Contributors Include:

- Lynn M. Etheredge
- David M. Eddy
- Sean R. Tunis & Colleagues
- Joel Kupersmith & Colleagues
- Paul J. Wallace
- Walter F. Stewart & Colleagues
- Perspectives from AHRQ, Kaiser, VA
- And Much More

HEALTH AFFAIRS

A Rapid-Learning Health System

Web-Exclusive Collection 2007

A Supplement To *Health Affairs*

www.healthaffairs.org

Published by Project HOPE

[This page is intentionally blank]

HEALTH AFFAIRS

A Rapid-Learning Health System

Web-Exclusive Collection 2007

A Supplement To *Health Affairs*

Project HOPE *Bethesda, Maryland*

The People-to-People Health Foundation, Inc.
(Project HOPE), Millwood, Virginia
HEALTH AFFAIRS Editorial Offices: Bethesda, Maryland

Published by Project HOPE
7500 Old Georgetown Road, Suite 600
Bethesda, Maryland 20814-6133
Tel: 301-656-7401 Fax: 301-654-2845
http://www.healthaffairs.org

A Rapid-Learning Health System: Web-Exclusive Collection 2007 (A Supplement To Health Affairs) / John K. Iglehart, Editor
ISSN 0278-2715 (print), ISSN 1544-5208 (electronic)
Printed and bound in the United States of America.

The paper used in this publication meets minimum requirements of the American National Standard for Information Sciences—Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984.
Printed on recycled paper.

Contents

A Rapid-Learning Health System
Web-Exclusive Collection 2007
A Supplement To *Health Affairs*

EDITOR'S NOTE:
Web-Exclusive papers in this reprint volume were originally published online 26 January 2007. Here is the suggested citation style for *Health Affairs* Web Exclusives:

J.B. Perlin and J. Kupersmith, "Information Technology and the Inferential Gap," *Health Affairs* 26, no. 2 (2007): w193-w194 (published online 26 January 2007; 10.1377/hlthaff.26.2.w192).

Rapid Learning: Getting Technology Into Practice

PROLOGUE: Amid persistent concerns about performance and quality, the health sector remains ambivalent about electronic health records (EHRs). Champions of accelerated adoption of health information technology (IT) have been unable to generate a groundswell of demand, despite excellent arguments for health IT's potential to save money, improve quality, and transform care. It may be, though, that the strongest argument for speeding IT adoption is still largely below the radar.

The dramatic pace of biomedical innovation has dazzled America but created nagging tensions as well. Our insatiable demand for new drugs and technologies is driving unsustainable growth in health spending. An explosion of new knowledge has strained clinicians' learning capacity and fostered subspecialization and fragmentation of care. Clinical research and regulatory capabilities are swamped with urgent questions about the safety and effectiveness of new treatments.

But on scattered islands within the dominant system, promising approaches to managing innovation are beginning to surface, and their foundation is the EHR. In organizations such as the Veterans Health Administration (VHA), Kaiser Permanente, and the Geisinger Health System, the richness of data capture in fully deployed patient record systems is enabling clinicians and researchers to answer practical questions about safety, effectiveness, and cost more efficiently than the traditional process of randomized clinical trials (RCTs) possibly could.

The implications of these approaches for the future of "rapid learning" are spelled out in an overview paper by Lynn Etheredge. "An inadequate knowledge base limits initiatives to improve health system performance," Etheredge writes. "With large, computer-searchable databases, studies that would now take years will be doable, at low expense, in a matter of weeks, days, or hours." Case studies accompanied by commentaries explore how EHR database research is being used at the VHA (for diabetes research and care), Kaiser (for cancer research and care), and Geisinger (to close the "inferential gap" between RCTs and real-world clinical decisions). David Eddy offers his vision for a health system that will use predictive models from large, merged databases of EHRs to advance the biomedical sciences as well as clinical care. Sean Tunis and colleagues suggest strategies to use large new government clinical care databases to support Medicare coverage decisions, comparative effectiveness studies, and postmarket drug safety surveillance.

The rapid-learning efforts described here were originally presented at a March 2006 conference in Washington, D.C., organized by Etheredge and *Health Affairs* and sponsored by the Robert Wood Johnson Foundation. The publication of the papers is also supported by Kaiser Permanente and the federal Agency for Healthcare Research and Quality.

DOI 10.1377/hlthaff.26.2.106

A Rapid-Learning Health System

What would a rapid-learning health system look like, and how might we get there?

by Lynn M. Etheredge

ABSTRACT: Private- and public-sector initiatives, using electronic health record (EHR) databases from millions of people, could rapidly advance the U.S. evidence base for clinical care. Rapid learning could fill major knowledge gaps about health care costs, the benefits and risks of drugs and procedures, geographic variations, environmental health influences, the health of special populations, and personalized medicine. Policymakers could use rapid learning to revitalize value-based competition, redesign Medicare's payments, advance Medicaid into national health care leadership, foster national collaborative research initiatives, and design a national technology assessment system. [*Health Affairs* 26, no. 2 (2007): w107–w118 (published online 26 January 2007; 10.1377/hlthaff.26.2.w107)]

NATIONAL HEALTH POLICIES THAT USE government price setting and market competition are running out of steam. Neither approach is doing well at increasing the value Americans get from their "highest-in-the-world" medical expenditures. And neither government price setting nor markets has been effective at dealing with rising health costs driven by new technologies. With advances in human genome research and a doubling of the National Institutes of Health (NIH) budget in recent years, to $28 billion per year, an even faster stream of new products and therapies might be emerging. Strategies to encourage rapid learning—to quickly develop new evidence for daily medical practice and policy—might be able to increase the value of health care, assess such new technologies, and avoid draconian cost cutting.

■ **Electronic records.** Electronic health record (EHR) databases now being built by large organized delivery systems will dramatically expand the nation's research capacity. They will make it possible to include clinical experience from tens of millions of patients annually in computer-searchable databases for collaborative research. For example, the Department of Veterans Affairs (VA) is building a research database drawing on more than eight million patient records from its Veterans Health Information Systems and Technology Architecture (VistA) system. In the private sector, Kaiser Permanente is leveraging its EHR investments to create a national research

Lynn Etheredge (lynneth1@aol.com) is a consultant with the Rapid Learning Project at the George Washington University in Washington, D.C.

DOI 10.1377/hlthaff.26.2.w107

database from its eight million enrollees; the Geisinger Health System in Pennsylvania is developing a similar capability, drawing on two million patients.

■ **Beyond EHRs.** The Cancer Research Network, sponsored by the National Cancer Institute (NCI) with the HMO Research Network, includes eleven health systems with more than ten million enrollees in a new national model for virtual research organizations.[1] The Vaccine Safety Datalink network at the Centers for Disease Control and Prevention (CDC) includes seven health maintenance organizations (HMOs) and six million patient records.[2] The American Medical Group Association, whose membership of multispecialty practices treats more than fifty million patients annually, has started a collaborative database with 1.5 million records and plans for expansion to 8–15 million patient records in the next three to five years.[3] These new research capacities augment research registries already being sponsored by physician specialty societies; the NIH's Roadmap plans for genetic research, clinical trials databases, and clinical research networks; and new prescription drug databases at the Centers for Medicare and Medicaid Services (CMS).

■ **Potential for rapid learning.** EHR research will not replace randomized clinical trials as a definitive research method for specific issues, but it will offer the capacity for real-time learning from the experience of tens of millions of people and will greatly increase the ability to generate and test hypotheses. If a rapid-learning strategy that uses all of the research and data capabilities that can be brought to bear is successful, the U.S. lead in biomedical science could be matched by its learning about how to deliver the most value for its health care spending.

This paper offers a national health policy perspective on these rapid-learning opportunities. First, it briefly reviews how current national policies leave major gaps and too many uncertainties in the evidence base for clinical care. Second, it discusses how a rapid-learning strategy could fill some of these knowledge gaps quickly. Third, it describes a new generation of national health policies to leverage such new knowledge into rapid advances in health system performance. Although there are many questions yet to be answered about a national rapid-learning strategy, extraordinary new research capabilities—from investments of billions of dollars and years of work on EHR systems design, implementation, and databases; predictive computer models; and software development—will be emerging over the next few years. It seems worthwhile to invest public and private resources in discovering the potential benefits and lessons of their use.

Knowledge Gaps And Uncertainties

Is national health policy now advancing the scientific evidence base for clinical practice as rapidly as possible? An approximate answer is that the United States leads the world in basic biomedical research. But there are stunning gaps in clinical research and the knowledge base for evidence-based medicine.

■ **Diffused responsibility.** The knowledge gaps for evidence-based medicine result mostly from how national science policy has allocated responsibilities among

the NIH, biotechnology-based industries, and the Food and Drug Administration (FDA). In the broad schema of national science policy, the federal government has had the lead responsibility for support of basic research, primarily through grants to investigators at academic institutions. Private-sector biotechnology industries are expected to lead in turning this knowledge into new drugs, devices, and biological agents for clinical use. The FDA's primary responsibilities focus on requiring evidence of safety and effectiveness through carefully designed clinical trials before these products enter the market. The required FDA clinical trials (Phase III) typically include several hundred to several thousand patients. The FDA's approval standard for effectiveness is characterized by a leading pharmaceutical expert as follows: "A new drug must merely be slightly better than placebo in achieving a surrogate outcome over a few months, in modest numbers of highly selected patients."[4] After that point, the diffusion of new technologies and the accumulation of knowledge about how well they work have come to depend greatly on industry-financed research and marketing. By 2002, 80 percent of clinical trials were reported to be financed by the biotechnology industries.[5]

■ **Concerns about clinical trials.** If we are to make evidence-based medicine a paradigm for U.S. national health policy, we have to examine and improve the evidence base, its gaps, and its biases.[6] Clinical trials usually study selected younger-adult populations in carefully controlled circumstances. The evidence base is weakest in measuring the real-world effectiveness of the $2 trillion in U.S. health expenditures for typical patients seen in clinical practice—such as seniors and disabled people among the Medicare and Medicaid programs' eighty-five million enrollees. Scientific proof of effectiveness and safety are not required before medical and surgical procedures become part of clinical practice. Few studies effectively compare different therapies with each other for various patient groups.

For the most part, clinical trials report average effectiveness, yet there often are wide variations in treatment benefits among patient subpopulations. Moreover, there are growing concerns about (1) the extent to which the evidence published in research literature is financed by industry, (2) authors' conflicts of interest, (3) the dependence of leading professional journals on drug company advertising, (4) failure to report all clinical trial studies, and (5) failure to publish research with negative results. About two-thirds of clinical trials are reported to be done by for-profit research companies paid by the drug industry.[7] Widely cited initial findings have been found by subsequent research to be erroneous or overstated.[8] Many major technologies and clinical practices diffuse, over decades, into much broader and diverse patient populations, sometimes including millions or tens of millions of patients, than studied at initial entry.[9]

■ **Consequences of underinvestment.** The consequences of public underinvestment in evidence-based research are increasingly seen in the national media. Vioxx has become a prominent example of inadequate testing and monitoring for longer-term safety. It was an unwelcome surprise for tens of millions of women

"The biotechnology industries could benefit from public support for large EHR databases with genetic information."

worldwide when hormone replacement therapy was found to have previously undocumented risks. The federal government's "food pyramid" nutritional advice has needed major revisions. There are recent studies showing that the previous knowledge base was in error about low-fat diets, vitamin E, calcium and vitamin D, and so on. All of these are examples of instances where a rapid-learning strategy could have produced much better knowledge, years—even decades—earlier.

■ **Genetics-based medicine.** The advent of genetics-based medical science also requires rethinking traditional strategies for evidence development. It is known that drugs can vary widely in benefits and risks based on patients' genetic characteristics.[10] To maximize health care value, today's pharmacopeias and new drugs might need to be evaluated for many patient subpopulations. This will be enormously expensive and time-consuming if it requires premarket entry clinical trials many times larger than the FDA's current requirements.

■ **Overall system performance.** An inadequate knowledge base also limits initiatives to improve health system performance. Quality measures and pay-for-performance (P4P) incentives can now be applied only to a relatively small fraction of medical care. Similarly, health plans need more scientific and professional consensus to change practice patterns and advance evidence-based care. Physicians and patients could become a more effective force for driving health system value if there were greater certainty about the benefits and risks of treatment options.

In the past, there have been concerns that governmental initiatives to upgrade the clinical evidence base would involve far more regulatory requirements, large expenses, or long delays in technology diffusion.[11] The rapid-learning strategies that will now be possible will avoid many of these faults. With large, computer-searchable EHR databases and new research software, studies that would now take years will be doable, at low expense, in a matter of weeks, days, or hours. Rather than slowing the appropriate use of new technologies, a successful rapid-learning strategy could accelerate progress with more-definitive information for personalized medicine and greater professional and patient confidence in evidence-based guidelines. The biotechnology industries could benefit from public support for large EHR databases with genetic information, which involves far too great an expense for most individual companies to afford to construct and maintain on their own. Evidence from off-label uses in these data banks could provide important indicators about potential new uses of drugs.

Rapid-Learning Opportunities

Other papers in this collection discuss development history of rapid-learning capabilities—for example, at Kaiser Permanente, the VA, and the U.S. Department

of Health and Human Services (HHS)—and their potential use for breakthroughs in biomedical science and in clinical areas such as diabetes and cancer.[12] Here I focus on how rapid learning could provide comparable knowledge breakthroughs in several national health policy areas.

■ **Why are health care costs increasing?** For many years, health care costs and increases have been a major topic in health policy discussions. Yet the federal government does not yet have an adequate database to examine these rising costs. For example, national estimates that "intensity" of care—most likely new health care technology—accounts for 30–50 percent of annual spending increases are residual estimates; they are derived by subtracting population increases and inflation from health care spending increases.[13] We do not know what changes in clinical practices, for what patients, and for what conditions account for rising costs; nor do we know the extent to which such increases are from newly introduced technologies or diffusion of older technologies. Thus, it is very difficult to know for sure what we are buying, or to know whether or not the increases reflect evidence-based science. New national databases, with tens of millions of patient records, could quickly provide far more accurate information on these issues.

■ **What are the comparative benefits and risks of prescription drugs?** Seniors, children, women, minorities, and patients with comorbidities—the largest groups of prescription drug users—are frequently underrepresented in clinical trial data. There are also knowledge gaps about long-term benefits and side effects after the end of clinical trials—for instance, for chronic disease medications. The real-world effectiveness of drug therapies might differ, for better or worse, from the results in carefully managed clinical trials. As Jerry Avorn has noted, "We often cannot determine that a drug will turn out to be more effective or safe than its alternatives until it's been used for some time by large numbers of typical patients."[14] All of these concerns add to the need for new rapid-learning initiatives.

A comprehensive research agenda should be able to fill in important evidence gaps about prescription drugs. Drugs' effectiveness in population groups that were not adequately represented in randomized clinical trials could be checked with real-world data. New drugs could be assessed for effectiveness and safety, on a real-time basis, as they diffused into broader use. Heterogeneity in treatment responses could be analyzed to identify possible genetic and other still-unknown factors about diseases and their treatment and to design new clinical trials. Genetic information could be incorporated into EHR records where research has identified a target set of genes implicated in the benefits and risks from a particular drug. With such a system, Vioxx-type problems might be spotted quickly, in the first one or two years, and possibly identified as adversely affecting only small fractions of patients with specific biomarkers. Given the great importance of drug therapies for the Medicare aged and disabled populations, national health policy could make particularly good use of new databases for the Medicare prescription drug program.

■ **What is the evidence base for procedures?** Medical and surgical procedures are the largest component of health care spending, but without FDA-type scientific testing, many questions exist about their safety and effectiveness. There are wide variations in rates at which procedures are used for different populations and in different areas. One major research need is for definitive comparisons of minimally invasive surgery versus standard surgery. Physicians and patients could also use more-definitive knowledge about issues of medical versus surgical treatment. EHR systems can report data about all of these issues, making it feasible to learn from large numbers of patient experiences, rapidly and relatively inexpensively.

■ **What explains variations in health care spending and use?** Medical care practices differ widely among states, small geographic regions, and cities and even among leading hospitals and within health plans. Studies using EHR databases should allow much better analyses of variations, such as the extent to which they reflect health status differences. The appropriate federal policies could differ depending on these research results.

■ **How do environmental factors affect disease patterns?** Many health problems, including some cancers, birth defects, and heart disease, have concentrations in specific geographic areas. This suggests that environmental factors might add an as-yet-unknown factor, such as low-level chemicals in the water supply leading to birth defects or microbes that could produce clogging of arteries. Research epidemiologists could make good use of EHR databases for studying areas with high and low prevalence of health problems and correlating these findings with environmental data. Genetic information could be incorporated into EHR databases to assess damage from environmental factors.

■ **How can the health of minorities and special-needs groups be improved?** The gaps in our evidence base disadvantage minorities and people with special needs, who are often underrepresented in FDA-required clinical trials and industry-financed studies. Yet there are many health issues that need to be understood for minority populations, such as higher incidence of certain diseases and differing treatment patterns. Many Medicaid patients have special needs. For example, Medicaid enrolls many children with serious, long-term, and expensive disabilities, such as autism, cerebral palsy, cystic fibrosis, hemophilia, HIV/AIDS, sickle cell anemia, spina bifida, muscular dystrophy, epilepsy, and mental retardation. Their care is often fragmented among specialists and lacks good coordination; primary care also is often lacking. Large-scale EHR databases for such groups could provide valuable new resources to document current care and its successes and failures, engage more researchers, advance the evidence base, and reengineer patterns of care.

■ **What does this mean for patients like me?** One of the insights from studies of patient decision making is that people are much more interested in information about "patients like me" than in general statistics about risk factors and treatment options. But physicians and patients are often uncertain about treatment choices. Search software for large EHR databases could be designed to provide more-relevant information. A patient or physician could enter personal profile information and pose questions about the health experiences of and outcomes for large numbers of similar patients. If genetic information were included in EHR databases, it would be possible to move much more quickly toward a future of personalized medical care, where "patients like me" questions can be answered with greater accuracy.

■ **Components of a rapid-learning system.** A national rapid-learning system could include many databases, sponsors, and research networks. Its databases could be organized by enrolled populations (private health plans, VA, Medicare, Medicaid), providers (multispecialty clinics, academic health centers, specialist registries), conditions (disease registries), technologies (drug safety and efficacy studies, outcomes research), geographic areas (the Framingham Heart Study), age cohorts (the National Children's Study), minority populations (human genome studies), and other ways. With national EHR data, registry, connectivity, reporting, and privacy protection standards, all EHR systems could be compatible and capable of multiple uses; information in one data bank could be supplemented with that from another.

■ **Patient confidentiality.** A national legal and regulatory framework, the Health Insurance Portability and Accountability Act (HIPAA), exists for computerized medical records, and research organizations nationwide are required to have institutional review boards (IRBs) to approve and oversee studies that use individually identifiable data.[15] These laws and regulations also allow the development and use of research databases with non–personally identifiable information.[16] Rapid-learning organizations have been able to carry out research programs within current laws and regulations, and some EHR research data banks exclude personal identifiers. So far, a rapid-learning strategy does not seem to require changes in confidentiality laws; however, these issues are likely to attract ongoing scrutiny.

National Policies To Support Rapid Learning And To Use The Information Generated

National investment in rapid-learning initiatives would enable a new generation of health policies to realize the benefits of our expanding knowledge base.

■ **Reenergized markets.** The professionals and organizations that adopt EHRs and use EHR research databases could emerge as leaders for market-driven change. EHR use has advanced quickly, particularly among HMOs and multispecialty practices. A leading EHR software company, for example, reported that its EHRs were in use for fifty-eight million people as of November 2006.[17] EHR research databases could be the key technology for advancing the evidence base, and EHRs could be the key management technology for applying it. By supporting rapid learning, national health policy could reenergize competitive markets, led by those health plans, multispecialty clinics, hospitals, and physicians that make best use of EHRs.

To realize such competitive benefits for Medicare, investment in rapid-learning initiatives for Medicare's populations will be needed. Medicare health plans and multispecialty clinics with EHR research databases, such as Kaiser Permanente

"With a rapid-learning strategy, it will be possible to develop a Medicare payment system based on evidence-based protocols."

and Geisinger, could be partners. In the public sector, the VA could have a similar role, particularly in assessing new technologies, refining evidence-based protocols, and conducting practical clinical trials. The VA has a growing Medicare-age service population, it is a national leader in EHR systems and research databases, and it has affiliations with 107 medical schools and a research program of about $1.5 billion annually. Another federal player, the Agency for Healthcare Research and Quality (AHRQ), has a national network of evidence evaluation centers, including centers to assess research for Medicare, Medicaid, and the State Children's Health Insurance Program (SCHIP). By partnering with the VA, health plans, and others, Medicare could create a stronger evidence base for its beneficiaries, physicians, hospitals, clinics, and health plans.

■ **New payment policies.** When Medicare's diagnosis-related group (DRG) and resource-based relative value scale (RBRVS) payment policies were adopted, they were designed as national payment schedules, with various adjustments, to replace Medicare's original payment methods that reimbursed each hospital on the basis of its costs and each physician on the basis of his or her billed charges. The main objective was a rational payment system, where the federal government could decide on payment amounts. Quality of care was assumed to be something that payers did not need to worry about. Today, however, we know that quality variations must be a central concern for future national health policy. The health system delivers care whose quality meets professional standards only about half the time.[18] Yet Medicare pays the same amount for care of below-average quality as for the best care. Quality report cards and P4P are important initiatives to adjust the payment amounts, but they now cover only a fraction of medical care.

With a rapid-learning strategy, it will be possible to develop a new Medicare payment system based on evidence-based protocols. With this approach, health care providers would be paid for delivering courses of treatment that have been scientifically proven to achieve the best outcomes and for documenting, such as through EHRs, that they provided such care. Health care that did not meet scientifically validated standards would be paid for at a lower rate.[19] Such a new system will require fast-track research to assess alternative evidence-based protocols. If national health policymakers intend to end mediocrity and outdated practices, Medicare's $400 billion of annual spending can be leveraged to advance science-based medicine.

Organizing Medicare's payment requirements around evidence-based protocols could be a powerful strategy for advancing clinical science. Pediatric oncology is often cited for its rapidly improving patient outcomes.[20] In this field, most children with cancer are in clinical trials with defined protocols of care, and treatment results are rapidly reported to other specialists, so that the knowledge base expands from the experience of every patient. In contrast, only a small percentage of Medicare cancer patients are now part of clinical trials, and Medicare does not require reporting on cancer treatment protocols and results. A Medicare rapid-learning system could be designed, however, to learn continuously from the experiences of its forty-three million beneficiaries.

■ **Medicaid as a rapid-learning leader.** Medicaid and SCHIP could be one of the highest-payoff areas for new EHR-based learning strategies and program management. Not only are these, combined, the nation's largest health program, in terms of enrollees (more than fifty million) and costs (about $350 billion this year), but they also cover large numbers of the poorest, sickest patients, including many minorities. The Medicare/Medicaid dual-eligible population of seven million accounts for over 30 percent of spending for Medicare's enrollees and more than 40 percent of Medicaid spending; they need much more attention for better care.[21] The two programs are also the nation's largest payer for pregnancy care and a leading payer of care for children; the frail elderly; and people with serious, long-term disabilities. Asthma, chronic mental illness, alcoholism, drug abuse, and AIDS are also health problems of particular importance to Medicaid. However, no state Medicaid program is yet using EHR systems, and it is unrealistic to think that Medicaid-participating public hospitals, clinics, emergency rooms, and physicians are going to be able to buy commercial EHR products.

Medicaid and SCHIP could become national leaders in EHR adoption and in use of EHR research databases. A rapid and low-cost way to implement an EHR strategy for them would be for states to create EHRs for enrollees on a central computer server that could be accessed by their Medicaid providers. The National Health Service (NHS) in England is a leading example of this implementation strategy. State EHR choices could include the VA's VistA system, which is already in the public domain and sponsored by HHS as a low-cost option for physicians and clinics, as well as EHR systems offered by Medicaid's contractors. Since federal law already offers a 90 percent federal match for computerized Medicaid management information systems, a national strategy to encourage states to adopt EHRs could move quickly. A centralized EHR system could be an enormous advance in a state's capabilities to understand and manage Medicaid and SCHIP. Medicaid/SCHIP research databases could become a valuable resource for advancing clinical science for people enrolled in these programs. The federal government could invest in partnerships with leadership states—perhaps up to three states initially—to develop comprehensive EHR systems and databases and to become "rapid-learning laboratories" for SCHIP and Medicaid nationwide.

■ **National clinical trials databases.** A rapid-learning strategy could create national computer-searchable databases from NIH-funded and FDA-required clinical trials to make these studies available for further scientific scrutiny and convenient use by other researchers. There have been discussions of such HHS initiatives—and

even ideas for global networks of clinical trials data—but only modest progress to date.[22] At the FDA, for example, a typical new drug application (NDA) involves the delivery of hundreds of thousands of paper records, which often arrive by the truckful. How much simpler and user-friendly it would be for everyone if the FDA required these data to be reported in EHR-type formats for computer analysis. The NIH could require that all NIH-funded clinical trials, as well as the patient care at the NIH's selected national clinical care centers, be reported in EHR-type formats to national research databases. Researchers who try to reach conclusions from multiple studies now must engage in "meta-analyses" of reported statistics rather than being able to analyze combined data sets. Biomedical research could advance more rapidly, and these databases could also prove useful to evaluate new evidence-based protocols and new technologies, since the NIH funds leading-edge work at academic institutions. As the real-world analyses from EHR rapid-learning databases raise questions that require definitive answers, more public investment will be needed in clinical trials, particularly "practical clinical trials" that target important patient care issues.[23]

■ **National assessments of new technologies.** Medicare's recent proposals for "coverage with evidence development" (CED) offer a useful set of ideas for rapid learning about all new technologies. After FDA approval of a new drug or device or the introduction of a new procedure, Medicare (and other payers) would finance coverage for its broader use only when a required set of information was reported to computerized national EHR research registries. Research on these registries and other data would then be used for developing guidelines for use. The federal government and the private sector could jointly develop a national research plan for each new technology. The plan would identify needed research and suggest rapid-learning strategies, using EHR databases and other approaches; it would also set a target date for pulling together answers to these questions and for considering a subsequent research plan.[24] Learning as much as possible, as soon as possible, seems a sensible way to realize the full value of new technologies.

Next Steps

All of the research questions and new policies discussed in this paper are candidates for national initiatives. Growth in spending by Medicare and Medicaid—largely driven by use of new technologies and growing populations of beneficiaries as the baby boomers retire—is among the compelling fiscal reasons for new research. Together, these public programs already consume $600 billion of federal expenditures annually and enroll more than eighty-five million people. Medicare has embarked on a costly new drug benefit; there are many questions to be answered about the best use of drug therapies in senior citizens. Medicare and Medicaid populations are often excluded from the clinical trials evidence base. If these trends continue, without rapid learning about new technologies, the budget costs will be truly extraordinary.

"The key short-term issues for advancing a rapid-learning strategy are leadership, developing specific research programs, and funding."

The key short-term issues for advancing a rapid-learning strategy are leadership, developing specific research programs, and funding. The HHS health agencies and the VA could be public-sector (and funding) leaders; in the private sector, today's rapid-learning networks and organizations, professional societies, and others that invest in EHR systems and develop EHR research capabilities will best be able to lead the advance of evidence-based care. A list of HHS agency-specific initiatives might include expanding on the NCI's Cancer Research Network with NIH networks for heart disease and diabetes; a broad expansion of AHRQ's research to address issues related to Medicare prescription drugs, Medicaid, national health spending, socioeconomic and racial disparities, effectiveness, and quality; expanding the CDC's Vaccine Safety Datalink network and the FDA's postmarket surveillance into a national FDA/CDC program for the evaluation of drug safety and efficacy, including pharmacogenomics; starting national EHR research programs for Medicaid special-needs populations; integrating NIH and FDA clinical studies into national computer-searchable databases; expanding Medicare's evidence development requirements for coverage into a prototype national EHR-based system for evaluating new technologies; developing rapid-learning evidence for Medicare "best practice" protocols and new payment systems in several areas; and integrating new evidence and EHR database search capabilities into the National Library of Medicine's Web site for consumers and professionals, Medline Plus (http:// medlineplus.gov). HHS investments could start at $50 million per year. Over the next five years, as research networks develop and research programs expand, new funding could rise to $300 million annually.

THERE IS A COMPELLING PUBLIC INTEREST to advance the scientific knowledge for health care as rapidly as possible. There is much to be learned quickly about the best uses of current technologies. For the longer term, a national goal could be for the U.S. health care system to learn about the best uses of new technologies at the same rate at which it produces those new technologies.

The Robert Wood Johnson Foundation provided support for this paper through a grant to the George Washington University.

NOTES

1. E.H. Wagner et al., "Building a Research Consortium of Large Health Systems: The Cancer Research Network," *Journal of the National Cancer Institute: Monographs* no. 35 (2005): 3–11.
2. Centers for Disease Control and Prevention, "Vaccine Safety Datalink Project (VSD)," 23 October 2006, http://www.cdc.gov/od/science/iso/research_activities/vsdp.htm (accessed 7 November 2006).
3. American Medical Group Association, "Anceta/AMGA Collaborative Data Warehouse: Background," December 2006, http://www.amga.org/QMR/Anceta/bg_anceta.asp (accessed 12 December 2006).
4. J. Avorn, *Powerful Medicines: The Benefits, Risks, and Costs of Prescription Drugs* (New York: Knopf, 2004), 380.
5. J. Abramson, *Overdosed America: The Broken Promise of American Medicine* (New York: HarperCollins, 2004), 95.
6. For history, see D.M. Eddy, "Evidence-Based Medicine: A Unified Approach," *Health Affairs* 24, no. 1 (2005): 9–17. For prescription drugs, see J.P. Kassirer, *On the Take: How Medicine's Complicity with Big Business Can Endanger Your Health* (New York: Oxford University Press, 2004); Avorn, *Powerful Medicines*; and Abramson, *Overdosed America*.
7. Abramson, *Overdosed America*, 95.
8. J.P. Ioannidis, "Contradicted and Initially Stronger Effects in Highly Cited Clinical Research," *Journal of the American Medical Association* 294, no. 2 (2005): 218–228.
9. On technology, see R.A. Deyo and D.L. Patrick, *Hope or Hype: The Obsession with Medical Advances and the High Cost of False Promises* (New York: AMACOM, 2005).
10. W.E. Evans and H.L. McLeod, "Pharmacogenomics—Drug Disposition, Drug Targets, and Side Effects," *New England Journal of Medicine* 348, no. 6 (2003): 538–549.
11. A.B. Cohen et al., *Technology in American Health Policy* (Ann Arbor: University of Michigan Press, 2004); and P.J. Neumann, *Using Cost-Effectiveness Analysis to Improve Health Care: Opportunities and Barriers* (New York: Oxford University Press, 2005).
12. A listing of these papers is available at http://content.healthaffairs.org/cgi/content/full/hlthaff.26.2.w107/DC2.
13. For example, see C. Smith et al., "National Health Spending in 2004: Recent Slowdown L
ed by Prescription Drug Spending," *Health Affairs* 25, no. 1 (2006): 186–196.
14. Avorn, *Powerful Medicines*, 381.
15. See U.S. Department of Health and Human Services, "Office for Civil Rights—HIPAA," 2 November 2006, http://www.hhs.gov/ocr/hipaa (accessed 7 November 2006).
16. L. Rybowski, "Protecting the Confidentiality of Health Information," National Health Policy Forum background paper (Washington: George Washington University, July 1998).
17. See the Epic Systems home page, http://www.epicsystems.com/index.php.
18. E.A. McGlynn et al., *The First National Report Card on Quality of Care in America* (Santa Monica: RAND, 2006).
19. Refining each DRG to pay for scientifically validated "best practices" would be one path toward a full evidence-based protocol system.
20. CureSearch, "Frequently Asked Questions," http://www.curesearch.org/our_research/index_sub.aspx?id=1488 (accessed 2 July 2006).
21. Kaiser Commission on Medicaid and the Uninsured, "Dual Eligibles: Medicaid's Role in Filling Medicare's Gaps," March 2004, http://www.kff.org/medicaid/7058.cfm (accessed 14 December 2006).
22. National Cancer Institute, *Restructuring the National Cancer Clinical Trials Enterprise*, June 2005, http://integratedtrials.ncl.nih.gov/ict/CTWG_report_June2005.pdf (accessed 12 December 2006); and National Institutes of Health, Office of Portfolio Analysis and Strategic Initiatives, "NIH Roadmap," 25 October 2006, http://nihroadmap.nih.gov/index.asp (accessed 7 November 2006).
23. S.R. Tunis, D.B. Stryer, and C.M. Clancy, "Practical Clinical Trials: Increasing the Value of Clinical Research for Decision Making in Clinical and Health Policy," *Journal of the American Medical Association* 290, no. 12 (2003): 1624–1632.
24. Avorn, *Powerful Medicines*, 383–384.

PERSPECTIVE

The Gap Between Evidence And Practice

We still have much to learn about practice and patient factors that affect clinical outcomes.

by Louise Liang

ABSTRACT: Despite the urgent call to action made by the Institute of Medicine's (IOM's) *Crossing the Quality Chasm* report in 2001, several fundamental issues in health care remain largely unaddressed. Although a number of organizations have addressed many of the system-level factors cited in the report, we have much to learn about practice and patient factors that affect clinical outcomes. Now we have new opportunities to further improve health care by learning from the data available in electronic health record databases and, perhaps more importantly, to better understand the human behavior of caregivers and patients necessary to improve health care quality. [*Health Affairs* 26, no. 2 (2007): w119–w121 (published online 26 January 2007; 10.1377/hlthaff.26.2.w119)]

MORE THAN FIVE years ago, the Institute of Medicine (IOM) issued its urgent call for fundamental change in *Crossing the Quality Chasm*.[1] It identified four key factors underlying the quality chasm: (1) the increasing complexity of science and technology, (2) the rise in the incidence of chronic conditions, (3) a poorly organized delivery system, and (4) a lack of critical information technology (IT) supports. In 2003 Elizabeth McGlynn and colleagues further defined the quality gap, reporting that the U.S. health care system delivers evidence-based care to patients only 55 percent of the time.[2] Yet despite the initial attention these reports received and agreement by health care leaders and policymakers across the nation that improvements were required, the issues of overuse, underuse, and misuse are largely unaddressed.

Even in integrated systems of care that have implemented electronic health records (EHRs), the improvement in clinical outcomes is not as dramatic as predicted by *Crossing the Quality Chasm*. Even though a number of organizations have addressed many of the system-level factors cited above, we have much to learn about practice and patient factors that affect clinical outcomes. Now we have new opportunities to further improve health care by learning from the data available in these electronic databases and, perhaps more importantly, to better understand the human behavior of caregivers and patients necessary to greatly improve health care quality.

Lynn Etheredge provides a thorough assessment of the many opportunities for leveraging EHRs to rapidly advance our evidence base for clinical care.[3] EHRs certainly have the potential to fill major knowledge gaps about health care expenditures, the benefits and risks of drugs and procedures, geographic variation, and environmental health influences as never before. However, it is not clear that barriers to improving health care outcomes will be overcome by additional clinical evidence,

Louise Liang (louise.liang@kp.org) is senior vice president, Quality and Systems Support, at Kaiser Permanente in Oakland, California.

DOI 10.1377/hlthaff.26.2.w119

since we have largely failed to improve clinical outcomes using the evidence we already have.

Consider the Institute for Healthcare Improvement's (IHI's) 100,000 Lives Campaign. The campaign was created in December 2004 by IHI chief executive officer Don Berwick out of deep frustration with the lack of progress since the IOM issued its call to action in 2001. The campaign did not introduce new evidence or practices. Instead, it highlighted known practices by packaging and promoting them to generate focus and attention among hospital staff. One of the interventions focused on preventing surgical site infection (SSI). The Centers for Disease Control and Prevention (CDC) issued guidelines for prevention of SSI in 1999. Yet SSI still accounted for an estimated 780,000 hospital-acquired infections at the start of the campaign.[4]

The 100,000 Lives Campaign enrolled more than 3,000 hospitals, representing an estimated 75 percent of U.S. hospitals beds, which implemented up to six evidence-based and life-saving interventions. In eighteen months, the campaign prevented an estimated 122,300 deaths through the implementation of these six evidence-based practices across the country.[5]

This experience and others demonstrate the prevailing disconnect between what we know to be effective and what we practice daily. It is estimated that it takes on average seventeen years for proven medical advances to be incorporated into common practice, with the exception of new technologies and pharmaceuticals.[6] Rapid learning via EHRs can make a major contribution to the understanding of clinician and patient behavior that will unlock this critical challenge of converting clinical evidence into better outcomes. With this focus, I highlight and add to the opportunities identified in Etheredge's paper.

■ Identifying new evidence. The first opportunity is EHR systems' ability to rapidly, efficiently, and more comprehensively identify new evidence and gather new knowledge. As Etheredge explains, EHRs have the potential to take over where clinical trials and evidence-based research leave off, by providing real-world evidence of drugs' and treatments' effectiveness across subpopulations and over longer periods of time. EHR systems have the capacity to identify outcomes, unexpected complications, and side effects very quickly. An EHR can monitor whether a medication is performing unexpectedly in a physician's patient population, well before the established regulatory machinery can identify a potential problem or initiate a response.

■ Diffusing new evidence. The second opportunity is the use of EHRs to accelerate and understand the diffusion of new evidence. The length of time it takes a new evidence-based practice to become part of daily practice should be drastically reduced with EHRs. As new guidelines, tools, or alerts are integrated with the EHR, new practices and evidence can be in the hands of clinicians within hours instead of years. The information can be delivered in the form of decision-support tools such as practice guidelines or patient safety alerts automatically triggered at the point of care. Clinicians' burden to learn and apply all the latest evidence is lifted so they can focus on critical human factors related to changing behavior and ultimately outcomes of care.

■ Changing clinicians' behavior. The third important opportunity is the use of EHRs to efficiently monitor the actual use of evidence and the improvement of outcomes. If we can understand why clinicians change their practices and design content within the EHR to meet their needs, we may finally move the *Quality Chasm* agenda forward. Until now, we have applied more clinical evidence, clinician education, and recently physician measurement and incentives, all without systematic success.

EHRs can help identify the combination of clinical evidence, tools, clinical circumstances, and patient information that supports a clinician in the of use evidence to produce better, safer outcomes. The dynamic interaction between the workflow of a clinician's practice, the tools imbedded in the EHR, and the availability of evidence at the point of care is poorly understood. Knowledge about this critical nexus, coupled with performance feedback, may be the key to major improvement. However, this is only half of the equation.

■ Patients as partners in care. Clinicians do not change outcomes alone but can do so in partnership with patients. Throughout most of the history of health care, we have focused on the clinician side of the equation. We have invested far less effort working with patients to understand how to leverage their participation in their own care. This represents an enormous untapped opportunity.

Services for patients with chronic diseases now account for 75 percent of total U.S. health care spending.[7] Our ability to effectively support these patients at work and at home is critically important for managing national health care costs as well as for improving clinical outcomes. EHRs with robust patient interfaces can provide patients with information, tools, and communication venues, including secure messaging, to help them make informed decisions about their care. With the right tools, informed patients can collaborate with clinicians to define and implement shared care plans, including self-care, home monitoring, and prevention. Identifying specific patient factors that drive behavior change and therapeutic adherence is essential to leveraging the clinician-patient partnership. Understanding which approaches most effectively help patients change behavior, and ultimately improve outcomes, will greatly advance our progress toward closing the quality gap.

■ Moving the agenda forward. It has been apparent for many years that EHRs offer a singular opportunity for breakthrough improvement in clinical outcomes, appropriate resource use, and patient safety. Rapid learning related to practical clinical trials, diffusion of best practices, and behavioral change by clinicians and patients will also be essential steps on the pathway to those ends.[8] Among the many opportunities available, the area of the most potential and the most need is comparative clinical effectiveness research. Current health care research does not identify the most effective treatment and approaches in real-world settings. To deliver better clinical outcomes, we need to understand the complex interactions of everyday clinical choices made by thousands of clinicians and millions of patients. The vast information available through EHRs allows us to efficiently access the experience and related outcomes of millions of clinical decisions. We will have a longitudinal view that can bring unforeseen results, positive and negative, to the surface in both planned and unplanned clinical situations. We will also have the breadth of data to understand small subpopulations as never possible before. We have found the resources to fund many lesser advances and tools in health care practice. What is a more powerful clinical tool than knowledge?

NOTES

1. Institute of Medicine, *Crossing the Quality Chasm: A New Health System for the Twenty-first Century* (Washington: IOM, 2001).
2. E.A. McGlynn et al., "The Quality of Health Care Delivered to Adults in the United States," *New England Journal of Medicine* 348, no. 26 (2003): 2635–2645.
3. L.M. Etheredge, "Rapid Learning: A New Approach to Diffusing Research into Practice," *Health Affairs* 26, no. 2 (2007): w107–w118 (published online 26 January 2007; 10.1377/hlthaff.26.2.w107).
4. A.J. Mangram et al., "Guideline for Prevention of Surgical Site Infection, 1999," *Infection Control and Hospital Epidemiology* 20, no. 4 (1999): 250–278.
5. A.D. Hackbarth, C.J. McCannon, and D.M. Berwick, "Interpreting the 'Lives Saved' Result of IHI's 100,000 Lives Campaign," *Joint Commission Benchmark* 8, no. 5, September/October 2006, http://www.ihi.org/NR/rdonlyres/A3552730-6841-4269-A783-BF73E125401E/0/JointCommissionBenchmarkSeptOct06.pdf (accessed 20 December 2006).
6. R. Larson, "Medical Advances Can Outpace Doctors: Retraining Not Enforced, Critics Say," *Washington Times*, 21 March 1999.
7. Centers for Disease Control and Prevention, "Chronic Disease Overview," 18 November 2005, http://www.cdc.gov/nccdphp/overview.htm (accessed 20 November 2006).
8. S.R. Tunis, D.B. Stryer, and C.M. Clancy, "Practical Clinical Trials: Increasing the Value of Clinical Research for Decision Making in Clinical and Health Policy," *Journal of the American Medical Association* 290, no. 12 (2003): 1624–1632.

PERSPECTIVE

Moving Closer To A Rapid-Learning Health Care System

Establishing a culture of learning while providing care will take collaboration among the many participants in the U.S. health system.

by Jean R. Slutsky

ABSTRACT: This Perspective discusses activities that are necessary for developing a rapid-learning health system. Recognition of the central role that patients play in the successful evolution of such a system will help ensure that the goals of the transformation are met. Understanding the trade-offs of using a less controlled form of research to inform health care decision making and making necessary investments in methodology and translation will help secure the success of continuous-learning research. Major public policy interest in promoting health information technology and in getting more value for health care spending creates a framework for moving ahead. [*Health Affairs* 26, no. 2 (2007): w122–w124 (published online 26 January 2007; 10.1377/hlthaff.26.2.w122)]

THE FUTURE WORLD proposed by Lynn Etheredge is very exciting—exhilarating, really—to imagine and explore.[1] Etheredge carefully describes a health system in change, largely because of the intersection of consistently disappointing health outcomes and gains in outputs of health information technology (IT).[2] Visionary leaders—local and national, public and private—have led the charge to embrace these new realities. Still, some difficult uncertainties exist that will require thoughtful and careful navigation. The risk of not moving forward further justifies the investment in creating environments of "learning while doing."

Etheredge describes an unfortunate mismatch between remarkable biomedical research discoveries and the evidence base for clinicians and policymakers. There is no question that innovations developed as the result of remarkable discoveries in biomedical research have dramatically changed the course of many diseases. What is disappointing is that many health care interventions are not supported by evidence of clinical effectiveness in actual clinical settings (often called the "real world")—evidence needed to determine if and when they should be used and for whom. As Etheredge points out, decisionmakers, including patients, are often in the dark about the comparative effectiveness of many interventions used to treat or prevent illness.

■ Things we need to keep in mind. In establishing a learning health system, we need to keep the following in mind.

Recognize that the patient is paramount. This is critical to the success of a learning health care system. We have made great progress moving toward a time where personalized health care is more than a wistful vision. The ultimate goal is better patient health outcomes and value for our health care dollar—better health outcomes for the most reasonable cost. Patients have choices: which doctor to choose, which hospital to go to, which treatment to get. Yet often patients cannot make these decisions rationally because the information on which to base the decision does not exist or is not understandable. Patients also fuel the research enterprise. Although government and the private sector fund the studies that explore the basis for disease and cure, it is patients who assume the risks and benefits of enrolling in clinical trials and other studies.

Trust is essential. What is required is trust that patients will not be labeled inappropriately or harmed in any way through the use of their data; trust that health care organizations won't lose their individuality and competitive edge by collaborating with other patient care organizations; and trust that researchers will make good and relevant use of the data and that industry will provide adequate insight into their study data without accusations of inappropriate involvement. Without patients' participation in research studies, our knowledge base would be meager and narrow. Recruitment is difficult, especially in underserved and culturally diverse populations.[3] Establishing trust that participating in research studies will further our knowledge of what works and for whom is a primary goal. This raises the question of whether study findings can ethically be kept secret from other researchers and patients themselves. Etheredge is correct in stating that without these results, negative or positive, the whole story will never be known.

"Investments in methodology often take a backseat to funding for hypothesis-driven studies."

Agreeing to agree on some things is elemental. For example, a simple thing such as what age group constitutes adolescence is often inconsistent across studies, medical records, and so on. How do we define "elderly"? The agreement on simple categories and descriptions (what some disciplines call an "ontology") would increase the usability and relevance of data repositories in research.[4] But this is a complicated and tricky business. It takes—and will continue to take—extraordinary cooperation among groups to address this issue, to realize the full potential of a learning health care system.

Learn what we can about the risks of findings that come from studies other than the randomized clinical trial. Correct interpretation of research findings in light of different study designs is essential. Assessing how results can be generalized to different patients, whether the study design is appropriate for the question being asked, and other factors must be taken into account.[5] We need to make investments in better understanding of how to interpret findings while investigating different methodologies for making these studies as rigorous as they can be. Investments in methodology often take a backseat to funding for hypothesis-driven studies. Developing a credible and actionable learning health care system will depend heavily on how interpretable the findings are found to be.

■ The Effective Health Care program. The importance of getting more value for health care in the United States is underscored by Congress's appropriation of $15 million in each of the past two years to the Agency for Healthcare Research and Quality (AHRQ) to undertake comparative effectiveness research. The comparative effectiveness research was authorized under Section 1013 of the Medicare Prescription Drug, Improvement, and Modernization Act (MMA) of 2003. The infrastructure and research created under this authorization will contribute to the objectives for establishing a learning health system.

AHRQ's Effective Health Care program is committed to establishing a methodological research foundation for studies done using existing databases. In early 2007, a manual, *Establishing Registries for Evaluating Patient Outcomes*, will be published. The manual, coauthored by a mix of scientists and specialists from academe, industry, and government, will be a practical document for establishing, maintaining, and interpreting patient registries.

Jean Slutsky (jean.slutsky@ahrq.hhs.gov) is director of the Center for Outcomes and Evidence, Agency for Healthcare Research and Quality, in Rockville, Maryland.

DOI 10.1377/hlthaff.26.2.w122

The Effective Health Care program establishes two other mechanisms that promote the concepts of learning health systems. First, the program encourages participation from all interested participants in the health care system. This insures the relevancy of the program as well as the actual research. Research questions and draft reports are put on a public Web site (http://www.effectivehealthcare.ahrq.gov) for public comment. Comment is also solicited on topics for further research. Finally, the John M. Eisenberg Clinical Decisions and Communications Science Center was established to promote the translation of complex scientific findings for health care decisionmakers, including policymakers, clinical decisionmakers, and patients. The Eisenberg Center will tackle issues of innumeracy, interpretation, and communication of risk, and also methods for developing decision aids.

THIS IS A UNIQUE POINT in time for the U.S. health care system. Our collective investments in health IT are poised to provide access to provider and patient data in a more comprehensive and less labor-intensive manner than ever before. Quality measurement will be much easier to execute and track in real time. Understanding how different treatments work in patients who are reflective of actual patients, rather than those who participate in controlled trials, will help interpret more rationally what the benefits and harms might be and what trade-offs we are willing to accept. It is also likely that our understanding of unintended consequences, both good and bad, of different treatments will become clearer. Establishing a culture of learning while providing care will take collaboration among the many participants in the health care system, both private and public. The potential benefits are anticipated to go well beyond our initial expectations.

The views expressed in this paper are those of the author and do not necessarily reflect the official position of the U.S. Department of Health and Human Services.

NOTES

1. L.M. Etheredge, "Rapid Learning: A New Approach to Diffusing Research into Practice," *Health Affairs* 26, no. 2 (2007): w107–w118 (published online 26 January 2007; 10.1377/hlthaff.26.2.w107).
2. Agency for Healthcare Research and Quality, *2005 National Healthcare Disparities Report*, Pub. no. 06-0017 (Rockville, Md.: AHRQ, December 2005); and AHRQ, *2005 National Healthcare Quality Report*, Pub. no. 06-0018 (Rockville, Md.: AHRQ, December 2005).
3. J.G. Ford et al., *Knowledge and Access to Information on Recruitment of Underrepresented Populations to Cancer Clinical Trials*, Evidence Report/Technology Assessment no. 122 prepared by the Johns Hopkins University Evidence-based Practice Center under Contract no. 290-02-0018, Pub. no. 05-E019-2 (Rockville, Md.: AHRQ, June 2005).
4. T.R. Gruber, "A Translation Approach to Portable Ontologies," *Knowledge Acquisition* 5, no. 2 (1993): 199–220.
5. S.M. Teutsch, M.L. Berger, and M.C. Weinstein, "Comparative Effectiveness: Asking the Right Questions, Choosing the Right Method," *Health Affairs* 24, no. 1 (2005): 128–132.

Linking Electronic Medical Records To Large-Scale Simulation Models: Can We Put Rapid Learning On Turbo?

The Archimedes model offers one example of how mathematical modeling can assist in medical decision making.

by David M. Eddy

ABSTRACT: One method for rapid learning is to use data from electronic medical records (EMRs) to help build and validate large-scale, physiology-based simulation models. These models can than be used to help answer questions that cannot be addressed directly from the EMR data. Their potential uses include analyses of physiological pathways; simulation and design of clinical trials; and analyses of clinical management tools such as guidelines, performance measures, priority setting, and cost-effectiveness. Linking the models to EMR data also facilitates tailoring analyses to specific populations. The models' power and accuracy can be improved by linkage to comprehensive, person-specific, longitudinal data from EMRs. [*Health Affairs* 26, no. 2 (2007): w125–w136 (published online 26 January 2007; 10.1377/hlthaff.26.2.w125)]

A CENTRAL COMPONENT OF RAPID LEARNING is the use of data from electronic medical records (EMRs) to help address clinical and administrative questions that would otherwise go unanswered. There are two main ways EMRs can do this. The first is simply to use the data in the EMR to look up the answers. The other is to use the data to build mathematical models and then use the models to answer the questions. The models can vary from a single equation to deeper, more comprehensive models that use hundreds of equations to, in essence, try to simulate everything important that happens in a health care system, from the underlying physiology to logistics and costs. The latter type of a large-scale simulation model obviously involves much more work than simply looking up an answer in the database or building a simpler model. But once built, such a model can help answer a considerably broader range of questions. This paper describes how large-scale, physiology-based simulation models can increase the use

David Eddy (eddyaspen@yahoo.com) is founder and medical director of Archimedes Inc. in San Francisco. Archimedes was founded to improve the quality and efficiency of health care by using advanced mathematics and computing methods to build realistic simulation models of physiology, diseases, and health care systems.

DOI 10.1377/hlthaff.26.2.w125

of EMRs for rapid learning, and it illustrates their use with a particular model called Archimedes.

The Need For Rapid Learning From EMRs

The need for better methods to answer questions is obvious. The modern practice of medicine is bristling with questions, ranging from the medical director of a managed care organization wanting to know the effects of improving control of blood glucose in diabetics, to the head of medical benefits at a large corporation wanting to know why the company's health costs went up in double digits three years in a row, to Medicare wanting to design a pay-for-performance (P4P) program and needing to know how much to pay for how much performance on which indicators, to Jane Smith wanting to know how much her risk of a stroke will go down if she loses ten pounds.

In theory, all of these questions could be answered if we could conduct enough evaluation studies and clinical trials. But the feasibility of that approach is severely limited by several facts: the long times required to observe long-term outcomes; the large sample sizes needed to sort out real results from random "noise"; the large number of options that need to be evaluated; the need to disrupt current practices to implement each of the options; the reluctance of physicians and patients to participate; the narrowness and artificiality that trial designs often require; the fact that the answer one gets in one setting does not necessarily apply to other settings; the high cost of the studies (typically in the tens of millions of dollars, and often in the hundreds of millions); the rapid changes in technologies and practices that make answers obsolete almost as soon as they come in; and the fact that even if we can set a study in motion, we will still need to know what to do while we wait for the results. If it is feasible to answer a question with evaluation studies and clinical trials, then by all means we should do that. But the unfortunate fact is that empirical studies are not feasible for the great majority of decisions that need to be made.

The Appeal Of EMRs

From this list of limitations, it is easy to see the attractiveness of EMRs. The data are already collected; it would seem that all we have to do is sort through them and pull out the answers. For a simple example, to determine the proportion of people with diagnosed diabetes who have poorly controlled blood glucose (hemoglobin A1c [HbA1c] levels greater than 7 percent), a medical director could go into the database, identify everyone with a diagnosis of diabetes, look up their HbA1c levels, and calculate the proportion with HbA1c levels above 7 percent. Such an approach is fast and inexpensive. It is also real; the data in EMRs represent real patients getting real treatments and having real outcomes in real practice settings. The hope for EMRs and rapid learning is well summarized in the commonly expressed idea that EMRs should provide opportunities for "natural experiments." Properly used and interpreted, EMR data used in this way can help answer many important questions. Several good examples are described elsewhere in this collection of *Health Affairs* papers.[1]

The Limitations Of EMRs For Rapid Learning

Unfortunately, with the simplicity, low expense, and speed come some limitations and pitfalls. For an example, suppose that after identifying people with diabetes and HbA1c levels above 7 percent, the medical director now wants to know whether treating them with a combination of an oral drug, metformin, and a new injectable drug called exenatide has any affect on the incidence of myocardial infarction (MI) in this group. Ideally, the medical director would like to have a randomized controlled trial (RCT) in which people with HbA1c levels greater than 7 percent are randomized to receive metformin and exenatide or not, and then followed for ten years to observe MI rates in both groups. The randomization would ensure that the two groups were statistically identical in all ways except the treatment, so any difference in MI rates could be attributed to the treatment. There is no clinical trial that does this. Could we use the natural experiment of the EMR to reproduce such a trial?

Unfortunately, we cannot. Consider the steps. We would have to go back in the EMR records ten years and try to identify two groups of people: a group that had HbA1c levels above 7 percent and were put on metformin and exenatide, and another group that had HbA1c levels greater than 7 percent but were not put on any drug. At this point we encounter the first problem: We can do this only for treatments that have been in use for at least ten years, and exenatide is new. This approach is useless for new or even recent treatments. But set that aside and imagine that we can find the two groups of people we need. Now the second problem arises: We need to ensure that those put on metformin and exenatide had the same pretreatment HbA1c levels as those not so treated. This is very unlikely, because people are put on drugs for the very reason that they have abnormal HbA1c values. Suppose we set that problem aside as well and imagine that the two groups we found had the same initial levels of HbA1c. We immediately encounter a third problem: The two groups also need to be the same in all other ways that might affect their chances of having an MI. That is, they must have the same age distribution, sex proportion, race/ethnicity mix, weight, proportion of smokers, blood pressure, high-density lipoprotein (HDL) level, glucose level, use of aspirin, and so forth. This is extremely unlikely. After all, whatever caused them to be put on metformin and exenatide (an alert physician?) would also likely cause them to be treated for hypertension, hyperglycemia, obesity, tobacco addiction, or anything else they might have. If they had been treated for any one of these, we would not be able to determine if any difference in MIs was due to the use of metformin and exenatide versus any of the other risk factors. Thus, the approach of trying to "find" the desired clinical trial in the data set will not work. We might try a differ-

"The pure look-it-up approach is not good for determining what would happen if we did something differently."

ent approach, such as trying to find people who are matched in all important ways except the use of the drugs. But by the time we list all of the important variables that can affect heart attack rates and try to find people who are matched on all of them, we would end up with exceedingly few matched pairs, if any. If we reduce the list of the variables to those we need to match, we would find more pairs, but we would never know how the unmatched variables biased the results. There are other approaches we might try, but they all suffer from some bias or other.

In addition to the impossibility of analyzing new treatments, the confounding effects of the unmatched variables, the difficulty of finding a sufficient number of people, and the need for long follow-up times, the look-it-up approach has other problems. They include cohort effects (people who were being observed ten years ago are different from people being observed today); variations in and gradual drifting of practice patterns and performance levels; changes in technologies such as the introduction of new tests and treatments; changes in definitions of diseases and outcomes; turnover in the population; and difficulty of transferring results from one setting to another. In general, the pure look-it-up approach is good for observing what happened in the past and what is happening now, but it is not good for determining what will happen in the future, or what would happen if we did something differently, or the merits of different options.

The Role Of Models

To answer the latter set of questions, we need to turn to the second way of using EMR data to achieve rapid learning: use the data to build models, and then use the models to answer the questions. The concept is illustrated with an analogy. Imagine that an investigator is interested in the time it takes to get from New York to Boston. Imagine a data set that tracks cars as they pass tollbooths and other checkpoints, among other things. The investigator can answer the question by going to the data set, identifying all of the cars that traveled from New York to Boston, noting the times of departure and arrival, subtracting, and averaging the results. To go further, the investigator might report the results separately for cars that left in the daytime (between 6 a.m. and 6 p.m.) and those that left at night. That would answer the question; with the look-it-up method, that would likely be the end of it. But we can squeeze much more out of the data set if we use it to build a model. For example, if we were careful to collect the right data, we could start with the model "distance = rate × time" and then go on to add the effects of number of lanes, time of day, weather, congestion, accidents, tollbooths, radar, direction signs, curves, road surface, and so on. That model could then be used to answer not only the original question but also any number of other questions, such as the

effects of adding a carpool lane or designing a freeway around New Haven.

The use of models in this fashion is not new. When Boeing designs a new airplane, it faces scores of questions about the design of the wing, shape of the body, weight, engine power, and other variables. It does not build prototypes of all possible combinations, fly them, and get the pilots' assessments. Boeing uses the principles of physics and data from wind tunnels to build mathematical representations of the planes and then flies them inside a computer. It does conduct experiments. But the purpose of the experiments is not to get an answer to one particular question; it is to collect data to build models that can answer hundreds of questions. Virtually every other sector of our economy uses mathematical models to manage complexity and uncertainty. Architects use them to design buildings, engineers use them to calculate the trajectories of satellites, UPS uses them to calculate optimal transportation routes, and Intel uses them to design computer chips. Even in medicine, computed tomography (CT) scans, radiation doses, nurses' schedules, decisions to buy versus lease hospital beds, and the funds in which physicians invest their money are all based on models. So the idea of using data from EMRs to build models and then use the models to answer a variety of questions has a long history of success in a variety of fields, including medicine.

Types Of Models

Data from EMRs along with other sources can be used to help build a wide variety of models. At one end of the spectrum, in terms of simplicity, are single-equation models such as regression equations to help adjust for confounding factors ("risk adjustment") or to help project a patient's future costs based on the patient's history. Both of these are common and well-accepted types of models. But models can be used for considerably more than that. Properly built and validated, they can be used to help analyze physiological processes; design guidelines, performance measures, and the "what-to-do" parts of disease management programs; design the "how-to-do-it" parts of disease management programs, case management protocols, and continuous quality improvement (CQI) projects; forecast logistics, utilization, costs, and cost-effectiveness; set clinical priorities and design strategic goals; prioritize or combine performance measures; analyze the effects of multiple diseases (comorbidities), syndromes that affect multiple organ systems, drugs that have multiple effects, and combinations of drugs; address questions of timing, such as screening, frequency of follow-up visits, or how long a medication should be tried before the dose is changed; and help design and predict clinical trials.

For these types of applications, one needs larger and deeper models. A rule of thumb is that if one wants to answer a question that involves a particular variable, then the model needs to include that variable. And if one wants to explore the effect on particular outcomes of changing variables, the model needs to include the variables that are to be changed, the outcomes of interest, and the pathways that

"If one is interested in a broad and deep range of questions, one needs a broad and deep model."

connect them. For example, if we want a model to help individual patients and clinicians make decisions; to help design guidelines, performance measures, and the "what-to-do" parts of disease management programs; and to be credible, the model should start at the level of physiological and clinical detail at which clinicians think. Essentially, it should encompass all of the biological variables that physicians consider to be important in the management of their patients, and it should relate them in a natural way. This level of detail is also required to help analyze the physiological processes underlying diseases and their treatments and to help design, interpret, and extend clinical trials. If we want a model to address the issues that arise in the design of the "how-to-do-it" parts of disease management programs, case management protocols, and CQI projects, the model should include care processes, logistics, and behaviors at an equally high level of detail. If we want a model to provide credible information about logistics and cost-effectiveness, it should be able to track the use of resources such as facilities, personnel, visits, admissions, equipment, and their costs. If we want a model to help set clinical priorities, design strategic goals, and prioritize or combine performance measures, the model should span all types of interventions (primary prevention, screening, diagnosis, treatment, secondary prevention, and support care) and span multiple diseases using the same methodology. A broad span is also required to address patients who have multiple diseases (comorbidities), syndromes that affect multiple organ systems, drugs that have multiple effects, and combinations of drugs. If we want a model to address questions of timing—such as screening, frequency of follow-up visits, or how long a medication should be tried before the dose is changed—the model should function in continuous time and be able to address events that can change as rapidly as minute by minute, or as slowly as year by year. Finally, if we want a model to be credible, the model should be able to simulate the most important epidemiological studies and clinical trials at the level of clinical detail at which they are designed and reported, and to match or predict their results within the appropriate confidence limits. In general, if one is interested in a broad and deep range of questions, one needs a broad and deep model.

Archimedes: An Example Of A Large-Scale, Physiology-Based Model

The feasibility of building models like this can be illustrated with the Archimedes model, which has been developed over the past ten years by a team led by Len Schlessinger and me.[2] Briefly, the core of the model is a set of ordinary and differential equations that represent physiological pathways at a level of detail roughly comparable to that found in general medical textbooks, clinical trials, and patient

charts. The model includes pathways relating to diabetes, congestive heart failure, coronary artery disease, stroke, hypertension, obesity, and metabolic syndrome in a single integrated model; and asthma in a stand-alone model. Cancers of the lung, breast, and colon are being added. Other conditions will be added in the future.

Stemming from this core are additional equations that represent the development of signs and symptoms; patients' behaviors in seeking care and complying with treatments; clinical events such as telephone consultations, office visits, and admissions; tests and treatments (including errors, side effects, and complications, to the extent they are known); physicians' behaviors and performance; logistics and utilization (for example, emergency department visits, admissions); resources; costs; and measures of quality of life.

Because the model begins with differential equations, time is continuous, biological variables are continuously interacting functions of time, any event can occur at any time, and the timing of events is as condensed or drawn out as occurs in reality. Clinical outcomes are defined in terms of the underlying variables, which enables the model to incorporate different definitions and changes in definitions. The effects of interventions are modeled at the level of the underlying biology. The inclusion of multiple organ systems and diseases as part of a single physiology enables accurate representation of the effects of syndromes, comorbidities, multiple drugs, and drugs with multiple effects. The model is programmed in an object-oriented language called Smalltalk and runs on a grid of computers using distributed computing methods.

The objective is to create a virtual world that can do the things listed four paragraphs above. The virtual world of Archimedes consists of thousands of virtual people, all of whom have virtual physiologies, can get virtual diseases, and have virtual behavior. To represent real populations, Archimedes can create copies or "clones" of real people drawn from the settings in which the questions are being asked, using person-specific data from surveys and data sets such as the National Health and Nutrition Examination Surveys (NHANES), health risk appraisals, personal health records, and, of course, EMRs. It does this at the level of detail captured in the survey, including, if available, demographic characteristics, physical examination results, behavior, family history, medical conditions, biological variables, medical history, symptoms, current medications, and so forth. Archimedes uses the information provided on each person (for example, characteristics, lab values, medications, symptoms, medical history) to calculate the hidden chronic damage these factors imply, such as the degree of stenosis in coronary arteries.

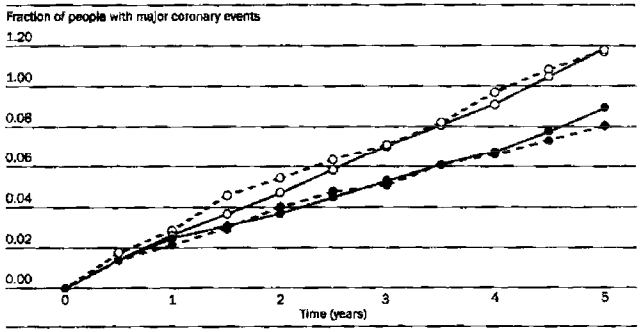
The model is constructed so that it can be tailored to different settings, distinguished by specific populations (for example, age, sex, race/ethnicity, behavior, risk factors, incidence rates); clinical practice guidelines and levels of performance; and costs. Examples of possible settings are a managed care organization, the U.S. population, the employees of General Motors, Medicare, Colorado Medicaid, people covered by Blue Cross of Michigan, and the uninsured of Los Angeles.

Example Of An Application

The potential uses of a large-scale simulation model can be illustrated with an example. Suppose a medical director is interested in the effects of giving simvastatin 40 mg, a cholesterol-lowering drug, to men and women ages 40–80 who have had a previous MI or currently have angina, or who have had a bypass graft or angioplasty. We have already seen that this type of question, as common as it is, is beyond the capability of the look-it-up approach for EMRs. We could use a model like Archimedes to select virtual people in the virtual world who meet those criteria, as well as other characteristics we might want to specify (for example, average HDL level, proportion taking aspirin), randomize them to receive simvastatin or not (delivered by virtual physicians through virtual visits, and so forth), follow them for ten years in virtual time, let the patients' virtual physiologies progress (for example, virtual atherosclerosis developing in virtual coronary arteries), and record the occurrences of virtual MIs. The results of such a virtual trial are shown as the dashed lines in Exhibit 1.

One can appreciate the range of potential uses of models like this by listing some of the questions the Archimedes model has been used to address recently: (1) What would be the effect of a cure for insulin resistance (insulin resistance is the major cause of diabetes and the metabolic syndrome)? How would such a cure compare with simply getting all patients and physicians to faithfully follow existing guidelines for control of risk factors such as hypertension, high cholesterol,

EXHIBIT 1
Comparison Of Model Results With Those Of An Actual Clinical Trial: Cumulative Probability Of Major Coronary Events In The Heart Protection Study



SOURCE: D.M. Eddy and L. Schlessinger, "Validation of the Archimedes Diabetes Model," *Diabetes Care* 26, no. 11 (2003): 3102–3110.
NOTES: Comparison of results predicted by the Archimedes model (dashed lines) versus actual results (solid lines) from the Heart Protection Study (HPS). Open circles are placebo group; solid circles are treated group.

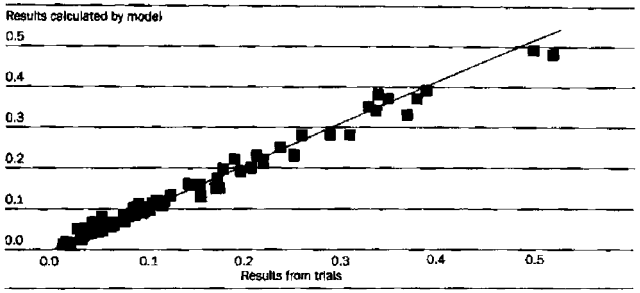
and smoking? (2) What would be the effect on biological variables and clinical outcomes of a drug that decreases weight by 5 percent? For what populations would such a drug be best indicated if the goal is to decrease heart attacks? To decrease net costs? (3) How does delivering insulin through inhalation compare with the current method of injection in people with Type 1 diabetes? What are the implications of the different effects on HbA1c for downstream clinical outcomes? What trial should we conduct to determine the appropriate indications? (4) We want a calculator that will tell people their risk of diabetes and its complications, taking into account not only the usual Framingham-type variables (sex, age, systolic blood pressure, total cholesterol/HDL cholesterol, whether they smoke, whether they have diabetes) but also duration and severity of diabetes, medical history (for example, previous MI), past treatments, past and current weight, current symptoms and complications, and current medications. (5) What are the relative effects of raising performance from the average to the ninetieth percentile level for each of the Health Plan Employer Data and Information Set (HEDIS) measures for cardiovascular disease, diabetes, congestive heart failure, and tobacco use? If we can focus on only three measures, which are the most important? (6) We've done a Phase II trial that showed the effects of [a particular drug] on [a particular set of biomarkers]. What would a Phase III trial show for clinical outcomes? What would happen in different populations (that is, different indications)? What is the optimal size and duration for a trial? (7) We want to combine two drugs. What will be the combination drug's effects in populations that have different initial risks, different current treatments, and different current degrees of control? (8) We are contemplating [a new approach to the treatment of diabetes]. We want to know its effects on clinical outcomes and costs. Will it save money? If so, by when?

Validation Of Large-Scale Models

A critical aspect of using models, large or small, is validating them. Large-scale simulation models provide a helpful method for accomplishing this; their accuracy can be checked by using the virtual world to simulate activities that have happened in the real world and then comparing the results. For example, the application just described—the effects of simvastatin 40 mg on MIs in people at high risk of cardiovascular disease—has been studied in a real clinical trial called the Heart Protection Study (HPS).[3] The real results are shown in Exhibit 1 as solid lines. Here it is important to understand that the HPS was not used to build the Archimedes model; it is being used only to test it. The close concordance between the real and virtual results builds confidence in the model's accuracy.

This approach of comparing predicted results against real results can be repeated for a range of trials that span the types of questions the model will be used to address. For example, thus far the Archimedes model has simulated about three dozen clinical trials relating to diabetes and coronary artery disease. Exhibit 2

EXHIBIT 2
Comparison Of Model Results With Those Of An Actual Clinical Trial For Seventy-Four Validation Exercises



SOURCE: D.M. Eddy and L. Schlessinger, "Validation of the Archimedes Diabetes Model," *Diabetes Care* 26, no. 11 (2003): 3102–3110.
NOTES: Actual results (horizontal axis) versus calculated results (vertical axis) from seventy-four independent validation exercises involving eighteen randomized controlled trials relating to diabetes and cardiovascular disease. Forty of the validation exercises are independent; no data from the trial were used to build or calibrate the model. Each square represents a particular outcome of a particular arm of a particular clinical trial (outcomes vary from arm to arm and include morbidity, mortality, and various clinical markers). Perfect accuracy would show all of the squares on the diagonal line. Deviations from this line are attributable to sampling variations in the real clinical trials as well as any inaccuracies in the model.

presents a scatter plot showing the accuracy of the first seventy-four exercises (counting the different arms and outcomes of a trial as separate exercises).[4] More than half of the validations are independent, which means that no information from the real trials was used to build the model. Overall, the correlation between the real and predicted results is 0.98. For the independent validations, the correlation is 0.96. Although these results certainly do not prove that every new analysis will be accurate, they do indicate that it is possible to build large-scale, physiology-based models that reflect the available evidence from clinical and epidemiological studies reasonably well.

Using Data From EMRs

The quality of a model depends directly on the quality of the data available to build it. To build a large-scale, physiology-based simulation model like Archimedes, one would ideally like to have access to several data sets that are large (from many people), person-specific (reporting the measurements and events of each person separately), comprehensive (recording of all the variables of interest), and longitudinal (tracking each person's measurements and events over time—preferably five to ten years). Until very recently, this has been extremely difficult, if not impossible, to achieve. A model like Archimedes typically has to be constructed from published data—epidemiological studies, basic physiology studies, large surveys, and clinical trials. These may be supplemented by single-point-in-time

"Access to EMR data will make large-scale models like Archimedes much more powerful and accurate."

(not longitudinal) data from a public survey like NHANES and by occasional access to person-specific data from narrowly defined studies of particular physiological pathways, but the majority of the model is built from published sources.

Published data are limited by two main facts. First, they are aggregated; they give the average results in the population without reporting the values for individual people. Second, they report what the investigators are interested in, which is not always what a modeler needs. These limitations can be severe. To appreciate this, let us revisit the example of travel times from New York to Boston. Suppose we are interested in how rainy conditions affect travel time. A published summary of the data set might report the average time it takes cars to go from one point to another and the average number of inches of rainfall. It is easy to see that this is not sufficient to answer the question. To determine the relationship between travel time and rainfall, we need data on each car—the amount of rain it was experiencing and its times at each point on the route. This illustrates only one way in which the object-specific longitudinal data can be used. There are many others. Compared with using only reported averages, this type of object-specific information is like moving from black and white TV to color, or from a car radio to high-end hi-fi.

Data from EMRs could be extremely helpful in filling this gap. First, they could be used to develop a much more accurate understanding of biological pathways—how different biological variables relate to one another in complex physiological processes. Second, such data would greatly improve our understanding of individual variations—how the values of biological variables are distributed across individuals and how they change or progress within individuals over time. Third, they can help us understand how medicine is actually practiced, not just in idealized settings like clinical trials but in real settings of patient care. Fourth, they can be used to develop person-by-person copies of specific populations. This would enable analyses to be tailored to particular settings so that the one-size-fits-all assumptions we now have to make could be discarded. A fifth use of EMR data is validation. A model like Archimedes is now validated against clinical trials. Data from EMRs would enable the validations in more realistic settings.

For these and other reasons, we can expect that access to EMR data will make large-scale models like Archimedes much more powerful and accurate. It is important to understand, however, that EMR data are not a panacea; they have their own limitations that will affect their value and building models. To list just three, they are not particularly useful for determining the effects of interventions, for reasons already given; important data may be miscoded or missing; and they typically do not include research variables that are needed to build deeper models of physiological pathways.

Conclusions

The practice of medicine is extraordinarily complicated. The number of questions we face far outstrips our current methods for getting answers. Clinical trials and other well-designed empirical studies are essential for helping us understand human physiology, disease, and the effects of tests and treatments. They are our fundamental anchor to reality. However, it is not feasible to answer all questions with empirical studies. For questions that cannot be answered with empirical studies, mathematical models can be useful. In particular, large-scale, physiology-based models can help answer a wide variety of questions ranging from physiological pathways; to clinical trials; to management tools such as guidelines, performance measures, priority setting, forecasting, and cost-effectiveness. The Archimedes model has demonstrated the feasibility of building these types of models. The main limitation of these types of models at this time is the quality of data available to build and validate them. Patient-specific, comprehensive, longitudinal data from EMRs can greatly improve the quality of data available for building these types of models. Together, the linkage of data from EMRs and large-scale, physiology-based models could open up a promising new way to improve the quality and efficiency of medical care.

NOTES

1. A listing of these papers is available at http://content.healthaffairs.org/cgi/content/full/hlthaff.26.2.w107/DC2
2. L. Schlessinger and D.M. Eddy, "Archimedes: A New Model for Simulating Health Care Systems—The Mathematical Formulation," *Journal of Biomedical Informatics* 35, no. 1 (2002): 37–50. D.M. Eddy and L. Schlessinger, "Archimedes: A Trial-Validated Model of Diabetes," *Diabetes Care* 26, no. 11 (2003): 3093–3101; and D.M. Eddy and L. Schlessinger, "Validation of the Archimedes Diabetes Model," *Diabetes Care* 26, no. 11 (2003): 3102–3110.
3. Heart Protection Study Collaborative Group "MRC/BHF Heart Protection Study of Antioxidant Vitamin Supplementation in 20,536 High-Risk Individuals: A Randomised Placebo-Controlled Trial," *Lancet* 360, no. 9326 (2002): 23–33.
4. Eddy and Schlessinger, "Validation of the Archimedes Diabetes Model."

PERSPECTIVE

Archimedes: A Bold Step Into The Future

A promising new model for improving care, which deserves a solid policy foundation as it moves into use.

by John R. Lumpkin

ABSTRACT: The increasing adoption of electronic health records (EHRs) enables the development of new tools to guide clinical research, clinical protocol development, and national policy formulation. Archimedes is an example of a new generation of tools that go beyond identifying past problems with medical devices and pharmaceuticals or failures with health care delivery to predicting potential problems and identifying new treatments and approaches that can improve care. Although the arrival of this new generation of tools raises some concerns, the tools' great potential for improving care must be carefully considered. [*Health Affairs* 26, no. 2 (2007): w137-w139 (published online 26 January 2007; 10.1377/hlthaff.26.2.w137)]

THE WORLD OF health information technology (IT) has changed dramatically in the past five years. Electronic health records (EHRs) and the associated support systems have catapulted from obscurity into the stratosphere of presidential attention. Although adoption rates among individual physicians remain low, EHRs' ability to reduce errors and improve quality, coupled with the high priority placed on adoption by national leaders, can be expected to lead to increasing adoption over the next few years.[1] EHRs along with the redesign of care delivery can have a big impact on quality through the implementation of decision support. This technology enables the EHR to provide situational knowledge to the clinician in the form of reminders or suggested diagnostic or therapeutic steps.

■ **Current uses of health data.** In the nineteenth century, pioneers such as John Snow and Florence Nightingale demonstrated the power of using data derived from individual health experiences to answer health questions affecting large groups. Snow used the data to identify the cause of an outbreak of cholera, and Nightingale used the data to drive improvements in military hospitals. Large-scale longitudinal studies, such as the Framingham Heart Study, have demonstrated the power of large databases to identify health risks and demonstrate the effectiveness of interventions.[2] Yet despite their proven effectiveness, the construction of these large databases has been restricted by their very large costs. The collection of data from paper-based clinical records is labor-intensive, time-consuming, and expensive.

To serve their primary function, EHRs accumulate large quantities of data in a readily accessible electronic format. In an interoperable environment, data that are stored in many

John Lumpkin (jlumpki@rwjf.org) is senior vice president of the Robert Wood Johnson Foundation in Princeton, New Jersey; he also is director of the foundation's Health Care Group. Prior to that, he was the director of the Illinois Department of Public Health, a cabinet-level post in Illinois. He has served as chairman of the National Committee on Vital and Health Statistics and is a fellow of the American College of Medical Informatics.

DOI 10.1377/hlthaff.26.2.w137

locations are assembled as needed to aid clinical decision making by the patients and their clinicians. Just as clinical data in a paper-based world can drive public health interventions, quality improvement, and research, the data in EHRs can be used for a number of similar secondary purposes.

Several papers in this collection note how large amounts of clinical data made accessible by electronic health information systems can enable rapid assessment of drug and treatment protocol effectiveness as well as identification of unsuspected adverse outcomes.[3] The important work reflected in these papers shows how current tools can be improved with access to more data in a timely fashion.

■ **Promise of Archimedes model.** In contrast, David Eddy's development of the Archimedes simulation model represents the beginning of a new phase of rapid learning enabled by the health information revolution.[4] Traditional studies use data to determine what happened at some point in the past. In the case of an adverse side effect of a pharmaceutical, analysis of large volumes of clinical data can rapidly identify findings that might not be apparent with relatively small clinical trials. Analysis of large clinical data sets can also determine trends in patterns of care that result from the adoption of treatment protocols or policy changes. In each instance, the data analysis determines what happened in the past to guide actions in the future. The patients affected are real people who must live with real consequences. The key innovation in the development of simulation models like Archimedes is that alternative approaches to care and health care policies can be tested before they are implemented. Clinical trials can be performed in an electronic environment indicating usefulness of new techniques and treatments. The important difference with the simulation model is that deaths and injuries or failed improvements are happening to virtual people, not real ones.

As such, Archimedes is to health IT what the first amphibian that crawled out of the primordial swamp was to evolution of human beings: an important evolutionary step and one that foreshadows future wonders. In 1969, the ARPANET was just a tool that a bunch of geeks at four universities used to share research data. Few, if any, could have envisioned that this network would grow into what we know as the Internet. No one could have imagined all the ways the Internet would be used to advance communications, commerce, entertainment, and research. In fact, the Internet has become the base infrastructure that enables interoperability in health IT systems. Archimedes gives us that glimpse into a future where health information is easily accessible to be manipulated in real time to improve the health and health care of all Americans.

■ **Concerns.** The importance of this development must be tempered by concerns that must be addressed by the development of a simulation modeling system like Archimedes. Archimedes is built on the records of the real-world experiences of individuals over time. The raw material that enables the building of a model like Archimedes is information about real people and real health experiences. A longitudinal record has to be built based on identifying data that belong to an individual, linking those data over time and geography. As the model is built, any connection between a real person and the string of information is severed as the data are deidentified and aggregated. Because Archimedes is used to simulate the impact of an intervention of a population of people, the structure and use of the model inherently protect individuals from identification. The use of data in this way is called secondary use. A recent report from the American Medical Informatics Association (AMIA) notes:

> Secondary use of health data can enhance health care experiences for individuals, expand knowledge about disease and appropriate treatments, strengthen understanding about the effectiveness and efficiency of our health care systems, support public health and security goals, and aid businesses in meeting the needs of their customers. Yet, access to and secondary use of data poses complex ethical, political, technical, and social challenges.[5]

Privacy. The most immediate social challenge is the protection of privacy. The deliberate or unintentional release of individually identifiable health information can have devastating affects on an individual. The trust that exists in the health care system is based in part on the belief that private information will be kept private. Although Archimedes itself is protective of privacy, very clear and stringent procedures must be adopted in the construction and maintenance of the model to assure that individual privacy is protected. The AMIA report recommends that addressing the policy issues related to secondary data should be high on the national policy agenda. Archimedes' ultimate usefulness will depend on a firm policy foundation for the secondary use of clinical data.

Assumptions and biases. The second concern is more conceptual. If Archimedes or similar modeling systems are as powerful as the early indications indicate they are, progress in identifying problems and finding solutions can be greatly facilitated. Sizable resources can be preserved by avoiding dead ends before actual research and clinical trials are performed. Promising lines of research can be tested and implemented efficiently and effectively. However, the construction of any model reflects the assumptions and biases of those who build it. Today those assumptions and biases would have minimal impact on an Archimedes-based assessment of an innovation. Over time, though, the assumptions and biases built into the model might become less valid as new research accumulates and knowledge advances. It will be important for the model's ongoing usefulness that the inherent biases and assumptions be reevaluated regularly.

Future innovation. Finally, as innovative as Archimedes may be, its existence may stifle future innovation if research funders rely too heavily on Archimedes as the arbitrator of the usefulness of exploring lines of research. Innovation comes in the form of unconventional approaches to conventional problems. As Archimedes and similar simulation models become conventional, their ability to assess unconventional approaches will need to be used with caution.

THOSE CONCERNS ASIDE, the development of Archimedes as described in Eddy's paper represents an important development in the world of health IT. As a powerful tool for the present and a harbinger of things to come, it represents a new age in the application of IT to preserving and restoring health.

NOTES

1. A.K. Jha et al., "How Common Are Electronic Health Records in the United States? A Summary of the Evidence," *Health Affairs* 25, no. 6 (2006): w496–w507 (published online 11 October 2006; hlthaff.25.w496).
2. National Heart, Lung, and Blood Institute, National Institutes of Health, "Framingham Heart Study," http://www.nhlbi.nih.gov/about/framingham (accessed 3 November 2006).
3. A listing of these papers is available at http://content.healthaffairs.org/cgi/content/full/hlthaff.26.2.w107/DC2.
4. D.M. Eddy, "Linking Electronic Medical Records to Large-Scale Simulation Models: Can We Put Rapid Learning on Turbo?" *Health Affairs* 26, no. 2 (2007): w125–w136 (published online 26 January 2007; 10.1377/hlthaff.26.2.w125).
5. C. Safran et al., "Toward a National Framework for the Secondary Use of Health Data," 8 September 2006, http://www.amia.org/inside/initiatives/healthdata.asp (click the link for "Final Report" PDF; accessed 27 November 2006).

Federal Initiatives To Support Rapid Learning About New Technologies

The federal government is in a unique position to generate information essential to rapid learning in health care.

by Sean R. Tunis, Tanisha V. Carino, Reginald D. Williams II, and Peter B. Bach

ABSTRACT: Health policy and financing reforms place increasing emphasis on the ability of doctors and patients to make informed, cost-conscious care decisions. The federal government is supporting new initiatives in Medicare to increase the supply of reliable information on the benefits and risks of health care technologies. Medicare also is working with the Agency for Healthcare Research and Quality (AHRQ) to evaluate the comparative effectiveness of prescription drugs and other items or services. The value of these efforts will depend on coordination among individuals and institutions in the public and private sectors; clarity about focus, purpose, and priorities; and adequate and reliable long-term funding. [*Health Affairs* 26, no. 2 (2007): w140–w149 (published online 26 January 2007; 10.1377/hlthaff.26.2.w140)]

MANY CURRENTLY POPULAR APPROACHES to improving quality, reducing cost, and expanding access to care depend on doctors' and patients' ability to make informed, cost-conscious decisions about their health care. This popularity evolved in large part from the expanding body of evidence demonstrating inefficient and inconsistent use of health care services and unsustainable trends in health care spending. If patients, clinicians, and other decisionmakers are responsible for making critical health care decisions, they need access to reliable information on the quality, outcomes, and costs of care. Despite considerable progress in producing this kind of information over the past several decades, much more is required to meet these decisionmakers' needs. This paper addresses federal initiatives, particularly those pursued by the Centers for Medicare and Medicaid Services (CMS), designed to support the efficient development of information about the benefits, risks, and costs of health care technologies.

Sean Tunis (sean.tunis@netzero.net) is director of the Center for Medical Technology Policy in San Francisco. Tanisha Carino is director of the Center on Evidence-Based Medicine at Avalere Health LLC in Washington, DC. Reginald Williams is a manager there. Peter Bach is an associate attending physician in the Department of Epidemiology and Biostatistics at the Memorial Sloan Kettering Cancer Center in New York City.

DOI 10.1377/hlthaff.26.2.w140

Although many federal programs fit within this broad scope, we focus on recent efforts that feature several methodological approaches to learning about the outcomes of health care interventions after they have obtained regulatory approval. The CMS aims to support many different analytic needs by developing an integrated database that will include linked information from claims submitted to the CMS by physicians, hospitals, prescription drug plans (PDPs), and other Medicare providers. The Effective Health Care program, under the direction of the Agency for Healthcare Research and Quality (AHRQ), uses systematic literature reviews and analysis of routinely collected clinical and administrative data, primarily from organizations with electronic medical record (EMR) systems. The CMS also supports high-quality effectiveness studies by linking Medicare reimbursement to required protocol-driven collection of prospective clinical data.

This paper emphasizes programs that involve a major role for the CMS, but many important initiatives are also under way at other federal health programs. In the private sector, health plans and other for-profit and nonprofit entities are also pursuing initiatives, in some cases offering valuable lessons that inform the federal initiatives. The objectives, progress, and limitations of these initiatives offer insight into how the federal government can support the development of better evidence for health care decision making and what will be necessary to ensure that these programs achieve their goals. Although it is clear that these efforts depend on the data, resources, and efforts of the CMS, AHRQ, and other federal agencies, the programs' ultimate value will be determined by the degree of engagement by many stakeholders, including researchers, product developers, health plans, employers, medical professionals, and patient organizations.

The CMS's Integrated Data Strategy

On 1 January 2006 the CMS launched its largest expansion of the Medicare program since its inception in 1965, Medicare Part D, adding outpatient prescription drugs to its list of benefits. This coverage change expanded access to prescription drugs for many Medicare beneficiaries and increased the federal government's role as a payer of pharmacotherapy from 2 percent in 2005 to an expected 28 percent in 2006.[1] In addition, the drug benefit provides the CMS, researchers, and the public with a clearer picture of the overall health care experience of the U.S. elderly and disabled populations and with an infusion of new information on the use, safety, and, potentially, effectiveness of medications in these populations.

■ **Medicare's drug data strategy.** On 11 May 2005 the CMS issued a White Paper on its new drug data strategy.[2] It outlined the CMS's intention to integrate new drug claims data from Medicare Part D with the medical data from Medicare Parts A and B. The stated goal of doing so is to support the programmatic missions of both the CMS and the Food and Drug Administration (FDA) by supporting the FDA's postmarketing surveillance activities and the CMS's goal of providing evidence on drugs and drug use for a broad range of conditions. These conditions include off-

"The CMS's goal to integrate electronic health information for all beneficiaries will require the cooperation of plans and providers."

label uses and evidence of drugs' effectiveness and safety in specific subpopulations, such as the elderly and those with multiple chronic conditions.

This initiative signaled the federal government's intent for Medicare data, and other data, to play a more prominent role in the design of new policies and programs of the CMS and its sister agencies. Researchers have already identified the availability of new Medicare Part D data and their integration with existing medical data on the Medicare population as a potential panacea for some of the most pressing problems facing health policy.[3] The high-profile withdrawal of Vioxx from the market, potential problems with other COX-2 inhibitors, and the recall of some models of implantable cardioverter-defibrillators (ICDs) have raised some concerns about the FDA's current capacity to proactively monitor the safety of medical products used in the United States.[4] Beyond these postmarketing safety questions, a linked data set reflecting use of inpatient, outpatient, nursing home, prescription drug, and other services and medical supplies would provide many important insights into the quality, safety, effectiveness, efficiency, and cost-effectiveness of beneficiaries' care.

■ **Obstacles to the initiative.** The CMS and the scientific community face several obstacles in moving the Medicare integrated data initiative forward. Logistical issues include the sheer mechanical challenge of integrating vast amounts of drug claims data with other Medicare claims files and compiling those data into longitudinally analyzable files. The CMS's goal to integrate electronic health information for all Medicare beneficiaries will also require the cooperation of the private health plans and providers that hold this information. Business confidentiality issues also could prove problematic, as the integrated data may be sufficient to reveal the underlying cost structure of a PDP or Medicare Advantage (MA) plan. Finally, the CMS will have to address privacy and confidentiality issues, such as the nature of institutional review board (IRB) oversight of the research that the CMS and others plan to conduct using these data.

A number of methodological issues will arise as well. For instance, one of the major objectives of data analysis is to focus on drug usage and monitoring rare adverse events. Ensuring against false-positive findings in this context will require a great deal of effort. Such findings are a natural by-product of multiple analyses; in a drug surveillance system that will rely on vast groups of individual researchers, the number of analyses and hypotheses will be very great, and the number of false-positive findings will be as well. Using claims data to understand causal or associative relationships will remain a challenge, regardless of the richness of the data source. There will continue to be debate about whether the information captured on either a Medicare claim or an electronic record accurately reflects the experience of the patient. One sensible initial approach to both of these concerns is to establish robust and uniform standards for data definitions, analyses, and distinct analytic subsets of the claims data, so that findings obtained through analyses of one data set could be validated in a separate data set. This approach would make it easier for the private sector, academe, and the regulatory bodies to share a common understanding regarding the meaning of claims data analyses.

■ **Ensuring optimal uses of the data.** The CMS is working with AHRQ on several studies to explore the optimal uses of Medicare claims data and other observational data. Researchers are attempting to develop frameworks and algorithms to use existing and future Medicare data—particularly Medicare Part D claims—to explore drug safety issues and comparative effectiveness research. Researchers also are conducting studies to validate the accuracy of administrative data using medical records.[5] These studies address some of the limitations of claims data and should lead to more effective use of Medicare prescription drug data. Methods and standards for cost-effectiveness analysis developed by AHRQ and others will also be increasingly important in evaluating the comparative effectiveness of drugs and other health care interventions, particularly as patients and clinicians become more accountable for the outcomes and costs of their care. On 18 October 2006, the CMS published a proposed rule on use of Medicare outpatient prescription drug benefit (Part D) claims data for research, analysis, reporting, and public health functions by the Department of Health and Human Services, federal oversight agencies, and other researchers.

AHRQ's Effective Health Care Program

To learn more about prescription drugs and other health care services, in October 2005 AHRQ formally unveiled the Effective Health Care program, authorized by Section 1013 of the Medicare Prescription Drug, Improvement, and Modernization Act (MMA) of 2003.[6] In MMA, Congress mandated that AHRQ conduct comparative effectiveness research focused on the needs of the Medicare, Medicaid, and State Children's Health Insurance Program (SCHIP) populations. According to MMA, comparative effectiveness research should address the outcomes and effectiveness of health care items and services (including prescription drugs) and gaps in clinical research.

AHRQ responded to the need for evidence to make informed decisions by developing a research program to increase such knowledge. The Effective Health Care program (1) synthesizes knowledge in comparative effectiveness reports; (2) generates knowledge through rapid-turnaround research using deidentified data on patients; and (3) translates knowledge into targeted products for health care decisionmakers.[7] AHRQ developed a parallel research infrastructure to fulfill these aims. The program uses existing Evidence-based Practice Centers (EPCs) and created the new Developing Evidence to Inform Decisions about Effectiveness (DEcIDE) network and the John Eisenberg Clinical Decisions and Communica-

"The DEcIDE network advances the development of scientific methods to refine new methods to analyze data."

tions Science Center (Decision Sciences Center).[8]

For the Effective Health Care program, EPCs conduct comparative effectiveness reviews that highlight what is known about health care technologies while identifying research gaps where more information is needed and unanswered questions that future research should address. The DEcIDE network, comprising thirteen research centers, fulfills item 2 above using Medicare's claims data; EHR data; pharmacy records; private health insurance data; and disease, procedure, or device registries. In addition, the DEcIDE network advances the development of scientific methods to synthesize research and refine new methods to analyze data.[9] With adequate resources, this network could ultimately contribute much new knowledge through prospective studies of important questions identified by systematic reviews and observational studies.

■ **Collaborative process.** AHRQ's approach to implementing the Effective Health Care program involves stakeholders collaborating in a variety of ways. Experts and stakeholders recommend topics for research, which are reviewed and prioritized by an interagency steering committee with representatives from AHRQ, the CMS, the FDA, and the U.S. Department of Health and Human Services (HHS) Office of the Secretary.[10] AHRQ also established a Stakeholder Group for the Effective Health Care program, which provides input on evidence gaps, development of research questions, and methodological issues related to the program. The program also offers opportunities for input on the development of future research topics, key questions, and draft reports by allowing the public to submit comments to AHRQ. However, AHRQ has faced criticism for not disclosing how the program incorporates these comments into its decision making.

As of July 2006, three comparative effectiveness reviews have been released: management strategies for gastroesophageal reflux disease (GERD), noninvasive diagnostic tests for evaluating breast abnormalities, and epoetin and darbepoetin for managing anemia in patients undergoing cancer treatment. The report on noninvasive diagnostic tests highlighted the limited amount of good evidence to support the use of diagnostic imaging as an alternative to biopsy in the evaluation of breast abnormalities, prompting a strong negative reaction from the imaging community.[11] Because these findings challenge conventional clinical thinking, this report could serve to promote shared understanding among clinicians, product developers, and research funding organizations about the critical gaps in evidence and, ideally, lead to the design and execution of studies to address those gaps.

■ **Challenges ahead.** Although AHRQ's Effective Health Care program provides an important opportunity to learn more about the effectiveness of health care technologies, many challenges lie ahead. For the program to be a success, AHRQ

must (1) identify research gaps on comparative effectiveness; (2) develop a research agenda that addresses these gaps, including prospective clinical trials; (3) translate research findings into information that decisionmakers can act upon; and (4) balance the need for adequate stakeholder input with the importance of producing timely and objective findings.[12] These challenges are common to all organizations engaged in evaluating the comparative effectiveness of medical technologies, and past efforts highlight the importance of robust methods, adequate funding, and political support.

Linking Medicare Payment To Prospective Data Collection

The federal government also supports the development of new information about the benefits and risks of health care services by linking payment for services to the collection of prospective clinical, demographic, and economic data. Medicare has taken this approach on a range of mechanisms and recently on the generation of high-quality data about specific new and emerging medical technologies.

■ **Early clinical trials policy.** An early manifestation of Medicare policies supporting the conduct of clinical trials was the 2000 national coverage decision on Medicare payment for the routine costs of clinical trials.[13] Although the policy was designed primarily to support Phase III clinical trials of anticancer drugs, it was written to accommodate a wide range of drug and device trials. No formal evaluation of this policy has been conducted, but anecdotal reports indicate that some trials have proceeded more rapidly because of the reassurance that the costs of routine care will be covered among participating institutions. Other anecdotes have suggested that ambiguities in the policy have hindered its implementation. In July 2006 the CMS announced that it is reconsidering its clinical trials policy and seeks to expand the policy to include other types of clinical research.[14]

■ **Coverage for specific technologies.** The first example of Medicare coverage for a specific technology linked to beneficiaries' enrollment in a specific clinical trial occurred in 1995 through a national coverage decision (NCD) on lung volume reduction surgery (LVRS), a surgical treatment for severe emphysema.[15] The CMS and the National Heart, Lung, and Blood Institute (NHLBI) agreed to collaborate on the National Emphysema Treatment Trial (NETT), a randomized, controlled trial that ran for more than five years, enrolled more than 1,000 patients, and cost more than $100 million. Final study results demonstrated that quality of life improved for some patients, while a small subset of patients experienced increased survival from surgical intervention. As a result of the NETT, clinicians gained an understanding of the negative outcomes associated with the surgery; fewer than 500 procedures have been performed, possibly sparing many Medicare beneficiaries from harm, while also saving the program sizable unnecessary expenditures.[16]

Medicare coverage of LVRS offered a model by which to cover certain high-impact, promising technologies while they undergo further prospective evaluation. Since that time, the CMS has issued policies linking prospective data collec-

"The foremost requirement is recognition that the government is in a unique position to generate much information."

tion to coverage for the use of fluorodeoxyglucose positron-emission tomography (FDG-PET) for suspected dementia and for diagnostic use in oncology, for the use of ICDs in patients at high risk of sudden cardiac death, and for certain off-label uses of drugs approved for colorectal cancer.[17]

■ **Coverage with evidence development.** More recently, Medicare has taken steps to formalize its coverage policy approach under coverage with evidence development (CED), which is explained in a draft guidance document from April 2005.[18] The guidance document provides a description of Medicare's efforts to support the rapid development of high-quality evidence through coverage policy. Numerous public comments on the draft guidance were submitted to the CMS, and a revised guidance document was released in July 2006. It defines two forms of CED: coverage with appropriateness determination (CAD) and coverage with study participation (CSP). The CAD policy will be applied to items and services that are determined reasonable and necessary for Medicare coverage, and additional data collection would ensure that beneficiaries receiving the item or service meet criteria specified in the NCD. Under the CSP policy, inadequate evidence exists to conclude that an item or service is reasonable and necessary for Medicare coverage, but additional clinical research data would help clarify the benefit to Medicare beneficiaries, and then the item or service could be covered.[19]

Medicare's initial experiences with CED are already beginning to produce a large body of data reflecting the clinical experiences of patients using recently developed technologies. Certain refinements to this policy are already under way, including an effort for more explicit and systematic selection of study topics. Furthermore, future CED will be increasingly attentive to the ability of the chosen study designs to answer the most important clinical questions. It is unclear whether the ICD or PET oncology registries will provide the breadth of reliable information on the effectiveness of these technologies for making refinements in coverage policy for specific patient subpopulations. But even these examples demonstrate the potential for a constructive collaboration between multiple stakeholders to engage in the development of better evidence relevant to the care of Medicare beneficiaries. Improving the scientific value and operational simplicity of these early CED efforts will depend on the active participation of private-sector experts and stakeholders, particularly patients and medical professionals.[20]

■ **Other Initiatives.** Since Medicare's work in CED began, several other entities have launched initiatives that are likely to assist in further refining the approach within Medicare and perhaps help increase the engagement of other stakeholders. The Institute of Medicine (IOM) has convened a Roundtable on Evidence-Based Medicine to discuss a broad portfolio of issues related to rapid creation of evidence

for decisionmakers. A recent paper has proposed the establishment of a large national program to support comparative effective reviews and clinical trials to provide payers with better information for coverage and payment decisions.[21] In addition, a new private-sector initiative, the Center for Medical Technology Policy, funded by the California HealthCare Foundation and the Blue Shield of California Foundation, is bringing together public- and private-sector stakeholders to select and design prospective studies of new and emerging health care technologies.[22]

Other Federal Initiatives That Support Rapid Learning

■ **Veterans Affairs.** The Department of Veterans Affairs (VA) has focused for many years on the rapid and efficient development of high-quality evidence about health care technologies. The VA system is much like the government-sponsored, organized health systems in other countries. Its leaders have recognized the importance of an adequately supported applied research capacity to the effective and efficient use of their resources. There has been much progress at the VA in developing and using EMRs and quality improvement methods, and the quality and efficiency of care have improved dramatically.[23]

■ **Office of the National Coordinator.** The Office of the National Coordinator of Health Information Technology (ONCHIT) and the National Health Information Infrastructure will be critical to rapid learning. The federal government plays a vital role in setting standards, thereby supporting the development of vigorous, market-based products. Stakeholders should pay greater attention to the capacity of standard systems to gather information that is useful in measuring and reporting on quality of care and in supporting clinical research applications. In the future, EMR systems, in tandem with Medicare's integrated data strategy, have the opportunity to create valuable information that the federal government can use to make decisions about the allocation of resources to address critical research gaps and prudently purchase health care services.

■ **National Institutes of Health.** The NIH focus on translational research and the clinical research Roadmap initiative are also potentially important in rapid learning. It is important and useful for the NIH to be supporting the development of infrastructure within the health care system to conduct prospective pragmatic and effectiveness studies rapidly and efficiently.

Concluding Comments

The federal government has undertaken a number of initiatives with the objective to produce reliable information, rapidly and efficiently, about the benefits, risks, and costs of alternative health care interventions. Some of these approaches boast a longer record by which to evaluate their likely contributions, while others offer only limited data by which to assess impact. Each initiative illustrates the unique potential of the federal government to make meaningful progress toward these goals, and each also offers insights into the challenges and opportunities.

The foremost requirement is recognition that the federal government is in a unique position to generate large amounts of information essential to improving the appropriateness of decisions and the quality of care. Furthermore, the use of public dollars to finance public health care programs places a unique responsibility on federal policymakers to take the steps necessary to rapidly learn and improve. For this to become reality, several requirements must be met. First, these efforts need to be better coordinated and systematically directed at high-priority issues and questions. Second, there will need to be ongoing and meaningful engagement with private-sector experts and stakeholders. Third, it will be important to recognize the applied nature of the research agenda and that decision-makers' information needs should be a major factor in research priorities and study design, even if these needs might not be the most interesting questions to investigate from an academic standpoint. Finally, these programs will require sufficient attention and resources to achieve their considerable potential. Adequate and sustained funding, clear messages about the importance of these programs, and dedicated staff with these efforts as their primary assignment (protected from the pressures of vested interests) will be required.

Each of the initiatives described above has great potential, but their actual contributions to improved quality and more efficient care delivery will depend on stronger coordination among individuals and institutions in the public and private sectors; more clarity about focus, purpose, and priorities; better staffing; and additional dedicated and reliable funding. The private sector can also contribute greatly on its own to rapid learning in health care; however, no private-sector entities have the size, visibility, and resources potentially available to the CMS and other federal health programs.

The authors thank Lynn Etheredge and John Iglehart for their vision in bringing together researchers and policymakers to discuss the subject of rapid learning in health care. The authors have ongoing consulting relationships with public and private life sciences companies, public and private payers, trade associations, and federal agencies. None of these relationships represents direct conflicts of interest related to the content of this paper.

NOTES

1. S. Heffler et al., "U.S. Health Spending Projections for 2004–2014," *Health Affairs* 24 (2005): w74–w85 (published online 23 February 2005; 10.1377/hlthaff.w5.74).
2. Centers for Medicare and Medicaid Services, "Medicare Prescription Drug Data Strategy: Improving Evidence for Patient Care through the Medicare Prescription Drug Benefit," 11 May 2005, http://www.medicarerxguide.com/MDRG/2005/CMS_Data_Strategy_for_Part_D.pdf (accessed 25 April 2006).
3. R. Platt and A. Ommaya, "A Beneficial Side Effect of the Medicare Drug Benefit," *New England Journal of Medicine* 353, no. 26 (2005): 2742–2743.
4. Center for Devices and Radiological Health, *Ensuring the Safety of Marketed Medical Devices: CDRH's Medical Device Postmarket Safety Program*, 18 January 2006, http://www.fda.gov/cdrh/postmarket/mdpi-report.pdf (accessed 25 April 2006).
5. See Agency for Healthcare Research and Quality, "Work in Progress," http://effectivehealthcare.ahrq.gov/products/progress.cfm (accessed 30 October 2006).
6. AHRQ, "AHRQ Launches New 'Effective Health Care Program' to Compare Medical Treatments and Help Put Proven Treatments into Practice," Press Release, 29 September 2005, http://www.ahrq.gov/news/press/pr2005/effectivepr.htm (accessed 19 December 2006).
7. Ibid.
8. J.R. Slutsky and C.M. Clancy, "The Agency for Healthcare Research and Quality's Effective Health Care Program: Creating a Dynamic System for Discovering and Reporting What Works in Health Care," *American Journal of Medical Quality* 20, no. 6 (2005): 358–360.
9. AHRQ, "Fact Sheet: Centers for Education and Research on Therapeutics (CERTs)," Pub. no. 02-P025, April 2004, http://www.ahrq.gov/clinic/certsovr.pdf (accessed 19 December 2006).
10. AHRQ, "Medicare Prescription Drug, Improvement, and Modernization Act of 2003, Section 1013: Suggest Priority Topics for Research," *Federal Register* 69, no. 79 (2004): 22045–22047.
11. American College of Radiology, "2/9/2006—AHRQ Breast Imaging Report Confused and Inaccurate," Press Release, http://www.acr.org/s_acr/doc.asp?CID=2540&DID=23431 (accessed 19 December 2006).
12. As of November 2006, the Effective Health Care program had not undertaken any prospective comparative effectiveness clinical trials.
13. Institute of Medicine, *Extending Medicare Reimbursement in Clinical Trials* (Washington: National Academies Press, 2000).
14. CMS, "NCA Tracking Sheet for Clinical Trial Policy (CAG-00071R)," 10 July 2006, http://www.cms.hhs.gov/mcd/viewtrackingsheet.asp?id=186 (accessed 30 October 2006).
15. T. Catino, S. Sheingold, and S. Tunis, "Using Clinical Trials as a Condition of Coverage: Lessons from the National Emphysema Treatment Trial," *Clinical Trials* 1, no. 1 (2004): 108–114.
16. G. Kolata, "Medicare Says It Will Pay, but Patients Say 'No Thanks'," *New York Times*, 3 March 2006.
17. B.P. Carey, "Medicare Coverage of PET for Cervical Cancer," *Journal of the American College of Radiology* 3, no. 1 (2006): 19–22; M.B. McClellan and S.R. Tunis, "Medicare Coverage of ICDs," *New England Journal of Medicine* 352, no. 3 (2005): 222–224; and CMS, "Decision Memo for Anticancer Chemotherapy for Colorectal Cancer (CAG-00179N)," 28 January 2005, http://www.cms.hhs.gov/mcd/viewdecisionmemo.asp?id=90 (accessed 3 November 2006).
18. CMS, "Draft Guidance for the Public, Industry, and CMS Staff: Factors CMS Considers in Making a Determination of Coverage with Evidence Development," 7 April 2005, http://www.cms.hhs.gov/coverage/download/guidanceced.pdf (accessed 12 December 2006).
19. CMS, "National Coverage Determinations with Data Collection as a Condition of Coverage: Coverage with Evidence Development," 12 July 2006, http://www.cms.hhs.gov/mcd/ncpc_view_document.asp?id=8 (accessed 3 November 2006).
20. S.R. Tunis, "A Clinical Research Strategy to Support Shared Decision Making," *Health Affairs* 24, no. 1 (2005): 180–184.
21. G.R. Wilensky, "Developing a Center for Comparative Effectiveness Information," *Health Affairs* 25 (2006): w572–w585 (published online 7 November 2006; 10.1377/hlthaff.25.w572).
22. See the Center for Medical Technology Policy home page, http://www.cmtpnet.org.
23. See J. Kupersmith et al., "Advancing Evidence-Based Care for Diabetes: Lessons from the Veterans Health Administration," *Health Affairs* 26, no. 2 (2007): w156–w168 (published online 26 January 2007; 10.1377/hlthaff.26.2.w156).

PERSPECTIVE

Challenges Ahead For Federal Technology Assessment

The government's rapid-learning initiatives, although vital for technology assessment, will face challenges on several fronts.

by **Peter J. Neumann**

ABSTRACT: Sean Tunis and colleagues provide an excellent critique of current federal activities to assess new medical technology. These efforts generally do not involve primary data collection but rather reflect attempts to better synthesize existing information, to make conditional coverage decisions based on the data, and to increase coordination among government agencies. Many challenges remain on analytical, logistical, legal, and political fronts. Researchers and analysts should be more precise about what "rapid learning" means and strive to measure performance. Efforts are also needed to prioritize research, to communicate it to decisionmakers, to involve stakeholders in the process, and to include cost-effectiveness information. [*Health Affairs* 26, no. 2 (2007): w150–w152 (published online 26 January 2007; 10.1377/hlthaff.26.2.w150)]

ONE DOESN'T ALWAYS see the words "federal" and "rapid" in such close proximity as in the title of the paper by Sean Tunis and colleagues, so one approaches their paper with a mixture of curiosity and skepticism.[1] The piece, however, provides an excellent critique of current government activities to assess new medical technology. Although the authors sidestep a few thorny issues, they generally make a compelling case that federal policymakers are actively and in some cases creatively attempting to improve efficiency as they attempt to balance rigor and timeliness in technology assessment. Tunis and colleagues catalogue the federal initiatives by type of activity, but they might have divided them into one of three categories: making better use of existing information; developing conditional coverage strategies; and improving coordination across agencies.

■ Making better use of existing information. Although Tunis and colleagues don't call attention to it, their review underscores an underappreciated phenomenon: Much of the recent activity in evidence policy does not involve primary data collection. Rather, it involves systematically reviewing existing data through evidence syntheses and data-mining techniques, and strengthening available databases. The initiatives of the Agency for Healthcare Research and Quality (AHRQ), stemming from Section 1013 of the Medicare Prescription Drug, Improvement, and Modernization Act (MMA) of 2003, are a prime example. AHRQ has carried out its statutory mandate admirably, through existing infrastructure, such as the Evidence-based Practice Centers (EPCs), and newly created programs.

Over the years, systematic reviews have evolved from impressionistic and unstructured assessments to more rigorous and methodical evaluations.[2] Many challenges remain for the discipline, including how to select studies and pool data from nonrandomized sources; how to weigh trade-offs about risks and benefits; how to prioritize research; how to communicate results to decisionmakers; and how best to involve stakeholders in the process. Another challenge pertains to how and when to include cost-effectiveness information. Tunis and colleagues argue that methods and standards for cost-effectiveness analysis, developed by AHRQ and others, will also be increasingly important. However, previous attempts at advancing this approach have been vigorously resisted, and it is unclear how the authors expect it to develop.

"The real question is not whether learning is more rapid, but whether decisions based on that learning are more rapid."

■ Developing conditional coverage strategies. Medicare's new guidance on coverage with evidence development (CED) in some ways represents old wine in new bottles. Almost all Medicare national coverage decisions over the years have come with conditions, and as Tunis and Steven Pearson noted in a previous paper, some, such as the National Emphysema Treatment Trial (NETT) and prophylactic use of implantable cardioverter-defibrillators (ICDs), have required explicit evidence development.[3] What is different are the attempts to clarify and codify the policy, to provide legal justification, and to establish a framework for the collection of prospective data in the context of coverage determinations.[4] The CED guidance reflects some forward movement on the part of federal officials, who have long struggled to unchain themselves from simple yes-no decision rules for new technologies and to ground their authority in existing law. In theory, it should provide more flexibility. In practice, it will present challenges in figuring out exactly how and when to use CED. It also will not remove hard choices, and expectations about the policy's impacts should be appropriately modest.

Another point bears mentioning: The discussion about CED is an offshoot of a larger debate about evidence that includes the Food and Drug Administration's (FDA's) regulatory approval procedures. The FDA's drug evaluation decisions reflect a kind of CED policy that involves approving drugs while continuing to monitor adverse-event data and oversee Phase IV commitments (although not as rigorously as many of its critics would like). FDA officials should watch the unfolding Medicare CED experience closely as they consider ways to allow more flexibility in clinical trial design and approval decisions.

■ Improving coordination across agencies. Calling for better coordination across government is a favored strategy for policymakers and politicians looking to improve efficiencies—and to appease constituencies. Recent federal initiatives for coordinating national security agencies have also been motivated by a "rapid learning" philosophy.

Improved coordination offers promise for technology assessment. Medicare data can help the FDA understand risks and benefits of therapies after they have been approved. The Centers for Medicare and Medicaid Services (CMS) can communicate with the FDA earlier in the regulatory approval process to anticipate data needs for reimbursement. AHRQ can bring expertise to both parties. In practice, different statutory authority, constituencies, and cultures across agencies make this difficult to accomplish. Also, Tunis and colleagues highlight opportunities for coordination and cooperation with the private sector, but they do not elaborate on how this would happen.

■ What does "rapid" mean? For all of the discussion about rapid learning and faster-turnaround research, Tunis and colleagues' paper and the government initiatives they describe do not address what "rapid" means, except to leave the hazy impression that technol-

Peter Neumann (pneumann@tufts-nemc.org) is director of the Center for the Evaluation of Value and Risk in Health at the Institute for Clinical Research and Health Policy Studies, Tufts–New England Medical Center, and a professor of medicine at Tufts University Medical School in Boston, Massachusetts.

DOI 10.1377/hlthaff.26.2.w150

ogy assessments will be made available more quickly than in the past. Researchers and analysts should be more precise about such matters and strive to measure performance.

The real question is not whether learning is more rapid, but whether decisions based on that learning are more rapid. There is evidence, for example, that when national coverage decisions are referred to the Medicare Coverage Advisory Committee (MCAC) or require a formal health technology assessment, review times are delayed by several months.[5] There is also evidence that when evidence is good, the CMS makes decisions more quickly.[6] MMA set stricter timelines for Medicare coverage decisions. It will be important to measure CMS performance against these standards and to analyze the impact of decisions on Medicare costs and health outcomes.

An alternative approach involves reframing the question. The issue facing any decision-maker is not whether learning is rapid, per se, but rather whether existing information is sufficient to make a decision. Decisions are made when the costs of collecting additional information exceed the benefits. Analysts should pursue the further development of value of information techniques for deciding whether to adopt or reject an intervention based on existing information or whether additional information is warranted.[7]

■ **The political subtext.** Finally, there is an ideological dimension to the discussion about the federal government's rapid-learning initiatives that Tunis and his colleagues only hint at and that is seldom mentioned in debates about evidence. The initiatives presume that the federal government has a unique and vital role in conducting technology assessment and disseminating findings. To many observers, this seems self-evident, but it runs headlong into a view held by some stakeholders that federal technology assessment can have a darker side: that it places bureaucrats between doctors and patients, that it represents one-size-fits-all medicine, that it jeopardizes physician autonomy, and that there are risks with having a single assessor making judgments.

The potential conflicts are mentioned in passing by Tunis and his colleagues (they note that the AHRQ report on noninvasive diagnostic tests for evaluating breast abnormalities received a negative reaction in the imaging community). CMS policymakers have attempted to avoid controversies by making innocuous-sounding statements that Medicare should help ensure that the right drug gets to the right person in the right setting. They have also assiduously avoided denying coverage based on costs. AHRQ for its part has striven to gingerly manage its constituencies. Navigating these shoals may be the biggest challenge of all.

NOTES

1. S.R. Tunis et al., "Federal Initiatives to Support Rapid Learning about New Technologies," *Health Affairs* 26, no. 2 (2007): w140–w149 (published online 26 January 2007; 10.1377/hlthaff.26.2.w140).
2. D.M. Fox, "Evidence of Evidence-Based Health Policy: The Politics of Systematic Reviews in Coverage Decisions," *Health Affairs* 24, no. 1 (2005): 114–122.
3. S.R. Tunis and S.D. Pearson, "Coverage Options for Promising Technologies: Medicare's 'Coverage with Evidence Development'," *Health Affairs* 25, no. 5 (2006): 1218–1230.
4. Ibid.
5. P.J. Neumann et al., "Medicare's National Coverage Decisions, 1999–2003: Quality of Evidence and Review Times," *Health Affairs* 24, no. 1 (2005): 243–254.
6. Ibid.
7. K. Claxton, J.T. Cohen, and P.J. Neumann, "When Is Evidence Sufficient?" *Health Affairs* 24, no. 1 (2005): 93–101.

PERSPECTIVE

Speed Bumps, Potholes, And Tollbooths On The Road To Panacea: Making Best Use Of Data

Using new data resources effectively will require considerable investment in infrastructure and in training those who use them.

by Richard Platt

ABSTRACT: Electronic health databases promise to transform both the assessment of health care delivery and our understanding of treatments' safety and effectiveness. To achieve these goals, it will be necessary to (1) recognize limits on inferring causality; (2) protect confidentiality while allowing important societal gain; (3) link health data back to the individual patient; (4) obtain additional information from medical records; (5) understand ways in which electronic data can misrepresent reality; and (6) create the infrastructure, expertise, and resources to use the data. Realizing databases' potential will require long-term commitment and investment beyond the maintenance of the databases themselves. [*Health Affairs* 26, no. 2 (2007): w153–w155 (published online 26 January 2007; 10.1377/hlthaff.26.2.w153)]

THE IMPENDING availability of vast electronic health information databases, including electronic medical records (EMRs) and payers' administrative data, has raised the expectation of using this information for purposes beyond their intended use of supporting patient care and provider compensation. Two of these uses are assessing the quality of care and developing new knowledge about the safety, effectiveness, and comprehensive costs and benefits of different treatments.

The value of this information is anticipated by Sean Tunis and colleagues in their description of federal initiatives, particularly Medicare's intention to integrate data on drug dispensing with information about diagnosis and treatment, and to make those data available for quality assessment and research (a "potential panacea").[1] Because the opportunity is so great, it will be important to be clear about the actions and policies needed to achieve the benefits.

■ **Have realistic expectations.** The most important requirement is to avoid inappropriately attributing causal relationships to simple associations, even very strong ones. The risk is pervasive but will be particularly great in comparing the effectiveness of different regimens. For example, patients assigned drugs with the same or overlapping indications might differ in critical ways that obscure the actual relationship between treatment regimen and outcome. That is, if clinicians believe that a particular regimen is superior, they might prescribe it preferentially to their sickest patients. This phenomenon, confounding by indication, can result in a superior regimen's appearing to have worse outcomes. The importance of this and many other types of confounding, selection,

Richard Platt (Richard_Platt@harvard.edu) is professor and chair, Department of Ambulatory Care and Prevention, Harvard Medical School and Harvard Pilgrim Health Care, in Boston, Massachusetts.

DOI 10.1377/hlthaff.26.2.w153

and bias depends on the specific question and the way in which the data are used. It is possible to adjust for some differences among treatment groups, but the problem cannot always be well addressed or even recognized.

■ **Establish appropriate privacy protections.** Even without identifiers such as name, Social Security number, or exact address, many people can be identified by combining information in medical claims files with additional information. Some of this additional information, such as birth date, might be publicly available; other information, such as the date on which a person visited a particular medical care provider, might be available to an array of people, including the person's acquaintances. Knowing the date and location of a single medical encounter provides a route to reidentification because only one person with a particular birth date is likely to receive a specific service—for example, a hearing test—at a particular institution on a specific date. Therefore, someone with access to the complete claims file can identify all of the recipients of hearing tests at the particular institution on the date in question and then cross-link to a demographic file to find the one with the birth date of the target person. Identification of the target's specific claim provides that person's unique medical record identifier; it is then straightforward to return to the claims data to discover all they contain about that person, including surgery, mental health care, and drugs dispensed, even if these services were provided years apart, in different facilities. For this reason, it will be necessary to maintain oversight over access to, and the uses of, these nominally deidentified data.

On the other hand, it will be important to avoid restrictions that make it difficult to use the information in beneficial ways. This could mean modifying the Health Insurance Portability and Accountability Act (HIPAA) to permit these uses of protected health information, or to define evaluation of safety and effectiveness to be consistent with HIPAA's current allowance of operational activities. Additionally, institutional review boards that oversee protection of research subjects will need to appreciate the unique aspects of an inquiry that has many features of routine quality assurance activities and that requires the inclusion of all recipients of medical care.

■ **Preserve the linkage of health data to external information.** The answers to many questions require more information about individuals than exists in the electronic records. For example, studies of survival or of birth outcomes require linkage to death or birth registries. Thus, it will be important for the databases to use unique identifiers that can be linked to people's actual identities when there is an approved need to do so.

■ **Affirm the appropriateness of accessing full medical records.** A related need is for access to a small number of full medical records to confirm coded information or to collect additional data. For example, in a study of drug-induced gastroduodenal bleeding, we were able to confirm 167 well-documented events among 1,041 hospitalizations with suggestive diagnoses by reviewing endoscopy and other reports in the full-text records.[2] Typically, a study of several million people's electronic health information will depend on the availability of a few hundred full medical records. For the databases to fulfill their potential, it is necessary to link these few individuals to their medical records and for clinicians and patients to accept the appropriateness of requests for access to them.

■ **Understand data anomalies.** Data collected for patient care or administration could yield misleading results when used for other purposes. In some cases, the data are simply not precise enough to answer the question of interest. Examples include inability to discriminate between incident and recurrent conditions and the use of a single Healthcare Common Procedure Coding System (HCPCS) code for many different diagnostic tests. Even when coding systems are nominally precise enough, there can be much variation in the use of specific codes. As an example, while developing an EMR-based system to alert public health officials about potential bioterrorism events or outbreaks of natural disease, we discovered that an apparent excess of shock and

death events in one metropolitan area was an artifact of a large group practice's custom of assigning the ICD-9 code 799.9, "other ill-defined and unknown causes of morbidity and mortality," during telephone triage calls when the responder could not make a more specific diagnosis.[3]

It is also important to recognize that the evolution of data systems might greatly skew results. These changes could be undocumented or unknown to users of the data. For example, an increase from one year to the next in the number of diagnoses captured per encounter caused the public health alerting system described above to appear to have found an early signal of a severe regional influenza outbreak in Denver in 2003; in fact, the signal occurred at about the same time as other detection methods had.[4] In this case, the clinicians' coding practices had not changed, but an intermediate data-processing routine that was not apparent to the public health users had been upgraded. These discontinuities are particularly likely when medical record systems are modified, when data warehouses are reconfigured, when compensation rules change, or when clinicians are asked to provide new kinds of information. Since all of these attributes will be in flux for the foreseeable future, one must be especially mindful of their potential impact on assessments of trends over time or comparisons of practices or outcomes involving multiple care settings.

■ **Make the necessary investments.** Making effective use of the new data resources will require considerable investment in infrastructure and in training and support of those who use them. These investments will be needed for new physical computing systems that can support analyses that are not now required for administrative purposes or clinical care. For example, it might be necessary to group individual medical encounters into episodes of illness, or to construct measures that span several years of care. The need for capacity to link to external data has already been mentioned. Perhaps most important, we will need well-trained investigators who can use the data well.

NEW INFORMATION about medical care and its outcomes can transform our understanding of which therapies work, and for whom, and whether effective therapies are used when they are indicated. However, simply having the data won't take us to this better place. For that we will need to agree on the rules of the road and ensure that we have drivers, vehicles, and fuel to take us to there.

This work was supported by a grant (AHRQ U18HS10391) to the HMO Research Network Center for Education and Research on Therapeutics (CERTs) from the Agency for Healthcare Research and Quality.

NOTES

1. S.R. Tunis et al., "Federal Initiatives to Support Rapid Learning about New Technologies," *Health Affairs* 26, no. 2 (2007): w140–w149 (published online 26 January 2007; 10.1377/hlthaff.26.2.w140).
2. J.G. Donahue et al., "Gastric and Duodenal Safety of Daily Alendronate," *Archives of Internal Medicine* 162, no. 8 (2002): 936–942.
3. W.K. Yih et al., "National Bioterrorism Syndromic Surveillance Demonstration Program," *Morbidity and Mortality Weekly Report* 53 Supp. (2004): 43–49.
4. D.P. Ritzwoller et al., "Comparison of Syndromic Surveillance and a Sentinel Provider System in Detecting an Influenza Outbreak—Denver, Colorado, 2003," *Morbidity and Mortality Weekly Report* 54 Supp. (2005): 151–156.

Advancing Evidence-Based Care For Diabetes: Lessons From The Veterans Health Administration

A highly regarded EHR system is but one contributor to the quality transformation of the VHA since the mid-1990s.

by Joel Kupersmith, Joseph Francis, Eve Kerr, Sarah Krein, Leonard Pogach, Robert M. Kolodner, and Jonathan B. Perlin

ABSTRACT: The Veterans Health Administration (VHA) is a unique laboratory for using the electronic health record (EHR) to transform health care and accelerate discovery. This is particularly evident in the care of veterans with diabetes, who constitute a quarter of those served by the VHA. Although EHRs have enabled rapid learning, additional factors were necessary, including the lead participation of clinician-investigators, accountability through performance measurement, a delivery system focused on population health, and favorable economic externalities. "Off-the-shelf" technology is unlikely to generate similar benefits if these attributes are not in place. [*Health Affairs* 26, no. 2 (2007): w156–w168 (published online 26 January 2007; 10.1377/hlthaff.26.2.w156)]

AS THE LARGEST INTEGRATED DELIVERY SYSTEM (IDS) in the United States, the Veterans Health Administration (VHA) serves 5.3 million patients annually across nearly 1,400 sites of care. Although its patients are older, sicker, and poorer than the general U.S. population, the VHA's performance now surpasses that of other health systems on standardized quality measures.[1] These advances are related in part to the VHA's leadership in the development and use of electronic health records (EHRs). In this paper we describe the VHA's health information infrastructure and factors that made it possible, illustrating its impact on research in and care of diabetes, one of most prevalent conditions among veterans. We also describe how next-generation EHRs will facilitate veteran-centered care and continued improvement. We emphasize the human and system characteristics essential to the transformation of VHA care.

Joel Kupersmith (Joel.Kupersmith@va.gov) is chief research and development officer at the Department of Veterans Affairs (VA) in Washington, D.C. Joseph Francis is deputy chief research and development officer. Eve Kerr is research coordinator, Diabetes Quality Enhancement Research Initiative (QUERI), at the VA. Sarah Krein is co-research coordinator, Diabetes QUERI. Leonard Pogach is clinical coordinator, Diabetes QUERI. Robert Kolodner is the VA chief medical information officer. Jonathan Perlin is chief medical officer and senior vice president, Quality, at the Hospital Corporation of America in Nashville, Tennessee; he is the former undersecretary for health at the VA.

DOI 10.1377/hlthaff.26.2.w156

Historical Context Of Health Care For U.S. Veterans

Adding computers to a delivery system unprepared to leverage the advantages of health information can create inefficiency and other negative outcomes.[2] In contrast, during the period of time in which the VHA deployed its EHR system, the number of veterans seen increased from fewer than three million to nearly five million, while costs per patient and numbers of full-time employees per patient both decreased.[3] To understand how this could be possible, it is important to highlight historical and organizational factors that were important to the adoption of the VHA's EHR system.

Health care in the VHA is the product of decades of innovation. In 1930 Congress consolidated programs for U.S. veterans in the Veterans Health Administration under the Department of Veterans Affairs (VA, then known as the Veterans Administration). Facing more than one million returning troops following World War II, the VHA partnered with U.S. medical schools, gaining access to faculty and trainees and adding research and education to its statutory missions. That bold move created an environment uniquely suited to rapid learning. The VHA now has affiliations with 107 medical schools and trains almost 90,000 physicians and associated health professionals annually.

The VHA was originally based on inpatient care, and administrative and legal factors created inefficiency and inappropriate use. By the 1980s the VHA's public image was poor. In 1995, facing scrutiny from Congress, the VHA reorganized into twenty-two integrated care networks. Incentives were created for providing care in the most appropriate setting, and legislation established universal access to primary care. Those changes resulted in a reduction of 40,000 inpatient beds and an increase of 650 community-based care sites. Evidence-based practice guidelines and quality measures were adopted, and safeguards were put in place for vulnerable groups such as the mentally ill and those needing chronic care, while the VHA's performance management system held senior managers accountable for evidence-based quality measures. All of these changes created a strong case for robust information systems and spurred dramatic improvements in quality.[4]

VistA: The VHA's Electronic Health Record System

Because the VHA was both a payer and a provider of care, its information system was developed to support patient care and its quality with clinical information, rather than merely to capture charges and facilitate billing. In the early 1980s the VHA created the Decentralized Hospital Computer Program (DHCP), one of the first EHR systems to support multiple sites and health care settings. DHCP developers worked incrementally with a network of VHA academic clinicians

across the country, writing and testing code locally and transmitting successful products electronically to other sites, where they could be further refined. Over time, the group had created a hospital information system prototype employing common tools for key clinical activities. The system was launched nationally in 1982, and by 1985 the DHCP was operational throughout the VHA system.

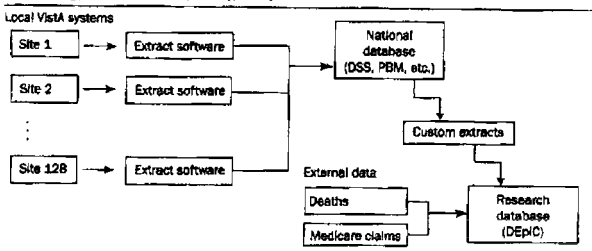
The DHCP evolved to become the system now known as the Veterans Health Information Systems and Technology Architecture (VistA), a suite of more than 100 applications supporting clinical, financial, and administrative functions. Access to VistA was made possible through a graphical user interface known as the Computerized Patient Record System (CPRS). With VistA/CPRS, providers can securely get access to patient information at the point of care and, through a single interface, update a patient's medical history, place orders, and review test results and drug prescriptions. Because VistA also stores medical images such as x-rays and photographs directly in the patient record, clinicians have access to all of the information needed for diagnosis and treatment. As of December 2005, VistA systems contained 779 million clinical documents, more than 1.5 billion orders, and 425 million images. More than 577,000 new clinical documents, 900,000 orders, and 600,000 images are added each workday—a wealth of information for the clinician, researcher, and health care administrator.

This brief description is not meant to minimize the inertia encountered with any major organizational change. Many clinicians (particularly those with poor keyboarding skills) initially resisted use of the EHR system. Convincing them otherwise took several approaches. Most important was involving clinicians at the onset. This meant working incrementally to ensure usability and integration of the EHR system with clinical processes. Both local and national supports were created: For example, local "super-users" were designated to champion the project; and a national "Veterans Electronic Health University" facilitated collaboration among local, regional, and national sponsors of EHR rollout. National performance measures, as well as the gradual withdrawal of paper records, made EHR use an inescapable reality. Finally, because economic costs to clinicians were blunted by a salaried environment, other beneficial effects (such as reductions in time wasted searching for missing paper records) emerged. Over time, staff came to view VistA/CPRS as indispensable for good clinical care.[5]

Leveraging The EHR: Diabetes Care In The VHA

■ **From individual records to population insights.** VistA/CPRS allows clinicians to access and generate clinical information about their individual patients, but additional steps are needed to yield insights into population health. Structured clinical data in the EHR can be aggregated within specialized databases, providing a rich source of data for VHA administrators and health services researchers (Exhibit 1). Additionally, unstructured text data, such as clinicians' notes, can be reviewed and abstracted electronically from a central location. This is of particular benefit to

EXHIBIT 1
Current Data Flow And Aggregation In The Veterans Health Information Systems And Technology Architecture (VistA), Simplified



SOURCE: Veterans Affairs (VA) Information Resource Center.
NOTES: This exhibit shows the resources and flow of the data most often used by Veterans Health Administration (VHA) researchers for national studies. Most data originate from the VistA system. To build a national database of health care information, each application must extract data from each of 128 separate local VistA sites. At the host site, specialized software cleans, translates, and loads the extract data into the national database. Researchers use custom extract software routines to access national databases and may combine VHA data with data from external sources such as Medicare claims data and the National Death Index. DSS is decision-support system (VHA clinical utilization data). PBM is pharmacy benefit manager. DEpiC is Diabetes Epidemiological Cohort.

researchers: VHA multisite clinical trials and observational studies are facilitated by immediate 100 percent chart availability. Furthermore, the VHA has invested in an External Peer Review Program (EPRP), in which an independent external contractor audits the electronic text records to assess clinical performance using evidence-based performance criteria. Finally, data derived from the EHR can be supplemented by information from other sources, such as Medicare utilization data or data from surveys of veterans.

Diabetes care in the VHA illustrates the advantages of a national EHR system supported by an intramural research program. Much of the work that follows has been supported by the VA Office of Research and Development through its Health Services Research and Development and Quality Enhancement Research Initiative (QUERI) programs.[6]

■ **Understanding disease burden: the VHA diabetes registry.** The VHA was an early leader in using EHRs for a national diabetes registry containing clinical elements as well as administrative data. While the VHA's EHR system made a diabetes registry possible, operationalizing data transfer and transforming those data into useful information did not come automatically or easily. In the early 1990s the VHA began extracting clinical data from each local VHA database into a central data repository. By 2000 the VHA diabetes registry contained data on nearly 600,000 patients receiving care in the VHA system, including drugs, test results, blood pressures, and vaccinations. This information has subsequently been merged with Medicare claims data to create the VHA's Diabetes Epidemiology Cohort (DEpiC).[7]

Seventy-three percent of diabetic veterans are eligible for Medicare, and 59 percent of dual eligibles (Medicare and Medicaid) use both systems. Adding Medicare administrative data results in less than 1 percent loss to follow-up; although those data are not as rich as the clinical information in the VHA's EHR system, their addition fills gaps in follow-up, complication rates, and resource use.[8] Combined VHA and Medicare data also reveal a prevalence of diabetes among veterans exceeding 25 percent. The impact of the diabetic population on health spending is considerable, including total inpatient spending (VHA plus Medicare) of $3.05 billion ($5,400 per capita) in fiscal year 1999.[9]

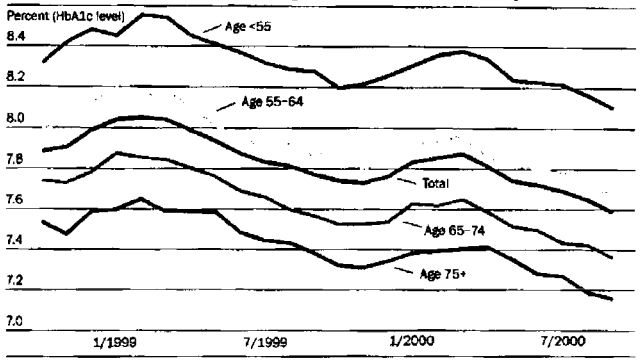
The rich clinical information made possible through the EHR system yields other insights. For instance, the VHA has identified a high rate of comorbid mental illness (24.5 percent) among patients with diabetes and is using that information to understand the extent to which newer psychotropic drugs, which promote weight gain, as well as mental illness itself contribute to poor outcomes.[10] The influence of sex and race/ethnicity can also be more fully explored using EHR data.[11]

Delineating and tracking diabetic complications are also facilitated by the EHR system. For example, using EHR clinical data allows identification of early chronic kidney disease in one-third of veterans with diabetes, fewer than half of whom have renal impairment indicated in the record.[12] The VHA is able to use the EHR system to identify patients at high risk for amputation and is distributing that information to clinicians to better coordinate their care.[13]

■ **EHR-enabled approaches to monitoring quality and outcomes.** Traditional quality report cards may provide incentives to health providers to disenroll the sickest patients.[14] The VHA's EHR system provides a unique opportunity to construct less "gameable" quality measures that assess how well care is managed for the same person over time for diseases such as diabetes, for which metrics of process quality, intermediate outcomes, and complications (vision loss, amputation, renal disease) are well defined. Using the VHA diabetes registry, longitudinal changes within individual patients can be tracked. In Exhibit 2, case-mix-adjusted glycosylated hemoglobin (HbAlc) values among veterans with diabetes decreased by −0.314 percent (range −1.90 to 1.03, $p < 0.0001$) over two years, indicating improved glycemic control over time, rather than simply the enrollment of healthier veterans.[15] These findings provide a convincing demonstration of effective diabetes care.

Longitudinal data have other important uses. For instance, knowledge of prior diagnoses and procedures can distinguish new complications from preexisting conditions. This was shown to be the case for estimates of amputation rates among veterans with diabetes, which were 27 percent lower once prior diagnoses and procedures were considered. Thus, longitudinal data better reflect the effectiveness of the management of care and can help health systems avoid being unfairly penalized for adverse selection.[16] Longitudinal EHR data are also important for evaluating the safety and effectiveness of treatments, which are critical insights for national formulary decisions.

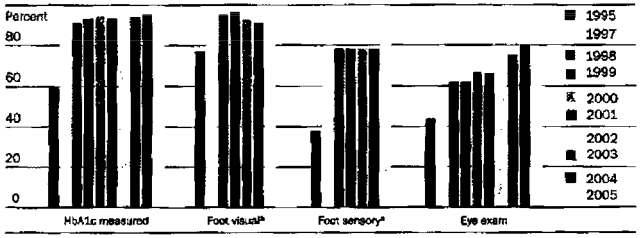
EXHIBIT 2
Trends in Mean Glycosylated Hemoglobin (HbA1c) Levels Among Veterans Health Administration (VHA) Clinic Users, By Age Category, October 1998–September 2000



SOURCE: Diabetes Quality Enhancement Research Initiative (QUERI).
NOTE: Additionally, a regression model that adjusts for clustering (patient and facility) and seasonal effects was used to confirm the downward linear trend in monthly HbA1c levels overall (−0.013, $p < 0.0001$) and minimal differences in this trend by each age category ($p = 0.492$)

■ **Advancing evidence-based care.** Exhibit 3 shows the trends in the VHA's national performance scorecard for diabetes care based on EHR data. In addition to internal benchmarking, this approach has compared VHA performance with that of commercial managed care.[17] These performance data are obtained by abstracting the electronic chart; the completion of a national Health Data Repository with aggre-

EXHIBIT 3
Diabetes Process Quality in The Veterans Health Administration (VHA), Selected Years 1995–2005



SOURCE: Based on results from the VHA External Peer Review Program.
NOTE: Results are for VHA primary care outpatients with diabetes mellitus.
[a] Data for 2004 and 2005 are not provided.

gated relational data will eventually support automatic queries about quality and outcomes ranging from the individual patient to the entire VHA population.

The richness of EHR data allows the VHA to refine its performance measures. VHA investigators were able to demonstrate that annual retinal screening was inefficient for low-risk patients and inadequate for those with established retinopathy.[18] The VHA therefore modified its performance metrics and is developing an approach to risk-stratified screening that will be implemented nationally.

The greatest advantage of EHRs in the VHA system is their ability to improve performance by influencing the behavior of patients, clinicians, and the system itself. For instance, the VHA's diabetes registry has been used to construct performance profiles for administrators, clinical managers, and clinicians. These profiles included comparisons of facilities and identified the proportion of veterans with substantial elevations of HbA1c, cholesterol, and blood pressure. Patient lists also facilitated follow-up with high-risk patients. The EHR system also allows consideration of clinicians' actions to intensify therapy in response to that risk level (such as having a cholesterol medication started or increased when low-density lipid, or LDL, cholesterol is elevated). This approach credits clinicians with providing optimal treatment and informs them about what might be required to improve care.[19]

Data from the EHR system and diabetes registry also demonstrate the critical importance of defining the level of accountability in reporting the quality of diabetes care. EHR data show that for most measures in the VHA system, only a small fraction (2 percent or less) of the variance is attributable to individual primary care providers (PCPs) and that PCP profiling will be inaccurate unless panel sizes are very large (200 diabetics or more). In contrast, much more variation (12–18 percent) is attributable to overall performance at the site of care, a factor of relevance for the design of approaches to rewarding quality. Use of EHR data also highlights the important influence of organizational and system factors on providers' adherence to guidelines.[20]

The EHR system can identify high-risk populations and can facilitate targeted interventions. For instance, poor blood pressure control contributes greatly to cardiovascular complications, the most common cause of death in diabetics. In the VHA, investigators are working with pharmacy leaders to identify gaps in medication refills or lack of medication titration and thereby identify patients with inadequate blood pressure control because of poor medication adherence or inadequate medication intensification. Once identified, those patients can be assigned proactive management by clinical pharmacists integrated into primary care teams and trained in behavioral counseling.[21] Other approaches being tested and evaluated using EHR data are group visits, peer counseling, and patient-directed electronic reminders.

VistA/CPRS provides additional tools to improve care at the point of service. For example, PCPs get reminders about essential services (such as eye exams or

influenza vaccinations) at the time they see the patient, and CPRS functions allow providers and patients to view trends in laboratory values and blood pressure control. Perhaps most importantly, the VHA's EHR system allows for effective care coordination across providers, to communicate patients' needs, goals, and clinical status as well as to avoid duplication of services.

■ Care coordination and telehealth for diabetes. In-home monitoring devices can collect vital data for high-risk patients from the home and transmit those data to a care coordinator who can make early interventions that might prevent the need for institutional intervention.[22] Such an approach is possible only with an EHR. Based on promising pilot data as well as needs projections, the VHA has implemented a national program, Care Coordination through Home Telehealth (CCHT).[23]

Information technology (IT) also supports cost-effective access to specialized services. The VHA recently piloted the use of digital retinal imaging to screen for diabetic retinopathy and demonstrated that it could be a cost-effective alternative to ophthalmoscopy for detecting proliferative retinopathy.[24] Diabetic retinopathy is not only a preventable complication but also a biomarker for other end-organ damage (for example, kidney damage). In October 2005 the VHA began implementing a national program of teleretinal imaging, to be available on VistA/CPRS and for use by clinicians and researchers. In the future, computerized pictorial analysis and new tools for mining text data across millions of patient records have the potential to transform the clinical and research enterprise by identifying biomarkers of chronic illness progression.

Limits Of The EHR System In The VHA

Although the VHA has one of the most sophisticated EHR systems in use today, VistA is not a single system, but rather a set of 128 interlinked systems, each with its own database—that is, a decentralized system with central control. This limits its ability to make queries against all of a patient's known data. In addition, lack of standardization for laboratory values such as glycosylated hemoglobin and other data elements creates challenges for aggregating available data for administrative and research needs. The VHA diabetes registry, although a product of the EHR system, took years of effort to ensure data integrity.

A national data standardization project is under way to ensure that data elements are compliant with emerging health data standards and data management practices. Extracting data from free-text data fields—a challenge for all electronic records—will be addressed by defining moderately structured data elements for public health surveillance, population health, clinical guidelines compliance, and performance monitoring. Mapping of legitimate local variations to standard representations will allow easier creation of longitudinal registries for a variety of conditions.

Studies with the VHA's EHR system as well as others' systems have shown that electronic reminders, although effective at changing providers' behavior, have limited benefit, in part as a result of "signal overload" and other human-factor con-

"Through the MHV Web portal, veterans can securely view and manage their personal health records online."

straints.[25] The VA Office of Research and Development is funding work testing advanced decision-support systems that rely on artificial intelligence–based systems working in the background that integrate a broader range of patient clinical information with up-to-date care guidelines to provide tailored recommendations (for example, optimal choice of diabetes therapy). The hope, as yet unproven, is that such systems, by reducing provider burden, can improve process quality and patient outcomes.

The care of diabetes is complex and demanding, and delivering all indicated services might require more time than is typically available in a follow-up visit.[26] Studies of the impact of the EHR system on workflow and efficiency in the VHA and other settings have shown conflicting results.[27] Although it is unlikely that having EHRs saves time during the office encounter, downstream benefits such as better care coordination, reduction of duplicative and administrative tasks, and new models of care (such as group visits) translate into a "business case" when the reimbursement structure favors population management.

Creating Patient-Centered, Community-Based Care: My HealthEVet

The VHA's quality transformation since 1996 involved shifting from inpatient to integrated care. The next phase will involve empowering patients to be more actively engaged and moving care from the clinic to the community and home. Again, health IT has been designed to support the new delivery system.

My HealthEVet (MHV) is a nationwide initiative intended to improve the overall health of veterans and support greater communication between VHA patients and their providers. Through the MHV Web portal, veterans can securely view and manage their personal health records online and can get access to health information and electronic services. Veterans can request copies of key portions of their VHA health records and store them in a personal "eVAult," along with self-entered health information and assessments, and can share this information with their health care providers and others inside and outside the VHA. The full functionality of MHV, available online at http://www.myhealth.va.gov, will help patients plan and coordinate their own care through online access to care plans, appointments, laboratory values, and reminders for preventive care. Research itself can be facilitated by MHV: Patients will be able to identify ongoing clinical studies for which they are eligible to enroll, communicate with investigators via encrypted e-mail, have their outcomes tracked through computer-administered "smart surveys," and even provide suggestions for future studies. In addition, the effectiveness of patient-centered care can be evaluated.

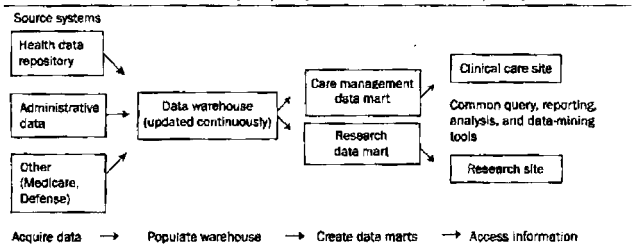
The Twenty-First-Century EHR

The next phase of VistA/CPRS will feature open-source applications and relational database structures. One benefit of the conversion will be easier access to national stores of clinical data through a unified Health Data Repository (HDR) that will replace the current 128 separately located VistA systems. The HDR is under construction; it now contains records from nearly sixteen million patients, with more than 900 million vital-sign recordings and 461 million prescriptions.

Additionally, a clinical observations database linked to Systematized Nomenclature of Medicine (SNOMED) terms and semantic relationships will greatly expand the scope of data available for research data-mining activities. Improved decision-support capabilities will help clinicians provide care according to guidelines and understand situations where it is appropriate to deviate from guidelines. The reengineered EHR will also link orders and interventions to problems, greatly increasing the VHA's clinical data-mining capabilities.

To support the delivery of consistent information to all business units, the VHA is developing a Corporate Data Warehouse (CDW, Exhibit 4), which will include the HDR as the primary source of clinical data but also encompass other administrative and financial data sets (including Medicare data) to create a unified view of veterans' care. Among other things, the CDW will supplement the capabilities of VistA by providing an integrated analytical system to monitor, analyze, and disseminate performance measures. This will assist population-based health services research by offering standardized data across all of the subjects it contains, tools for rapidly performing hypothesis testing, and ease of data acquisition. Unlike the VHA's current diabetes registry, which has been labor-intensive to create and maintain, future registries based on the CDW will be easier to construct and up-

EXHIBIT 4
Veterans Health Administration (VHA) Corporate Data Warehouse (CDW) Architecture



SOURCE: VHA Office of Information.
NOTES: The VHA's future Corporate Data Warehouse will combine health data, administrative data, and externally derived data for all patients seen in the VHA. This structure will facilitate the creation of automated data marts to facilitate care management (using patient identifiers) as well as research with appropriately deidentified data.

date. The CDW will eventually facilitate personalized medicine by allowing the linkage of genomic information collected from veterans to information on longitudinal outcomes. These changes will introduce more central control than was present during the early days of the VHA's EHR system, but clinicians and researchers will continue their involvement in developing innovations.

Conclusions And Policy Implications

The VHA has been an EHR innovator, developing a clinically rich system from the ground up that has become so integrated into the delivery of care and the conduct of research that one cannot imagine a veterans' health system without it. However, many factors in addition to the EHR system contributed to the VHA's quality transformation, including a culture of academician-clinicians that valued quality, scientific evidence and accountability (for which the EHR became an organizer and facilitator); the presence of embedded researchers who were active clinicians, managers, policymakers, and developers of VistA/CPRS; and a research infrastructure that could be applied to this topic.[28] Although the data structures are complex and sometimes flawed, they are, because of their user origins, effectively linked to the needs of clinicians and researchers, who in turn incorporate their input into the further evolution of the VHA's EHR system.

The design of the VHA system also ensures that overall incentives are aligned to realize EHRs' beneficial externalities. The VHA benefits, for instance, by being able to eliminate duplicative test ordering when veterans seek care at different facilities.[29] The cost of maintaining the EHR system amounts to approximately $80 per patient per year—roughly the amount saved by eliminating one redundant lab test per patient per year.[30] The VHA also benefits greatly by being an interactive, permeable entity in a free-market system: The VHA is an enrollment system, not an entitlement program or a safety-net provider, and thus has incentives for maintaining high satisfaction and perceived value among those it serves.

For patient care management, the VHA's EHR system has developed an infrastructure for collecting and organizing information from which a diabetes database (DEpiC) evolved to provide valuable information related to disease prevalence, comorbidities, and costs that is necessary for quality improvement, systemwide planning, and research. Longitudinal within-cohort assessment, made possible by EHRs, is a major advance in attaining precise measures of quality that mitigate the effects of adverse patient selection.

Home telehealth linked to EHRs has made possible novel patient-provider interactions of which the care coordination and teleretinal imaging initiatives are among the earliest prototypes. This approach has the capacity to expand care delivery to many others, and the benefits are not limited to the homebound: A new generation of Internet-savvy veterans will appreciate round-the-clock access to health care the same way they do for instant messaging and shopping. MHV, which is in its launch phase, is part of the future plan to give veterans control over

their health and includes many possibilities for research.

One more important EHR-enabled initiative has the capacity to greatly change the practice of medicine: adding genomics information to the medical record. With its EHR database, the VHA has an opportunity to identify the genetic correlates of disease and drug response, which may transform medical practice from a process of statistical hunches to one of targeted, personalized care.

Because of the vastly larger scale of the health care enterprise and the changing needs of veterans, the VHA's focus now has models in place to shift to issues involving clinical decision support, content standardization, and improved interaction among patients, VHA providers, and other systems. These capabilities are made possible by the VHA's EHR system. The VHA experience could provide a model for how federal health policies can help the United States bridge its "quality chasm." As we have described, this transformation involves far more than simply installing VistA/CPRS (or any other EHR system). The primary lesson the VHA can offer other health systems is to emphasize the necessary clinical and organizational factors needed for successful EHR implementation and to link those factors to a research and quality infrastructure capable of using electronic health information for discovery and improvement.

The authors thank the following people in the Department of Veterans Affairs (VA) and Veterans Health Administration (VHA) for their invaluable assistance: Jack Bates, VHA Office of Information; Adam Darkins, Office of Care Coordination; Linda Fischetti, HealthePeople, VHA Office of Information; Denise Hynes, VA Information Resource Center; Donald R. Miller, Diabetes Epidemiology Cohorts (DEpiC) Center for Health Quality, Outcomes, and Economic Research; and Syed Tirmizi, VHA Office of Information. The views presented in this paper are those of the authors and do not necessarily represent the views of the Department of Veterans Affairs.

NOTES

1. S.M. Asch et al., "Comparison of Quality of Care for Patients in the Veterans Health Administration and Patients in a National Sample," *Annals of Internal Medicine* 141, no. 12 (2004): 938–945; E.A. Kerr et al., "Diabetes Care Quality in the Veterans Affairs Health Care System and Commercial Managed Care: The TRIAD Study," *Annals of Internal Medicine* 141, no. 4 (2004): 272–281; and A.K. Jha et al., "Effect of the Transformation of the Veterans Affairs Health Care System on the Quality of Care," *New England Journal of Medicine* 348, no. 22 (2003): 2218–2227.
2. D.U. Himmelstein and S. Woolhandler, "Hope and Hype: Predicting the Impact of Electronic Medical Records," *Health Affairs* 24, no. 5 (2005): 1121–1123.
3. D.C. Evans et al., "Effect of the Implementation of an Enterprise-Wide Electronic Health Record on Productivity in the Veterans Health Administration," *Health Economics, Policy, and Law* 1, no. 2 (2006): 163–169; and J.B. Perlin, R.M. Kolodner, and R.H. Roswell, "The Veterans Health Administration: Quality, Value, Accountability, and Information as Transforming Strategies for Patient-Centered Care," *American Journal of Managed Care* 10, no. 11, Part 2 (2004): 828–836.
4. Jha et al., "Effect of the Transformation"; and Perlin et al., "The Veterans Health Administration."
5. S.H. Brown et al., "VistA—U.S. Department of Veterans Affairs National-Scale HIS," *International Journal of Medical Informatics* 69, no. 2–3 (2003): 135–156.
6. S.L. Krein et al., "Department of Veterans Affairs' Quality Enhancement Research Initiative for Diabetes Mellitus," *Medical Care* 38, no. 6, Supp. 1 (2000): 138–148.
7. D.R. Miller, M.M. Safford, and L.M. Pogach, "Who Has Diabetes? Best Estimates of Diabetes Prevalence in the Department of Veterans Affairs Based on Computerized Patient Data," *Diabetes Care* 27, no. 2 Supp. (2004): B10–B21.

8. Donald Miller, VA Center for Health Quality Outcomes and Economic Research, personal communication, 10 March 2006.
9. L. Pogach and D.R. Miller, "Merged VHA-Medicare Databases: A Tool to Evaluate Outcomes and Expenditures of Chronic Diseases" (Poster Session, VHA Health Services Research and Development Conference, 15 February 2006).
10. S.M. Frayne et al., "Disparities in Diabetes Care: Impact of Mental Illness," *Archives of Internal Medicine* 165, no. 22 (2005): 2631–2638; and F. Cunningham et al., "Antipsychotic-Induced Diabetes in Veteran Schizophrenic Patients," *Pharmacoepidemiology and Drug Safety* 12, Supp. (2003): S154.
11. M. Safford et al. "Disparities in Use of Lipid-Lowering Medications among People with Type 2 Diabetes Mellitus," *Archives of Internal Medicine* 163, no. 8 (2003): 922–928.
12. E.F. Kern et al., "Failure of ICD-9-CM Codes to Identify Patients with Comorbid Chronic Kidney Disease in Diabetes," *Health Services Research* 41, no. 2 (2006): 564–580.
13. Jeffrey Robbins, VHA national program director for podiatry, personal communication, 17 February 2006.
14. T.P. Hofer et al., "The Unreliability of Individual Physician 'Report Cards' for Assessing the Costs and Quality of Care of a Chronic Disease," *Journal of the American Medical Association* 281, no. 22 (1999): 2098–2105.
15. W. Thompson et al., "Assessing Quality of Diabetes Care by Measuring Longitudinal Changes in Hemoglobin A1c in the Veterans Health Administration," *Health Services Research* 40, no. 6, Part 1 (2005): 1818–1835.
16. C.L. Tseng et al., "Use of Administrative Data to Risk Adjust Amputation Rates in a National Cohort of Medicare-Enrolled Veterans with Diabetes," *Medical Care* 43, no. 1 (2005): 88–92.
17. Kerr, "Diabetes Care Quality"; and C.T. Sawin et al., "Diabetes Process and Outcome Measures in the Department of Veterans Affairs," *Diabetes Care* 27, no. 2 Supp. (2004): B90–B94.
18. R.A. Hayward et al., "Causes of Preventable Visual Loss in Type 2 Diabetes Mellitus: An Evaluation of Suboptimally Timed Retinal Photocoagulation," *Journal of General Internal Medicine* 20, no. 5 (2005): 467–469.
19. E.A. Kerr et al., "Building a Better Quality Measure: Are Some Patients with 'Poor Quality' Actually Getting Good Care?" *Medical Care* 41, no. 10 (2003): 1173–1182.
20. S.L. Krein et al., "Whom Should We Profile? Examining Diabetes Care Practice Variation among Primary Care Providers, Provider Groups, and Health Care Facilities," *Health Services Research* 37, no. 5 (2002): 1159–1180.
21. H.M. Choe et al., "Proactive Case Management of High-Risk Patients with Type 2 Diabetes Mellitus by a Clinical Pharmacist: A Randomized Controlled Trial," *American Journal of Managed Care* 11, no. 4 (2005): 253–260.
22. M. Hudileston and R. Kobb, "Emerging Technology for At-Risk Chronically Ill Veterans," *Journal of Healthcare Quality* 26, no. 6 (2004): 12–15, 24.
23. N.R. Chumbler et al., "Evaluation of a Care Coordination/Home-Telehealth Program for Veterans with Diabetes: Health Services Utilization and Health-Related Quality of Life," *Evaluation and the Health Professions* 28, no. 4 (2005): 464–478.
24. P. Conlin et al., "Framework for a National Tele-Retinal Imaging Program to Screen for Diabetic Retinopathy in the Veterans Health Administration," *Journal of Rehabilitation Research and Development* 43, no. 6 (2006): 741–748.
25. A.X. Garg et al., "Effects of Computerized Clinical Decision Support Systems on Practitioner Performance and Patient Outcomes: A Systematic Review," *Journal of the American Medical Association* 293, no. 10 (2005): 1223–1238.
26. M.L. Parchman, R.L. Romero, and J.A. Pugh, "Encounters by Patients with Type 2 Diabetes—Complex and Demanding: An Observational Study," *Annals of Family Medicine* 4, no. 1 (2006): 40–45.
27. J.M. Overhage et al., "Controlled Trial of Direct Physician Order Entry: Effects on Physicians' Time Utilization in Ambulatory Primary Care Internal Medicine Practices," *Journal of the American Medical Informatics Association* 8, no. 4 (2001): 361–371.
28. S. Greenfield and S.H. Kaplan, "Creating a Culture of Quality: The Remarkable Transformation of the Department of Veterans Affairs Health Care System," *Annals of Internal Medicine* 141, no. 4 (2004): 316–318; and J.B. Perlin, "Transformation of the U.S. Veterans Health Administration," *Health Economics, Policy and Law* 1, no. 2 (2006): 99–105.
29. J.D. Kleinke, "Dot-Gov: Market Failure and the Creation of a National Health Information Technology System," *Health Affairs* 24, no. 5 (2005): 1246–1262.
30. Perlin, "Transformation."

Reshaping Cancer Learning Through The Use Of Health Information Technology

Integrated health IT systems like Kaiser Permanente's complement networked learning with the capture of clinically detailed data.

by Paul J. Wallace

ABSTRACT: In this paper I describe Kaiser Permanente's experience with health information technology (IT) in cancer care. Health IT holds the potential to accelerate learning in cancer care by comprehensively capturing rich patient data and supporting optimally standardized care. Rapid learning in cancer depends on simultaneously working toward universal technical and data standards and taking intermittent steps to reconcile variations in legacy systems through data-mapping and information-sharing initiatives. [*Health Affairs* 26, no. 2 (2007): w169–w177 (published online 26 January 2007; 10.1377/hlthaff.26.2.w169)]

HEALTH INFORMATION TECHNOLOGY (IT) is widely viewed as having the potential to help transform health care in the United States.[1] Efficiency and safety savings from interoperable electronic health record (EHR) systems have been estimated at $142–$371 billion and occur through reduced adverse drug events and more systematic use of disease management and preventive care recommendations.[2] These savings can certainly be found within cancer care, the direct costs of which were estimated at $74 billion in 2005.[3]

However, as significant as those savings might be, health IT also offers a less commonly noted benefit: It can expand the breadth and accelerate the pace of learning within cancer care. There is no clinical area likely to be better served by rapid learning. Options for cancer detection, diagnosis, and treatment evolve ever more quickly, and the emerging areas of genomics, proteomics, and nanotechnology will continue to evolve into the future.[4] Health IT can transform—in terms of pace, scale, and scope—the process of answering important cancer-related inquiries. Its potential to do so lies in its inherent capacity to comprehensively capture rich patient data and to directly support care standardization. Kaiser Permanente's experience in cancer care and research provides an early perspective on the opportunities that health IT affords for rapid learning.

Paul Wallace (Paul.Wallace@kp.org) is medical director, Health and Productivity Management Programs, and senior adviser, Care Management Institute and KP-Healthy Solutions, at Kaiser Permanente in Oakland, California.

DOI 10.1377/hlthaff.26.2.w169

The Kaiser Permanente Experience With Health IT

■ **Cancer care.** Kaiser Permanente, the largest U.S. not-for-profit integrated health care delivery system, has more than 8.5 million members in eight geographic regions. Medicare covers roughly 900,000 of them. Physicians are employed by professional partnerships or corporations in each region that contract with the not-for-profit Kaiser Foundation Health Plan to provide and arrange necessary medical care for members. Kaiser Permanente addresses all health care needs for adult and pediatric members, including preventive, routine, specialty, emergency, and inpatient care; ancillary testing; pharmacy and rehabilitative services; and home care.

Approximately 40,000 Kaiser Permanente members receive a new cancer diagnosis each year; roughly 250,000 have had a cancer diagnosis at some point. Kaiser Permanente delivers cancer care for members across the continuum from screening and early detection through diagnosis, treatment, and follow-up to survivorship and end-of-life care.

■ **Cancer research.** All regions maintain research centers addressing cancer and other clinical areas, collectively employing more than a hundred full-time scientists and a thousand staff members. They participate in a growing number of multisite, cancer-focused research networks, including the National Cancer Institute (NCI)–sponsored Cancer Research Network (CRN), the Center for Education and Research in Therapeutics (CERT), the HMO Research Network (HMORN), and the Cancer Care Outcomes Research and Surveillance (CanCORS) group.

Research activities also permeate the clinical arena; many Permanente physicians and delivery system personnel administer clinical trials, participate in epidemiologic and health services research investigations, and serve on institutional review boards (IRBs) and research committees. Several hundred Kaiser Permanente members are enrolled each year in NCI Cooperative Group trials.

■ **Health IT at Kaiser Permanente: KP HealthConnect.** In 2002 Kaiser Permanente contracted with Epic Systems Corporation to create and implement a programwide integrated EHR, KP HealthConnect. Features include inpatient and outpatient clinical decision support; Web-based access for patients and providers; pharmacy and clinical laboratory support and reporting; emergency department (ED) management; scheduling and billing; and interfaces to other systems, such as archiving and population care management. A Web site, http://www.kp.org, allows members to access portions of their medical record, send secure messages to clinicians, schedule appointments, and refill medications, as well as to find health information, advice, and tools such as health risk appraisals and interventions for smoking cessation, stress management, and dietary and exercise counseling.

An Epic-based system similar to KP HealthConnect has been implemented in Kaiser Permanente's Northwest region for more than a decade. KP HealthConnect itself is at least partially implemented in all KP regions; within two years it will support care in all clinical settings.

The Potential For Rapid Learning In Cancer Care

Accelerating the pace of cancer-related inquiry requires increasing the number of questions being concurrently addressed and minimizing the time needed to generate answers. More numerous and rapid knowledge turnarounds require high-quality data that are readily available at low expense.[5] Ideally, those data (1) consist of accurate and complete digital capture of details about care as it is delivered and (2) are comparable across individuals and populations because excess variability in care processes is minimized. Thus, rapid learning in cancer care relies on two health IT–enabled prerequisites: vastly improved data capture and support for care standardization.

■ **Data capture.** KP HealthConnect reframes the possibilities for rigorously observing the care of individuals and populations over time. Information about clinical and operational events is captured as analyzable data, and all aspects of care become accessible to researchers and for quality improvement. Detailed individual-level data make it possible to tap into and combine the experiences of each patient and to maximize learning.

Data are generous for all patients, encounters, and providers, longitudinally and across the entire population of members with cancer. Over time, the long-term health consequences of cancer survivorship—on the prevalence and outcome of chronic disease, for instance—can be elucidated. More immediately, the comprehensive capture of information on all patients with cancer, combined with robust search mechanisms, allows for population-level identification of trends in occurrence, outcomes, and comorbidities, among other things.

Leveraging health IT to accelerate learning requires shifting emphasis from historical research practices focused on finding and managing scarce, expensive data to effectively and efficiently managing abundant data available at greatly reduced unit cost. New concerns about technical and semantic interoperability and governance of data use accompany this shift.

Interoperability. Interoperability is a foundational component of health IT.[6] Making maximal use of rich data stores assumes that multiple IT systems are fully aligned and interactive within and across organizations. This assumption is somewhat tenuous, given that universal technical and data standards have yet to emerge in this country and might not do so for some time.[7]

Even as industry standards evolve, Kaiser Permanente has invested heavily in achieving internal technical interoperability by purchasing integrated software for core systems, such as EHRs, scheduling, and database management, from a single vendor. A commensurate investment in semantic interoperability incorporates an internally developed medical terminology solution, Convergent Medical Technology (CMT), into content management. CMT is semantically congruent with the Systematized Nomenclature of Medicine, Clinical Terminology (SNOMED CT) (and with laboratory Logical Observation Identifiers Names and Codes, or LOINC, and First DataBank drug terminology), which has been adopted by soft-

ware suppliers, government entities, and health care organizations in more than thirty countries.[8]

Data use and governance. Existing processes for preserving confidentiality and protecting human subjects are products of an era when the use of research data within health care organizations was relatively rare and more easily controlled. Health IT dramatically increases both the abundance of and access to data through advanced databases and user-friendly query tools, precipitating the need for more robust and shared accountability between research and clinical care for policies regarding how data can and should be used, and by whom.

Kaiser Permanente investigators are identifying requirements for a national research database (NRDB) that will capitalize on the growing access to comprehensive health IT data and address data governance issues. The planned user community includes Kaiser Permanente researchers from all regions and their collaborators from other organizations and government agencies. Specific elements are still under development, but proposed requirements for the NRDB include the following: (1) Permission to access regional NRDB patient data for research purposes can only be granted by the associated IRB; (2) each region will retain ultimate control of and responsibility for access to regional members' data; and (3) separate access channels for researchers and providers/administrators will simplify compliance with Health Insurance Portability and Accountability Act (HIPAA) privacy and security rules.

■ **Care standardization.** Historically, the comparability of care necessary for making valid inferences has largely been achieved through clinical trials. Health IT can support comparability of care through increased enrollment in and operational support for formal trials. It can also foster comparability by making care more consistent with standardized treatment guidelines and protocols for all patients.

Broad comparability of care allows new kinds of inquiry, such as how well prior trial lessons have been generalized into care and the impact of interventions on entire populations, including those who are ineligible for participation in formal trials. Health IT can enable both the efficacy-focused study of the comparative effectiveness of clinical trials and the measurement of interventions' overall effectiveness in various populations.

■ **Increased clinical trial awareness, enrollment, and operational support.** A recent CRN study of barriers to physician participation in clinical trials identified a need for improved intra-organization communication about trial availability.[9] Regular computer use is associated with higher rates of patient recruitment into clinical trials, and KP HealthConnect is used to promote awareness of available trials.[10] Regional Kaiser Permanente clinical trial programs, which participate in national Cooperative Group and commercially funded protocols, use local Web pages to provide information about available trials and eligibility requirements, resources for patient evaluation and recruitment for studies, and reports on medical center and individual investigator accrual of individual patient enrollments in clinical trials.

"To the extent that health IT can improve efficiency, more patients can be included in trials within existing financial support."

Decision supports for specific trials, including study-associated care plans and treatment and follow-up charting templates, are also in use.

Research trials involve sizable overhead costs; accurate accounting of trial-related activities ensures appropriate billing and reimbursement. To the extent that health IT can improve the efficiency of trial administration and billing, more patients can be included in trials within existing financial support.

■ **Increased use of treatment protocols outside of formal clinical trials.** Through decision supports embedded at the point of care, health IT can increase the use of standardized treatment protocols and guidelines for patients not in formal clinical trials. Examples include the following: (1) Templates for comprehensive cancer treatment care plans based on prior or concurrent trials, clinical guidelines, and standardized chemotherapy regimens: These support consistent use of chemotherapy and adjunct therapies such as anti-emetic drugs and hematopoietic growth factors. In the Northwest region, where EpicCare use predated current versions of KP HealthConnect by several years, EHR-embedded standardized templates addressing more than sixty common chemotherapy protocols and follow-up testing are used to manage most patients receiving chemotherapy. (2) Flow sheets to track clinical findings over time: Examples include blood counts and laboratory chemistries for patients undergoing cyclic chemotherapy. (3) Intrusive alerts capable of interrupting clinician workflow: These may notify clinicians about prevention opportunities, such as breast and cervical cancer screening, and address risks such as adverse drug events, redirecting care to a preferred alternative. Alerts can also notify clinicians of an atypical or worrisome finding, such as a questionable result on an x-ray or screening test.

Our experience is that KP HealthConnect must be enhanced to fully support key aspects of cancer care, such as transmitting and tracking pharmaceutical orders and documenting chemotherapy administration. As noted above, Kaiser Permanente's Northwest region, with an EpicCare-based system, has used standardized chemotherapy templates for approximately ten years as decision support in conjunction with an Internet-based repository of standardized treatment and available investigational protocols, patient educational materials, and other tools. The majority of patients receive standardized protocols. However, transmission of chemotherapy orders to the pharmacy or nursing infusion stations required manual workarounds and parallel documentation in legacy IT systems and on paper.

Two other regionally developed health IT oncology applications supported cancer care processes such as chemotherapy ordering and administration and also included patient information such as treatment calendars and educational materi-

als. Practicing oncologists evolved the first, Case Management for Medical Oncology with Laboratory and Outcomes Tracking (CAMMOLOT). Regional pharmacy informatics resources developed the second, Computerized Oncology Practice System (COPS). Both became widely used in their respective regions.

The Northwest region's use of embedded templates and other decision support, although fully interoperable with other aspects of KP HealthConnect, did not cover the necessary range of oncology services. Both CAMMOLOT and COPS had more robust support for isolated oncology care but were not interoperable with KP HealthConnect.

A proposed oncology module for KP HealthConnect would capture and expand on the capabilities of all three regionally developed applications through hundreds of specific design requirements. These include multiple patient-safety-check processes; medication-related safety checks; standardized treatment plans and order sets; integrated decision support and treatment administration rules; full documentation of clinical trial plans, assessments, and interventions; display of chemotherapy regimen costs; and overall utilization and cost reporting. But even without fully interoperable support, much progress in supporting cancer care at Kaiser Permanente has been made through existing combinations of compatible technology-based systems and workflow accommodations as appropriate.

The Shifting Landscape Of Cancer Research With Health IT

The combination of the wealth of data and increased standardization of care supported by health IT reshapes the landscape of research in cancer. More usable data will be available on more patients in more clinical situations than ever before. Health IT offers new opportunities to address key research questions in coordinated and complementary ways.

■ **Collaborative research: beyond a single system of cancer care.** A national health information network is an explicit priority at the level of the federal government, and the adoption of health IT systems is slowly increasing.[11] It is highly likely that widespread health IT implementation and interoperability will eventually support wide-ranging collaborative cancer research activities across organizations and sites of care. However, health IT systems are far from ubiquitous, particularly outside integrated care models such as KP HealthConnect. Collaborative research in cancer care can—and must—occur before the ideal state of implementation and interoperability is reached. Important initiatives do exist, however, demonstrating how the research capabilities of integrated HIT systems such as KP HealthConnect can be complemented and extended by networking with other care providers and settings.

Cancer Research Network. The CRN includes six Kaiser Permanente regions and other HMORN institutions.[12] Its goal is to improve research on cancer prevention, early detection, treatment, long-term care, and postdiagnosis monitoring in health care delivery systems by facilitating data sharing. Data are pooled across plans,

"Rapid learning about cancer supported by KP HealthConnect can and should apply beyond Kaiser Permanente's borders."

time periods, and individual patients and include information on demographics, health plan eligibility, tumor registry, inpatient and ambulatory care use, medication dispensing, laboratory tests, and imaging procedures. Local programmers and analysts control data, thereby minimizing privacy and proprietary concerns.

Major completed CRN studies include evaluations of prophylactic mastectomy and studies of organizational factors related to incidence of advanced-stage breast and cervical cancer; system-level approaches to increasing participation in cancer clinical treatment trials; cancer screening; and the use of smoking-cessation interventions within systems.[13] The research centers also respond to ad hoc inquiries for population-based data on cancer incidence and treatment patterns.

Data warehouse. CRN member organizations have disparate EHR systems. However, the CRN has developed a broadly applicable standardized data structure to create a virtual data warehouse (VDW), allowing rapid learning in the absence of full functional interoperability. CRN leaders set research priorities and develop a standard set of variable names and structures that can be used across the member organizations. Seven major types of standardized data files developed to date are demographics, health plan eligibility, tumor registry, inpatient and ambulatory care use, medication dispensing, laboratory tests, and imaging procedures.[14] CRN site data managers write programs mapping local data onto standardized names and structures, and CRN editing programs detect any problems with the data. VDW users are CRN site analysts; the warehouse is virtual because the data remain at the individual organizations and are only centralized when a specific research use has received required institutional approvals.[15]

An example of the use of the VDW is the HMOs Investigating Tobacco 2 (HIT2) study. A team of tobacco-cessation experts developed a concept map for defining coding rules consistent with national tobacco-cessation guidelines. These concepts formed the foundation for the development of a natural language coding program (MediClass). Each of four HIT2 sites—three Kaiser Permanente regions and Harvard Pilgrim—have developed local procedures for extracting relevant patient data and putting them into a standardized format for MediClass analysis. An assessment of MediClass revealed that it adequately replaced human coders for assessing clinician adherence to tobacco-cessation guidelines.[16]

Biomedical Informatics Grid. Another NCI-sponsored networking initiative in which Kaiser Permanente is not directly involved further demonstrates the potential for increasing the scale and pace of learning in cancer. The cancer Biomedical Informatics Grid (caBIG) was formed in 2004 and is an open-source, open-access, voluntary information network enabling cancer researchers to share tools, standards, data, applications, and technologies according to agreed-upon standards

and needs.[17] Membership includes NCI-designated cancer centers, Clinical Trials Cooperative Groups, and additional NCI programs. The caBIG mission is to promote and facilitate the use of a set of common data standards; a sharable, interoperable infrastructure; tools for applying information associated with cancer research and care; and data sharing among appropriate individuals and organizations, with safeguards for privacy and security.[18]

caBIG explicitly recognizes that interoperability is "not an all or nothing process, [and] that software systems can exist along a continuum of interoperability levels."[19] Although interoperable systems are a highly desirable ideal state, rapid learning in cancer in the near term depends on simultaneously striving for the ultimate technical solution of universal technology and data standards and taking intermediate steps to reconcile the differences between systems through data-mapping and information-sharing projects such as those of the CRN and caBIG.

■ **Clinically detailed data.** Integrated health IT systems such as KP HealthConnect complement networked learning with the universal capture of data that are rich in clinical detail. Populationwide depictions of care are available; so is analysis at the level of individual patients' interactions with care processes. Rapid learning about cancer supported by KP HealthConnect can and should apply beyond Kaiser Permanente's borders.

Nearly a quarter-million Kaiser Permanente members will be diagnosed with cancer during the next five years. The robust data available through KP HealthConnect on members will rival, in terms of clinical detail if not in absolute numbers, existing cancer research databases. In comparison, an average of 33,600 people were enrolled in NCI trials during 2001–2003.[20] With broad inclusion of members of all ages and geographic representation in nine states and the District of Columbia, Kaiser Permanente membership reflects the larger U.S. population.

ACHIEVING FULL DATA CAPTURE AND INTEROPERABILITY in the U.S. health care system is unlikely in the near future. However, this does not preclude the expansion and acceleration of learning in cancer now. Rapid learning beyond that available through conventional clinical trials and surveillance may be achieved by expanding the breadth of ongoing observations across diverse sites and legacy systems using networking and data mapping and combining these results with the clinical depth attainable through more comprehensive data capture from integrated systems such as that of Kaiser Permanente.

The author expresses his appreciation to Daniel Colley, Nagendra Tirumali, Robert Hiatt, David Campen, Robert Dolin, Joseph Terdiman, Larry Walker, Ted Palen, and Michael Johnson for their helpful comments, suggestions, and related presentations and to Jennifer Green for manuscript production and editorial support.

NOTES

1. White House, "Transforming Health Care: The President's Health Information Technology Plan," 2004, http://www.whitehouse.gov/infocus/technology/economic_policy200404/chap3.html (accessed 27 June 2006).

2. R. Hillestad et al., "Can Electronic Medical Record Systems Transform Health Care? Potential Health Benefits, Savings, and Costs," *Health Affairs* 24, no. 5 (2005): 1103–1117.
3. American Cancer Society, "Costs of Cancer," 9 February 2006, http://www.cancer.org/docroot/MIT/content/MIT_3_2X_Costs_of_Cancer.asp (accessed 14 December 2006).
4. National Cancer Institute, "The Nation's Investment in Cancer Research: Integrating Cancer Science," http://plan2007.cancer.gov/IntegratingCancer.shtml (accessed 15 December 2006).
5. A.S. Grove, "Efficiency in the Health Care Industries," *Journal of the American Medical Association* 294, no. 4 (2005): 490–492.
6. R. Kaushal et al., "The Costs of a National Health Information Network," *Annals of Internal Medicine* 143, no. 3 (2005): 165–173.
7. R.H. Dolin and A.M. Wiesenthal, "National Health Information Network Cost and Structure," *Annals of Internal Medicine* 144, no. 2 (2006): 144.
8. R.H. Dolin et al., "Kaiser Permanente's Convergent Medical Terminology," *Medinfo* 11, Part 1 (2004): 346–350; and SNOMED International, "Global Reach," http://www.snomed.org/about/global.html (accessed 28 June 2006).
9. M.C. Hornbrook et al., "Building a Virtual Cancer Research Organization," *Journal of the National Cancer Institute Monographs*, no. 35 (2005): 12–25.
10. C.P. Somkin et al., "Organizational Barriers to Physician Participation in Cancer Clinical Trials," *American Journal of Managed Care* 11, no. 7 (2005): 413–421.
11. U.S. Department of Health and Human Services, "Development and Adoption of a National Health Information Network," *Federal Register* 69, no. 219 (2004): 65599–65601; and Medical Records Institute, *Seventh Annual Survey of Electronic Health Records Trends and Usage for 2005*, http://www.medrecinst.com/files/ehrsurvey05.pdf (accessed 4 August 2006).
12. Hornbrook et al., "Building a Virtual Cancer Research Organization."
13. L.J. Herrinton et al., "Efficacy of Prophylactic Mastectomy in Women with Unilateral Breast Cancer: A Cancer Research Network Project," *Journal of Clinical Oncology* 23, no. 19 (2005): 4275–4286; A.M. Geiger et al., "A Population-Based Study of Bilateral Prophylactic Mastectomy Efficacy in Women at Elevated Risk for Breast Cancer in Community Practices," *Archives of Internal Medicine* 165, no. 5 (2005): 516–520; S.H. Taplin et al., "Reason for Late-Stage Breast Cancer: Absence of Screening or Detection, or Breakdown in Follow-up?" *Journal of the National Cancer Institute* 96, no. 20 (2004): 1518–1527; W.A. Leyden et al., "Cervical Cancer in Women with Comprehensive Health Care Access: Attributable Factors in the Screening Process," *Journal of the National Cancer Institute* 97, no. 9 (2005): 675–683; J.G. Zapka et al., "Processes of Care in Cervical and Breast Cancer Screening and Follow-up—The Importance of Communication," *Preventive Medicine* 39, no. 1 (2004): 81–90; Somkin et al., "Organizational Barriers"; K.V. Goins et al., "Implementation of Systems Strategies for Breast and Cervical Cancer Screening Services in Health Maintenance Organizations," *American Journal of Managed Care* 9, no. 11 (2003): 745–755; V.P. Quinn et al., "Tobacco-Cessation Services and Patient Satisfaction in Nine Nonprofit HMOs," *American Journal of Preventive Medicine* 29, no. 2 (2005): 77–84; and L.I. Solberg et al., "Tobacco Control Efforts in Managed Care: What Do the Doctors Think?" *American Journal of Managed Care* 10, no. 3 (2004): 193–198.
14. Hornbrook et al., "Building a Virtual Cancer Research Organization."
15. G. Hart et al., "The Cancer Research Network's Virtual Data Warehouse," http://www.centerforhealthstudies.org/scires/rposter/04hmorn/04_hart_HMORN.pdf (accessed 4 August 2006).
16. B. Hazlehurst et al., "Natural Language Processing in the Electronic Medical Record Assessing Clinician Adherence to Tobacco Treatment Guidelines," *American Journal of Preventive Medicine* 29, no. 5 (2005): 434–439.
17. NCI, "cancer Biomedical Informatics Grid (caBIG): Fact Sheet," 9 March 2004, http://www.cancer.gov/newscenter/pressreleases/caBIGQandA (accessed 22 November 2006).
18. NCI, "caBIG FAQs," http://cabig.nci.nih.gov/faqs/index.html#mission (accessed 15 December 2006).
19. NCI, "caBIG: Bronze Compatibility Certification," 7 April 2006, http://cabig.nci.nih.gov/guidelines_documentation/bronze (accessed 22 November 2006).
20. NCI, "NCI-Supported Cancer Clinical Trials: Facts and Figures," 28 March 2006, http://www.cancer.gov/clinicaltrials/facts-and-figures/page2 (accessed 19 June 2006).

PERSPECTIVE

Health Information Technology: Does It Facilitate Or Hinder Rapid Learning?

Can health IT help the fragmented delivery system, outside of Kaiser Permanente, achieve more "rapid learning" in cancer care?

by L. Gregory Pawlson

ABSTRACT: Health information technology presents major challenges as well as opportunities in creating care that fulfills the Institute of Medicine's aims of being safe, timely, effective, efficient, equitable, and patient-centered. This commentary examines the barriers that relate directly to the collection and use of information in practice, and it explores some of the possible solutions. Only through concerted efforts involving major changes in the way we collect, store, analyze, and use information related to the care of patients with cancer, paired with active practice redesign and reimbursement reform, are we likely to achieve the substantial progress envisioned by Paul Wallace. [*Health Affairs* 26, no. 2 (2007): w178–w180 (published online 26 January 2007; 10.1377/hlthaff.26.2.w178)]

UNDERSTANDING AND improving quality of care for patients with cancer has been noted to be challenging.[1] Some of the factors that create these challenges include rapid changes in treatment modalities, small-sample-size problems imposed by the relatively low incidence of specific types of cancer, and heterogeneous treatment approaches during the multiple phases of cancer care (screening, initial treatment, surveillance, secondary treatment, palliative care, hospice). In adult cancer care, there is the additional issue of fragmentation of care by multiple, loosely connected practitioners. For example, a woman with breast cancer might be treated initially by a surgeon, then by a radiation oncologist, then a medical oncologist, and then be followed by a general internist, all of whom operate independently. This fragmentation in treatment also contributes to conflicting and competing guidelines and a lack of consensus about the "best" approach; incomplete data present in any one practice, data system, or medical record; and confused communications between providers and between providers and patients. Finally, there are major delays both in learning about the effects (both good and bad) of new treatments and, as with other diseases, in the dissemination of treatments that have been shown to be of substantial benefit.[2]

As well illustrated in the paper by Paul Wallace, health information technology (IT) could provide a critically important set of tools to increase the speed and depth of our ability to measure, guide, and improve the care of people with cancer.[3] Wallace clearly implies that rapid learning is linked with better care. However, the promise of health IT in fostering rapid learning and better quality will not be realized without major changes in the way the information is recorded and used. Substantial changes are needed in what, when, and how clinical data are recorded and in the systems of care in which the enhanced information is used.

Greg Pawlson (Pawlson@ncqa.org) is executive vice president of the National Committee for Quality Assurance in Washington, D.C.

DOI 10.1377/hlthaff.26.2.w178

A growing number of studies and reviews of health IT implementation suggest that health IT, in the absence of careful attention to the context of its implementation and use, may provide little or no benefit.[4] Moreover, the organization (Kaiser Foundation Health Plan and Kaiser Permanente) in which Wallace practices is ideally positioned to derive maximal benefit from the introduction of health IT in rapid learning related to cancer care. Kaiser Permanente, and the small number of other large U.S. integrated practices, enjoy the presence of a highly evolved, multispecialty, and primarily salaried medical staff; this workforce, along with multiple systems related to medical management and feedback, is likely to result, as Wallace predicts, in much more rapid learning and improvement in care.[5] But what about the vast majority of practices in small offices across the country? Can Americans afford the substantial loss in quality from fragmented care and its added cost as well as the cost of health IT? Can health IT help the fragmented delivery system, outside of Kaiser Permanente, achieve more "rapid learning" in the care of people with cancer?

"The highest barrier to rapid learning is the poor quality and lack of standardization in the data recorded in most clinical treatment records."

One of the well-established pathways to learning in clinical practice is the use of the gold standard: randomized clinical trials. Comparisons between the clinical trial, usual practice, and what health IT—if well designed and implemented—could provide will help us understand the key factors that might lead to accelerated learning through the use of health IT. In this commentary I briefly examine each of these areas in terms of the barriers to health IT's resulting in rapid learning and improved quality of care in patients with cancer.

■ **Selecting patients.** The use of health IT would allow practices to identify groups of patients with a given type and phase of cancer and to group those patients for purposes of analysis with others within a practice or between practices. If the right information were encoded and shared, every patient identified with new-onset, early-stage breast cancer, with given characteristics (for example, positive or negative lymph nodes, drug and dose), could be followed along with other similar patients for initial response, complications, and outcomes. As Wallace notes, the key barriers to overcome in this area are largely the legal, technical, and political issues related to data exchange.

■ **Protocols and data.** Two related barriers, the use of standardized protocols and the data recorded, loom as substantial obstacles. The more variation there is in the treatments used in the care of people with a given type or stage, or in other factors, the more difficult it is to come up with any coherent information, even where there are substantial data to use for adjustment. Given the high cost and toxicity of most of the drugs used in cancer treatment and evidence of under- or overtreatment, the use of standardized protocols for most patients for most types of cancer would likely both improve quality and reduce resource use. Perhaps the availability of health IT—and specifically electronic health records (EHRs) with embedded protocols for cancer treatment—coupled with strong payment incentives for following the protocols, might actually drive us in the right direction.

■ **Data from treatment records.** Arguably the highest barrier to rapid learning at this point is the poor quality and lack of standardization in the data recorded in most clinical treatment records. In stark contrast to the highly standardized, evidenced-based data recorded in clinical trials, each clinician treating a patient decides independently what data are "important" enough to put into the clinical record. Again, EHRs offer the potential of such tools as drop-down data forms or embedded

data recording. However, little time and attention have been given to creating data forms that specify and prompt collection of the data that are most critical and important both in caring for the patient and in measuring and improving care. Health IT will achieve relatively little if we persist in allowing "free-text" nonstandardized or coded entries.

■ **Follow-up.** Another difference that appears important between learning from health IT and from clinical trials is that of the completeness of follow-up. Even clinical trials do not provide complete follow-up of all patients; however, they are vastly superior to the fragmented data that exist for most cancer patients. Although some data elements—primarily claims—are present at the insurer level, clinical data needed to adequately assess care of cancer patients are often spread across multiple charts of multiple providers. Kaiser as insurer and Permanente as well-integrated medical group has a unique opportunity to bring all of the critical data together, but even Kaiser Permanente is still far from the ideal. With the exception of few pilot projects, there has been little mapping of how data from the medical records of multiple unconnected physicians and multiple insurers will be brought together in any workable, coherent fashion. Again, there is some promise that health IT could, through prompts and reminders and the exchange of data between clinicians, bring together the right data, at the right time, in the right format, but there is a huge gap to be bridged.

■ **Analysis and reporting.** The issue of analysis and reporting, although not trivial, might be the most easily solved. Once fairly complete data from multiple patients are available, a variety of approaches might be used to emulate the analytical approaches and reporting mechanisms used in clinical trials. There is still a need for more standardization in how comorbidities might be handled and what formats are most effective in producing change, but these do not seem to be major barriers.

■ **Privacy provisions.** Finally, and not addressed in any detail in Wallace's paper, is the continued need for clarification of data-use limitations and privacy provisions in the Health Insurance Portability and Accountability Act (HIPAA). If clinical information is to be used for what could arguably be termed research as well as rapid learning, would review by an institutional review board or other patient privacy and protection issues need to be addressed?

NOTES

1. National Cancer Policy Board, *Ensuring Quality Cancer Care* (Washington: National Academies Press, 1999); and S.B. Clauser, "Use of Cancer Performance Measures in Population Health: A Macro-Level Perspective," *Journal of the National Cancer Institute: Monographs* no. 33 (2004): 142–154.
2. C. Enfant, "Shattuck Lecture: Clinical Research to Clinical Practice—Lost in Translation?" *New England Journal of Medicine* 349, no. 9 (2003): 868–874.
3. P.J. Wallace, "Reshaping Cancer Learning through the Use of Health Information Technology," *Health Affairs* 26, no. 2 (2007): w169–w177 (published online 26 January 2007; 10.1377/hlthaff.26.2.w169).
4. C. Delpierre et al., "A Systematic Review of Computer-Based Patient Record Systems and Quality of Care: More Randomized Clinical Trials or a Broader Approach?" *International Journal for Quality in Health Care* 16, no. 5 (2004): 407–416; P.J. O'Connor et al., "Impact of an Electronic Medical Record on Diabetes Quality of Care," *Annals of Family Medicine* 3, no. 4 (2005): 300–306; T.D. Sequist et al., "A Randomized Trial of Electronic Clinical Reminders to Improve Quality of Care for Diabetes and Coronary Artery Disease," *Journal of the American Medical Informatics Association* 12, no. 4 (2005): 431–437; and B. Kralj et al., "The Impact of Computerized Clinical Reminders on Physician Prescribing Behavior: Evidence from Community Oncology Practice," *American Journal of Medical Quality* 18, no. 5 (2003): 197–203.
5. R.R. Gillies et al., "The Impact of Health Plan Delivery System Organization on Clinical Quality and Patient Satisfaction," *Health Services Research* 41, no. 4 (2006): 1181–1199; and S.M. Shortell et al., "An Empirical Assessment of High-Performing Medical Groups: Results from a National Study," *Medical Care Research and Review* 62, no. 4 (2005): 407–434.

Bridging The Inferential Gap: The Electronic Health Record And Clinical Evidence

Emerging tools can help physicians bridge the gap between knowledge they possess and knowledge they do not.

by Walter F. Stewart, Nirav R. Shah, Mark J. Selna, Ronald A. Paulus, and James M. Walker

ABSTRACT: Most clinical decisions involve bridging the inferential gap: Clinicians are required to "fill in" where they lack knowledge or where no knowledge yet exists. In this context we consider how the inferential gap is a product, in part, of how knowledge is created, the limits to gaining access to such knowledge, and the variable ways in which knowledge is translated into decisions. We consider how electronic health records (EHRs) will help narrow this gap by accelerating the creation of evidence relevant to everyday practice needs and facilitating real-time use of knowledge in practice. [*Health Affairs* 26, no. 2 (2007): w181–w191 (published online 26 January 2007; 10.1377/hlthaff.26.2.w181)]

IN THIS PAPER WE EXAMINE the inferential gap common to everyday practice: the gap between the paucity of what is proved to be effective for selected groups of patients versus the infinitely complex clinical decisions required for individual patients. Clinicians engage in information gathering and interpretation; they implicitly or explicitly bridge the gap every day to care for their patients. The breadth of the inferential gap varies according to available knowledge, its relevance to clinical decisions, access to the knowledge (that is, what the physician actually knows at the time of a clinical decision), the variable ways in which knowledge is interpreted and translated into a decision, the patient's needs and preferences, and a host of other factors. Clinicians are required to fill in where their knowledge (or knowledge itself) falls short. These issues are increasingly important for an aging U.S. population where clinical decisions must consider the patient's entire complement of comorbidities, genetic predispositions, and preferences.[1]

We consider two fundamental means by which electronic health records (EHRs) will narrow this gap. First, EHRs will facilitate the creation of evidence

The authors are with the Geisinger Health System in Danville, Pennsylvania. Walter Stewart (wfstewart@geisinger.edu) is the director of the Center for Health Research, Geisinger Clinic. Nirav Shah is an associate investigator at the center. Mark Selna is associate chief medical officer, Division of Clinical Effectiveness. Ronald Paulus is chief technology and innovation officer; James Walker is chief medical information officer.

DOI 10.1377/hlthaff.26.2.w181

that is directly relevant to everyday clinical decisions.[2] Second, EHRs will greatly increase real-time access to knowledge in the practice setting.[3] We consider five specific scenarios that are relevant to these fundamental shifts in the creation and use of clinical evidence. In the long term, we believe that EHRs will offer a novel approach to the creation of clinical knowledge, in which observing, intervening, and creation of clinical evidence are part of the normal clinical encounter.

We first briefly review the evolution of medical evidence and its limitations.

Medical Evidence: How Relevant To Patient Care?

Historically, the formal creation of clinical knowledge began with case reports; it then evolved to more sophisticated observational studies, then to controlled trials, and, ultimately, to the randomized controlled trial (RCT). To some degree, observational studies and RCTs are complementary. The primary strength of systematic observations (that is, scheduled collection of data) in relatively unselected populations is that the resulting evidence is highly relevant to a broad spectrum of patients (that is, generalizable). However, inference from such studies is prone to challenges with interpretation from bias and confounding.[4] The use of randomization—the key design feature of RCTs—minimizes these biases.

Traditionally, knowledge regarding quality of care was based on clinical experience, case studies, and reasoning from physiological first principles. This knowledge sometimes led to improvements dramatic enough to convince even the most skeptical observers of a treatment's benefits (for example, Jenner's smallpox vaccination): There was minimal variability in outcomes and the cause-effect relationship appeared obvious. The relative dearth of knowledge at the time, the narrowness and simplicity of the questions being asked, the low expectations of patients and other interested parties, and the high benefit-to risk ratios justified the use of observational methods as a means of discovery. In the early twentieth century, though, when medical knowledge creation itself emerged as a nascent "industry," research methods evolved to include more comprehensive observational studies and, ultimately, the RCT.[5] The increasing complexity of clinical questions and treatment options, diversity of potential clinical outcomes and their alternative scientific explanations, as well as growing concerns with bias and confounding all contributed to a demand for a more rigorous and reliable approach to answering questions, for which the RCT emerged as the gold standard. However, the application of the RCT has evolved in response to the above demands with increasingly narrow-focused interventions applied to increasingly selected populations.[6]

As the dominant method for creating clinical knowledge, the RCT is not typically used to address questions directly relevant to the practice setting; for cost, practical, and logistical reasons, the traditional means of using the RCT in this context is not sensible. Instead, RCTs have been and will continue to be used to test for treatment benefits in highly selected populations with a low comorbid disease burden. More specifically, RCTs are used primarily to define whether a

"In the past three decades, questions that have to be answered in making clinical decisions have become more intricate."

drug or intervention is beneficial in an artificially optimized clinical setting, not whether it makes sense for most patients or is suitable for one patient subgroup but not another. Strict selection criteria exclude high-risk patients. Furthermore, commonly used run-in phases prior to randomization select for the most adherent patients. The RCT can be easily manipulated to influence what is observed and discoverable.[7] Even seemingly trivial changes in inclusion criteria, intervention characteristics, or follow-up duration (for example, the recent publicity over the cutoff date for the Vioxx trial's cardiac events) can lead to different conclusions.[8] For example, RCTs on the treatment of mild hypertension and coronary disease have alternatively shown both benefit and harm, as a result of the differential patient exclusion criteria among different trials.[9]

Other important changes have also occurred that contribute to growing challenges with the inferential gap. In the past three decades, patient populations have become older and more heterogeneous, and the questions that have to be answered in making clinical decisions have become more intricate. Patients with complex medical needs tend not to be eligible for RCTs. Consequently, for the growing population of patients with multiple comorbidities, medication intolerances, limited cognition, and diverse insurance coverage, the knowledge needed to support objective clinical decision making is largely nonexistent.[10]

To some degree, systematic observational studies—especially in clinical settings—address some of the limits to RCT-based evidence.[11] Outcomes in "real-world" clinical practice reflect the clinician's knowledge, skills, preferences, and interaction with the patient as well as patient factors (such as self-management, adherence, and willingness and ability to pay) and system features (such as care-coordination assets and ease of access to care). Systematic observations of these and other factors in unselected populations address a key need: generalizability. The primary challenge, however, is dealing with the increased likelihood of confounding, specifically confounding by indication (that is, when a medical condition both triggers the use of a specific treatment and is associated with a risk of the outcome under study).[12] Confounding by indication is inherent to many clinical decisions. For example, disease severity can influence the choice of treatments, making it difficult to separate the indication for a treatment from the risks/benefits of the treatment itself. This problem is commonplace when a new drug comes to market. For example, because of preestablished practice patterns, a new class of medication might implicitly be used only for patients with more-severe disease (for example, uncontrolled hypertension). Analysis of the benefits of the new drug class versus older classes will be confounded because, in part, the treatment decision will link disease severity to when a drug came to market.[13]

EHRs And Clinical Evidence

The creation of evidence using traditional research designs is extremely time-consuming, costly, and, as previously noted, limited. We consider five different scenarios (Exhibit 1) for how EHRs will likely influence the traditional research paradigm. The scenarios are presented in sequence from ones that primarily offer logistical advantages to those directly relevant to clinical practice needs. Together, these scenarios describe unique features of research enabled by the advent of EHR-based clinical practice. Namely, the process of engaging in practice-based research will motivate improvements in data quality and the specificity of questions that can be answered, which will in turn influence the ability to monitor and improve patient outcomes and vice versa.

■ **Facilitating practice-based RCTs.** There are regulatory, logistical, and cost

EXHIBIT 1
Profiles Of Scenarios Involving The Electronic Health Record (EHR)

Scenario	Dominant application	Example	Strengths/advantages	Disadvantages
Facilitating RCTs in practice	Traditional practice-based RCTs	Large multicenter clinical trials	Acceleration of RCT timeline, understanding of selection bias	Favor EHR adopters
Retrospective analysis of EHR data	Retrospective analysis of typical clinical interventions that have been in use for several years; etiology, diagnosis, and prognosis studies	Evaluate effectiveness of different treatments and combinations of treatment for the management of hypertension	Highly relevant to clinical practice; offers the only means for comparative treatment analysis for the same outcome	Measurement error, bias, confounding, confounding by indication, and completeness of data
Testing of self-sustaining health services models	Use of patient-completed questionnaires, systematic use of behavioral health interventions, patient self-management training, systematic adherence to care per guidelines	For example, like the 1960s HIP study of mammography; introducing a new diagnostic by varying eligibility[a]	Testing of new models of care must fit the workflow of clinical practice; if the model improves outcomes and is successful, then the "translation to practice" challenge is simultaneously addressed and solved	Requires a strong alignment of the research, IT, and practice groups
Real-time use of decision support and all patient data in practice	In everyday practice to directly integrate CPGs and other codified knowledge with practice decisions and to make use of data on other patients as needed to tailor treatment for a particular patient	What is the best medication to add to an already complex regimen for a patient with multiple comorbidities and insurance restrictions on treatment options?	Provides clinical guidance in complex situations; directly provides the physician with knowledge as needed; makes effective use of longitudinal data on other patients	Limits to knowledge and the above methodological limits to use of retrospective data
Randomizing uncertainty in real time	Clinical decision making in situations of equipoise	What is the real-world NNT for various statins?	Simple to do and replicate in numerous settings	Generalizability might be overestimated if relying on one population

SOURCE: Derived from the authors' own work.

NOTES: RCT is randomized controlled trial. IT is information technology. CPG is clinical practice guideline. NNT is number needed to treat.

[a] See S. Shapiro et al., "Current Results of the Breast Cancer Screening Randomized Trial: The Health Insurance Plan (HIP) of Greater New York Study," in *Screening for Breast Cancer*, ed. N.E. Day and A.B. Miller (Toronto: Hans Huber, 1988), 3–15.

challenges in conducting RCTs that together have greatly increased trials' average time to completion. For a number of reasons, expansion of the number of ambulatory practices with EHRs will mitigate various causes of delay. First, pretrial analysis of various inclusion/exclusion criteria against an EHR database will improve the protocol used to optimize safety, increase the number of eligible patients, and speed enrollment. Second, the growth of EHR-based practices will influence the "reach" of clinical trials and help move from the typical "high recruiting" clinics to a broader population. Third, access to EHR data will provide a more specific understanding of one dimension of the inferential gap: differences between patients who choose to participate in trials versus peers who would have qualified but opted out versus those who were excluded. Lastly, the efficiency of data capture and data quality monitoring is likely to improve as EHRs and electronic data capture (EDC) tools become increasingly integrated, providing for seamless transmission of data from the EHR to a digital case report form, billing record, or real-time adverse-event alert.

Several years ago we at Geisinger Health System embarked on a systems approach to clinical trials, providing centralized support services (for example, protocol review, contracting, institutional review board [IRB] application and management, and trial setup and monitoring) to both primary and specialty care. In the past year we have begun to use EDC tools to seamlessly extract data from the EHR to the digital case report form. Although we believe that EDC tools offer tactical advantages in the management of individual trials, they also address a fundamental challenge at a system level in accurately tracking progress in multiple trials (that is, number of patients, specific visits completed, specific data captured). Finally, implementation and use of EDC tools at a system level will likely facilitate data collection and study management for traditional epidemiologic and health services research. Clinical departments at Geisinger may develop their own resources; however, the centralized resources we are developing are offered as a competitive option. Increasingly, the system-level resources and capabilities exceed what is possible within any given research unit or clinical department.

■ **Retrospective analysis.** Given the relative ease with which access to longitudinal EHR data can be gained, retrospective data analysis will be a dominant focus of interest in addressing questions of treatment benefit and harm, in mining for new treatment indications, and in answering the complex but common questions that arise in practice. We consider a few examples of retrospective data analyses with a specific focus on blood pressure—a common and important clinical measure.

Policymakers have an interest in the relative value of different treatment regimens for managing blood pressure.[14] Access to this type of comparative information will address a fundamental gap relevant to the cost and effectiveness of various drug formulary designs. Although meta-analysis applied to traditional RCT data might reveal comparative information on treatment benefits across RCTs, conclusions are almost always uncertain, given numerous methodological challenges (for example, heterogeneity in measurement protocols, populations, and

treatment regimens) and previously noted limitations.[15]

For individual patients, the study of treatment response in practice and interactions between treatment status and patient traits will advance clinical practice guidelines toward the numerous and complex issues common to everyday practice. Treatment decisions for a particular patient will likely be influenced by a host of system (for example, formulary options), socioeconomic, treatment (for example, adverse events and response), clinical, genetic, and other factors. Making treatment decisions with evidence relevant to these and other nuances of needs and individual attributes both closes the inferential gap and supports a more patient-centered approach to care.

High blood pressure. A large population example we have begun to consider concerns the management of high blood pressure in the elderly. Since 1991, large-scale studies suggest that antihypertensive drugs that greatly lower diastolic blood pressure in men older than age seventy-five and possibly in older women are associated with higher mortality rates, despite their apparent benefit in lowering systolic pressure.[16] The relative impact on systolic and diastolic pressure likely differs by drug class, some of which have not been evaluated; comorbidities; and other risk factors such as smoking status. Today we still know relatively little about who is actually at risk, aside from the broad category definitions such as males older than age seventy-five. We will be using comprehensive longitudinal EHR data (that is, sequential blood pressures, medications prescribed, nonfatal and fatal events, diagnoses, smoking status, and so on) to investigate the relative safety and benefits of antihypertensive medications relative to probable risk factors. The resulting knowledge will be used to develop decision-support logic for blood pressure management in older patients.

Data limitations. Limitations to the quality and completeness of EHR data will be the "Achilles' heel" that constrains evidence that can be extracted from retrospective analysis. In research settings, rigorous standards are used to measure and record data. In practice, though, this is not always the case. Ultimately, we believe that data quality and completeness challenges will have to be resolved through standards of practice that satisfy different stakeholders. For example, improving the quality of blood pressure data would require that the numerous clinical "habits" that result in bias and error are minimized.[17] Research protocols designed for this purpose are not logistically or financially suitable for practice settings. Rather, a more sensible standard of practice could mean that blood pressure is measured only by automated cuffs and obtained sequentially at defined intervals, and that results are directly transmitted (perhaps wirelessly) to the EHR. Improvements to data quality will accelerate as common interests are identified among key stakeholders (that is, clinical effectiveness monitoring, clinical operations, and research) and, in particular, where new methods improve data quality without imposing a burden on the practice and, more likely, where such methods improve efficiency.

Notions of data completeness—a routine obsession of research protocols—poses challenges for retrospective analysis. The schedule for when and what data are collected on patients in practice must, by necessity, be linked to what is sensible for appropriate care. In general, patients with more health problems will have more visits and more data. EHRs can facilitate and improve the likelihood that a patient is seen when appropriate (for example, via automated visit reminder letters). But the notion of completeness itself raises numerous questions. Are there minimal data needs (for example, height, repeated measures of weight, blood pressure, or lipids) for any retrospective analysis? Does the optimal schedule for collecting data differ for each clinical measure? What design, analytic, and inference challenges are created when the amount of data is related to a patient's health status? These and other questions will pose challenges, as previously noted, to aligning the interests of stakeholders regarding the data to be routinely collected during encounters. In particular, the data needs of researchers, practitioners, and those who manage quality of care will create tensions that can best be balanced through protocols that both improve the completeness and standardization of data captured and reduce the cost of obtaining such data. Such technical solutions already exist in a number of areas and are used with increasing frequency at Geisinger. For example, we have begun to develop workflow and data capture models for patient-completed questionnaires. Computerized order entry of prescriptions, tests, and procedures that require selection from a menu of predefined options can be set to require that one or more diagnostic codes be selected, indicating the intention behind the order. When used properly, predefined order sets, structured notes with defaults, and consultation templates standardize the content and organization of data input and can even enable structured, codified data capture. Even these rudimentary EHR protocols have the potential to contribute meaningful evidence that will complement knowledge gained from randomized trials.

Within the bounds of these limitations, retrospective analyses of EHR data offer enormous potential value and will inevitably advance methods relevant to causal inference. It is likely that a body of science (for example, validation studies of established previous findings and knowledge of the potential influence of different confounders) will emerge to focus on just this issue as it has in other areas and will give way to standards of practice relevant to the interpretation of EHR-based evidence. Specific methods will ultimately need to be developed linking the types of questions to the analytic methods most suitable to answer them.[18]

■ **Translation of health services models to practice.** The products of research on a new health services protocol rarely get widely translated into practice. Inherent constraints to traditional practice settings limit both the usability and the complexity of protocols that can be tested and the sustainability, exportability, and scalability of proven solutions. The EHR-based practice offers a paradigm shift for how research moves beyond these traditional constraints. We specifically consider the use of patient-completed questionnaires to exemplify differences in the method

and impact of research.

Questionnaires have been developed to facilitate diagnosis, improve patient-physician communication, standardize patient-reported outcomes, and possibly save time.[19] Many excellent questionnaires are used in research; few are actually used in practice. Simple protocols (for example, patients completing a self-scoring questionnaire) do not do enough to influence outcomes in a multistep care process. Complex, idealized protocols might influence outcomes but are costly to deploy and not logistically feasible outside of a research framework. The failure to make effective use of questionnaires speaks to practical limits in a paper-based environment and how this tempers what is sensible to imagine.

Several years ago we began to examine how questionnaires could be used, specifically focusing on our Pediatric Neurodevelopmental Clinic, for the diagnosis and management of disorders such as autism and language delay. In this project, parents completed a seven-page questionnaire before the encounter, with a specific focus on rudimentary questions—which did not require a physician—on eight domains of behavior. Questionnaire data were digitized and imported to EHR templates. The challenge was in designing a process that would allow the physician to rapidly interpret parent-reported data during a clinical encounter and build on these data with standardized structured and semistructured probes. Finally, access to structured patient and physician data was used to generate highly tailored after-visit summaries, including educational material. It is impossible to develop a workable model like this in a paper-based world.

Although the model we created was narrowly focused, the lessons learned are generalizable to clinical settings where behavioral or symptom-based conditions are commonplace. Questionnaires are more valuable in an EHR environment than they can ever be in a paper-based environment. Some values (such as access to structured data from patient) are common to both environments; however, most are unique to an EHR environment (efficiency, seamless use of data and links to structured probes, tracking outcomes over time, and tailored patient education). The EHR provides the means to make effective use of questionnaire data in creating a practical workflow model (that is, meaningful display of data and higher-level probing), solving a fundamental barrier in paper-based clinics. The questionnaire is part of and motivates thinking about a sustainable systems-based approach to new health services models—ingredients that are essential to exportability and scalability of new solutions.

■ **Decision support: CPGs in real time and more.** To facilitate access to clinical knowledge, efforts have been under way over the past decade to codify what we know through the development of clinical practice guidelines (CPGs). This activity is essential to but insufficient for translating knowledge into practice. Sophisticated clinical decision-support technology in combination with EHRs will be required to make effective use of CPGs.

CPGs, based on explicit methods for summarizing established evidence with

"Patient data themselves will become a critical practice asset, motivating the need to generate high-quality, complete data."

expert clinical consensus, have expanded at an accelerating rate.[20] Today there are 1,970 active CPGs listed in the U.S. National Guideline Clearinghouse.[21] This codification of knowledge is essential to increasing its usefulness. However, CPGs represent only one step in a complex process to translate knowledge into practice.

For a number of reasons, CPGs have had only a modest impact on care. First, relatively primitive means such as publication and education are used to promulgate them. Second, clinicians do not have the time to begin to learn even a fraction of them. Third, CPGs rarely attempt to describe how to operationalize the recommended tasks (that is, CPGs represent knowledge but not accountable, manageable work flows). Even if complete, relevant, perfect, and codified knowledge existed today, there is no effective means of accessing such knowledge in real time even with existing EHRs. What we are lacking is a decision-support capability that can assimilate detailed relevant information about the patient, evaluate such data in real time against existing knowledge, and then yield recommendations that the physician can act upon.

We recently completed a pilot project to determine how such a process could be created with a specific focus on cardiovascular risk management in primary care. The process itself led to specific protocols for ordering measures to fully determine a patient's cardiovascular risk, including questionnaire data on behavioral risk factors. A decision-support rules "engine," external to the EHR, was used to extract patient data in real time, evaluate the data in relation to rules, and generate and return a recommended order. This project demonstrated that it is technically possible to create a real-time decision-support workflow that translates CPGs into practice. It also revealed a conceptual and practical challenge: Translation of all CPGs into practice will result in too much care. CPGs often do not recognize the confounding aspects and logistic complexities of comorbid conditions. For example, the blind application of all relevant CPGs to a typical hypothetical elderly patient would lead to twelve prescribed medications, costs of $400 per month, and numerous potential side effects.[22] Codified knowledge does not exist that can guide decision priorities among diverse sets of CPGs.

■ **Determining optimal management protocol: clinically ranked data.** Finally, we expect advances in the use of retrospective data analysis to influence clinical decision support. The real-time use of longitudinal EHR data to guide clinical decision making will be a logical extension of retrospective data analysis but will represent a conceptual leap—well beyond traditional notions of evidence and decision support. In this framework, patient data themselves will become a critical practice asset, motivating the need to generate high-quality, complete data. For example, consider the needs of a patient with hypertension, diabetes, and atrial fibril-

lation. Different classes of antihypertensive medication might be recommended if each disease were clinically considered in isolation: thiazide for hypertension, angiotensin-receptor blocker (ARB) for diabetes, and a beta-blocker for atrial fibrillation. Traditional efficacy evidence cannot provide explicit guidance on a decision that seeks to simultaneously optimize the relative benefits of each medication, along with ease of the regimen, formulary coverage, interactions with other medications, importance of side effects, and patient frailty.[23] A logical extension of retrospective analysis of EHR data will be the real-time, rapid processing of longitudinal, population-based EHR data to determine the optimal management protocol given a patient's overall profile and individual preferences. In the example above, the EHR might simplify decision making by presenting data in clinically relevant rankings, taking into account known factors from earlier experience with the patient and others similar patients. For example, a patient's strong preference for generic medications and lower copayments (even at the expense of more frequent dosing) might lead the EHR's ranking of medications to list twice-daily generic medications over once-daily brand-name medications.

■ **Randomizing decision uncertainty: randomized database studies.** The inferential gap in medicine will continue to be an everyday occurrence where the needs and questions asked are always changing and evolving. An everyday solution will be required to meet the perpetually growing demand for new knowledge in medicine. One solution might be to use the power of the EHR to randomize clinical decisions in the face of uncertainty and to evaluate the outcomes accordingly.[24]

In the patient scenario described above, there will inevitably and frequently be true clinical equipoise between two or more potential decisions. In such a situation, a real-time protocol might be embedded within the system to randomly prioritize one decision pathway or another. Over time, the knowledge created by such mini-RCTs could shed light on many questions that could not be addressed by traditional randomized trials or observational studies.

Concluding Comments

In this paper we have characterized a notion labeled the "inferential gap" and considered the future role of the EHR in closing this gap. In part, the gap is the product of knowledge being created at a faster rate than we can use it, and, importantly, clinical questions growing at a faster rate than can be answered through traditional research methods. We recognize that numerous challenges exist in the widespread application and effective use of EHRs. We expect that these challenges will be minimized as the technology, the data, and their application in practice evolve. From our own recent experience, it is clear that an EHR-based practice environment engenders an unavoidable shift in thinking about clinical evidence and how to create and use it and, importantly, a loss of distinction between clinical practice, quality management, and the creation of knowledge.

The authors appreciate the input of Marc Gourevitch, who commented on earlier drafts of this paper.

NOTES

1. T. Bodenheimer, E.H. Wagner, and K. Grumbach, "Improving Primary Care for Patients with Chronic Illness," *Journal of the American Medical Association* 288, no. 14 (2002): 1775–1779.
2. R.B. Haynes, R.S. Hayward, and J. Lomas, "Bridges between Health Care Research Evidence and Clinical Practice," *Journal of the American Medical Informatics Association* 2, no. 6 (1995): 342–350.
3. R.B. Haynes, "Using Informatics Principles and Tools to Harness Research Evidence for Patient Care: Evidence-Based Informatics," *Medinfo* 9, Part 1 Supp. (1998): 33–36.
4. K.J. Rothman and S. Greenland, *Modern Epidemiology* (Philadelphia: Lippincott Williams and Wilkins, 1998).
5. A. Flexner, "Medical Education in the United States and Canada: From the Carnegie Foundation for the Advancement of Teaching, Bulletin Number Four, 1910," *Bulletin of the World Health Organization* 80, no. 7 (2002): 594–602; and J.P. Vandenbroucke, "A Short Note on the History of the Randomized Controlled Trial," *Journal of Chronic Disease* 40, no. 10 (1987): 985–987.
6. N. Black, "Why We Need Observational Studies to Evaluate the Effectiveness of Health Care," *British Medical Journal* 312, no. 7040 (1996): 1215–1218.
7. K.M. Mattocks and R.I. Horwitz, "Placebos, Active Control Groups, and the Unpredictability Paradox," *Biological Psychiatry* 47, no. 8 (2000): 693–698.
8. J.M. Drazen, "Cox-2 Inhibitors—A Lesson in Unexpected Problems," *New England Journal of Medicine* 352, no. 11 (2005): 1131–1132.
9. S. MacMahon et al., "Blood Pressure, Stroke, and Coronary Heart Disease, Part 1: Prolonged Differences in Blood Pressure: Prospective Observational Studies Corrected for the Regression Dilution Bias," *Lancet* 335, no. 8692 (1990): 765–774.
10. S.H. Greenblatt, "Limits of Knowledge and Knowledge of Limits: An Essay on Clinical Judgment," *Journal of Medical Philosophy* 5, no. 1 (1980): 22–29.
11. R.I. Horwitz et al., "Developing Improved Observational Methods for Evaluating Therapeutic Effectiveness," *American Journal of Medicine* 89, no. 5 (1990): 630–638.
12. Greenblatt, "Limits of Knowledge and Knowledge of Limits."
13. M. Salas, A. Hofman, and B.H. Stricker, "Confounding by Indication: An Example of Variation in the Use of Epidemiologic Terminology," *American Journal of Epidemiology* 149, no. 11 (1999): 981–983.
14. P.J. Neumann, "Emerging Lessons from the Drug Effectiveness Review Project," *Health Affairs* 25 (2006): w262–w271 (published online 6 June 2006; 10.1377/hlthaff.25.w262).
15. Ibid.
16. F. Forette et al., "Does the Benefit of Antihypertensive Treatment Outweigh the Risk in Very Elderly Hypertensive Patients?" *Journal of Hypertension* 18, no. 3 Supp. (2000): S9–S12.
17. U.O. Andersen, J.H. Henriksen, and G. Jensen, "Sources of Measurement Variation in Blood Pressure in Large-Scale Epidemiological Surveys with Follow-up," *Blood Pressure* 11, no. 6 (2002): 357–365.
18. J.D. Seeger, P.L. Williams, and A.M. Walker, "An Application of Propensity Score Matching using Claims Data," *Pharmacoepidemiology and Drug Safety* 14, no. 7 (2005): 465–476; and N.R. Shah, "What Is the Best Evidence for Making Clinical Decisions?" *Journal of the American Medical Association* 284, no. 24 (2000): 3127–3128.
19. S.M. Downs et al., "Human and System Errors, Using Adaptive Turnaround Documents to Capture Data in a Busy Practice," *American Medical Informatics Association Annual Symposium Proceedings* (2005): 211–215.
20. D.A. Grimes et al., "The Good Clinical Practice Guideline: A Bronze Standard for Clinical Research," *Lancet* 366, no. 9480 (2005): 172–174.
21. See the National Guideline Clearinghouse home page, http://www.guideline.gov.
22. C.M. Boyd et al., "Clinical Practice Guidelines and Quality of Care for Older Patients with Multiple Comorbid Diseases: Implications for Pay for Performance," *Journal of the American Medical Association* 294, no. 6 (2005): 716–724.
23. M.V. Maciosek et al., "Priorities among Effective Clinical Preventive Services Methods," *American Journal of Preventive Medicine* 31, no. 1 (2006): 90–96.
24. J.A. Sacristan et al., "Randomized Database Studies: A New Method to Assess Drugs' Effectiveness?" *Journal of Clinical Epidemiology* 51, no. 9 (1998): 713–715.

PERSPECTIVE

Information Technology And The Inferential Gap

Only by using new tools can we greatly expand the scientific basis of everyday clinical medicine.

by Jonathan B. Perlin and Joel Kupersmith

ABSTRACT: In medical practice, an "inferential gap" exists in two contexts: the nonapplication of relevant existing evidence, and the absence of evidence germane to a particular clinical situation. Randomized controlled trials are the current gold standard of evidence development; however, they suffer limits of generalization to the "real world." Conversely, observational studies might be more generalizable but are prone to bias, data inconsistency, and measurement error. The electronic health record offers hope for supporting the real-time presentation of information relevant to a clinical situation and serves as a platform for the conduct of large observational studies and novel quasi-experimental research. [*Health Affairs* 26, no. 2 (2007): w192–w194 (published online 26 January 2007; 10.1377/hlthaff.26.2.w192)]

IN THIS AGE OF SCIENCE it is perhaps surprising how often clinical decisions are made absent the application of existing knowledge and how often clinical judgment or intuition is required when no knowledge exists. Walter Stewart and colleagues examine an "inferential gap" that exists in two contexts: the nonapplication of relevant existing evidence and the absence of evidence germane to a particular clinical situation.[1]

■ **Different types of trials.** As to the evidence base for clinical practice, the explanatory randomized controlled trial (RCT) is the fundamental methodology. However, the limitations of RCTs—in particular, those related to their limited environment (carefully selected patients and physicians)—are well known. RCTs are basically designed to determine whether a new treatment can work. They might have limited generalizability of results to the average practice setting and insufficient data for many circumstances, especially when there is clinical complexity.[2]

Broadening of explanatory RCTs could be achieved in part by the inclusion of health care utilization and cost data and (at times faulty) subgroup analysis. In addition, by grouping studies and, it is hoped, achieving broad-spectrum populations, the technique of meta-analysis could make the RCT more generalizable as well as giving it greater weight of numbers. However, each RCT in meta-analysis might be similarly narrow, with the sum still not sufficiently representing a broad population.

Practical controlled trials (PCTs) compare effectiveness of treatments with somewhat different methodology than explanatory RCTs, in more representative and routine practice settings and populations, and they aim at clinically relevant dilemmas.[3] They have more generalizable results and are closer to

Jonathan Perlin (Jonathan.Perlin@HCAHealthcare.com) is chief medical officer and senior vice president, quality, at HCA Healthcare in Nashville, Tennessee. Joel Kupersmith is chief research and development officer at the Department of Veterans Affairs in Washington, D.C.

DOI 10.1377/hlthaff.26.2.w192

capturing care in the various settings in which it is ordinarily given. However, they share some challenges with RCTs such as time periods too short for some illnesses, difficulties with nonquantifiable data, and possible obsolescence during the study period.

Apart from methodological concerns, both RCTs and PCTs can be costly. To greatly narrow the inferential gap by answering many clinical questions or comparing therapies would be nearly impossible with available resources.

A variety of methodologies have been promulgated to make trials either more efficient or more generalizable while attempting not to sacrifice rigor. These include adaptive designs (decision-based, risk-based) using Bayesian techniques and elaborations such as clinician choice (randomization to a circumscribed selection of choices by clinicians), equipoise-stratified design, and cross-design synthesis.[4] Notably, although much methodological effort has been applied to the evidence base of evaluating treatments, much less has been applied to that of making diagnoses.

■ **Observational studies.** At what might be considered the other end of the scientific spectrum to RCTs are observational studies. These have generally been considered less scientifically rigorous depending on the specific approach and the specific observer. As Stewart and colleagues correctly identify, observational studies are beset by issues related to bias, measurement error, completeness of data, and so on. However, they have certain advantages over RCTs. They are "naturalistic" experiments in routine clinical settings rather than in settings that have been artificially created and are subject to distortion.[5] When available electronic health record (EHR) databases are used, observational studies are much more economical than entering patients in a trial; they do not depend on a future timeline, thus eliminating the issue of obsolescence; and if they comprise a spectrum of the population, the data are generalizable.

There are many sources of ambiguity in clinical practice, and the information tools available to clinicians—pen and paper 83 percent of the time, according to a recent study by David Blumenthal—provide little help in associating relevant evidence from the literature with clinical practice.[6] The relative absence of EHRs; standard definitions of electronic data elements; interoperability among EHR systems; and systematic organization of data into large, relational data sets make the systematic creation of new knowledge as a natural by-product of patient care difficult.

Some examples of analytic capability for large observational studies are emerging in health systems, such as the Geisinger environment that Stewart and colleagues describe or the U.S. Veterans Health Administration (VHA), where the health data of more than 5.3 million active patients are potentially available for retrospective analysis. As Stewart and colleagues note, the use of such data is potentially restricted by legal, ethical, regulatory, and cultural barriers as well as the described biases. Nevertheless, provisions in these environments do allow aggregated, retrospective analysis, and insights into medical care and health services delivery can be stunning. For example, using the EHR data to trend blood pressures at a VHA hospital for quality improvement purposes identified a novel finding of seasonal variation in blood pressure control among more than 10,000 hypertensive patients.[7]

■ **Other considerations.** It is not surprising to learn that large data sets generated by EHRs would be useful for retrospective studies. Stewart and colleagues describe the variety of challenges presented by the inconsistency of clinical context and definitions that might introduce bias and uncertainty into interpretation. The novel prospect of EHRs in generating a real-time random choice among treatment alternatives under consideration is a tantalizing application for advancing the rapidity of the conduct of RCTs.

There is growing discussion of novel applications that combine the best features of observational studies and RCTs. Methodological strategies involving cluster randomization of groups of patients under one or another rubric within the basic principle of equipoise (uncer-

tainty as to equivalent effectiveness between groups decipherable with adequate evidence) have been advocated for this purpose. For example, might algorithms be applied for retrospective randomization to create a retrospective pseudo-randomized trial? And might the results of such a study be "served" to a clinician choosing between therapeutic alternatives in the real-time context of patient care? Might there also be real-time guidance that supports the probabilistic reasoning necessary to surmise how close the actual patient is to the "patterned patient composite"? How will this learning and analytical engine support the clinician's and patient's risk tolerance, understanding that a "weak" inference might be hopeful in a critical situation, yet dangerous in routine care?

■ **Bringing health care up to speed.** Health care is behind other industries in knowledge generation from complex data inputs. The annoying Internet "spyware" phenomenon persists because very accurate demographic data can be inferred and applied for directed marketing. Someday, inferential engines might support more real-time delivery of evidence-based health services. As machine learning evolves, the clinician's and researcher's work could change from mining a data set with a hypothesis to evaluating the biological plausibility and the clinical relevance of a hypothesis that the computer generates.

Finally, we are in a state of flux as to the expansion of the base of scientific medicine. Although understandably cautious in rushing toward acceptance of new methodology, the clinical, scientific, and regulatory communities are working on these issues. There is a need for applicability of research results to a broader spectra of patients, including those with comorbidities and long-term illnesses, and the ability to address less easily quantifiable situations and study such things as care models—particularly models providing greater research efficiency.

MANY OBSTACLES EXIST between the creation of these incidental data and making them relevant to practice. However, it is clear that the ancient tools of paper-based health care delivery are a fundamental barrier to closing the inferential gap by supporting both access to relevant evidence and the generation of new evidence. Only by using the new tools can we greatly expand the scientific basis of everyday clinical medicine.

The views presented in this paper are those of the authors and do not necessarily represent the views of the Department of Veterans Affairs.

NOTES

1. W.E. Stewart et al., "Bridging the Inferential Gap: The Electronic Health Record and Clinical Evidence," *Health Affairs* 26, no. 2 (2007): w181–w191 (published online 26 January 2007; 10.1377/hlthaff.26.2.w181).
2. R.M. Califf and D.L. DeMets, "Principles from Clinical Trials Relevant to Clinical Practice: Part I," *Circulation* 106, no. 8 (2002): 1015–1021.
3. S.R. Tunis, D.B. Stryer, and C.M. Clancy, "Practical Clinical Trials: Increasing the Value of Clinical Research for Decision Making in Clinical and Health Policy," *Journal of the American Medical Association* 290, no. 12 (2003): 1624–1632.
4. P.W. Lavori et al., "Strengthening Clinical Effectiveness Trials: Equipoise-Stratified Randomization," *Biological Psychiatry* 50, no. 10 (2001): 792–801; and J.A. Sacristan et al., "Randomized Database Studies: A New Method to Assess Drugs' Effectiveness?" *Journal of Clinical Epidemiology* 51, no. 9 (1998): 713–715.
5. Sacristan et al., "Randomized Database Studies."
6. D. Blumenthal, "Adoption of EHRs: Where Are We, Where Are We Going, How Can We Know?" p. 44. Testimony to the American Health Information Community, U.S. Department of Health and Human Services, 12 September 2006, http://www.hhs.gov/healthit/documents/Presentations091206.pdf (accessed 4 December 2006).
7. R.D. Fletcher et al., "The Effect of Seasonal Variation of Blood Pressure when Analyzing Improvement and Treating Large Groups of Hypertensive Patients," *Journal of the American College of Cardiology* 43, no. 5 (2004): A:449.