priorities focusing on back problems (Table 5-13), which are listed among the most prevalent, most costly, most variable, and most morbid conditions (AHRQ, 2009a,b,c; Wennberg, 2009). Two of these topics focus on management and treatment strategies for low back pain and cervical spondylotic myelopathy (compression of the spinal cord) (MS-A–B), including identification of patient-specific biomarkers to help predict outcome and inform treatment strategies. The others focus on surgical and nonsurgical treatment strategies for cervical disc and neck pain (MS-C–D). The remaining topic in this research area addresses interventions to prevent disability and progression of osteoarthritis (MS-E).

TABLE 5-13 Musculoskeletal Disorders Priority Topics

MS-A	Establish a prospective registry to compare the effectiveness of treatment strategies for low back pain without neurological deficit or spinal deformity.
MS-B	Establish a prospective registry to compare the effectiveness of surgical and nonsurgical strategies for treating cervical spondylotic myelopathy (CSM) in patients with different characteristics to delineate predictors of improved outcomes.
MS-C	Compare the effectiveness of treatment strategies (e.g., artificial cervical discs, spinal fusion, pharmacologic treatment with physical therapy) for cervical disc and neck pain.
MS-D	Compare the effectiveness (e.g., pain relief, functional outcomes) of different surgical strategies for symptomatic cervical disc herniation in patients for whom appropriate nonsurgical care has failed.
MS-E	Compare the effectiveness of different treatment strategies in the prevention of progression and disability from osteoarthritis.

Infectious Diseases and Liver and Biliary Tract Disorders

Infectious diseases carry risks for infected patients and also constitute a significant public threat because they can be transmitted from person to person through a variety of mechanisms. Once detected, effective treatments can be applied and transmission of many infectious diseases can be mitigated. The committee's topics focus on screening for detection, interventions to reduce transmission, and clinical management of chronic infectious diseases (Table 5-14). The specific diseases highlighted by the committee's topics include methicillin-resistant *Staphylococcus aureus* (MRSA) (INFD-A), hepatitis C (INFD-B), human immunodeficiency virus (HIV) (INFD-C), and more generally hospital acquired infections (HAI) (INFD-D). Hospital acquired infections can be deadly if not treated properly—in fact, septicemia and pneumonia, two diseases commonly transmitted in hospital settings are among the most variably treated conditions according to the Dartmouth Atlas (Wennberg, 2009). Finding effective methods to reduce such infections is critically important to the health of the nation. Chronic infections with HIV and hepatitis C can now be treated so that people live decades. However, identifying optimal treatment strategies, particularly in African American populations and at-risk populations, such as intravenous drug users, require more research. Both infectious diseases generally, and HIV/AIDS in particular, are listed by AHRQ's Effective

TABLE 5-14 Infectious Disease and Liver and Biliary Tract Disorder Priority Topics

INFD-A	Compare the effectiveness of various screening, prophylaxis, and treatment interventions in eradicating methicillin-resistant *Staphylococcus aureus* (MRSA) in communities, institutions, and hospitals.
INFD-B	Compare the effectiveness of alternative clinical management strategies for hepatitis C, including alternative duration of therapy for patients based on viral genomic profile and patient risk factors (e.g., behavior-related risk factors).
INFD-C	Compare the effectiveness of HIV screening strategies based on recent Centers for Disease Control and Prevention recommendations and traditional screening in primary care settings with significant prevention counseling.
INFD-D	Compare the effectiveness of strategies (e.g., bio-patches, reducing central line entry, chlorhexidine for all line entries, antibiotic impregnated catheters, treating all line entries via a sterile field) for reducing healthcare-associated infections (HAI), including catheter-associated bloodstream infection, ventilator associated pneumonia, and surgical site infections in children and adults.

Health Care Program and *Healthy People 2010* as conditions of national importance (HHS, 2000; Whitlock et al., 2009).

Endocrinology and Metabolism Disorders and Geriatrics

Diabetes, which ranks among the most prevalent and most costly diseases throughout the nation, is associated with multiple comorbidities including heart disease, stroke, and obesity (AHRQ, 2009a,c). In addition, it is among the leading causes of morbidity and mortality (AHRQ, 2009b; Kung et al., 2008). Determining the effectiveness and cost effectiveness of alternative strategies to treat type 2 diabetes in adolescents and adults has the potential to dramatically improve health and reduce healthcare costs across the country. As such, the committee recommended it as a priority (ENDO-A), as did AHRQ's Effective Health Care Program and *Healthy People 2010* (AHRQ Effective Health Care Program, 2009; HHS, 2000) (Table 5-15).

As the baby boomer generation continues to age, it will be important to determine the effectiveness of strategies to reduce hip and vertebral fractures in patients both with and without osteopenia and osteoporosis. The committee concluded that falls, which are a contributing factor to fractures, should also be among its list of national priorities (ENDO-B–C).

Many older Americans take multiple medications on a routine basis. The committee recommends performing studies to evaluate the impact of

TABLE 5-15 Endocrinology and Metabolism Disorders and Geriatric Priority Topics

ENDO-A	Compare the effectiveness and cost-effectiveness of conventional medical management of type 2 diabetes in adolescents and adults, versus conventional therapy plus intensive educational programs or programs incorporating support groups and educational resources.
ENDO-B	Compare the long-term effectiveness of weight-bearing exercise and bisphosphonates in preventing hip and vertebral fractures in older women with osteopenia and/or osteoporosis.
ENDO-C	Compare the effectiveness of primary prevention methods, such as exercise and balance training, versus clinical treatments in preventing falls in older adults at varying degrees of risk.
ENDO-D	Compare the effectiveness of different disease management strategies in improving the adherence to and value of pharmacologic treatments for the elderly.

disease management strategies on the efficiency and value of pharmacological treatments (ENDO-D). There are multiple other topics that affect the elderly population; these topics are listed according to the specific organ system or disease area to which they pertain.

Birth and Developmental Disorders

The uncertainty surrounding the root causes of social-emotional disorders in infants and toddlers, as well as autism spectrum disorder, has resulted in a lack of effective treatment options for these individuals. As a result, AHRQ's Effective Health Care Program recommended this as a national priority area for CER (Whitlock et al., 2009). The final list includes two priority topics focused on identifying effective treatment strategies for these disorders (BDEV-A–B) (Table 5-16). With the remarkable improvement in survival and attendant costs for premature infants, the impact of support programs on child and family outcomes after a child is discharged from a neonatal intensive care unit (NICU) (BDEV-C) was felt to be of significant value. For specific topics related to pregnancy, refer to the Women's Health category.

Complementary and Alternative Medicine

The widespread use of complementary and alternative methodologies (including yoga, meditation, acupuncture, and nutriceuticals [CAM-A–C]) in managing a broad array of disorders (e.g., anxiety and depression, pain, cardiovascular risk factors, chronic diseases, other prevalent conditions)

TABLE 5-16 Birth and Developmental Disorders Priority Topics

BDEV-A	Compare the effectiveness of therapeutic strategies (e.g., behavioral or pharmacologic interventions, the combination of the two) for different autism spectrum disorders (ASD) at different levels of severity and stages of intervention.
BDEV-B	Compare the effectiveness of the co-location model (psychological and primary care practitioners practicing together) and usual care (identification by primary care practitioner and referral to community-based mental health services) in identifying and treating social-emotional and developmental disorders in children ages 0-3.
BDEV-C	Compare the effectiveness of diverse models of comprehensive support services for infants and their families following discharge from a neonatal intensive care unit.

TABLE 5-17 Complementary and Alternative Medicine Priority Topics

CAM-A	Compare the effectiveness of mindfulness-based interventions (e.g., yoga, meditation, deep breathing training) and usual care in treating anxiety and depression, pain, cardiovascular risk factors, and chronic diseases.
CAM-B	Compare the effectiveness of acupuncture for various indications using a cluster randomized trial.
CAM-C	Compare the effectiveness of dietary supplements (nutriceuticals) and usual care in the treatment of selected high-prevalence conditions.

provides the impetus to compare their effectiveness to more conventional approaches to care (Table 5-17).

Nutrition

Obesity is a growing epidemic with medical consequences that extend to multiple chronic conditions, such as diabetes, hypertension, heart disease, and arthritis. Within the medical community, there is currently uncertainty regarding effective strategies for preventing and treating obesity. The committee recommended priorities that compare strategies for improving social conditions to reduce obesity (NUTR-A), including various school policies (NUTR-B) (Table 5-18). Both of these priorities include a focus on populations with varying risk rates. Identifying effective methods for treating obese populations could significantly improve health in this country. As such, the committee recommends comparing the effectiveness of surgical procedures, such as bariatric surgery (gastric bypass), behavior modification, and medication (NUTR-C).

Racial and Ethnic Disparities

Disparities in access to care and in clinical outcomes between different populations were of considerable concern for the committee. Some minority populations, such as African Americans, Asian Pacific Islanders, Latinos, and Native Americans, have higher rates of chronic diseases and also experience greater barriers to obtaining care. Together, these factors contribute to creating disparities in health status and clinical outcomes. The committee recommends comparing the effectiveness of several strategies aimed at reducing these disparities, including community-based and multilevel interventions (RED-A), providing literacy sensitive disease management programs (RED-B), and strategies to improve engagement and retention (RED-C) (Table 5-19).

TABLE 5-18 Nutrition Priority Topics

NUTR-A	Compare the effectiveness of various strategies (e.g., clinical interventions, selected social interventions [such as improving the built environment in communities and making healthy foods more available], combined clinical and social interventions) to prevent obesity, hypertension, diabetes, and heart disease in at-risk populations such as the urban poor and American Indians.
NUTR-B	Compare the effectiveness of school-based interventions involving meal programs, vending machines, and physical education, at different levels of intensity, in preventing and treating overweight and obesity in children and adolescents.
NUTR-C	Compare the effectiveness of treatment strategies for obesity (e.g., bariatric surgery, behavioral interventions, pharmacologic treatment) on the resolution of obesity-related outcomes such as diabetes, hypertension, and musculoskeletal disorders.

Skin Disorders

Skin disorders across the country are widespread, cause a high degree of morbidity, and are among the most costly disorders in children and adolescents between ages 1 and 17 (AHRQ, 2009a,b,c). The committee's priorities on skin disorders include chronic conditions such as lower extremity wounds (common complications in patients with diabetes, peripheral vascular disease, and paralysis) (SKIN-A), and acne—specifically comparing the long-term safety and effectiveness of alternative treatments (SKIN-B) (Table 5-20). Another topic focused on reducing skin disease and comparing treatments to improve quality of life for chronic psoriatic disease (SKIN-C).

TABLE 5-19 Race and Ethnic Disparities Priority Topics

RED-A	Compare the effectiveness of interventions (e.g., community-based multilevel interventions, simple health education, usual care) to reduce health disparities in cardiovascular disease, diabetes, cancer, musculoskeletal diseases, and birth outcomes.
RED-B	Compare the effectiveness of literacy-sensitive disease management programs and usual care in reducing disparities in children and adults with low literacy and chronic disease (e.g., heart disease).
RED-C	Compare the effectiveness of different strategies to engage and retain patients in care and to delineate barriers to care, especially for members of populations that experience health disparities.

TABLE 5-20 Skin Disorders Priority Topics

SKIN-A	Compare the effectiveness of topical treatments (e.g., antibiotics, platelet-derived growth factor) and systemic therapies (e.g., negative pressure wound therapy, hyperbaric oxygen) in managing chronic lower extremity wounds.
SKIN-B	Compare the effectiveness of different long-term treatments for acne.
SKIN-C	Compare the effectiveness (including effects on quality of life) of treatment strategies (e.g., topical steroids, ultraviolet light, methotrexate, biologic response modifiers) for psoriasis.

Alcoholism, Drug Dependency, and Overdose

The harms of tobacco smoking are well known and well documented. Yet, roughly one-fifth of the nation's population continues to smoke. The committee recommended that a national priority for comparative effectiveness should be to examine alternative smoking cessation strategies in understudied populations such as minorities, individuals with mental illness, and adolescents (ADDO-A) (Table 5-21). The Cochrane Collaboration and *Healthy People 2010* also include tobacco use as national priorities (Doyle et al., 2005; HHS, 2000).

The increasing prevalence of abuse of and dependency on pain medications led the committee to recommend an examination of treatment and prescribing practices to reduce substance dependence for patients with non-cancer chronic pain and acute pain (ADDO-B).

Functional Limitations and Disabilities

While many of the committee's priority topics affect patients with disabilities, the following topics specifically address two populations: (1) the

TABLE 5-21 Alcoholism, Drug Dependency, and Overdose Priority Topics

ADDO-A	Compare the effectiveness of smoking cessation strategies (e.g., medication, individual or quitline counseling, combinations of these) in smokers from understudied populations such as minorities, individuals with mental illness, and adolescents.
ADDO-B	Compare the effectiveness of different opioid and non-opioid pain relievers, in different doses and durations, in avoiding unintentional overdose and substance dependence among subjects with acute and non-cancer chronic pain.

TABLE 5-22 Functional Limitations and Disability Priority Topics

DIS-A	Compare the effectiveness of the different treatments (e.g., assistive listening devices, cochlear implants, electric-acoustic devices, habilitation and rehabilitation methods [auditory/oral, sign language, and total communication]) for hearing loss in children and adults, especially individuals with diverse cultural, language, medical, and developmental backgrounds.
DIS-B	Compare the effectiveness of focused intense periodic therapy and usual weekly therapy in managing cerebral palsy in children.

hearing-impaired, and (2) children with cerebral palsy (Table 5-22). The committee recommended one priority focus on treatment strategies for hearing loss among those with diverse cultural/linguistic and medical/developmental backgrounds (DIS-A) and another on usual care compared to focused and intense periodic therapy sessions to manage symptoms related to cerebral palsy (DIS-B).

Eyes, Ears, Nose, and Throat Disorders

The committee included two topics on eye disorders: (1) comparing the effectiveness of alternative treatment strategies for diabetic retinopathy, macular degeneration, and retinal vein occlusion (EENT-A), and (2) comparing strategies for treatment of primary open-angle glaucoma (EENT-B), including a focus on minority populations (Table 5-23).

Kidney and Urinary Tract Disorders

The committee identified prostate cancer and renal replacement therapies as priority areas for comparative effectiveness research (Table 5-24). Because prostate cancer is the second leading cause of cancer death in men

TABLE 5-23 Ears, Eyes, Nose, and Throat Disorders Priority Topics

EENT-A	Compare the effectiveness of different treatment options (e.g., laser therapy, intravitreal steroids, anti-vascular endothelial growth factor [anti-VEGF]) for diabetic retinopathy, macular degeneration, and retinal vein occlusion.
EENT-B	Compare the effectiveness of treatment strategies for primary open-angle glaucoma (e.g., initial laser surgery, new surgical techniques, new medical treatments) particularly in minority populations to assess clinical and patient-reported outcomes.

TABLE 5-24 Kidney and Urinary Tract Disorders Priority Topics

KUT-A	Compare the effectiveness of management strategies for localized prostate cancer (e.g., active surveillance, radical prostatectomy [conventional, robotic, and laparoscopic], and radiotherapy [conformal, brachytherapy, proton-beam, and intensity-modulated radiotherapy]) on survival, recurrence, side effects, quality of life, and costs.
KUT-B	Compare the effectiveness (including survival, hospitalization, quality of life, and costs) of renal replacement therapies (e.g., daily home hemodialysis, intermittent home hemodialysis, conventional in-center dialysis, continuous ambulatory peritoneal dialysis, renal transplantation) for patients of different ages, races, and ethnicities.

(U.S. Cancer Statistics Working Group, 2009), the committee recommended that all aspects of managing the disease be studied (KUT-A).

Renal failure is among the leading causes of mortality across all age groups (Kung et al., 2008). It is also one of the most costly diseases in adults over 65 years of age (AHRQ, 2009a). As such, the committee recommended comparing alternative renal replacement therapies with an emphasis on determining the effectiveness differences among different ages, race, and ethnicities (KUT-B).

Oral Health

The committee recommended two priority topics within oral health for CER, one comparing prevention to surgery in adults with periodontal disease (ORAL-A), and the other in children comparing delivery model approaches for preventing dental caries (cavities) (ORAL-B) (Table 5-25).

Palliative and End-of-Life Care

Effective management and delivery of palliative and end-of-life care is a challenge as the elderly population grows in the United States. Palliative and

TABLE 5-25 Oral Health Priority Topics

ORAL-A	Compare the clinical and cost effectiveness of surgical care and a medical model of prevention and care in managing periodontal disease to increase tooth longevity and reduce systemic secondary effects in other organ systems.
ORAL-B	Compare the effectiveness of the various delivery models (e.g., primary care, dental offices, schools, mobile vans) in preventing dental caries in children.

TABLE 5-26 Palliative and End-of-Life Care Priority Topics

PELC-A	Compare the effectiveness of coordinated care (supported by reimbursement innovations) and usual care in long-term and end-of-life care of the elderly.
PELC-B	Compare the effectiveness of hospital-based palliative care and usual care on patient-reported outcomes and cost.

end-of-life care services must be effective for a variety of populations, and in a variety of environments, including hospitals, long-term care facilities, and homes. The committee specifically recommends research comparing strategies to improve delivery of long-term and end-of-life care, including reimbursement models to support coordinated care (PELC-A) and comparing hospital-based palliative care services with standard care to standard care alone (PELC-B) (Table 5-26).

Gastrointestinal System Disorders

Disorders of the upper gastrointestinal tract, such as gastroesophageal reflux disease (GERD), are among the most prevalent disorders in the nation, and they are particularly prevalent among the elderly (AHRQ, 2009c). They are also among the most costly conditions for infants less than 1 year old (AHRQ, 2009a). The committee specifically recommends the research of the effects of endoscopy on the management and outcomes of patients with GERD as a priority (GI-A) (Table 5-27).

Immune System, Connective Tissue, and Joint Disorders

Conditions of the immune system, connective tissue, and joints such as arthritis and connective tissue disorders are some of the most prevalent and costly diseases in all age groups, especially in the elderly (AHRQ, 2009a,c). Both AHRQ's Effective Health Care Program and *Healthy People 2010* list arthritis and non-traumatic joint disorders as national research priorities (HHS, 2000; Whitlock et al., 2009). The committee recommended comparing the effectiveness of different strategies, including biologics, in the treatment of these diseases (IMUN-A) (Table 5-28).

TABLE 5-27 Gastrointestinal System Disorders Priority Topics

GI-A	Compare the effectiveness of upper endoscopy utilization and frequency for patients with gastroesophageal reflux disease on morbidity, quality of life, and diagnosis of esophageal adenocarcinoma.

TABLE 5-28 Immune System, Connective Tissue, and Joint Disorders Priority Topics

IMUN-A	Compare the effectiveness of different strategies of introducing biologics into the treatment algorithm for inflammatory diseases, including Crohn's disease, ulcerative colitis, rheumatoid arthritis, and psoriatic arthritis.

Pediatrics

There are a variety of alternative and pharmacological treatments available for children with attention deficit hyperactivity disorder (ADHD), but more research is needed to compare their effectiveness. In fact, AHRQ's Effective Health Care Program lists ADHD as a national priority (Whitlock et al., 2009). The committee recommended more research that addresses the comparative effectiveness of these treatments in decreasing the symptoms of ADHD in children (PEDS-A) (Table 5-29). There are a number of other important pediatric topics that are discussed under the research area categories eyes, ears, nose and throat; functional limitations and disabilities; birth and developmental disorders; nutrition; and respiratory disease.

Respiratory Disease

Chronic Obstructive Pulmonary Disease (COPD) and asthma are among the most prevalent, most costly, and morbid conditions for all age groups (AHRQ, 2009a,c; Kung et al., 2008). Asthma is especially common in children and is the leading condition in terms of cost (AHRQ, 2009a). In addition, AHRQ's Effective Health Care Program lists asthma as a priority research area (Whitlock et al., 2009). The committee recommended alternative strategies for managing asthma be studied through CER (RD-A) (Table 5-30).

Trauma, Emergency Medicine, and Critical Care Medicine

Accidents are a leading cause of death for all ages in the United States, and trauma-related disorders are listed as one of the most prevalent and

TABLE 5-29 Pediatric Disorders Priority Topics

PEDS-A	Compare the effectiveness of various primary care treatment strategies (e.g., symptom management, cognitive behavior therapy, biofeedback, social skills, educator/teacher training, parent training, pharmacologic treatment) for attention deficit hyperactivity disorder (ADHD) in children.

TABLE 5-30 Respiratory Disorders Priority Topics

RD-A	Compare the effectiveness of an integrated approach (combining counseling, environmental mitigation, chronic disease management, and legal assistance) with a non-integrated episodic care model in managing asthma in children.

TABLE 5-31 Trauma, Emergency Medicine, and Critical Care Medicine Priority Topics

TEMC-A	Compare the effectiveness of treatment strategies (e.g., cognitive behavioral individual therapy, generic individual therapy, comprehensive and intensive treatment) for Post-traumatic Stress Disorder stemming from diverse sources of trauma.

costly (AHRQ, 2009a,c). While there are many disorders that arise from trauma and emergencies, the committee focused on the treatment of post-traumatic stress disorder (PTSD) in all populations and from all sources of trauma. With the large number of veterans returning from the wars in Iraq and Afghanistan, and increased recognition of the inadequacies of the nation's health system to effectively treat patients with mental health conditions, it is important to identify effective treatment strategies. The committee recommended that PTSD be studied as part of a balanced portfolio of CER (TEMC-A) (Table 5-31).

TIMELINESS AND LIMITATIONS OF THE COMMITTEE'S PRIORITY LIST

The committee believes that the priority list presented in this chapter is relevant to the needs and conditions of today. New questions in CER will continue to appear. However, the balance of topics across the portfolio, the correlation with established priorities by other groups, and the good fit between the topics and the pre-established, pre-specified criteria suggest that the process used by the committee was effective. As discussed in Chapters 4 and 6, this process requires modification if it is to be continued in the future.

Recognizing the dynamic nature of disease and the rapid technologic and therapeutic advancements in health care, these priorities may very well be answered by ongoing research or be superseded by new disorders in the near future. In fact, that is the committee's hope and expectation. Recognition of priorities and initiation of new research should provide answers to the clinical dilemmas identified. Therefore, an ongoing and active process

of priority setting using stakeholder input is imperative. The previous two chapters described systems for continuous stakeholder input, together with methodologies for identifying which of these topics deserve priority. However, the committee emphasizes the importance of repeating this exercise on a regular basis or of integrating aspects of the process described here into the routine determination of CER funding in order to sustain the effort to discover what works best and for whom.

REFERENCES

AHRQ (Agency for Healthcare Research and Quality). 2009a. *Medical Expenditure Panel Survey. Total Expenses for Conditions by Site of Service: United States*. http://www.meps.ahrq.gov/mepsweb/ (accessed March 10, 2009).

———. 2009b. Total number of events accounting for expenditures by site of service: United States, 2006. In *Medical Expenditure Panel Survey Component Data*.

———. 2009c. Total number of people accounting for expenditures (deduplicated) by site of service: United States, 2006. In *Medical Expenditure Panel Survey Component Data*.

AHRQ Effective Health Care Program. 2009. *Effective Health Care: Topic triage cover sheets*.

Doyle, J., E. Waters, D. Yach, D. McQueen, A. De Francisco, T. Stewart, P. Reddy, A. M. Gulmezoglu, G. Galea, and A. Portela. 2005. Global priority setting for Cochrane systematic reviews of health promotion and public health research. *Journal of Epidemiology and Community Health* 59:193-197.

HHS (Department of Health and Human Services). 2000. *Healthy People 2010: Understanding and Improving Health*. U.S. Government Printing Office. http://purl.access.gpo.gov/GPO/LPS4217 (accessed April 3, 2009).

Kung, H.-C., D. L. Hoyert, J. Xu, S. L. Murphy, and Division of Vital Statistics. 2008. *Deaths: Final Data for 2005. National Vital Statistics Reports*. National Center for Health Statistics.

NPP (National Priorities Partnership). 2008. *National priorities and goals*. Washington, DC: National Quality Forum.

U.S. Cancer Statistics Working Group. 2009. United States cancer statistics: 1999–2005 Incidence and mortality web-based report. Atlanta: Department of Health and Human Services, Centers for Disease Control and Prevention and National Cancer Institute.

Wennberg, J. E. 2009 (unpublished). *Recommendations to the Institute of Medicine on comparative effectiveness research priorities*. Submitted in response to a request from the Institute of Medicine Committee on Comparative Effectiveness Research Prioritization. The Dartmouth Institute for Health Policy and Clinical Practice.

Whitlock, E. P., S. A. Lopez, S. Chang, M. Helfand, M. Eder, and N. Floyd. 2009. *Identifying, selecting, and refining topics for comparative effectiveness systematic reviews: AHRQ and the Effective Health Care Program*. http://effectivehealthcare.ahrq.gov/repFiles/20090427IdenttifyingTopics.pdf (accessed June 5, 2009).

Appendix D

Comparative Effectiveness Research Priorities: *FCCCER Recommendations (2009)*

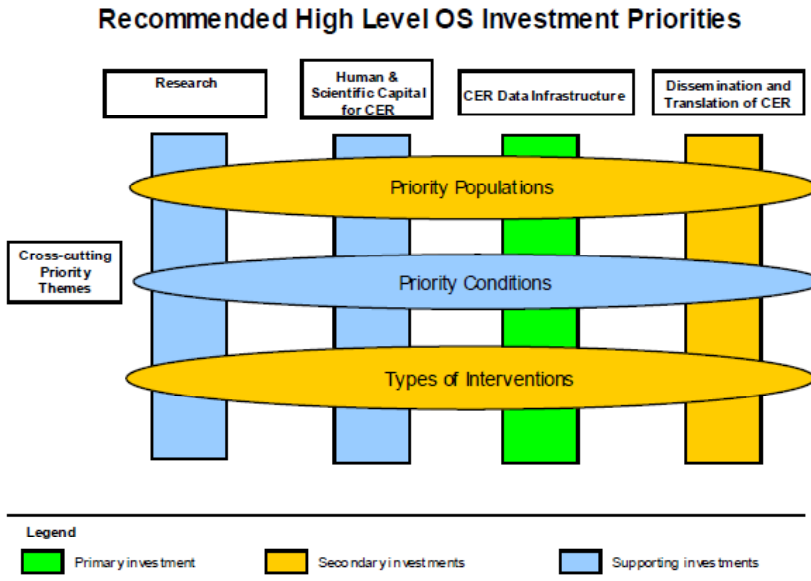
PRIORITY RECOMMENDATIONS FOR OFFICE OF SECRETARY CER FUNDS[1]

Using the strategic framework for CER discussed in Section V, and taking into consideration the unique role that OS funds can play in addressing high priority gaps, the Council developed a recommended high-level investment strategy for the use of the OS ARRA funds. The strategy has three different levels of priority recommendations for OS fund investments in the Core Activities and Cross-cutting Priority Themes in the CER framework (Figure 3).

- Primary investment. This area of investment should represent a large portion of the OS funds. It best fulfills the full range of prioritization criteria and requires scaled investment in order to be successful. The Council recommends that CER Data Infrastructure be the primary investment.
- Secondary investments. These areas should also receive significant investment. They are as critical to success in CER as the primary focus, but individually may require a smaller amount of funding to be successful. The Council recommends that Dissemination and

[1] Passage from the Federal Coordinating Council for Comparative Effectiveness Research *Report to the President and Congress, June 30, 2009*. Department of Health and Human Services.

Figure 3



Translation of CER, Priority Populations, and Priority Types of Intervention be secondary investments.

- Supporting investments. These areas should not be the major focus of OS funding as they do not fulfill the prioritization criteria as well as primary and secondary investments, but some funding may be necessary to support and enable investments in higher priority areas and fill identified gaps. The Council recommends that Human and Scientific Capital, Research, and Conditions receive supporting investments. It is important to note that these recommendations pertain only to OS funds; AHRQ, NIH, and VA have a history of significant investments in Research, Human and Scientific Capital, and Conditions.

The Council believes that this strategy and distribution of investments will best position the Secretary to:

- Respond to patient and physician demand for CER.
- Balance achieving near-term results with building longer-term opportunities.
- Capture the distinctive value of the Secretary's ARRA funds.

While it is the responsibility of the Office of the Secretary to operationalize this strategy, the Council's rationale for these recommendations is designed to help guide the Secretary in making specific investment decisions. The Council based its rationale for each level of investment in the strategy on the prioritization criteria described above, as well as representative examples of investment in each area proposed through the public comment process and by federal agencies.

Primary Investment

CER data infrastructure development is the most distinctive opportunity for OS ARRA funding. It requires a large, up-front infusion of capital to be successful that is unlikely to come from any source other than OS ARRA funds, making it ideal for this funding mechanism. It has broad potential impact, with the ability for resulting research to address conditions and populations captured in the primary data. Given the absence of comprehensive databases and data evaluation tools, there is significant demand from the patient, clinical, and public health communities for new, expanded data infrastructure and data access to support decision making. Finally, investments in data infrastructure have the potential to generate significant additional investments in two ways. First, some of these investments could take the form of public-private partnerships. Second, data infrastructure is a tool that, once developed, will result in new research conducted and/or funded by entities such as biomedical research organizations, payers, foundations, and health care providers.

The Council received proposals on a number of potentially promising initiatives related to data infrastructure, including but not limited to:

- Building, expanding, and linking longitudinal administrative claims databases.
- Linking administrative data with EHR-based or registry data.
- Expanding high-impact patient registries (e.g., collaborations with specialty organizations, SEER).
- Distributed data networks populated by EHRs in practice and provider settings.
- Expanding analysis of FDA and private sector data on drug and device trials and safety.

As the Office of the Secretary identifies specific opportunities in data infrastructure, the Council recommends that it consider most carefully those that:

- Expand access to existing resources, especially those currently managed by Federal agencies.
- Create scaled platforms by leveraging existing data and capabilities in the private sector.
- Capitalize on linkages between health IT investments and the potential for CERinfrastructure to develop evidence to inform decision making.
- Ensure that infrastructure is responsive to needs of patients, providers, and other decision makers—and not driven by what is most feasible.

The Council appreciates the relationship and need for coordination between CER and health IT (e.g., through a distributed network of EHRs) investments. As the Secretary develops HHS's full portfolio of ARRA investments, it will be critical to consider both CER and health IT holistically, not as policy silos, recognizing that success in CER is largely dependent on success in health IT and vice versa. With all data infrastructure investments, the government will need to ensure data security and privacy. Protecting security and privacy is key to maintaining the public's trust.

Secondary Investments

Secondary investments include a core area of investment—Dissemination and Translation of CER—and two cross-cutting themes—Priority Populations and Types of Intervention.

Dissemination, translation and adoption of CER is about realizing the benefits that comparative effectiveness research has to offer both patients and providers. While the breadth and depth of the near-term impact depends on what types of pilot programs the OS supports, the lessons and tools for translation developed by those pilots will be relevant to all.

The lack of reliable success in disseminating findings from CER in ways that translate into better health outcomes highlights the uncertainty and difficulty of this enterprise. However, dissemination and translation is essential to improving outcomes for patients and the link between evidence production and how best to get this information to physicians and patients in a way they understand is critical to capitalizing on the CER investment. Despite important efforts by the federal government, especially AHRQ, NIH, VA and DoD, the majority of current funding goes to building evidence as opposed to ensuring that the existing evidence base is utilized in

patient care and health systems management. This creates a unique role for OS ARRA funding. Investments in dissemination and translation programs also have the potential to generate additional investments, especially from providers, if private institutions elect to implement similar efforts or partner with the Federal Government on translation efforts.

There are a wide range of potential dissemination, translation and adoption programs that the OS could support, including:

- Investing in dissemination and translation of CER findings throughout the Federal delivery system.
- Dissemination and translation through partnerships with provider and/or patient organizations.
- Decision support and shared decision-making tools to provide information to clinicians and patients at the point of care.
- Developing standards for communication tools for patients and providers, (e.g., a patient-friendly simple scoring system).
- Partnering with an existing consumer media channel (e.g., Internet search engine or health information site) to expand patient access to existing CER data.
- Creating a National Patient Library with a primary focus on providing evidence to patients in easy-to-use and understandable formats.

The Council recommends that the Office of the Secretary consider the following in making investments in dissemination and translation:

- Investing in better understanding the most effective methods to disseminate and translate research findings to improve patient outcomes.
- Identifying opportunities both to develop tools for translation and to pilot implementation of these tools.
- Partnering with provider organizations in federal agencies, as well as in states and the private sector.
- Accounting for potential surrogate decision makers (e.g., families) and the context for decisions in patient-focused tools.
- Ensuring that programs address a specific need articulated by the implementing organization or the partner to ensure success and the sustainability ofdissemination activities.
- Focusing on developing standards for communication.
- Increasing understanding of the most effective methods to disseminate findings to clinicians and patients to inform decision making.

From an operational perspective, investments in the cross-cutting themes are somewhat distinct from investments in the core areas. Whereas funding for a core area might go to a project or organization focused on a specific activity, funding for a cross-cutting theme requires multiple coordinated investments and activities to be successful. Investments in these themes could cover some or all of the four core activities: research, data infrastructure, human and scientific capital, and dissemination and translation. These investments could involve a coordinated investment across HHS or the federal government, or they could be focused in academic centers, integrated delivery system organizations, private industry, or other nongovernmental entities. Collaborative efforts to inform and transform care will be essential to achieving meaningful impact across these cross-cutting themes.

Investments in specific populations, meanwhile, will help ensure that the benefits of CER are available to all. It can also focus CER efforts on populations with existing health disparities and worse outcomes. CER has the potential in some populations, such as racial and ethnic minorities, to fill critical gaps that, historically, efficacy research has left unaddressed.

The Council identified several populations for whom the Secretary should consider allocating CER funds:

- Racial and ethnic minorities
- Persons with disabilities
- Elderly
- Children
- Patients with multiple chronic conditions

Investment in specific types of interventions in a cross-cutting manner also presents a unique opportunity for the nation's health system. The Council has identified six specific interventions for the Secretary to consider that address large and varied populations, resulting in high potential impact, are areas of high clinical uncertainty, and are not being adequately addressed by other entities. They are:

- Medical and assistive devices (e.g., comparing rehabilitative devices).
- Procedures and surgery (e.g., evaluating surgical options or surgery versus medical management).
- Diagnostic testing (e.g., comparing imaging modalities for evaluating certain types of cancer).
- Behavioral change (e.g., developing and assessing smoking cessation programs).

- Delivery system strategies (e.g., testing two different discharge process care models on readmission rates or testing two different medical home models on preventing hospital admissions and improving quality of life).
- Prevention (e.g., comparing two interventions to prevent or decrease obesity, comparing strategies for reaching populations that do not access the healthcare system with prevention efforts).

Furthermore, the Council recommends that the Office of the Secretary consider the following in making investments in the cross-cutting themes of priority populations and types of interventions:

- Focusing on immediate, specific patient needs that can generate results.
- Concentrating on areas with cross-cutting gaps in research, data infrastructure, scientific capital, and/or translation.
- Building on promising systems and practices already in place, both within the government and in the private sector, and measuring results when scaled up and disseminated.
- Strongly encouraging coordination across the government and with entities outside of the government.

Supporting Investments

The Council recommends that the OS reserve some ARRA funding for Research, Human & Scientific Capital, and the Conditions cross-cutting theme. Because these investments and topics are the major foci of CER activities at NIH and AHRQ, both of which will likely utilize ARRA funds administered by those organizations for these purposes, they do not represent distinctive investment for OS funds. However, there will likely be targeted investments in these areas that could support other OS ARRA efforts, such as training new researchers in CER methods or addressing gaps not addressed elsewhere in the federal government.

In making these targeted investments, the Council recommends the Office of the Secretary consider:

- Focusing on areas that maximize the value of the Secretary's investments in other areas.
- Avoiding duplication of efforts with other agencies.

For all of the above investments, the Council recommends that the Office of the Secretary consider the portfolio of investments and where synergies exist to leverage one investment into multiple areas. For example,

a data infrastructure investment that can also be used for a cross-cutting priority theme would be of higher value than an investment that has more limited applications. Doing so will help to ensure that the funds allocated to the Office of the Secretary for CER will have a significant positive impact on the quality of patient care in the near term, and lay the foundations for continued improvements going forward.

Appendix E

Affordable Care Act (ACA) (2010) Provisions for the Patient-Centered Outcomes Research Institute (PCORI)

Subtitle D—Patient-Centered Outcomes Research

SEC. 6301. PATIENT-CENTERED OUTCOMES RESEARCH.

(a) IN GENERAL.—Title XI of the Social Security Act (42 U.S.C. 1301 et seq.) is amended by adding at the end the following new part:

"PART D—COMPARATIVE CLINICAL EFFECTIVENESS RESEARCH

"COMPARATIVE CLINICAL EFFECTIVENESS RESEARCH

"SEC. 1181. (a) DEFINITIONS.—In this section: 42 USC 1320e.

"(1) BOARD.—The term 'Board' means the Board of Governors established under subsection (f).

"(2) COMPARATIVE CLINICAL EFFECTIVENESS RESEARCH; RESEARCH.—

"(A) IN GENERAL.—The terms 'comparative clinical effectiveness research' and 'research' mean research evaluating and comparing health outcomes and the clinical effectiveness, risks, and benefits of 2 or more medical treatments, services, and items described in subparagraph (B).

"(B) MEDICAL TREATMENTS, SERVICES, AND ITEMS DESCRIBED.—The medical treatments, services, and items described in this subparagraph are health care interventions, protocols for treatment, care management, and delivery, procedures, medical devices, diagnostic tools, pharmaceuticals (including drugs and biologicals), integrative health practices, and any other strategies or items being used in the treatment, management, and diagnosis of, or prevention of illness or injury in, individuals.

"(3) CONFLICT OF INTEREST.—The term 'conflict of interest' means an association, including a financial or personal association, that have the potential to bias or have the appearance

of biasing an individual's decisions in matters related to the Institute or the conduct of activities under this section.

"(4) REAL CONFLICT OF INTEREST.—The term 'real conflict of interest' means any instance where a member of the Board, the methodology committee established under subsection (d)(6), or an advisory panel appointed under subsection (d)(4), or a close relative of such member, has received or could receive either of the following:

"(A) A direct financial benefit of any amount deriving from the result or findings of a study conducted under this section.

"(B) A financial benefit from individuals or companies that own or manufacture medical treatments, services, or items to be studied under this section that in the aggregate exceeds $10,000 per year. For purposes of the preceding sentence, a financial benefit includes honoraria, fees, stock, or other financial benefit and the current value of the member or close relative's already existing stock holdings, in addition to any direct financial benefit deriving from the results or findings of a study conducted under this section.

"(b) PATIENT-CENTERED OUTCOMES RESEARCH INSTITUTE.—

"(1) ESTABLISHMENT.—There is authorized to be established a nonprofit corporation, to be known as the 'Patient-Centered Outcomes Research Institute' (referred to in this section as the 'Institute') which is neither an agency nor establishment of the United States Government.

District of Columbia.

"(2) APPLICATION OF PROVISIONS.—The Institute shall be subject to the provisions of this section, and, to the extent consistent with this section, to the District of Columbia Nonprofit Corporation Act.

"(3) FUNDING OF COMPARATIVE CLINICAL EFFECTIVENESS RESEARCH.—For fiscal year 2010 and each subsequent fiscal year, amounts in the Patient-Centered Outcomes Research Trust Fund (referred to in this section as the 'PCORTF') under section 9511 of the Internal Revenue Code of 1986 shall be available, without further appropriation, to the Institute to carry out this section.

"(c) PURPOSE.—The purpose of the Institute is to assist patients, clinicians, purchasers, and policy-makers in making informed health decisions by advancing the quality and relevance of evidence concerning the manner in which diseases, disorders, and other health conditions can effectively and appropriately be prevented, diagnosed, treated, monitored, and managed through research and evidence synthesis that considers variations in patient subpopulations, and the dissemination of research findings with respect to the relative health outcomes, clinical effectiveness, and appropriateness of the medical treatments, services, and items described in subsection (a)(2)(B).

"(d) DUTIES.—

"(1) IDENTIFYING RESEARCH PRIORITIES AND ESTABLISHING RESEARCH PROJECT AGENDA.—

"(A) IDENTIFYING RESEARCH PRIORITIES.—The Institute shall identify national priorities for research, taking into account factors of disease incidence, prevalence, and burden in the United States (with emphasis on chronic conditions), gaps in evidence in terms of clinical outcomes, practice

variations and health disparities in terms of delivery and outcomes of care, the potential for new evidence to improve patient health, well-being, and the quality of care, the effect on national expenditures associated with a health care treatment, strategy, or health conditions, as well as patient needs, outcomes, and preferences, the relevance to patients and clinicians in making informed health decisions, and priorities in the National Strategy for quality care established under section 399H of the Public Health Service Act that are consistent with this section.

"(B) ESTABLISHING RESEARCH PROJECT AGENDA.—The Institute shall establish and update a research project agenda for research to address the priorities identified under subparagraph (A), taking into consideration the types of research that might address each priority and the relative value (determined based on the cost of conducting research compared to the potential usefulness of the information produced by research) associated with the different types of research, and such other factors as the Institute determines appropriate.

"(2) CARRYING OUT RESEARCH PROJECT AGENDA.—

"(A) RESEARCH.—The Institute shall carry out the research project agenda established under paragraph (1)(B) in accordance with the methodological standards adopted under paragraph (9) using methods, including the following:

"(i) Systematic reviews and assessments of existing and future research and evidence including original research conducted subsequent to the date of the enactment of this section.

"(ii) Primary research, such as randomized clinical trials, molecularly informed trials, and observational studies.

"(iii) Any other methodologies recommended by the methodology committee established under paragraph (6) that are adopted by the Board under paragraph (9).

"(B) CONTRACTS FOR THE MANAGEMENT OF FUNDING AND CONDUCT OF RESEARCH.—

"(i) CONTRACTS.—

"(I) IN GENERAL.—In accordance with the research project agenda established under paragraph (1)(B), the Institute shall enter into contracts for the management of funding and conduct of research in accordance with the following:

"(aa) Appropriate agencies and instrumentalities of the Federal Government.

"(bb) Appropriate academic research, private sector research, or study-conducting entities.

"(II) PREFERENCE.—In entering into contracts under subclause (I), the Institute shall give preference to the Agency for Healthcare Research and Quality and the National Institutes of Health, but only if the research to be conducted or managed under such contract is authorized by the governing statutes of such Agency or Institutes.

"(ii) CONDITIONS FOR CONTRACTS.—A contract entered into under this subparagraph shall require that the agency, instrumentality, or other entity—

"(I) abide by the transparency and conflicts of interest requirements under subsection (h) that apply to the Institute with respect to the research managed or conducted under such contract;

"(II) comply with the methodological standards adopted under paragraph (9) with respect to such research;

"(III) consult with the expert advisory panels for clinical trials and rare disease appointed under clauses (ii) and (iii), respectively, of paragraph (4)(A);

"(IV) subject to clause (iv), permit a researcher who conducts original research under the contract for the agency, instrumentality, or other entity to have such research published in a peer-reviewed journal or other publication;

"(V) have appropriate processes in place to manage data privacy and meet ethical standards for the research;

"(VI) comply with the requirements of the Institute for making the information available to the public under paragraph (8); and

"(VII) comply with other terms and conditions determined necessary by the Institute to carry out the research agenda adopted under paragraph (2).

"(iii) COVERAGE OF COPAYMENTS OR COINSURANCE.—A contract entered into under this subparagraph may allow for the coverage of copayments or coinsurance, or allow for other appropriate measures, to the extent that such coverage or other measures are necessary to preserve the validity of a research project, such as in the case where the research project must be blinded.

"(iv) REQUIREMENTS FOR PUBLICATION OF RESEARCH.—Any research published under clause (ii)(IV) shall be within the bounds of and entirely consistent with the evidence and findings produced under the contract with the Institute under this subparagraph. If the Institute determines that those requirements are not met, the Institute shall not enter into another contract with the agency, instrumentality, or entity which managed or conducted such research for a period determined appropriate by the Institute (but not less than 5 years).

"(C) REVIEW AND UPDATE OF EVIDENCE.—The Institute shall review and update evidence on a periodic basis as appropriate.

"(D) TAKING INTO ACCOUNT POTENTIAL DIFFERENCES.—Research shall be designed, as appropriate, to take into account the potential for differences in the effectiveness of health care treatments, services, and items as used with various subpopulations, such as racial and ethnic minorities, women, age, and groups of individuals with different comorbidities, genetic and molecular sub-types,

or quality of life preferences and include members of such subpopulations as subjects in the research as feasible and appropriate.

"(E) DIFFERENCES IN TREATMENT MODALITIES.—Research shall be designed, as appropriate, to take into account different characteristics of treatment modalities that may affect research outcomes, such as the phase of the treatment modality in the innovation cycle and the impact of the skill of the operator of the treatment modality.

"(3) DATA COLLECTION.—

"(A) IN GENERAL.—The Secretary shall, with appropriate safeguards for privacy, make available to the Institute such data collected by the Centers for Medicare & Medicaid Services under the programs under titles XVIII, XIX, and XXI, as well as provide access to the data networks developed under section 937(f) of the Public Health Service Act, as the Institute and its contractors may require to carry out this section. The Institute may also request and obtain data from Federal, State, or private entities, including data from clinical databases and registries.

"(B) USE OF DATA.—The Institute shall only use data provided to the Institute under subparagraph (A) in accordance with laws and regulations governing the release and use of such data, including applicable confidentiality and privacy standards.

"(4) APPOINTING EXPERT ADVISORY PANELS.—

"(A) APPOINTMENT.—

"(i) IN GENERAL.—The Institute may appoint permanent or ad hoc expert advisory panels as determined appropriate to assist in identifying research priorities and establishing the research project agenda under paragraph (1) and for other purposes.

"(ii) EXPERT ADVISORY PANELS FOR CLINICAL TRIALS.—The Institute shall appoint expert advisory panels in carrying out randomized clinical trials under the research project agenda under paragraph (2)(A)(ii). Such expert advisory panels shall advise the Institute and the agency, instrumentality, or entity conducting the research on the research question involved and the research design or protocol, including important patient subgroups and other parameters of the research. Such panels shall be available as a resource for technical questions that may arise during the conduct of such research.

"(iii) EXPERT ADVISORY PANEL FOR RARE DISEASE.—In the case of a research study for rare disease, the Institute shall appoint an expert advisory panel for purposes of assisting in the design of the research study and determining the relative value and feasibility of conducting the research study.

"(B) COMPOSITION.—An expert advisory panel appointed under subparagraph (A) shall include representatives of practicing and research clinicians, patients, and experts in scientific and health services research, health services delivery, and evidence-based medicine who have experience in the relevant topic, and as appropriate, experts

in integrative health and primary prevention strategies. The Institute may include a technical expert of each manufacturer or each medical technology that is included under the relevant topic, project, or category for which the panel is established.

"(5) SUPPORTING PATIENT AND CONSUMER REPRESENTATIVES.—The Institute shall provide support and resources to help patient and consumer representatives effectively participate on the Board and expert advisory panels appointed by the Institute under paragraph (4).

"(6) ESTABLISHING METHODOLOGY COMMITTEE.—

"(A) IN GENERAL.—The Institute shall establish a standing methodology committee to carry out the functions described in subparagraph (C).

"(B) APPOINTMENT AND COMPOSITION.—The methodology committee established under subparagraph (A) shall be composed of not more than 15 members appointed by the Comptroller General of the United States. Members appointed to the methodology committee shall be experts in their scientific field, such as health services research, clinical research, comparative clinical effectiveness research, biostatistics, genomics, and research methodologies. Stakeholders with such expertise may be appointed to the methodology committee. In addition to the members appointed under the first sentence, the Directors of the National Institutes of Health and the Agency for Healthcare Research and Quality (or their designees) shall each be included as members of the methodology committee.

Deadline.

"(C) FUNCTIONS.—Subject to subparagraph (D), the methodology committee shall work to develop and improve the science and methods of comparative clinical effectiveness research by, not later than 18 months after the establishment of the Institute, directly or through subcontract, developing and periodically updating the following:

"(i) Methodological standards for research. Such methodological standards shall provide specific criteria for internal validity, generalizability, feasibility, and timeliness of research and for health outcomes measures, risk adjustment, and other relevant aspects of research and assessment with respect to the design of research. Any methodological standards developed and updated under this subclause shall be scientifically based and include methods by which new information, data, or advances in technology are considered and incorporated into ongoing research projects by the Institute, as appropriate. The process for developing and updating such standards shall include input from relevant experts, stakeholders, and decisionmakers, and shall provide opportunities for public comment. Such standards shall also include methods by which patient subpopulations can be accounted for and evaluated in different types of research. As appropriate, such standards shall build on existing work on methodological standards for defined categories of health interventions and for each of the major categories of

comparative clinical effectiveness research methods (determined as of the date of enactment of the Patient Protection and Affordable Care Act).

"(ii) A translation table that is designed to provide guidance and act as a reference for the Board to determine research methods that are most likely to address each specific research question.

"(D) CONSULTATION AND CONDUCT OF EXAMINATIONS.—The methodology committee may consult and contract with the Institute of Medicine of the National Academies and academic, nonprofit, or other private and governmental entities with relevant expertise to carry out activities described in subparagraph (C) and may consult with relevant stakeholders to carry out such activities.

"(E) REPORTS.—The methodology committee shall submit reports to the Board on the committee's performance of the functions described in subparagraph (C). Reports shall contain recommendations for the Institute to adopt methodological standards developed and updated by the methodology committee as well as other actions deemed necessary to comply with such methodological standards. Recommendations.

"(7) PROVIDING FOR A PEER-REVIEW PROCESS FOR PRIMARY RESEARCH.—

"(A) IN GENERAL.—The Institute shall ensure that there is a process for peer review of primary research described in subparagraph (A)(ii) of paragraph (2) that is conducted under such paragraph. Under such process—

"(i) evidence from such primary research shall be reviewed to assess scientific integrity and adherence to methodological standards adopted under paragraph (9); and

"(ii) a list of the names of individuals contributing to any peer-review process during the preceding year or years shall be made public and included in annual reports in accordance with paragraph (10)(D). Lists. Public information.

"(B) COMPOSITION.—Such peer-review process shall be designed in a manner so as to avoid bias and conflicts of interest on the part of the reviewers and shall be composed of experts in the scientific field relevant to the research under review.

"(C) USE OF EXISTING PROCESSES.—

"(i) PROCESSES OF ANOTHER ENTITY.—In the case where the Institute enters into a contract or other agreement with another entity for the conduct or management of research under this section, the Institute may utilize the peer-review process of such entity if such process meets the requirements under subparagraphs (A) and (B).

"(ii) PROCESSES OF APPROPRIATE MEDICAL JOURNALS.—The Institute may utilize the peer-review process of appropriate medical journals if such process meets the requirements under subparagraphs (A) and (B).

"(8) RELEASE OF RESEARCH FINDINGS.—

"(A) IN GENERAL.—The Institute shall, not later than 90 days after the conduct or receipt of research findings under this part, make such research findings available Deadline.

to clinicians, patients, and the general public. The Institute shall ensure that the research findings—

"(i) convey the findings of research in a manner that is comprehensible and useful to patients and providers in making health care decisions;

"(ii) fully convey findings and discuss considerations specific to certain subpopulations, risk factors, and comorbidities, as appropriate;

"(iii) include limitations of the research and what further research may be needed as appropriate;

"(iv) not be construed as mandates for practice guidelines, coverage recommendations, payment, or policy recommendations; and

"(v) not include any data which would violate the privacy of research participants or any confidentiality agreements made with respect to the use of data under this section.

"(B) DEFINITION OF RESEARCH FINDINGS.—In this paragraph, the term 'research findings' means the results of a study or assessment.

"(9) ADOPTION.—Subject to subsection (h)(1), the Institute shall adopt the national priorities identified under paragraph (1)(A), the research project agenda established under paragraph (1)(B), the methodological standards developed and updated by the methodology committee under paragraph (6)(C)(i), and any peer-review process provided under paragraph (7) by majority vote. In the case where the Institute does not adopt such processes in accordance with the preceding sentence, the processes shall be referred to the appropriate staff or entity within the Institute (or, in the case of the methodological standards, the methodology committee) for further review.

Public information.

"(10) ANNUAL REPORTS.—The Institute shall submit an annual report to Congress and the President, and shall make the annual report available to the public. Such report shall contain—

"(A) a description of the activities conducted under this section, research priorities identified under paragraph (1)(A) and methodological standards developed and updated by the methodology committee under paragraph (6)(C)(i) that are adopted under paragraph (9) during the preceding year;

"(B) the research project agenda and budget of the Institute for the following year;

"(C) any administrative activities conducted by the Institute during the preceding year;

"(D) the names of individuals contributing to any peer-review process under paragraph (7), without identifying them with a particular research project; and

"(E) any other relevant information (including information on the membership of the Board, expert advisory panels, methodology committee, and the executive staff of the Institute, any conflicts of interest with respect to these individuals, and any bylaws adopted by the Board during the preceding year).

"(e) ADMINISTRATION.—

"(1) IN GENERAL.—Subject to paragraph (2), the Board shall carry out the duties of the Institute.

"(2) NONDELEGABLE DUTIES.—The activities described in subsections (d)(1) and (d)(9) are nondelegable.

"(f) BOARD OF GOVERNORS.—

"(1) IN GENERAL.—The Institute shall have a Board of Governors, which shall consist of the following members: Establishment.

"(A) The Director of Agency for Healthcare Research and Quality (or the Director's designee).

"(B) The Director of the National Institutes of Health (or the Director's designee).

"(C) Seventeen members appointed, not later than 6 months after the date of enactment of this section, by the Comptroller General of the United States as follows: Deadline.

"(i) 3 members representing patients and health care consumers.

"(ii) 5 members representing physicians and providers, including at least 1 surgeon, nurse, State-licensed integrative health care practitioner, and representative of a hospital.

"(iii) 3 members representing private payers, of whom at least 1 member shall represent health insurance issuers and at least 1 member shall represent employers who self-insure employee benefits.

"(iv) 3 members representing pharmaceutical, device, and diagnostic manufacturers or developers.

"(v) 1 member representing quality improvement or independent health service researchers.

"(vi) 2 members representing the Federal Government or the States, including at least 1 member representing a Federal health program or agency.

"(2) QUALIFICATIONS.—The Board shall represent a broad range of perspectives and collectively have scientific expertise in clinical health sciences research, including epidemiology, decisions sciences, health economics, and statistics. In appointing the Board, the Comptroller General of the United States shall consider and disclose any conflicts of interest in accordance with subsection (h)(4)(B). Members of the Board shall be recused from relevant Institute activities in the case where the member (or an immediate family member of such member) has a real conflict of interest directly related to the research project or the matter that could affect or be affected by such participation.

"(3) TERMS; VACANCIES.—A member of the Board shall be appointed for a term of 6 years, except with respect to the members first appointed, whose terms of appointment shall be staggered evenly over 2-year increments. No individual shall be appointed to the Board for more than 2 terms. Vacancies shall be filled in the same manner as the original appointment was made.

"(4) CHAIRPERSON AND VICE-CHAIRPERSON.—The Comptroller General of the United States shall designate a Chairperson and Vice Chairperson of the Board from among the members of the Board. Such members shall serve as Chairperson or Vice Chairperson for a period of 3 years. Designation.

"(5) COMPENSATION.—Each member of the Board who is not an officer or employee of the Federal Government shall be entitled to compensation (equivalent to the rate provided for level IV of the Executive Schedule under section 5315 of

title 5, United States Code) and expenses incurred while performing the duties of the Board. An officer or employee of the Federal government who is a member of the Board shall be exempt from compensation.

"(6) DIRECTOR AND STAFF; EXPERTS AND CONSULTANTS.—The Board may employ and fix the compensation of an Executive Director and such other personnel as may be necessary to carry out the duties of the Institute and may seek such assistance and support of, or contract with, experts and consultants that may be necessary for the performance of the duties of the Institute.

"(7) MEETINGS AND HEARINGS.—The Board shall meet and hold hearings at the call of the Chairperson or a majority of its members. Meetings not solely concerning matters of personnel shall be advertised at least 7 days in advance and open to the public. A majority of the Board members shall constitute a quorum, but a lesser number of members may meet and hold hearings.

"(g) FINANCIAL AND GOVERNMENTAL OVERSIGHT.—

"(1) CONTRACT FOR AUDIT.—The Institute shall provide for the conduct of financial audits of the Institute on an annual basis by a private entity with expertise in conducting financial audits.

"(2) REVIEW AND ANNUAL REPORTS.—

"(A) REVIEW.—The Comptroller General of the United States shall review the following:

"(i) Not less frequently than on an annual basis, the financial audits conducted under paragraph (1).

"(ii) Not less frequently than every 5 years, the processes established by the Institute, including the research priorities and the conduct of research projects, in order to determine whether information produced by such research projects is objective and credible, is produced in a manner consistent with the requirements under this section, and is developed through a transparent process.

"(iii) Not less frequently than every 5 years, the dissemination and training activities and data networks established under section 937 of the Public Health Service Act, including the methods and products used to disseminate research, the types of training conducted and supported, and the types and functions of the data networks established, in order to determine whether the activities and data are produced in a manner consistent with the requirements under such section.

"(iv) Not less frequently than every 5 years, the overall effectiveness of activities conducted under this section and the dissemination, training, and capacity building activities conducted under section 937 of the Public Health Service Act. Such review shall include an analysis of the extent to which research findings are used by health care decision-makers, the effect of the dissemination of such findings on reducing practice variation and disparities in health care, and the effect of the research conducted and disseminated on

innovation and the health care economy of the United States.

"(v) Not later than 8 years after the date of enactment of this section, the adequacy and use of the funding for the Institute and the activities conducted under section 937 of the Public Health Service Act, including a determination as to whether, based on the utilization of research findings by public and private payers, funding sources for the Patient-Centered Outcomes Research Trust Fund under section 9511 of the Internal Revenue Code of 1986 are appropriate and whether such sources of funding should be continued or adjusted.

"(B) ANNUAL REPORTS.—Not later than April 1 of each year, the Comptroller General of the United States shall submit to Congress a report containing the results of the review conducted under subparagraph (A) with respect to the preceding year (or years, if applicable), together with recommendations for such legislation and administrative action as the Comptroller General determines appropriate. Recommendations.

"(h) ENSURING TRANSPARENCY, CREDIBILITY, AND ACCESS.—The Institute shall establish procedures to ensure that the following requirements for ensuring transparency, credibility, and access are met: Procedures.

"(1) PUBLIC COMMENT PERIODS.—The Institute shall provide for a public comment period of not less than 45 days and not more than 60 days prior to the adoption under subsection (d)(9) of the national priorities identified under subsection (d)(1)(A), the research project agenda established under subsection (d)(1)(B), the methodological standards developed and updated by the methodology committee under subsection (d)(6)(C)(i), and the peer-review process provided under paragraph (7), and after the release of draft findings with respect to systematic reviews of existing research and evidence.

"(2) ADDITIONAL FORUMS.—The Institute shall support forums to increase public awareness and obtain and incorporate public input and feedback through media (such as an Internet website) on research priorities, research findings, and other duties, activities, or processes the Institute determines appropriate.

"(3) PUBLIC AVAILABILITY.—The Institute shall make available to the public and disclose through the official public Internet website of the Institute the following: Web posting.

"(A) Information contained in research findings as specified in subsection (d)(9).

"(B) The process and methods for the conduct of research, including the identity of the entity and the investigators conducing such research and any conflicts of interests of such parties, any direct or indirect links the entity has to industry, and research protocols, including measures taken, methods of research and analysis, research results, and such other information the Institute determines appropriate) concurrent with the release of research findings.

"(C) Notice of public comment periods under paragraph (1), including deadlines for public comments. Notice.

"(D) Subsequent comments received during each of the public comment periods.

"(E) In accordance with applicable laws and processes and as the Institute determines appropriate, proceedings of the Institute.

"(4) DISCLOSURE OF CONFLICTS OF INTEREST.—

"(A) IN GENERAL.—A conflict of interest shall be disclosed in the following manner:

"(i) By the Institute in appointing members to an expert advisory panel under subsection (d)(4), in selecting individuals to contribute to any peer-review process under subsection (d)(7), and for employment as executive staff of the Institute.

"(ii) By the Comptroller General in appointing members of the methodology committee under subsection (d)(6);

"(iii) By the Institute in the annual report under subsection (d)(10), except that, in the case of individuals contributing to any such peer review process, such description shall be in a manner such that those individuals cannot be identified with a particular research project.

Web posting.

"(B) MANNER OF DISCLOSURE.—Conflicts of interest shall be disclosed as described in subparagraph (A) as soon as practicable on the Internet web site of the Institute and of the Government Accountability Office. The information disclosed under the preceding sentence shall include the type, nature, and magnitude of the interests of the individual involved, except to the extent that the individual recuses himself or herself from participating in the consideration of or any other activity with respect to the study as to which the potential conflict exists.

"(i) RULES.—The Institute, its Board or staff, shall be prohibited from accepting gifts, bequeaths, or donations of services or property. In addition, the Institute shall be prohibited from establishing a corporation or generating revenues from activities other than as provided under this section.

"(j) RULES OF CONSTRUCTION.—

"(1) COVERAGE.—Nothing in this section shall be construed—

"(A) to permit the Institute to mandate coverage, reimbursement, or other policies for any public or private payer; or

"(B) as preventing the Secretary from covering the routine costs of clinical care received by an individual entitled to, or enrolled for, benefits under title XVIII, XIX, or XXI in the case where such individual is participating in a clinical trial and such costs would otherwise be covered under such title with respect to the beneficiary.".

(b) DISSEMINATION AND BUILDING CAPACITY FOR RESEARCH.—Title IX of the Public Health Service Act (42 U.S.C. 299 et seq.), as amended by section 3606, is further amended by inserting after section 936 the following:

42 USC 299b–37.

"SEC. 937. DISSEMINATION AND BUILDING CAPACITY FOR RESEARCH.

"(a) IN GENERAL.—

"(1) DISSEMINATION.—The Office of Communication and Knowledge Transfer (referred to in this section as the 'Office') at the Agency for Healthcare Research and Quality (or any other relevant office designated by Agency for Healthcare Research and Quality), in consultation with the National Institutes of Health, shall broadly disseminate the research findings that are published by the Patient Centered Outcomes Research Institute established under section 1181(b) of the Social Security Act (referred to in this section as the 'Institute') and other government-funded research relevant to comparative clinical effectiveness research. The Office shall create informational tools that organize and disseminate research findings for physicians, health care providers, patients, payers, and policy makers. The Office shall also develop a publicly available resource database that collects and contains government-funded evidence and research from public, private, not-for profit, and academic sources.

Public information.

"(2) REQUIREMENTS.—The Office shall provide for the dissemination of the Institute's research findings and government-funded research relevant to comparative clinical effectiveness research to physicians, health care providers, patients, vendors of health information technology focused on clinical decision support, appropriate professional associations, and Federal and private health plans. Materials, forums, and media used to disseminate the findings, informational tools, and resource databases shall—

"(A) include a description of considerations for specific subpopulations, the research methodology, and the limitations of the research, and the names of the entities, agencies, instrumentalities, and individuals who conducted any research which was published by the Institute; and

"(B) not be construed as mandates, guidelines, or recommendations for payment, coverage, or treatment.

"(b) INCORPORATION OF RESEARCH FINDINGS.—The Office, in consultation with relevant medical and clinical associations, shall assist users of health information technology focused on clinical decision support to promote the timely incorporation of research findings disseminated under subsection (a) into clinical practices and to promote the ease of use of such incorporation.

"(c) FEEDBACK.—The Office shall establish a process to receive feedback from physicians, health care providers, patients, and vendors of health information technology focused on clinical decision support, appropriate professional associations, and Federal and private health plans about the value of the information disseminated and the assistance provided under this section.

"(d) RULE OF CONSTRUCTION.—Nothing in this section shall preclude the Institute from making its research findings publicly available as required under section 1181(d)(8) of the Social Security Act.

"(e) TRAINING OF RESEARCHERS.—The Agency for Health Care Research and Quality, in consultation with the National Institutes of Health, shall build capacity for comparative clinical effectiveness research by establishing a grant program that provides for the training of researchers in the methods used to conduct such research, including systematic reviews of existing research and primary research such as clinical trials. At a minimum, such

training shall be in methods that meet the methodological standards adopted under section 1181(d)(9) of the Social Security Act.

"(f) BUILDING DATA FOR RESEARCH.—The Secretary shall provide for the coordination of relevant Federal health programs to build data capacity for comparative clinical effectiveness research, including the development and use of clinical registries and health outcomes research data networks, in order to develop and maintain a comprehensive, interoperable data network to collect, link, and analyze data on outcomes and effectiveness from multiple sources, including electronic health records.

"(g) AUTHORITY TO CONTRACT WITH THE INSTITUTE.—Agencies and instrumentalities of the Federal Government may enter into agreements with the Institute, and accept and retain funds, for the conduct and support of research described in this part, provided that the research to be conducted or supported under such agreements is authorized under the governing statutes of such agencies and instrumentalities.".

(c) IN GENERAL.—Part D of title XI of the Social Security Act, as added by subsection (a), is amended by adding at the end the following new section:

"LIMITATIONS ON CERTAIN USES OF COMPARATIVE CLINICAL EFFECTIVENESS RESEARCH

42 USC 1320e–1.

"SEC. 1182. (a) The Secretary may only use evidence and findings from research conducted under section 1181 to make a determination regarding coverage under title XVIII if such use is through an iterative and transparent process which includes public comment and considers the effect on subpopulations.

"(b) Nothing in section 1181 shall be construed as—

"(1) superceding or modifying the coverage of items or services under title XVIII that the Secretary determines are reasonable and necessary under section 1862(l)(1); or

"(2) authorizing the Secretary to deny coverage of items or services under such title solely on the basis of comparative clinical effectiveness research.

"(c)(1) The Secretary shall not use evidence or findings from comparative clinical effectiveness research conducted under section 1181 in determining coverage, reimbursement, or incentive programs under title XVIII in a manner that treats extending the life of an elderly, disabled, or terminally ill individual as of lower value than extending the life of an individual who is younger, nondisabled, or not terminally ill.

"(2) Paragraph (1) shall not be construed as preventing the Secretary from using evidence or findings from such comparative clinical effectiveness research in determining coverage, reimbursement, or incentive programs under title XVIII based upon a comparison of the difference in the effectiveness of alternative treatments in extending an individual's life due to the individual's age, disability, or terminal illness.

"(d)(1) The Secretary shall not use evidence or findings from comparative clinical effectiveness research conducted under section 1181 in determining coverage, reimbursement, or incentive programs under title XVIII in a manner that precludes, or with the intent to discourage, an individual from choosing a health care treatment based on how the individual values the tradeoff between extending the length of their life and the risk of disability.

"(2)(A) Paragraph (1) shall not be construed to—

"(i) limit the application of differential copayments under title XVIII based on factors such as cost or type of service; or

"(ii) prevent the Secretary from using evidence or findings from such comparative clinical effectiveness research in determining coverage, reimbursement, or incentive programs under such title based upon a comparison of the difference in the effectiveness of alternative health care treatments in extending an individual's life due to that individual's age, disability, or terminal illness.

"(3) Nothing in the provisions of, or amendments made by the Patient Protection and Affordable Care Act, shall be construed to limit comparative clinical effectiveness research or any other research, evaluation, or dissemination of information concerning the likelihood that a health care treatment will result in disability.

"(e) The Patient-Centered Outcomes Research Institute established under section 1181(b)(1) shall not develop or employ a dollars-per-quality adjusted life year (or similar measure that discounts the value of a life because of an individual's disability) as a threshold to establish what type of health care is cost effective or recommended. The Secretary shall not utilize such an adjusted life year (or such a similar measure) as a threshold to determine coverage, reimbursement, or incentive programs under title XVIII.".

(d) IN GENERAL.—Part D of title XI of the Social Security Act, as added by subsection (a) and amended by subsection (c), is amended by adding at the end the following new section:

"TRUST FUND TRANSFERS TO PATIENT-CENTERED OUTCOMES RESEARCH TRUST FUND

"SEC. 1183. (a) IN GENERAL.—The Secretary shall provide for the transfer, from the Federal Hospital Insurance Trust Fund under section 1817 and the Federal Supplementary Medical Insurance Trust Fund under section 1841, in proportion (as estimated by the Secretary) to the total expenditures during such fiscal year that are made under title XVIII from the respective trust fund, to the Patient-Centered Outcomes Research Trust Fund (referred to in this section as the 'PCORTF') under section 9511 of the Internal Revenue Code of 1986, of the following: 42 USC 1320e–2.

"(1) For fiscal year 2013, an amount equal to $1 multiplied by the average number of individuals entitled to benefits under part A, or enrolled under part B, of title XVIII during such fiscal year.

"(2) For each of fiscal years 2014, 2015, 2016, 2017, 2018, and 2019, an amount equal to $2 multiplied by the average number of individuals entitled to benefits under part A, or enrolled under part B, of title XVIII during such fiscal year.

"(b) ADJUSTMENTS FOR INCREASES IN HEALTH CARE SPENDING.—In the case of any fiscal year beginning after September 30, 2014, the dollar amount in effect under subsection (a)(2) for such fiscal year shall be equal to the sum of such dollar amount for the previous fiscal year (determined after the application of this subsection), plus an amount equal to the product of—

"(1) such dollar amount for the previous fiscal year, multiplied by

"(2) the percentage increase in the projected per capita amount of National Health Expenditures, as most recently published by the Secretary before the beginning of the fiscal year.".

(e) PATIENT-CENTERED OUTCOMES RESEARCH TRUST FUND; FINANCING FOR TRUST FUND.—

(1) ESTABLISHMENT OF TRUST FUND.—

(A) IN GENERAL.—Subchapter A of chapter 98 of the Internal Revenue Code of 1986 (relating to establishment of trust funds) is amended by adding at the end the following new section:

26 USC 9511.

"SEC. 9511. PATIENT-CENTERED OUTCOMES RESEARCH TRUST FUND.

"(a) CREATION OF TRUST FUND.—There is established in the Treasury of the United States a trust fund to be known as the 'Patient-Centered Outcomes Research Trust Fund' (hereafter in this section referred to as the 'PCORTF'), consisting of such amounts as may be appropriated or credited to such Trust Fund as provided in this section and section 9602(b).

"(b) TRANSFERS TO FUND.—

"(1) APPROPRIATION.—There are hereby appropriated to the Trust Fund the following:

"(A) For fiscal year 2010, $10,000,000.

"(B) For fiscal year 2011, $50,000,000.

"(C) For fiscal year 2012, $150,000,000.

"(D) For fiscal year 2013—

"(i) an amount equivalent to the net revenues received in the Treasury from the fees imposed under subchapter B of chapter 34 (relating to fees on health insurance and self-insured plans) for such fiscal year; and

"(ii) $150,000,000.

"(E) For each of fiscal years 2014, 2015, 2016, 2017, 2018, and 2019—

"(i) an amount equivalent to the net revenues received in the Treasury from the fees imposed under subchapter B of chapter 34 (relating to fees on health insurance and self-insured plans) for such fiscal year; and

"(ii) $150,000,000.

The amounts appropriated under subparagraphs (A), (B), (C), (D)(ii), and (E)(ii) shall be transferred from the general fund of the Treasury, from funds not otherwise appropriated.

"(2) TRUST FUND TRANSFERS.—In addition to the amounts appropriated under paragraph (1), there shall be credited to the PCORTF the amounts transferred under section 1183 of the Social Security Act.

"(3) LIMITATION ON TRANSFERS TO PCORTF.—No amount may be appropriated or transferred to the PCORTF on and after the date of any expenditure from the PCORTF which is not an expenditure permitted under this section. The determination of whether an expenditure is so permitted shall be made without regard to—

"(A) any provision of law which is not contained or referenced in this chapter or in a revenue Act, and

"(B) whether such provision of law is a subsequently enacted provision or directly or indirectly seeks to waive the application of this paragraph.

"(c) TRUSTEE.—The Secretary of the Treasury shall be a trustee of the PCORTF.

"(d) EXPENDITURES FROM FUND.—

"(1) AMOUNTS AVAILABLE TO THE PATIENT-CENTERED OUTCOMES RESEARCH INSTITUTE.—Subject to paragraph (2), amounts in the PCORTF are available, without further appropriation, to the Patient-Centered Outcomes Research Institute established under section 1181(b) of the Social Security Act for carrying out part D of title XI of the Social Security Act (as in effect on the date of enactment of such Act).

"(2) TRANSFER OF FUNDS.—

"(A) IN GENERAL.—The trustee of the PCORTF shall provide for the transfer from the PCORTF of 20 percent of the amounts appropriated or credited to the PCORTF for each of fiscal years 2011 through 2019 to the Secretary of Health and Human Services to carry out section 937 of the Public Health Service Act.

"(B) AVAILABILITY.—Amounts transferred under subparagraph (A) shall remain available until expended.

"(C) REQUIREMENTS.—Of the amounts transferred under subparagraph (A) with respect to a fiscal year, the Secretary of Health and Human Services shall distribute—

"(i) 80 percent to the Office of Communication and Knowledge Transfer of the Agency for Healthcare Research and Quality (or any other relevant office designated by Agency for Healthcare Research and Quality) to carry out the activities described in section 937 of the Public Health Service Act; and

"(ii) 20 percent to the Secretary to carry out the activities described in such section 937.

"(e) NET REVENUES.—For purposes of this section, the term 'net revenues' means the amount estimated by the Secretary of the Treasury based on the excess of— Definition.

"(1) the fees received in the Treasury under subchapter B of chapter 34, over

"(2) the decrease in the tax imposed by chapter 1 resulting from the fees imposed by such subchapter.

"(f) TERMINATION.—No amounts shall be available for expenditure from the PCORTF after September 30, 2019, and any amounts in such Trust Fund after such date shall be transferred to the general fund of the Treasury.".

(B) CLERICAL AMENDMENT.—The table of sections for subchapter A of chapter 98 of such Code is amended by adding at the end the following new item:

"Sec. 9511. Patient-centered outcomes research trust fund.".

(2) FINANCING FOR FUND FROM FEES ON INSURED AND SELF-INSURED HEALTH PLANS.—

(A) GENERAL RULE.—Chapter 34 of the Internal Revenue Code of 1986 is amended by adding at the end the following new subchapter:

"Subchapter B—Insured and Self-Insured Health Plans

"Sec. 4375. Health insurance.
"Sec. 4376. Self-insured health plans.
"Sec. 4377. Definitions and special rules.

"SEC. 4375. HEALTH INSURANCE. 26 USC 4375.

"(a) IMPOSITION OF FEE.—There is hereby imposed on each specified health insurance policy for each policy year ending after

September 30, 2012, a fee equal to the product of $2 ($1 in the case of policy years ending during fiscal year 2013) multiplied by the average number of lives covered under the policy.

"(b) LIABILITY FOR FEE.—The fee imposed by subsection (a) shall be paid by the issuer of the policy.

"(c) SPECIFIED HEALTH INSURANCE POLICY.—For purposes of this section:

Definition.

"(1) IN GENERAL.—Except as otherwise provided in this section, the term 'specified health insurance policy' means any accident or health insurance policy (including a policy under a group health plan) issued with respect to individuals residing in the United States.

"(2) EXEMPTION FOR CERTAIN POLICIES.—The term 'specified health insurance policy' does not include any insurance if substantially all of its coverage is of excepted benefits described in section 9832(c).

"(3) TREATMENT OF PREPAID HEALTH COVERAGE ARRANGEMENTS.—

"(A) IN GENERAL.—In the case of any arrangement described in subparagraph (B), such arrangement shall be treated as a specified health insurance policy, and the person referred to in such subparagraph shall be treated as the issuer.

"(B) DESCRIPTION OF ARRANGEMENTS.—An arrangement is described in this subparagraph if under such arrangement fixed payments or premiums are received as consideration for any person's agreement to provide or arrange for the provision of accident or health coverage to residents of the United States, regardless of how such coverage is provided or arranged to be provided.

"(d) ADJUSTMENTS FOR INCREASES IN HEALTH CARE SPENDING.—In the case of any policy year ending in any fiscal year beginning after September 30, 2014, the dollar amount in effect under subsection (a) for such policy year shall be equal to the sum of such dollar amount for policy years ending in the previous fiscal year (determined after the application of this subsection), plus an amount equal to the product of—

"(1) such dollar amount for policy years ending in the previous fiscal year, multiplied by

"(2) the percentage increase in the projected per capita amount of National Health Expenditures, as most recently published by the Secretary before the beginning of the fiscal year.

"(e) TERMINATION.—This section shall not apply to policy years ending after September 30, 2019.

26 USC 4376.

"SEC. 4376. SELF-INSURED HEALTH PLANS.

"(a) IMPOSITION OF FEE.—In the case of any applicable self-insured health plan for each plan year ending after September 30, 2012, there is hereby imposed a fee equal to $2 ($1 in the case of plan years ending during fiscal year 2013) multiplied by the average number of lives covered under the plan.

"(b) LIABILITY FOR FEE.—

"(1) IN GENERAL.—The fee imposed by subsection (a) shall be paid by the plan sponsor.

Definition.

"(2) PLAN SPONSOR.—For purposes of paragraph (1) the term 'plan sponsor' means—

"(A) the employer in the case of a plan established or maintained by a single employer,

"(B) the employee organization in the case of a plan established or maintained by an employee organization,

"(C) in the case of—

"(i) a plan established or maintained by 2 or more employers or jointly by 1 or more employers and 1 or more employee organizations,

"(ii) a multiple employer welfare arrangement, or

"(iii) a voluntary employees' beneficiary association described in section 501(c)(9), the association, committee, joint board of trustees, or other similar group of representatives of the parties who establish or maintain the plan, or

"(D) the cooperative or association described in subsection (c)(2)(F) in the case of a plan established or maintained by such a cooperative or association.

"(c) APPLICABLE SELF-INSURED HEALTH PLAN.—For purposes of this section, the term 'applicable self-insured health plan' means any plan for providing accident or health coverage if— Definition.

"(1) any portion of such coverage is provided other than through an insurance policy, and

"(2) such plan is established or maintained—

"(A) by 1 or more employers for the benefit of their employees or former employees,

"(B) by 1 or more employee organizations for the benefit of their members or former members,

"(C) jointly by 1 or more employers and 1 or more employee organizations for the benefit of employees or former employees,

"(D) by a voluntary employees' beneficiary association described in section 501(c)(9),

"(E) by any organization described in section 501(c)(6), or

"(F) in the case of a plan not described in the preceding subparagraphs, by a multiple employer welfare arrangement (as defined in section 3(40) of Employee Retirement Income Security Act of 1974), a rural electric cooperative (as defined in section 3(40)(B)(iv) of such Act), or a rural telephone cooperative association (as defined in section 3(40)(B)(v) of such Act).

"(d) ADJUSTMENTS FOR INCREASES IN HEALTH CARE SPENDING.—In the case of any plan year ending in any fiscal year beginning after September 30, 2014, the dollar amount in effect under subsection (a) for such plan year shall be equal to the sum of such dollar amount for plan years ending in the previous fiscal year (determined after the application of this subsection), plus an amount equal to the product of—

"(1) such dollar amount for plan years ending in the previous fiscal year, multiplied by

"(2) the percentage increase in the projected per capita amount of National Health Expenditures, as most recently published by the Secretary before the beginning of the fiscal year.

"(e) TERMINATION.—This section shall not apply to plan years ending after September 30, 2019.

26 USC 4377.

"SEC. 4377. DEFINITIONS AND SPECIAL RULES.

"(a) DEFINITIONS.—For purposes of this subchapter—

"(1) ACCIDENT AND HEALTH COVERAGE.—The term 'accident and health coverage' means any coverage which, if provided by an insurance policy, would cause such policy to be a specified health insurance policy (as defined in section 4375(c)).

"(2) INSURANCE POLICY.—The term 'insurance policy' means any policy or other instrument whereby a contract of insurance is issued, renewed, or extended.

"(3) UNITED STATES.—The term 'United States' includes any possession of the United States.

"(b) TREATMENT OF GOVERNMENTAL ENTITIES.—

"(1) IN GENERAL.—For purposes of this subchapter—

"(A) the term 'person' includes any governmental entity, and

"(B) notwithstanding any other law or rule of law, governmental entities shall not be exempt from the fees imposed by this subchapter except as provided in paragraph (2).

"(2) TREATMENT OF EXEMPT GOVERNMENTAL PROGRAMS.—In the case of an exempt governmental program, no fee shall be imposed under section 4375 or section 4376 on any covered life under such program.

"(3) EXEMPT GOVERNMENTAL PROGRAM DEFINED.—For purposes of this subchapter, the term 'exempt governmental program' means—

"(A) any insurance program established under title XVIII of the Social Security Act,

"(B) the medical assistance program established by title XIX or XXI of the Social Security Act,

"(C) any program established by Federal law for providing medical care (other than through insurance policies) to individuals (or the spouses and dependents thereof) by reason of such individuals being members of the Armed Forces of the United States or veterans, and

"(D) any program established by Federal law for providing medical care (other than through insurance policies) to members of Indian tribes (as defined in section 4(d) of the Indian Health Care Improvement Act).

"(c) TREATMENT AS TAX.—For purposes of subtitle F, the fees imposed by this subchapter shall be treated as if they were taxes.

"(d) NO COVER OVER TO POSSESSIONS.—Notwithstanding any other provision of law, no amount collected under this subchapter shall be covered over to any possession of the United States.".

(B) CLERICAL AMENDMENTS.—

(i) Chapter 34 of such Code is amended by striking the chapter heading and inserting the following:

"CHAPTER 34—TAXES ON CERTAIN INSURANCE POLICIES

"SUBCHAPTER A. POLICIES ISSUED BY FOREIGN INSURERS

"SUBCHAPTER B. INSURED AND SELF-INSURED HEALTH PLANS

"Subchapter A—Policies Issued By Foreign Insurers".

(ii) The table of chapters for subtitle D of such Code is amended by striking the item relating to chapter 34 and inserting the following new item:

"CHAPTER 34—TAXES ON CERTAIN INSURANCE POLICIES".

(f) TAX-EXEMPT STATUS OF THE PATIENT-CENTERED OUTCOMES RESEARCH INSTITUTE.—Subsection 501(l) of the Internal Revenue Code of 1986 is amended by adding at the end the following new paragraph: 26 USC 501.

"(4) The Patient-Centered Outcomes Research Institute established under section 1181(b) of the Social Security Act.".

SEC. 6302. FEDERAL COORDINATING COUNCIL FOR COMPARATIVE EFFECTIVENESS RESEARCH. 42 USC 2996–8 note.

Notwithstanding any other provision of law, the Federal Coordinating Council for Comparative Effectiveness Research established under section 804 of Division A of the American Recovery and Reinvestment Act of 2009 (42 U.S.C. 299b–8), including the requirement under subsection (e)(2) of such section, shall terminate on the date of enactment of this Act. Termination date.

Appendix F

Workshop Agenda

LEARNING WHAT WORKS
Infrastructure Required to Learn Which Care Is Best

A Learning Healthcare System Workshop
IOM Roundtable on Evidence-Based Medicine

July 30–31, 2008
Lecture Room, The National Academy of Sciences
Washington, DC 20001

Issues motivating the discussion

1. Substantial demand for greater insights into the comparative clinical effectiveness of clinical interventions and care processes to improve the effectiveness and value of health care.
2. Expanded interest and activity in the work needed—e.g., comparative effectiveness research, systematic reviews, innovative research strategies, clinical registries, coverage with evidence development.
3. Currently fragmented and largely uncoordinated selection of studies, study design and conduct, evidence synthesis, methods validation and improvement, and development and dissemination of guidelines.
4. Expanding gap in workforce with skills to develop data sources and systems, design and conduct innovative studies, translate results, and guide application.
5. Opportunities presented by the attention of recent initiatives and the increasing possibility for development of an entity and resources for expanded work on the comparative effectiveness of clinical interventions.
6. Growing appreciation of the importance of assessing the infrastructure needed for this work—e.g., workforce needs, data linkage and

improvement, new methodologies, research networks, technical assistance.

7. Desirability of a common venue to identify and characterize the need categories, begin to estimate the shortfalls, consider approaches to addressing the shortfalls, and identify priority next steps.

Discussion assumptions: (1) Resources will be available to expand work on the comparative effectiveness of medical interventions, and (2) a designated entity will exist with a formal charge to coordinate the expanded work.

Goal: Clarify the elements and nature of the needed capacity, solicit quantitative and qualitative assessments of the needs, and characterize in a fashion that will facilitate engagement of the issues by policy makers.

DAY ONE

8:00 **Welcome and Introductions**
John W. Rowe, Columbia University & IOM Roundtable on Evidence-Based Medicine

8:15 **Keynote: A Vision for the Capacity to Learn What Care Works Best**
What are the core elements of a robust and sustainable capacity for comparative effectiveness research? How do they relate to each other as a real infrastructure? What are the priorities and how might we build upon, link, and improve existing public and private system elements?
Mark B. McClellan, Brookings Institution & IOM Roundtable on Evidence-Based Medicine

8:45 **Session 1: The Work Required**
Each presenter will describe the nature of the activity and what is known about the current capacity, suggest an approach to determining the necessary capacity, give an "opening bid" estimate of what that need might be, and offer initial suggestions on policies or activities for progress.
Chair: Mark B. McClellan, Brookings Institution & IOM Roundtable on Evidence-Based Medicine

- The **Cost and Volume of Current Comparative Effectiveness Research**
  Erin Holve, AcademyHealth
- **Intervention Studies That Need to Be Conducted**
  Douglas B. Kamerow, RTI International
- **Clinical Data Sets That Need to Be Mined**
  Jesse A. Berlin, Johnson & Johnson
- **Knowledge Synthesis and Translation That Need to Be Applied**
  Richard A. Justman, UnitedHealthcare

[9:45 – 10:00 BREAK]

- **Methods That Need to Be Developed**
  Eugene H. Blackstone, Cleveland Clinic
- **Coordination and Technical Assistance That Need to Be Supported**
  Jean R. Slutsky, Agency for Healthcare Research and Quality

OPEN DISCUSSION (60 minutes)

11:30 LUNCH PRESENTATION: THE POTENTIAL RETURNS FROM EVIDENCE-DRIVEN HEALTH CARE
Gail R. Wilensky, Project HOPE

1:00 SESSION 2: THE INFORMATION NETWORKS REQUIRED
Each presenter will describe the nature of the systems required to capture and use data for insights on clinical effectiveness and to ensure effective and efficient review and dissemination of those insights. Presenters will also describe what is known about the current capacity, give a rough estimate of the necessary capacity, and offer initial suggestions on policies or activities for progress.
Chair: Kathleen A. Buto, Johnson &Johnson

- **Information Technology Requirements**
  Robert H. Miller, University of California at San Francisco
- **Data and Information Hub Requirements**
  Carol C. Diamond, The Markle Foundation

➢ **Integrative Vehicles Required for Evidence Review and Dissemination**
Lorne A. Becker, Cochrane Collaboration

OPEN DISCUSSION (45 minutes)

2:30 **Session 3: The Talent Required**
Session Three will review the competencies and workforce necessary to carry out comparative effectiveness work, including the presentation of initial estimates of the workforce needed in key areas, comments on the estimates from end users of such research, and insights on the implications for training programs.
Chair: Donald M. Steinwachs, Johns Hopkins University Bloomberg School of Public Health & IOM Roundtable on Evidence-Based Medicine

➢ **Comparative Effectiveness Workforce—Framework and Assessment**
William R. Hersh, Oregon Health and Science University

- **Reactor Panel**
Jean Paul Gagnon, sanofi-aventis
Bruce H. Hamory, Geisinger Health System
Steve E. Phurrough, Centers for Medicare & Medicaid Services
Robert J. Temple, Food and Drug Administration

OPEN DISCUSSION (30 minutes)

➢ **Panel Discussion: Training and Education Needs**
- **Clinical Epidemiology and Research**—*Eric B. Bass, Johns Hopkins University*
- **Health Services Research**— *Timothy S. Carey, University of North Carolina at Chapel Hill*
- **Informatics**—*Don E. Detmer, American Medical Informatics Association*
- **Guideline Development**— *Richard N. Shiffman, Yale University*
- **Knowledge Translation**—*David H. Hickam, Eisenberg Center*

OPEN DISCUSSION (30 minutes)

➢ **Toward an Integrated Enterprise—An International Case Study**
Sean R. Tunis, Center for Medical Technology and Policy

OPEN DISCUSSION (20 minutes)

5:15 HIGHLIGHTS OF DAY'S DISCUSSION
John W. Rowe, Columbia University & Roundtable on Evidence-Based Medicine

5:30 **RECEPTION**

DAY TWO

8:00 WELCOME AND RECAP OF THE FIRST DAY
J. Michael McGinnis, Institute of Medicine

8:15 SESSION 4: NEXT STEPS—PRIORITIES FOR IMPLEMENTATION
In this discussion of priorities, strategies, staging, and policies, each participant will offer comments that reflect on the nature (quantitative and qualitative) of the need in the relevant area, how to clarify the understanding on that count, and the possible approaches to ramping up. Specific consideration of how to take best advantage of existing manufacturer and related private sector capacity will also be discussed. Ten minutes of comments on each of the following areas, followed by open discussion.
Chair: Nancy H. Nielsen, American Medical Association & IOM Roundtable on Evidence-Based Medicine

➢ **Information Technology Platform Requirements**
Mark E. Frisse, Vanderbilt University
➢ **Data Resource Development and Analysis Improvement**
T. Bruce Ferguson, East Carolina Heart Institute
➢ **Research Infrastructure Improvement**
Daniel E. Ford, Johns Hopkins University
➢ **Transforming Health Professions Education**
Benjamin K. Chu, Kaiser Foundation Health Plan and Hospitals

- **Building the Training Capacity**
  Steven A. Wartman, Association of Academic Health Centers

PANEL & OPEN DISCUSSION (50 minutes)

[10:00 – 10:15 BREAK]

10:15 **Session 5: Public–Private Partnerships**
Session panelists will address the following questions: How might better advantage be taken of capacity in the private sector? What existing or planned efforts might best inform the development of needed infrastructure for comparative analyses and research (e.g., biomarkers consortium, data aggregation efforts, Food and Drug Administration Sentinel Network)?
Chair: W. David Helms, AcademyHealth

- *Carmella A. Bocchino, America's Health Insurance Plans*
- *Rachel E. Behrman, Food and Drug Association*
- *William Z. Potter, Merck Research Laboratories*

PANEL & OPEN DISCUSSION (45 minutes)

11:15 **Session 6: Getting Started**
This final session will be a facilitated open discussion of the priorities and strategies for building the needed infrastructure. There will be 30-45 minutes each on the following.
Chair: John W. Rowe, Columbia University & Roundtable on Evidence-Based Medicine

- **The Roadmap—Policies, Priorities, Strategies, and Sequencing**
  Stuart Guterman, The Commonwealth Fund
- **Quick Hits—Things That Can Be Done Now**
  W. David Helms, AcademyHealth
- **Building Support**
  Mary Woolley, Research!America

12:45 **Concluding Summary Remarks and Adjournment**
J. Michael McGinnis, Institute of Medicine

Planning Committee:

John W. Rowe (*Chair*), Mailman School of Public Health, Columbia University
Stuart H. Altman, Brandeis University
Kathleen A. Buto, Johnson & Johnson
Carolyn M. Clancy, Agency for Healthcare Research and Quality
W. David Helms, AcademyHealth
Mark B. McClellan, Brookings Institution
Peter R. Orszag, Congressional Budget Office

Appendix G

Biographical Sketches of Workshop Participants

Eric B. Bass, M.D., M.P.H., is a professor of medicine with joint appointments in the departments of epidemiology and health policy and management at the Johns Hopkins University (JHU). He is the director of the JHU Evidence-Based Practice Center, and codirector of the JHU Developing Evidence to Inform Decisions about Effectiveness Center. He was the editor of the *Journal of General Internal Medicine* for 5 years, and he is now the editor-in-chief of *Progress in Community Health Partnerships*. He served as the director of the general internal medicine fellowship at JHU for 15 years. He has extensive research experience in evidence-based medicine, with special interests in assessment of the effectiveness and costs of medical and surgical management strategies as well as in the assessment of patient preferences. He also has expertise in curriculum development, having served as a facilitator for the JHU Faculty Development Program and having coauthored a book on curriculum development. He has published more than 175 peer-reviewed articles as well as numerous evidence reports.

Lorne A. Becker, M.D., is emeritus professor at the State University of New York (SUNY) Upstate Medical University in Syracuse, New York. He serves as co-chair of the steering group of the Cochrane Collaboration, an international nonprofit healthcare organization, and is also convenor of the Cochrane Publishing Policy Group and Umbrella Reviews Working Group. He served as chief of family medicine at the Toronto Hospital and, until his retirement in 2004, he was chair of the Department of Family Medicine at SUNY Upstate Medical University. Dr. Becker received his M.D. from the University of Western Ontario, is a fellow of the American Academy

of Family Medicine and the College of Family Physicians of Canada, and is a recipient of the Curtis Hames Research Award from the Society of Teachers of Family Medicine. He is a member of the Research Committee of the World Organization of National Academies of Family Medicine and of the Guidelines International Network Advisory Board. Previously he has been a member of the board of the Society of Teachers of Family Medicine, on the editorial boards of the *Journal of Family Practice* and of *Family Medicine* (Oxford), and the coordinator of the Cochrane Primary Health Care Field.

Rachel E. Behrman, M.D., M.P.H., is the associate commissioner for clinical programs and director of the Office of Critical Path Programs of the Food and Drug Administration (FDA). In that capacity, she is responsible for developing, coordinating, and implementing policy and scientific programs aimed at innovating development and regulation of FDA-regulated products. She also oversees the Good Clinical Practice Program, which coordinates FDA policies on human subject protection and bioresearch monitoring. Dr. Behrman joined the FDA in 1989 and has held a number of positions in the Center for Drug Evaluation and Research, including medical reviewer and team leader in the Division of Antiviral Drug Products and deputy director of the Office of Medical Policy. Dr. Behrman is a board-certified internist and infectious disease subspecialist and a fellow of the American College of Physicians. She received her A.B. in mathematics from Washington University, her M.D. from Mt. Sinai School of Medicine, and her M.P.H. from the Johns Hopkins School of Hygiene and Public Health.

Jesse A. Berlin, Sc.D., received his doctorate in biostatistics from the Harvard School of Public Health in 1988. In 1989 he joined the faculty at the University of Pennsylvania, in a unit that became the Center for Clinical Epidemiology and Biostatistics, under the direction of Dr. Brian Strom. Dr. Berlin spent several years as director of biostatistics for the University of Pennsylvania Cancer Center, followed by assuming the role of faculty director of the Biostatistics and Epidemiology Consulting Center. At the end of the summer of 2004, Dr. Berlin left Penn to join Johnson & Johnson Pharmaceutical Research and Development, where he is currently vice president of pharmaco-epidemiology. He has authored or coauthored over 220 publications in a wide variety of clinical and methodological areas. Dr. Berlin has a great deal of experience in both the application of meta-analysis and the study of meta-analytic methods as applied to both randomized trials and epidemiology. He has served as a consultant on meta-analysis for the Australian government, and has served on two Institute of Medicine (IOM)

committees examining the association between exposure to chemicals contained in Agent Orange and risk of a wide variety of diseases.

Eugene H. Blackstone, M.D., is full-time head of clinical investigations at the Cleveland Clinic Heart and Vascular Institute (HVI) and a staff member at the Department of Thoracic and Cardiovascular Surgery, Quantitative Health Sciences, and Transplant Center. He is also a member of the HVI Executive Committee. In addition to generating new knowledge from clinical experiences, his specialty interests include novel mathematical models for analyzing time-related and longitudinal clinical outcomes, digital signal processing, mathematical models of the circulatory system, and semantic knowledgebase/database technology. He received undergraduate and medical degrees from the University of Chicago, followed by a research fellowship in mathematics, computer science, and statistics. He completed his medical internship at the University of Alabama at Birmingham (UAB) and served for 3 years as chief of the cardiovascular medicine branch at the U.S. Army Aeromedical Research Laboratory. He joined the UAB cardiac surgery faculty in 1972 as a full-time physician investigator and became the Cardiovascular Surgical Research Professor of Surgery in 1980. A major culmination of Dr. Blackstone's activities is the textbook *Cardiac Surgery*, now in its third edition. He joined the Cleveland Clinic in 1997 to head a multidisciplinary clinical research team focused on ischemic and valvar heart diseases, heart rhythm disturbances, heart failure, and benign and malignant diseases of the esophagus and lungs. He also leads a team of computer scientists in developing a novel semantic approach to data storage and use. Since 1987, Dr. Blackstone has been associate editor of the *Journal of Thoracic and Cardiovascular Surgery.* He has published over 450 articles and has received numerous awards, including a Distinguished Service Award from the University of Chicago in 1998 and the Maria and Sam Miller Professional Excellence Award for Scientific Achievement in Clinical Research in 2003. He is professor of surgery at the University of Toronto and at the Cleveland Clinic Lerner College of Medicine of Case Western Reserve University.

Carmella A. Bocchino, R.N., M.B.A., is a leading authority on identifying and promoting medical management strategies that advance public health goals and improve the overall quality of health care. As executive vice president of clinical affairs and strategic planning at America's Health Insurance Plans, Ms. Bocchino works with the executives of member organizations to develop innovative patient-centered medical management tools and clinical delivery models, foster private–public partnerships and advance an interconnected health care system. A registered professional nurse and former hospital administrator, Ms. Bocchino's clinical and public policy expertise

has been widely recognized by national and state lawmakers, policy makers, patient advocacy groups, employers, and throughout the health care community. She has been appointed to numerous private, state, and federal healthcare advisory committees, including the IOM Study of the Medicare End-Stage Renal Disease program, advisor to the RAND Health Sciences Program for the capitation study for the end-stage renal disease project, the Advisory Committee for Quality Improvement Standards for Managed Care, the planning committee establishing the National Quality Forum, and the Robert Wood Johnson Foundation (RWJF) Office of the National Coordinator for Health Information Technology Health Information Technology Adoption Initiative Expert Consensus Panel. She currently serves on the RWJF National Advisory Committee of Project Health Design: Rethinking the Power and Potential of Personal Health Records. Ms. Bocchino received her M.B.A. from Rutgers University, Graduate School of Management (Newark, NJ). She also has an undergraduate degree in human resources management from Upsala College and a nursing degree from Mountainside Hospital School of Nursing. Prior to her positions in health policy, Ms. Bocchino held administrative and clinical positions in critical care medicine and renal replacement therapy.

Kathleen A. Buto, M.P.A., is vice president for health policy, government affairs, at Johnson & Johnson. She has responsibility for providing policy analysis and developing positions on a wide range of issues, including the Medicare drug benefit, government reimbursement, coverage of new technologies, and regulatory requirements. In addition to reviewing how federal, state, and international government policies affect Johnson & Johnson products and customers, she is responsible for helping identify areas of opportunity for Johnson & Johnson to take leadership in shaping healthcare policy. Prior to joining Johnson & Johnson, Ms. Buto was a senior health advisor at the Congressional Budget Office, helping to develop the cost models for the Medicare drug benefit. Before that, she spent more than 18 years in senior positions at the Health Care Financing Administration, including deputy director of the Center for Health Plans and Providers, and associate administrator for policy. In these positions, she headed the policy, reimbursement, research, and coverage functions for the agency, as well as managing Medicare's fee-for-service and managed care operations. Ms. Buto received her B.A. from Douglass College and her M.P.A. from Harvard University.

Timothy S. Carey, M.D., M.P.H., was named director of the Cecil G. Sheps Center for Health Services Research in October 2000. A Sheps Center fellow and University of North Carolina (UNC) at Chapel Hill faculty member since 1986, Dr. Carey is a physician and health services researcher

with interests related to evidence-based medicine, access to care, health disparities, and medical outcomes. Several of his current research projects examine technology assessment, including the benefits and disadvantages of tube feedings in the frail elderly. He has also conducted a number of studies to examine the outcomes of care for low back pain, and he is codirector of the joint Research Triangle Institute/UNC Evidence-Based Practice Center examining the strength of the literature on a variety of medical and health policy topics. Dr. Carey is a Sarah Graham Kenan Professor in the departments of medicine and social medicine at UNC at Chapel Hill and was chief of the Division of General Medicine and Clinical Epidemiology at the UNC at Chapel Hill School of Medicine from 1991 to 2000. He is a frequent federal reviewer and has served as a member of the IOM's Monitoring Panel on Access to Care. He teaches at both the UNC schools of medicine and public health.

Benjamin K. Chu, M.D., M.P.H., is the Southern California regional president for Kaiser Foundation Health Plan and Hospitals, the nation's largest integrated health system, serving 8.7 million members in 10 states. His region is responsible for the care of 3.3 million patients. Before joining Kaiser, Dr. Chu was the president of the New York City Health and Hospitals Corporation, the largest municipal hospital system in the country, serving 1.3 million New York City residents with 11 hospitals, 5 skilled nursing facilities, and scores of outpatient facilities. Dr. Chu is an internist by training and has served as an associate dean and vice president for clinical affairs for the New York University (NYU) School of Medicine and as a senior associate dean for the Columbia College of Physicians and Surgeons. His experience with graduate medical education on the policy level, as the accountable executive for residency training programs in a large city hospital, and his experience working directly with medical students and residents—combined with his work with health systems' transformation—give him a unique perspective on the changes that will be needed in workforce training to adapt to a patient-centered, evidenced-based healthcare system.

Don E. Detmer, M.D., M.A., is president and chief executive officer of the American Medical Informatics Association. He is also professor of medical education in the Department of Public Health Sciences at the University of Virginia, and visiting professor at the Centre for Health Informatics and Multiprofessional Education, University College of London. Dr. Detmer is a member of the IOM as well as a lifetime associate of the National Academies. He is a fellow of American Association for the Advancement of Science (AAAS) as well as the American Colleges of Medical Informatics, Sports Medicine, and Surgeons. In addition to co-chairing the Blue

Ridge Academic Health Group, he chairs the board of MedBiquitous. He is treasurer of the Council of Medical Specialty Societies. Dr. Detmer is past chairman of the Board on Health Care Services of the IOM, the National Committee on Vital and Health Statistics, and the Board of Regents of the National Library of Medicine. He was a commissioner on the President's recent Commission on Systemic Interoperability. He chaired the 1991 IOM study, *The Computer-Based Patient Record*, and coedited the 1997 version of the same report. He was a member of the committee that developed the IOM reports *To Error Is Human* and *Crossing the Quality Chasm*. From 1999 to 2003 he was the Dennis Gillings Professor of Health Management at Cambridge University and is a lifetime member of Clare Hall College, Cambridge. His education includes a M.D. from the University of Kansas with subsequent training at the National Institutes of Health, the Johns Hopkins Hospital, Duke University Medical Center, the IOM, and Harvard Business School. His M.A. is from the University of Cambridge.

Carol C. Diamond, M.D., M.P.H., is the managing director of the Health Program of the Markle Foundation and chairs *Connecting for Health*, a public–private collaborative working to realize the full potential of information technology in health and health care in the United States. *Connecting for Health* engages more than 100 diverse organizations and institutions in an approach rooted in core values, including achieving medical excellence, fostering patient participation, and protecting personal privacy. Before joining the Markle Foundation, Dr. Diamond was president of U.S. Quality Algorithms (USQA), Aetna U.S. Healthcare's performance measurement affiliate. Prior to joining USQA, Dr. Diamond was a consultant for Johnson & Johnson Health Care Systems and the RWJF. Dr. Diamond sits on the American Academy of Family Physicians Public Advisory Board, the Electronic Health Record Safety Institute Advisory Board of the Geisinger Center for Health Research, and is a member of the IPRO Advisory Board for the Centers for Medicare & Medicaid Services (CMS) Doctor's Office Quality-Information Technology project in New York. Dr. Diamond earned her dual B.A./M.D. at the Medical School of SUNY at Brooklyn and her M.P.H. at the University of Medicine and Dentistry of New Jersey of Rutgers University.

T. Bruce Ferguson, Jr., M.D., is chairman of the Department of Cardiovascular Sciences at the East Carolina Heart Institute and the Brody School of Medicine at East Carolina University. He is a board-certified cardiothoracic surgeon who specializes in adult cardiothoracic surgery. He came to North Carolina from Louisiana, where he was Chief of Cardiac Surgery at Louisiana State University Health Sciences Center in New Orleans prior to Hurricane Katrina. While in Louisiana, he received funding from the Trans-

forming Healthcare Quality through Information Technology program of the Agency for Healthcare Research and Quality (AHRQ) to begin development of a longitudinal cardiovascular information system for the statewide Charity Hospital System population. He served for 6 years as the inaugural chair of the Society of Thoracic Surgeons Council on Quality, Research, and Patient Safety, which oversees all aspects of the society's national database efforts, in collaboration with the Duke Clinical Research Institute. He was principal investigator on the society's two clinical trials in quality improvement from 1999 through 2007, funded by AHRQ; this effort also led to the creation of the National Consortium of Clinical Databases. He is currently coprincipal investigator on the combined Duke–East Carolina University Clinical Site for the National Heart, Lung, and Blood Institute Cardiac Surgical Network, and he is principal investigator for the Clinical Research Skills Development Core. He is a fellow of the American Heart Association and is a member of the Quality Strategic Directions Committee and the Surgical Council for the American College of Cardiology. He received his degree in chemistry from Williams College, and received his M.D. degree from Washington University in St. Louis. He completed his training in general surgery and cardiothoracic surgery at Duke University Medical Center.

Daniel E. Ford, M.D., M.P.H., is a graduate of Cornell University and SUNY at Buffalo School of Medicine. He completed his internal medicine residency at the Johns Hopkins Hospital and received an M.P.H. from the Johns Hopkins School of Public Health. From 1985 to 1988 he was a fellow at the National Institute of Mental Health. He currently is the David M. Levine Professor of Medicine and Psychiatry at Johns Hopkins School of Medicine. Until he was appointed the vice dean for clinical investigation in 2005, he had been the director of the University Health Service since 1992. In 2007 he was appointed as director of the Johns Hopkins Institute for Clinical and Translational Research. Dr. Ford is one of the investigators that first documented the link between depression and subsequent heart disease. He has authored over 140 articles and book chapters. He has worked for many years on improving the quality of care for mental disorders in primary care and general medical settings. He was one of the coleaders for the Institute for Health Care Improvement Bureau of Primary Care Depression Collaborative. In addition, he currently is leading the evaluation for the RWJF Depression in Primary Care Program.

Mark E. Frisse, M.D., M.B.A., M.Sc., is a professor of biomedical informatics at Vanderbilt University. Working through the office of the Governor of the State of Tennessee he directs a federal- and state-sponsored health information exchange in the greater Memphis area with over 2 million records

covering the care of over 1 million individuals. He is active in a number of Markle Foundation efforts and served as a co-chair of the *Connecting for Health* Common Framework policy group developing model data-sharing agreements. He is a participant in the RWJF Project HealthDesign personal health records initiative. Dr. Frisse has worked in the academic sector, in consulting, and in pharmacy benefits management. He has served as a board member of the American Medical Informatics Association, the eHealth Initiative, the State of Tennessee Governor's eHealth Task Force, and SureScripts, LLC.

Jean Paul Gagnon, Ph.D., is director of public policy at sanofi-aventis in Bridgewater, New Jersey. He received a B.S. in pharmacy and a M.S. in pharmacy administration from the University of Connecticut and a Ph.D. in pharmacy administration from Ohio State University. He is a former professor and division head of pharmacy administration in the School of Pharmacy at UNC and has worked for sanofi-aventis for 17 years. He has written over 60 articles in peer-reviewed journals and has made presentations on a variety of issues, including Medicare, Part D, evidence-based medicine, and the effect of federal policy on the pharmaceutical industry and pharmacy practice. In 2002 he received the American Public Health Association Hugo H. Schaefer Award for outstanding volunteer contributions to the profession of pharmacy. From 1981 to 1982, he was an RWJF Health Policy Fellow in Washington, DC, and worked as a committee staff person on Energy and Commerce's Subcommittee on Health. He was chairman of the Health Outcomes Committee of the Pharmaceutical Research Manufacturers of America from 1997 until 2001, president of the American Association of Colleges of Pharmacy for the 1985–1986 year, served as treasurer and a board of trustee member of the American Pharmaceutical Association from 1991 until 1997, served as president of the International Association of Pharmaco-economics and Outcomes Research from 1996 until 1997 and was their treasurer from 1998 until 2004, and was a member of the board of trustees and treasurer for the U.S. Pharmacopeia from 1995 until 2005.

Stuart Guterman, Ph.D., has been senior program director for the Commonwealth Fund's Program on Medicare's Future since May 2005. He is responsible for the fund's research agenda on Medicare issues. He was director of the Office of Research, Development, and Information at CMS from 2002 to 2005, where he was responsible for research on Medicare and Medicaid issues, evaluation of Medicare and Medicaid programs, and developing, implementing, and conducting Medicare demonstrations. Prior to that Dr. Guterman was a senior analyst at the Congressional Budget Office, a principal research associate in the Health Policy Cen-

ter at the Urban Institute, and deputy director of the Medicare Payment Advisory Commission (MedPAC) (and its predecessor, the Prospective Payment Assessment Commission) from 1988 through 1999. Previously, Dr. Guterman was chief of institutional studies in the Health Care Financing Administration's Office of Research.

Bruce H. Hamory, M.D., is executive vice president and chief medical officer emeritus for Geisinger Health System. As such, he oversees the research activities of the Weis Center for Research and the Hood Center for Health Research as well as the medical education programs for the health system. As Geisinger's system chief medical officer from 1997 to 2008, he was responsible for guiding the activities of a 740-member multispecialty physician group practice in 40 locations serving 41 counties and the 3 Geisinger hospitals. These activities included compensation, quality and performance improvement, credentialing, research, and education. As one of the executive leaders for the system, he was involved directly in clinical operations, capital planning, and other issues for the clinic and the health system. Dr. Hamory is on the board of directors for the American Medical Group Association and serves on several national committees and panels concerned with improving the quality of medical care.

W. David Helms, Ph.D., is president and chief executive officer of AcademyHealth in Washington, DC, the professional society for health services researchers, policy analysts, and practitioners and a leading, nonpartisan resource for the best in health research and policy. Its programs are dedicated to stimulating the development, understanding, and use of the best available health services research and health policy information by public and private decision makers. In addition to leading AcademyHealth, Dr. Helms serves as president and chief executive officer of the Coalition of Health Services, AcademyHealth's advocacy arm. The coalition provides a unified voice for enhanced federal funding of health services research and health data to inform health policy and practice.

William R. Hersh, M.D., is professor and chair of the Department of Medical Informatics and Clinical Epidemiology in the School of Medicine at Oregon Health and Science University (OHSU) in Portland, Oregon. He is a leader and innovator in biomedical informatics both in education and research. Dr. Hersh has performed research in biomedical informatics in a number of areas. A current major interest of his is the health information technology (HIT) workforce, focusing on the personnel, their skills, and their training for implementing, innovating, and evaluating systems. Dr. Hersh is also active in clinical and translational research informatics. He serves as director of the biomedical informatics program of the Oregon

Clinical and Translational Research Institute and is chair of the National Informatics Steering Committee of the Clinical and Translational Science Awards program of the National Institutes of Health (NIH). His original focus of research was in the area of information retrieval, where he has authored over 100 scientific papers as well as the book, *Information Retrieval: A Health & Biomedical Perspective*. In education, Dr. Hersh serves as director of informatics educational programs at OHSU, where he has led the development of educational programs at the certificate, master's, and doctoral levels. He also initiated OHSU's efforts into distance learning for biomedical informatics. Most recently, he teamed up with the American Medical Informatics Association to launch the 10x10 program that aims to train 10,000 clinicians and others in informatics by the year 2010.

David H. Hickam, M.D., M.P.H., holds the rank of professor in the Department of Medicine at OHSU. He is a health services researcher and is director of the John M. Eisenberg Center at OHSU. He also serves as coprincipal investigator of the Health Services Research and Development Service Research Enhancement Award Program at the Portland Department of Veteran Affairs Medical Center. His research interests include clinical decision making, practice variation in both primary care and specialty settings, and patients' perceptions of their health care. The Eisenberg Center is part of the Effective Health Care Program of AHRQ and is charged with developing practical tools to assist consumers, clinicians, and policy makers in using clinical evidence for their decision making. This work uses a rigorous method of evidence translation and a product development process based on obtaining the perspectives of end users through qualitative research methods. Dr. Hickam directs an interdisciplinary team and has performed evidence translations on 15 clinical topics commissioned by AHRQ. Dr. Hickam was formerly a member of the editorial board of *Medical Decision Making*.

Erin Holve, Ph.D., M.P.H., M.P.P., is a senior manager at AcademyHealth. Dr. Holve heads AcademyHealth's work in professional development and continuing education, with a specific focus on analytic methods used in health services research. Prior to joining AcademyHealth, she worked as a consultant to AcademyHealth, creating the online methods resource, and she was part of the initial team that drafted core competencies for doctoral training in health services research. Previously, she worked for the Henry J. Kaiser Family Foundation as a senior policy analyst studying employer-sponsored health insurance and developed the online resource State Health Facts Online. She has also worked for the Department of Health and Human Services (HHS) on the Health Insurance Portability and Accountability Act and has consulted for local, regional, and national health policy organizations. She holds a Ph.D. in health services research from the Bloom-

berg School of Public Health at Johns Hopkins University and master's degrees in public health and public policy from the UC at Berkeley.

Richard A. Justman, M.D., is national medical director of UnitedHealthcare, a national health service delivery company. He works in the clinical advancement division. Dr. Justman is accountable for medical technology assessment, clinical support of pharmacy programs, and clinical support of benefit administration. He has been with UnitedHealthcare since 1993. He received his bachelor's degree from Cornell University and his M.D. degree from SUNY at Buffalo. He is board certified in pediatrics and received his postgraduate training at the University of Chicago Hospitals and Clinics and the Johns Hopkins Hospital. Dr. Justman practiced pediatrics in Minneapolis, Minnesota, for 15 years before joining UnitedHealthcare. He has served on the IOM Forum on Drug Discovery, Development, and Translation; the IOM Committee to Identify Highly Effective Clinical Services; the American Medical Association (AMA) Current Procedural Terminology codes-5 Project; the AMA Initiative to Transform Medical Education; and an expert panel developing an evidence report on diabetes education for children with type I diabetes, commissioned by AHRQ. Dr. Justman served on the AHRQ Stakeholders Panel for 2008 and 2009. He speaks frequently to external audiences on the use of clinical evidence to determine the safety and effectiveness of new and emerging medical treatments.

Douglas B. Kamerow, M.D., M.P.H., is a chief scientist at RTI International, a not-for-profit research company based in Research Triangle Park, North Carolina, where he leads research and evaluation projects in the areas of health policy, childhood obesity, preventive medicine, and evidence-based practice. Dr. Kamerow was formerly director of several key programs at AHRQ, including the Center for Practice and Technology Assessment and the Office of the Forum for Quality and Effectiveness in Health Care. He retired from the Public Health Service Commissioned Corps in 2001 with the rank of assistant surgeon general. He is a family physician and preventive medicine specialist with clinical, public health, and epidemiologic training. Dr. Kamerow is presently associate editor of *BMJ* (formerly *British Medical Journal*), for whom he writes a regular column on health policy issues; he previously served for 2 years as the *BMJ*'s U.S. editor. He is also a healthcare commentator for *All Things Considered* on National Public Radio. Finally, Dr. Kamerow is professor of clinical family medicine at Georgetown University in Washington, DC, where he teaches medical students and family medicine residents.

Mark B. McClellan, M.D., Ph.D., became the director of the Engelberg Center for Healthcare Reform at the Brookings Institution in July 2007. The center studies ways to provide practical solutions for access, quality,

and financing challenges facing the U.S. healthcare system. In addition, Dr. McClellan is the Leonard D. Schaeffer Chair in Health Policy Studies. Dr. McClellan has a highly distinguished record in public service and in academic research. He is the former administrator for CMS (2004 to 2006) and the former commissioner of the FDA (2002 to 2004). He also served as a member of the President's Council of Economic Advisers and senior director for healthcare policy at the White House (2001 to 2002). In these positions, he developed and implemented major reforms in health policy. Dr. McClellan was also an associate professor of economics and associate professor of medicine (with tenure) at Stanford University, from which he was on leave during his government service. He directed Stanford's Program on Health Outcomes Research and was also associate editor of the *Journal of Health Economics* and coprincipal investigator of the Health and Retirement Study, a longitudinal study of the health and economic status of older Americans. His academic research has been concerned with the effectiveness of medical treatments in improving health, the economic and policy factors influencing medical treatment decisions and health outcomes, the impact of new technologies on public health and medical expenditures, and the relationship between health status and economic well-being. Dr. McClellan is a member of the IOM of the National Academy of Sciences and a research associate of the National Bureau of Economic Research. A graduate of the University of Texas at Austin, Dr. McClellan earned his M.P.A. from Harvard's Kennedy School of Government in 1991, his M.D. from the Harvard–Massachusetts Institute of Technology (MIT) Division of Health Sciences and Technology in 1992, and his Ph.D. in economics from MIT in 1993.

J. Michael McGinnis, M.D., M.P.P., is a long-time contributor to national and international health policy leadership, now senior scholar at the IOM, and executive director of the IOM Roundtable on Evidence-Based Medicine. He is also an elected member of the IOM. He previously was senior vice president at the RWJF, and, unusual for political posts, held continuous appointment through the Carter, Reagan, Bush, and Clinton administrations, with responsibility for coordinating activities and policies in disease prevention and health promotion. Programs and policies created and launched at his initiative include the Healthy People process for setting national health objectives, the U.S. Preventive Services Task Force, the *Dietary Guidelines for Americans* (with the U.S. Department of Agriculture), the *Ten Essential Services of Public Health*, the RWJF Health and Society Scholars Program, the RWJF Young Epidemiology Scholars Program, and the RWJF Active Living family of programs. Internationally, he chaired the World Bank/European Commission Task Force on postwar

reconstruction of the health sector in Bosnia and worked both as field epidemiologist and state coordinator for the World Health Organization's successful smallpox eradication program in India.

Robert H. Miller, Ph.D., is professor of health economics in residence at the Institute for Health and Aging and in the Department of Social and Behavioral Sciences at UC at San Francisco. He received his doctorate in economics from the University of Michigan in 1987 and has been a faculty member since 1989. Dr. Miller conducts research on the economics of HIT and organizational change in ambulatory care settings, including health policy-oriented research on electronic health records (EHRs), patient–provider e-health capabilities, and community-wide electronic clinical data exchange. Dr. Miller is especially interested in EHR use for quality improvement in solo/small groups and community health centers. Dr. Miller is a member of the Joint Commission's HIT Advisory Panel and Lemetra's Quality Improvement Advisory Committee, was a member of the Connecting for Health Working Group on the economics of HIT, and is serving and has served as a member on HIT expert panels. Dr. Miller also has published on such topics as health maintenance organization (HMO) versus non-HMO plan performance, efforts by large medical groups to reduce medical injuries, the effects of managed care on physician practice change, as well as on acute and long-term care services for chronically impaired elders.

Nancy H. Nielsen, M.D., Ph.D., is an internist from Buffalo, New York, and current president of the AMA. She was elected speaker of the AMA House of Delegates in June 2003 and reelected in 2005. She is a delegate from New York and previously served two terms on the AMA Council on Scientific Affairs. Dr. Nielsen has also served as a member on the National Patient Safety Foundation Board of Directors, the Commission for the Prevention of Youth Violence, the Task Force on Quality and Patient Safety, the HHS Secretary's Advisory Committee on Regulatory Reform, and as the AMA representative to the National Quality Forum, Physicians Consortium for Performance Improvement, the Hospital Quality Alliance, and the Ambulatory Care Quality Alliance. She holds a Ph.D. in microbiology and received her M.D. from the SUNY School of Medicine and Biomedical Sciences in Buffalo, where she is clinical professor of medicine and senior associate dean for medical education. She has served as a trustee of SUNY and as a member of the board of directors of Kaleida Health—a five-hospital system in western New York. She is currently associate medical director for quality and interim chief medical officer at Independent Health Association, a major health insurer in New York.

Parashar B. Patel, M.P.A., joined Boston Scientific Corporation as vice president of health economics and reimbursement for the Healthcare Strategies and Programs group in 2003. Dr. Patel is responsible for the company's corporate and site health economics and reimbursement functions, chiefly focusing on the development and implementation of global strategic, reimbursement, and legislative initiatives. He is also closely involved in health economics analysis and outcomes research for the company. Prior to joining, Dr. Patel was deputy director of the Hospital and Ambulatory Policy Group in the Center for Medicare Management at CMS. The group was responsible for Medicare payment policy for a wide range of acute and ambulatory care services, including inpatient and outpatient hospital services and physician services. He has extensive experience in healthcare financing policy through his work with the American Association of Health Plans, the Office of (then) Senate Majority Leader George J. Mitchell, the U.S. Office of Management and Budget, and Connecticut's Medicaid agency. He holds a B.A. in political science and a master of public affairs from the University of Connecticut.

Steve E. Phurrough, M.D., M.P.A., C.P.E., is the director of the coverage and analysis group for CMS. Using evidence-based medicine principles, he assists in developing national policy on the appropriate devices, diagnostics, and procedures that should be provided by the Medicare program. Dr. Phurrough joined CMS in 2001 as the director of the division of medical and surgical services in the coverage and analysis group after completing a long and distinguished career in the U.S. Army. In addition to being a practicing family practitioner, his military career also included managing Department of Defense regional healthcare delivery systems, creating national and international healthcare policy for the U.S. Army, and developing practice guidelines. Dr. Phurrough received his M.D. from the UAB and an M.P.A. from the University of Colorado in Colorado Springs. He is board certified by the American Board of Family Practice and is a certified physician executive.

William Z. Potter, M.D., Ph.D., earned his degrees at Indiana University, after which he functioned in positions of increasing responsibility and seniority over the next 25 years at the NIH focused on translational neuroscience. While at the NIH, Dr. Potter was widely published and appointed to many societies, committees, and boards, a role that enabled him to develop a wide reputation as an expert in psychopharmacological sciences and championing the development of novel treatments for central nervous system (CNS) disorders. Dr. Potter left the NIH in 1996 to accept a position as executive director and research fellow at Lilly Research Laborato-

ries, specializing in the neuroscience therapeutic area and in 2004 joined Merck Research Laboratories as vice president of clinical neuroscience, then the newly created position of translational neuroscience in 2006. His experience at Lilly and Merck in identifying, expanding, and developing methods of evaluating CNS effects of compounds in human brain cover state-of-the-art approaches across multiple modalities. These include brain imaging and cerebrospinal fluid proteomics (plus metabolomics), as well as the development of more sensitive clinical, psychophysiological, and performance measures that allow a range of novel targets to be tested in a manner that actually addresses the underlying hypotheses. Dr. Potter has become a widely recognized champion for the position that more disciplined hypothesis testing of targets in humans is the best near-term approach to moving CNS drug development forward.

John W. Rowe, M.D., is currently a professor in the Department of Health Policy and Management at the Columbia University Mailman School of Public Health. From 2000 until his retirement in late 2006, Dr. Rowe served as chairman and chief executive officer of Aetna, Inc., one of the nation's leading health care and related benefits organizations. Before his tenure at Aetna, from 1998 to 2000, Dr. Rowe served as president and chief executive officer of Mount Sinai–NYU Health, one of the nation's largest academic healthcare organizations. From 1988 to 1998, prior to the Mount Sinai–NYU Health merger, Dr. Rowe was president of the Mount Sinai Hospital and the Mount Sinai School of Medicine in New York City. Before joining Mount Sinai, he was a professor of medicine and the founding director of the Division on Aging at the Harvard Medical School, as well as chief of gerontology at Boston's Beth Israel Hospital. He has authored over 200 scientific publications, mostly on the physiology of the aging process, including a leading textbook of geriatric medicine, in addition to more recent publications on healthcare policy. Dr. Rowe has received many honors and awards for his research and health policy efforts regarding care of elderly people. He was director of the MacArthur Foundation Research Network on Successful Aging and is coauthor, with Dr. Robert Kahn, of *Successful Aging* (Pantheon, 1998). Dr. Rowe currently leads the MacArthur Foundation's Initiative on an Aging Society. He was elected a member of the IOM and a fellow of the American Academy of Arts and Sciences. In addition, Dr. Rowe is a former member of MedPAC. He is also chairman of the board of trustees at the University of Connecticut and the Marine Biological Laboratory in Woods Hole, Massachusetts.

Richard N. Shiffman, M.D., M.C.I.S., is professor of pediatrics and associate director of the Center for Medical Informatics at Yale School of Medicine. Dr. Shiffman is a fellow of the American College of Medical

Informatics and the American Academy of Pediatrics (AAP). Dr. Shiffman serves on the AAP's Partners for Policy Implementation. He has served on several guideline development panels for national professional societies and on the AAP's Steering Committee on Quality Improvement and Management. In addition, Dr. Shiffman leads the GuideLines Into DEcision Support Project—an AHRQ-sponsored collaboration that is demonstrating transparent, systematic, and replicable processes for transforming guideline knowledge into computer-mediated decision support.

Jean R. Slutsky, P.A., M.S.P.H., has directed the Center for Outcomes and Evidence (COE) at AHRQ since June 2003. Prior to Ms. Slutsky's appointment as director of COE, she served as acting director of the Center for Practice and Technology Assessment at AHRQ. Most recently, Ms. Slutsky has implemented a comparative effectiveness research program that includes evidence synthesis, evidence generation, and evidence communication. The Effective Health Care Program is authorized under Section 1013 of the Medicare Prescription Drug, Improvement, and Modernization Act. Ms. Slutsky oversees the Evidence-Based Practice Center Program; Technology Assessment Program; extramural and intramural research portfolios concerning translating research into practice, outcomes, and effectiveness research, including pharmaceutical outcomes, and cost-effectiveness analyses; and the National Guideline, Quality Measures, and QualityTools Clearinghouses. She is a vice chair of the Guidelines International Network and a member of the editorial board of *Implementation Science*. Prior to becoming acting director of the Center for Practice and Technology Assessment, Ms. Slutsky served as project director of the U.S. Preventive Services Task Force, an internationally recognized panel of experts who make evidence-based recommendations on clinical preventive services. Ms. Slutsky received her Bachelor of Science degree at the University of Iowa, a Master's of Science in Public Health (health policy and administration) from UNC at Chapel Hill, and trained as a physician assistant at the University of Southern California.

Donald M. Steinwachs, Ph.D., is a professor in the Department of Health Policy and Management and director of the Health Services Research and Development Center at the Johns Hopkins University Bloomberg School of Public Health. Dr. Steinwachs's research seeks to identify opportunities to improve quality of health care and patient outcomes and, when feasible, evaluate promising quality improvement interventions. His previous research includes studies of medical effectiveness and patient outcomes for individuals with specific medical (e.g., asthma), surgical (e.g., cataract surgery), and psychiatric (e.g., schizophrenia) conditions. Dr. Steinwachs has contributed to the literature on the impact of managed care and payments

systems on access to care, quality, utilization, and cost. He was a codeveloper of the widely used adjusted clinical groups case mix adjustment. He has developed methods for measuring provider continuity, needs and unmet needs for care, and measures of the timeliness of care. He has a particular interest in the role of routine management information systems (MIS) as source of data for evaluating the effectiveness and cost of health care. This includes work on the integration of outcomes management systems with existing MIS in managed care settings. He is a member of the IOM and its Board of Health Care Services. A member of the National Committee on Vital and Health Statistics since 2002, he chairs the Subcommittee on Populations and serves on the executive committee. He also serves on the board of Mathematica Policy Research, Inc. Dr. Steinwachs holds a B.S. in engineering mathematics and an M.S. in systems engineering from the University of Arizona and a Ph.D. in operations research from Johns Hopkins University.

Robert J. Temple, M.D., is director of the Office of Medical Policy of the FDA's Center for Drug Evaluation and Research and is also acting director of the Office of Drug Evaluation I (ODE-I). ODE-I is responsible for the regulation of cardio-renal, neuropharmacologic and psychopharmacologic drug products. The Office of Medical Policy is responsible for regulation of promotion through the Division of Drug Marketing, Advertising, and Communication and for assessing quality of clinical trials. Dr. Temple has a long-standing interest in the design and conduct of clinical trials and has written extensively on this subject, especially on the choice of control groups in clinical trials, evaluation of active control trials, trials to evaluate dose–response, and trials using "enrichment" designs.

Sean R. Tunis, M.D., M.Sc., is the founder and director of the Center for Medical Technology Policy in San Francisco, where he works with healthcare decision makers, experts, and stakeholders to improve the value of clinical research on new and existing medical technologies. He consults with a range of domestic and international healthcare organizations on issues of comparative effectiveness, evidence-based medicine, clinical research, and technology policy. Through September 2005, Dr. Tunis was the director of the Office of Clinical Standards and Quality and chief medical officer at 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 Technology Assessment 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 biology and history of science from the Cornell University School of Agriculture, and an M.D. and M.A. in health services research from the Stanford University School of Medicine. Dr. Tunis did his residency training at UC at Los Angeles 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 schools of medicine.

Steven A. Wartman, M.D., Ph.D., M.A.C.P., became the third president of the Association of Academic Health Centers, based in Washington, DC, in July 2005. Before assuming this position, he was the executive vice president for academic and health affairs and dean of the school of medicine at the University of Texas Health Science Center in San Antonio. Prior to his tenure in San Antonio, Dr. Wartman held a number of other positions in academic medicine, including the Edward Meilman Distinguished Chairman of Medicine and physician-in-chief at Long Island Jewish Medical Center and professor of medicine at Albert Einstein College of Medicine. Dr. Wartman began his career at Brown University and Rhode Island Hospital where he founded the Division of General Internal Medicine and the General Internal Medicine Residency Program. A graduate of Cornell University, Dr. Wartman received both his M.D. and Ph.D. degrees from Johns Hopkins University. He was an RWJF Clinical Scholar at Johns Hopkins, a Henry Luce Scholar in Indonesia, and is a past president of the Society of General Internal Medicine. He is a board-certified internist, a sociologist, and a master of the American College of Physicians. His publications and interests lie in the areas of the structure and function of academic health centers, healthcare delivery, health policy, medical education, and academic leadership. He currently is a distinguished professor in the Department of Medicine at Georgetown University and is an adjunct professor of medicine at George Washington and Johns Hopkins universities.

Gail R. Wilensky, Ph.D., is an economist and a senior fellow at Project HOPE who analyzes and develops policies relating to healthcare reform

and to ongoing changes in the healthcare environment. Dr. Wilensky is a commissioner on the World Health Organization's Commission on the Social Determinants of Health and an elected member of the IOM of the National Academies where she served two terms on its governing council, is vice chair of the Maryland Health Care Commission, and serves as a trustee of the Combined Benefits Fund of the United Mineworkers of America and the National Opinion Research Center. She is an advisor to the RWJF and the Commonwealth Fund, past chair of the board of directors of AcademyHealth, and director on several corporate boards. From 1990 until 1992 she was administrator of the Health Care Financing Administration, directing the Medicare and Medicaid programs. She also served as deputy assistant to President George H.W. Bush for policy development, advising him on health and welfare issues from 1992 to 1993. From 1997 to 2001 she chaired MedPAC, which advises Congress on payment and other issues relating to Medicare, and from 1995 to 1997 she chaired the Physician Payment Review Commission. From 2001 to 2003 she co-chaired the President's Task Force to Improve Health Care Delivery for Our Nation's Veterans, which covered health care for both veterans and military retirees. In 2007 she was appointed to the President's Commission on Care for America's Returning Wounded Warriors and also as the co-chair of the Department of Defense task force on the future of military health care. Dr. Wilensky testifies frequently before congressional committees, acts as an advisor to members of Congress and other elected officials, and speaks nationally and internationally before professional, business, and consumer groups. She received a bachelor's degree in psychology and a Ph.D. in economics at the University of Michigan.

Mary E. Woolley, M.A., is the president of Research!America, the nation's largest not-for-profit alliance working to make research to improve health a higher national priority. Research!America's 500-plus organizational members represent the voices of 125 million Americans. Ms. Woolley is an elected member of the IOM and a fellow of the AAAS. She serves on several boards and committees, including the IOM Health Sciences Policy Board, the National Council for Johns Hopkins Nursing, and the board of overseers of the Harvard School of Public Health. She is a founding member of the board of associates of the Whitehead Institute for Biomedical Research. She has served as president of the Association of Independent Research Institutes, as editor of the *Journal of the Society of Research Administrators*, as a reviewer for the NIH and National Science Foundation, and as a consultant to several research organizations. Ms. Woolley has a 25-year editorial and publication history on science advocacy and research-related topics. She is a sought-after speaker and is frequently interviewed by science, news, and policy journalists.

Appendix H

Workshop Attendee List

Patricia Adams
National Pharmaceutical Council

Shilpa Amin
Agency for Healthcare Research and Quality

Carol Ashton
The Methodist Hospital

Maines Aviles-Santa
National Heart, Lung, and Blood Institute

Dennis Barbour
Society for Investigative Dermatology

Magda Barini-Garcia
Health Resources and Services Administration

Jane Barlow
Medco Health Solutions

Mercedes Barrs
Amylin Pharmaceuticals, Inc.

Eric B. Bass
Johns Hopkins University

James Beachy
American College of Cardiology

Lorne A. Becker
Cochrane Collaboration

Rachel Behrman
Food and Drug Administration

Jesse A. Berlin
Johnson & Johnson

Eugene H. Blackstone
Cleveland Clinic

Carolyn Bloch
Bloch Consulting Group

Meryl Bloomrosen
American Medical Informatics Association

Carmella A. Bocchino
America's Health Insurance Plans

Marilyn Sue Bogmer
Institute for the Study of Human Error

Randall Bovbjerg
The Urban Institute

Jeffrey Brady
Agency for Healthcare Research and Quality

Neal Brandes
U.S. Agency for International Development

Jennifer Bright
Society for Healthcare Epidemiology of America

Lynda Bryant-Comstock
GlaxoSmithKline

Shannon Bush
FasterCures

Kathleen A. Buto
Johnson & Johnson

Daniel Campion
AcademyHealth

Timothy S. Carey
University of North Carolina at Chapel Hill

Jenny Carloss
Avalere

Linda Carlson
EMD Serono, Inc.

Kristin Carman
American Institutes for Research

Betsy Carrier
National Association of Public Hospitals

Linda Carter
Johnson & Johnson

Kalipso Chalkidou
Johns Hopkins University

Stephanie Chang
Agency for Healthcare Research and Quality

Katherine Cherry
O'Neill Institute for National and Global Health Law

Benjamin K. Chu
Kaiser Foundation Health Plan and Hospitals

John Clarke
Drexel University

Perry Cohen
Parkinson Pipeline Project

Elaine Collier
National Center for Research Resources

Janet Craig
Clemson University and Health Sciences South Carolina

Robert Crane
Kaiser Permanente

Travis Crytzer
Association of American Medical Colleges

Frederick Curro
New York University

Donald DeNucci
National Institute of Dental and Craniofacial Research

Anuj Desai
Johnson & Johnson

Don E. Detmer
American Medical Informatics Association

Carol Diamond
The Markle Foundation

Louis Diamond
Thomson Reuters

Molla Donaldson
MSD Health

Anna Legreid Dopp
Office of U.S. Senator Joe Lieberman

Denise Dougherty
Agency for Healthcare Research and Quality

Gregory Downing
Department of Health and Human Services

Subash Duggirala
Centers for Medicare & Medicaid Services

Eva DuGoff
Office of U.S. Senator Ron Wyden

Maggie Elestwani
Memorial Hermann–Texas Medical Center

Haim Erder
Forest Research Laboratories

Henry Ernstthal
Ernstthal & Associates

Lynn Etheredge
George Washington University

Irene Farkas-Conn
University of Chicago (formerly)

Laurie Feinberg
Department of Health and Human Services

T. Bruce Ferguson, Jr.
East Carolina Heart Institute, Brody School of Medicine

Reuven Ferziger
Johnson & Johnson

Rosemarie Filart
National Institutes of Health

Contessa Fincher
EMD Serono, Inc.

Jack Fitzgibbons
CarePath, Inc.

Anamarie Ford
Association for Community Affiliated Plans

Daniel E. Ford
Johns Hopkins School of Medicine

Raymond Formanek
Food and Drug Administration

Steven Fox
Agency for Healthcare Research and Quality

Seth Frazier
Geisinger Health System

Kathleen Frisbee
Department of Veterans Affairs

Mark E. Frisse
Vanderbilt University

Jean Paul Gagnon
sanofi-aventis

Janice Genevro
Agency for Healthcare Research and Quality

Barry Gershon
Wyeth Pharmaceutical

Udi Ghitza
National Institute on Drug Abuse, National Institutes of Health

David Gollaher
California Healthcare Institute

Stephen Gorshow
(Great-West) CIGNA

Linda Greenberg
Agency for Healthcare Research and Quality

Mary Greene
Booz Allen Hamilton

Stuart Guterman
The Commonwealth Fund

Bruce H. Hamory
Geisinger Health System

Anthony Hayward
National Institutes of Health

Jan Heinrich
Health Policy R&D

Amy Heller
American College of Cardiology

W. David Helms
AcademyHealth

William R. Hersh
Oregon Health and Science University

David H. Hickam
Eisenberg Center

Giselle Hicks
National Breast Cancer Coalition Fund

Larry Holmes
A.I. duPont Hospital for Children

Erin Holve
AcademyHealth

Mark Horn
Pfizer, Inc.

Jane Horvath
Merck and Co., Inc.

Ardis Hoven
American Medical Association

Anna Schwamlein Howard
American Association for Retired Persons

Chris Izui
BlueCross BlueShield Association

Deborah Jaffe
National Cancer Institute

Len Jokubaitis
Ortho McNeil Janssen Scientific Affairs, LLC

Aranthan Jones II
U.S. Congress/Majority Whip

Diane Jones
Microsoft

Richard A. Justman
UnitedHealthcare

Douglas B. Kamerow
RTI International

Toshiko Kaneda
Population Reference Bureau

Randee Kastner
Center for Medical Technology Policy

Elisabeth Kato
Hayes, Inc.

Maureen Kelly
Biogen Idec

James Kirby
Agency for Healthcare Research and Quality

Neil Kirschner
American College of Physicians

Mary Jane Koren
The Commonwealth Fund

Mahesh Krishnan
Amgen

Steven Krosnick
National Cancer Institute

Crystal Kuntz
Astellas

Arnold Kuzmack
Food and Drug Administration

Mollie Lane
U.S. Senate Finance Committee

Lisa Lang
National Library of Medicine

Michael Lauer
National Heart, Lung, and Blood Institute

Renee Laughlin
American College of Cardiology

Teresa Lee
Advanced Medical Technology Association

Jean LeMasurier
Gorman Health Group

Daniel Leonard
National Pharmaceutical Council

Kenneth Lin
Agency for Healthcare Research and Quality

Keith Lind
American Association for Retired Persons

Ramon Llamas
Men's Health Network, University of Southern California

David Longnecker
Association of American Medical Colleges

Jacqueline Lorenc
Eli Lilly & Co.

Serena Lowe
EMD Serono, Inc.

Bryan Luce
United BioSource Corporation

Danica Marinac-Dabic
Food and Drug Administration

Norman Marks
Food and Drug Administration

Karen Matsuoka
Ways and Means/Health Subcommittee, U.S. House of Representatives

Elliot Maxwell
Johns Hopkins University

Mark B. McClellan
Brookings Institution

Kathleen McCormick
Science Applications International Corporation

Scott McKenzie
Centocor Ortho Biotech Outcomes Research

Kathryn McLaughlin
America's Health Insurance Plans

Erik Mettler
Food and Drug Administration

Mary Metzger
University North Carolina at Charlotte

Nancy Miller
National Institutes of Health

Robert H. Miller
University of California at San Francisco

Wilhelmine Miller
George Washington University School of Public Health and Health Services

Carol Monaco
American Osteopathic Association

Hazel Moran
Mental Health America

Barbara Myklebust
George Washington University

Nancy H. Nielsen
American Medical Association

John O'Donnell
AstraZeneca

Kavita Patel
Office of U.S. Senator Edward Kennedy

Parashar B. Patel
Boston Scientific Corporation

Ramesh Patel
RAND Corporation

Steve Pearson
Institute for Clinical and Economic Review

Steve Pelletier
Editorial Consulting

Eleanor M. Perfetto
Pfizer, Inc.

Doris Peter
Consumer Reports

Raymond Petryshyn
National Cancer Institute

Steve E. Phurrough
Centers for Medicare & Medicaid Services

William Pilkington
Cabarrus Health Alliance

Susan Pingleton
University Healthsystem Consortium

Particia Pittman
AcademyHealth

William Z. Potter
Merck Research Laboratories

Eva Powell
National Partnership for Women and Families

Charlene Quinn
University of Maryland School of Medicine

Meaghan Quinn
Association of American Medical Colleges

Greg Raab
Raab & Associates, Inc.

Susan Raetzman
Thomson Reuters

Shaalini Ramanadhan
Genentech

Nancy Ray
Medicare Payment Advisory Commission

Brian Raymond
Kaiser Permanente Institute for Health Policy

Eugene Rich
National Institutes of Health

Patrick Richard
George Washington University

Nancy Ridenour
U.S. House of Representatives

John C. Ring
American Heart Association

Kenneth Rogers
A.I. duPont Hospital

Lou Rossiter
Schroeder Center for Healthcare Policy

John W. Rowe
Columbia University

Jody Sachs
National Institutes of Health

Stephanie Safdi
CSP

Francois Sainfort
University of Minnesota

Rajni Samavedam
Booz Allen Hamilton

Karen Sanders
American Psychiatric Association

Michael Sayre
National Center for Research Resources

Jyme Schafer
Centers for Medicare & Medicaid Services

Adam L. Scheffler
American Medical Association

David Schulke
American Health Quality Association

Sven Seyffert
Medtronic

Gail Shearer
Consumers Union

Richard N. Shiffman
Yale School of Medicine

George Silberman
Cancer Policy Group, LLC

Dee Simons
Biogen Idec

Ken Skodacek
Capital Medical Device Consulting

Jean R. Slutsky
Agency for Healthcare Research and Quality

Nancy Smith
American College of Cardiology

Patricia Smith
Alliance of Community Health Plans

James Sorace
Department of Health and Human Services

Steven Sparenborg
National Institute on Drug Abuse

Donald M. Steinwachs
Johns Hopkins University

Mark Stewart
American College of Cardiology

Sally Stiebel
Mehlman Vogel Castagnetti

Lisa Summers
National Partnership for Women & Families

Andy Swire
Amgen

Betty Tai
National Institute on Drug Abuse

Robert J. Temple
Food and Drug Administration

Mae Thamer
Medical Technology & Practice Patterns Institute

Van Thompson
New York University College of Dentistry

Sean R. Tunis
Center for Medical Technology Policy

Yvonne Vargas
Federal government

Don Vena
The EMMES Corporation

Robert Wah
Computer Science Corporation

Paul Wakim
National Institute on Drug Abuse

Mary Walchak
Becton, Dickinson, and Company

Steven A. Wartman
Association of Academic Health Centers

Joel Weintraub
Hofstra Law School

Harlan Weisman
Johnson & Johnson

Larry Westfall
Johnson & Johnson

Kimberly Westrich
National Pharmaceutical Council

Danielle Whicher
Center for Medical Technology Policy

Anne White-Olson
National Institutes of Health

Gail R. Wilensky
Project HOPE

Deb Williams
Baxter Healthcare

Scott Williams
Men's Health Network

Deborah Willis-Fillinger
Health Resources and Services Administration

Mary Woolley
Research!America

Nelda Wray
The Methodist Hospital

Bill Wright
Healthcare Navigators

John Yeh
University of Buffalo, U.S. Agency for International Development

Laura Zick
Eli Lilly and Company

Joan Zlotnik
Institute for the Advancement of Social Work Research

OTHER PUBLICATIONS IN THE LEARNING HEALTH SYSTEM SERIES

The Learning Healthcare System

Evidence-Based Medicine and the Changing Nature of Health Care

Leadership Commitments to Improve Value in Health Care: Finding Common Ground

Value in Health Care: Accounting for Cost, Quality, Safety, Outcomes, and Innovation

Redesigning the Clinical Effectiveness Research Paradigm: Innovation and Practice-Based Approaches

Clinical Data as the Basic Staple of Health Learning: Creating and Protecting a Public Good

The Healthcare Imperative: Lowering Costs and Improving Outcomes

Engineering a Learning Healthcare System: A Look at the Future

The Digital Infrastructure for the Learning Health System: The Foundation for Continuous Improvement in Health and Health Care