Cancer Causes Control (2006) 17:861–869
DOI 10.1007/s10552-006-0028-4

REVIEW ARTICLE

The public health informatics infrastructure: anticipating its role in cancer

Edward H. Shortliffe · Edward J. Sondik

Received: 16 August 2005 / Accepted: 27 March 2006

Abstract Cancer information and surveillance, historically conducted with manual data collection and submission, are viewed increasingly as inherently dependent on the effective application of information science. One challenge is to use information technology (IT) in a manner that improves cancer-related decision-making and ultimately the quality of care that is offered to patients with cancer. In this article we begin by envisioning a future view of IT-supported surveillance and care that can be made available for application in cancer and its management. We then ask what barriers need to be overcome and what forces are at work that may help us in our efforts to effect the necessary changes. Our future vision for surveillance and information, although appealing and widely shared, requires major cultural change, financial investment, and logistical planning. Competition in the medical marketplace, coupled with fiscal pressures affecting providers and health systems, suggests that leadership for regional and national coordination will need to come from elsewhere—and likely from governments.

E. H. Shortliffe (✉)
Department of Biomedical Informatics, Columbia University, New York, NY, USA
e-mail: shortliffe@dbmi.columbia.edu

E. J. Sondik
National Center for Health Statistics, Hyattsville, MD, USA

E. J. Sondik
Centre for Public Health Informatics, Centers for Disease Control and Prevention, Atlanta, GA, USA

Keywords Information technology · Health information infrastructure · Cancer care and surveillance · Electronic medical records · Acceptance of computing technology

Introduction

Cancer information and surveillance, historically conducted with manual data collection and submission, are increasingly viewed as being inherently dependent on the effective application of information science. One of our challenges is to use information technology (IT) in a manner that improves cancer-related decision-making and ultimately the quality of care that is offered to patients with cancer. There is no shortage of ideas on how to use technology in ways that will improve medical practice, provide more information to the patients, and enhance the quality of the information and guidance that is available. Our challenge for biomedical informatics in general, and for its application to cancer care and surveillance in particular, is to leverage the evolving technology and communications infrastructure in a way that is cost-effective, that supports research, and that recognizes and encourages the development of standards and of the cultural changes that will be required. Our specific challenge in cancer care and surveillance is to use our knowledge, and the new and current information sources that will surely be available, more effectively to identify cancer occurrence, risk factors, patterns of care and care quality, and quality of life outcomes for cancer patients and their families.

The existing clinical culture reveals a variety of attitudes regarding the role that IT can and should play in patient care. Some practitioners are encouraged by progress in clinical computing and believe that technology can and will enhance both the efficiency with which they practice and the quality of the care that they deliver. Others are concerned that the technology may adversely affect the relationship between physicians and their patients, potentially dehumanizing the care process and encouraging impersonal approaches to the deep human problems that emerge in the context of disease and its management [1]. Others worry about whether they can personally develop the skills and attitudes necessary to utilize IT effectively. They may even worry about how they will be judged by their patients, or potential patients, based on either their use of, or failure to use, IT in their practices.

It would be misleading to suggest that the world of medicine has been unaffected by the IT revolution that has been touching all other aspects of society in the last few decades. We know, for example, that health information is among the most frequently sought search categories on the World Wide Web,[1] and consumers have been increasingly empowered to participate actively in both treatment and prevention. Similarly, it is a rare physician who does not use computer systems to obtain laboratory or radiology results for patients; increasingly, physicians also are using computers to submit prescriptions or to access drug information. The typical modern hospital is filled with computers. The devices are evident on every nursing unit, in outpatient clinics, and throughout specialty units such as oncology, cardiology, and radiology.

With the President of the United States calling for electronic medical records (EMRs) for all Americans within a decade, and a newly created Office of the National Coordinator for Health Information Technology (ONCHIT) in the Office of the Secretary of Health and Human Services, the conclusion we can draw is that the potential for a ''revolutionary'' change in how medicine is practiced is at hand. Those of us who have worked with health care IT are pleased by the recent attention that has been directed at this topic, both within government and in the private sector. The unfulfilled promise of IT in support of health and health care has been clear to some of us for many years, and those in the field often have been dismayed to see a widening gap between the implementation of IT solutions to pressing problems in other segments of society contrasted with their limited penetration into health care settings. On the other hand, a variety of factors recently have combined to heighten our awareness of what is possible and of the need for active intervention and promotion of solutions. We need to place such optimism in context by assessing current capabilities as well as what still remains to be realized, while asking what barriers exist that have prevented optimal progress to date. In this article we begin by envisioning a future view of IT-supported clinical data collection and care. We then ask what barriers need to be overcome and what forces are at work that may help us in our efforts to effect the necessary changes.

[1] Sources include, for example, prevention guidelines from the CDC including recommended immunization schedules for children; the Preventive Services Task Force guidelines; American Cancer Society, National Cancer Institute; the US Department of Health and Human Services Healthy People objectives and many more. Medline Plus from the National Library of Medicine is another popular resource intended for use by the public.

A model of integrated cancer care and surveillance

We start by envisioning one model of how cancer surveillance, prevention, and care could be influenced by information and communications technology a decade or so from now. The model we present below is in no way specific to cancer data collection and clinical care; it is intended to underlie the way in which *all* care could be managed and measured in the future. The underlying technology for this vision, the ''informatics infrastructure,'' is derived from work in the field of biomedical informatics, the scientific discipline that deals with the storage, retrieval, sharing, and optimal use of biomedical information, data, and knowledge for problem solving and decision-making. Inherently interdisciplinary, biomedical informatics in turn builds on component disciplines such as computer science, decision science, cognitive science, and information science.

Biomedical informatics touches on all basic and applied fields in biomedical science and is closely tied to modern information technologies, notably in the areas of computing and communication. All work in the field is motivated by biomedical applications, ranging from basic biological work (*bioinformatics*) and work with tissues and organs (*structural informatics* or *imaging informatics*) to systems that address the needs of whole organisms (*clinical informatics*) or populations of individuals (*public health informatics*) (see Chapter 1 in [2]).

Over the past 30 years or so, work on the key elements of our future vision has emerged from the informatics research and practice communities as well as the world of computer science and

telecommunications. Key components of the relevant infrastructure include networking technologies, clinical databases, EMRs, and a wide variety of other general and biomedical elements that can be integrated to enhance oncology practice and cancer surveillance.

Imagine, then, the day when *all* providers, regardless of practice setting (hospitals, emergency rooms, small offices, community clinics, military bases, multispecialty groups, etc.) use EMRs in their medical practices to assist both with patient care and to provide patients with counsel on illness prevention. The full impact of this use of electronic resources will occur when data from all such records are pooled in regional and national surveillance databases (Fig. 1), mediated through connectivity with the Internet. The challenge, of course, is to find a way to integrate data from such diverse practice settings, especially since it is inevitable that multiple vendors and system developers will be active in the marketplace, competing to provide value-added capabilities that will excite and attract the practitioners for whom their EMR product is intended.

The practical need to pool and integrate clinical data from such diverse resources and systems emphasizes the practical issues that must be addressed if this vision is to be achieved. Interestingly, most of the potential barriers are logistical, political, and financial rather than technical in nature:

- *Encryption of data*: Concerns regarding privacy and data protection require that Internet transmission of clinical information occur only if those data are encrypted, with an established mechanism for identifying and authenticating individuals before they are allowed to decrypt the information for surveillance or research use.

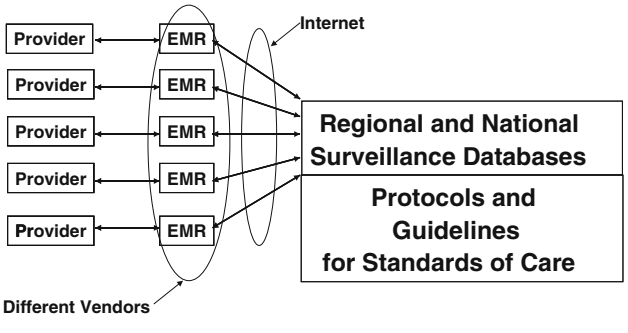

Fig. 1 A future vision of surveillance databases, in which clinical data are pooled in regional and national repositories through a process of data submission that occurs over the Internet (with attention to privacy and security concerns as discussed in the text). When information is effectively gathered, pooled, and analyzed, there are significant opportunities for feeding back the results of derived insights to practitioners at the point of care

- *HIPAA-compliant policies*: The privacy and security rules that resulted from the 1996 Health Insurance Portability and Accountability Act (HIPAA) do not prohibit the pooling and use of such data, but they do lay down policy rules and technical security practices that must be part of the solution in achieving the vision we propose.
- *Standards for data transmission and sharing*: Sharing data over networks requires that all developers of EMRs and clinical databases adopt a single set of standards for communicating and sharing information. The de facto standard for such sharing, Health Level 7 (HL7) is widely used but still not uniformly adopted, implemented, or utilized.
- *Standards for data definitions*: A uniform ''envelope'' for digital communication, such as HL7, does not assure that the *contents* of such messages will be understood or standardized. The pooling and integration of data requires the adoption of standards for clinical terminology and for the schemas used to store clinical information in databases.
- *Quality control and error checking*: Any system for accumulating, analyzing, and utilizing clinical data from diverse sources must be complemented by a rigorous approach to quality control and error checking. It is crucial that users have faith in the accuracy and comprehensiveness of the data that are collected in such repositories, because policies, guidelines, and a variety of metrics can be derived over time from such information.
- *Regional and national surveillance databases*: Any adoption of the model in Fig. 1 will require mechanisms for creating, funding, and maintaining the regional and national databases that are involved. Only if such records are truly ubiquitous will the data that are gathered be able to provide rigorous surveillance information, informed not only by events but by a realistic assessment of their frequency in the broadest possible populations. The role of state and Federal governments will need to be clarified, and the political issues addressed (including the concerns of some members of the populace that any government role in managing or analyzing their health data may have societal repercussions that threaten individual liberties, employability, and the like).

With the establishment of such databases, and a robust system of Internet integration with EMRs, summary information can flow back to providers to enhance their decision-making at the point of care

(Fig. 1). This assumes standards that allow such information to be integrated into the vendor-supplied products that the clinicians use in their practice settings. These may be EMRs or, increasingly, *order-entry systems* that clinicians use to specify the actions that they want to have taken for the treatment or management of their patients. Furthermore, as is shown in Fig. 1, the databases can help to support the creation of evidence-based guidelines, or clinical research protocols, which can be delivered to practitioners through the feedback process. Thus we envision a day in which clinicians, at the point of care, will receive integrated, non-dogmatic, supportive information regarding:

- Recommended steps for health promotion and disease prevention;
- Detection of syndromes or problems, either in their community or more widely;
- Trends and patterns of public health importance;
- Clinical guidelines, adapted for execution and integration into patient-specific decision support rather than simply provided as text documents;
- Opportunities for distributed (community-based) clinical research, whereby patients are enrolled in clinical trials and protocol guidelines are in turn integrated with the clinicians' EMR to support protocol-compliant management of enrolled patients.

Implementing the National Health information infrastructure

Large provider organizations, including hospitals and distributed health systems, routinely use networking technology as the infrastructure on which they build their computer-based communications channels (Fig. 2). With departmental computer systems (e.g., radiology, clinical lab, microbiology, and pharmacy) connected to the network, institutions generally collect and store data in a central clinical data repository. Over time, as this repository becomes more and more comprehensive, it effectively becomes an EMR. Clinicians access the patient data in such repositories using a variety of methods, ranging from tethered workstations installed in offices or nursing stations to handheld wireless devices such as PDAs or tablet computers. Clerical staff members use the same network to enter and access information, and sometimes patients are invited to enter their histories, to access educational materials, or even to review their personal clinical data over such networks. Data may be submitted to research databases, and the users of the network typically have access to library resources or to administrative or financial systems. The integration of such resources within an organization depends on a robust ''enterprise intranet.'' The implementation and maintenance of an advanced network is one of the fiscal and organizational challenges faced by complex provider institutions.

In the outpatient setting, both small and large networks are becoming commonplace (Fig. 3). Within an ambulatory practice, physicians and other personnel may have several computers networked together and sharing data from an EMR system. The full utility of the system depends on gateways from these local networks into the Internet because that is where the patients and business associates (such as pharmacies and clinical laboratories) increasingly access and provide information. Several EMR products provide specialized web interfaces so that patients can access their physician's practice for purposes ranging from appointment scheduling to review of laboratory results and drug lists.

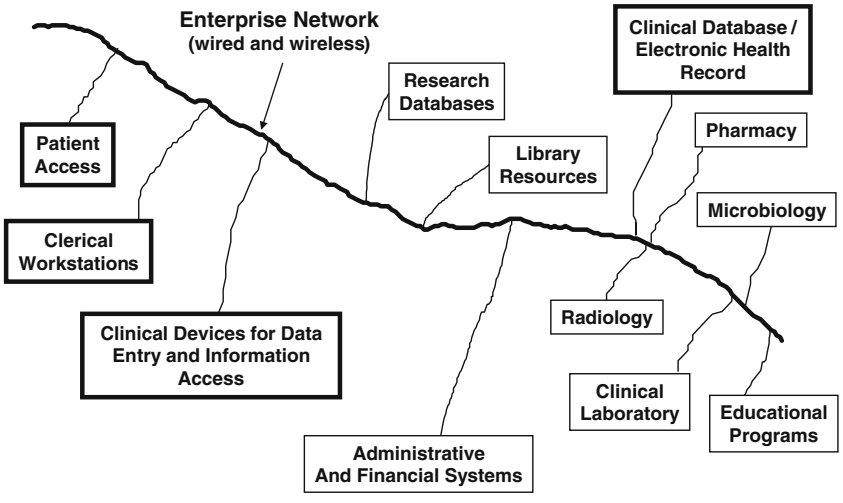

Fig. 2 We already live in an era when large hospitals and health care systems have implemented widespread networking technologies that allow diverse systems and users to communicate with one another within their organization. The resulting *enterprise Intranet* faces challenges of connectivity and integration that are a microcosm of what the larger community experiences in trying to link EMRs and other clinical systems from different organizations

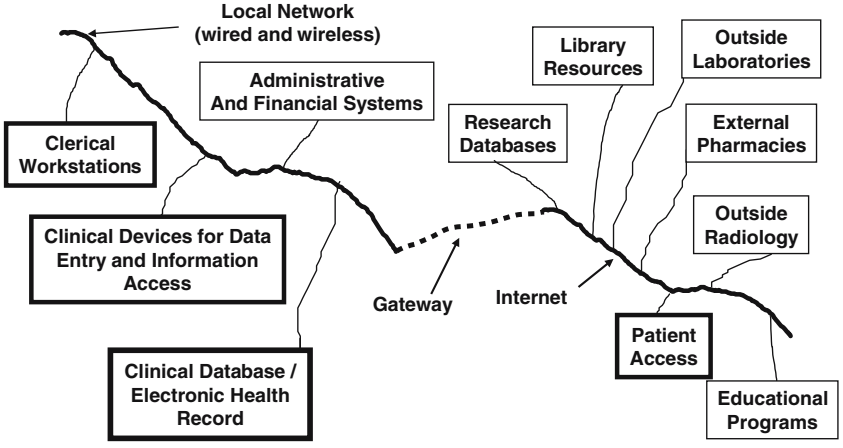

Fig. 3 Communications networks are increasingly found in outpatient practice settings, including small office practices, but much of their value is enhanced when they are linked through gateways to the Internet and to information resources, organizations, and individuals beyond their own doors

If we return to the future vision of Fig. 1, it should be clear that the pooled databases that we wish to build depend on the submission of data over the Internet from clinical databases that reside in large organizations (Fig. 2) and outpatient practice settings (Fig. 3). Furthermore, the delivery of information to these settings depends on an infrastructure that supports the integration of decision-support elements with the records and order-entry systems used in these same practice environments. This leads to our major claim: *We must tap into these clinical data, as a byproduct of routine patient encounters, if we want to create shared research and surveillance databases* (Fig. 1). If the submission of data for research or monitoring purposes requires an extra step, or special effort by busy clinicians, the process likely will fail, regardless of the good intentions of practitioners. Furthermore, this extra step should not be necessary. We can build integrated systems built on standards that allow automated data submission and collection via the Internet in a secure, responsible, and confidential manner.

Thus we recognize that the vision laid out in Figures 1 and 2, and its eventual application in the cancer domain, depends on the creation of a National Health Information Infrastructure (Fig. 4) that links all practices and practitioners in the country, offering them value in terms of access to information, decision support when desired, communication channels with patients and colleagues, and even support for their business operations (e.g., by online submission of invoices to payers, which carries the potential for error checking and real-time verification that will, in turn, greatly shorten the payment cycle for accounts receivables). The idealistic model we envision addresses a large number of the serious problems facing our health care system, ranging from error prevention and reduction in practice variation to reduced administrative costs and enhanced efficiency. The public health system, including cancer surveillance, will be only one of the many beneficiaries of such a transition.

The cycle of information flow in clinical care

The concepts outlined above lead to a composite model of cyclical information flow in the future, as is

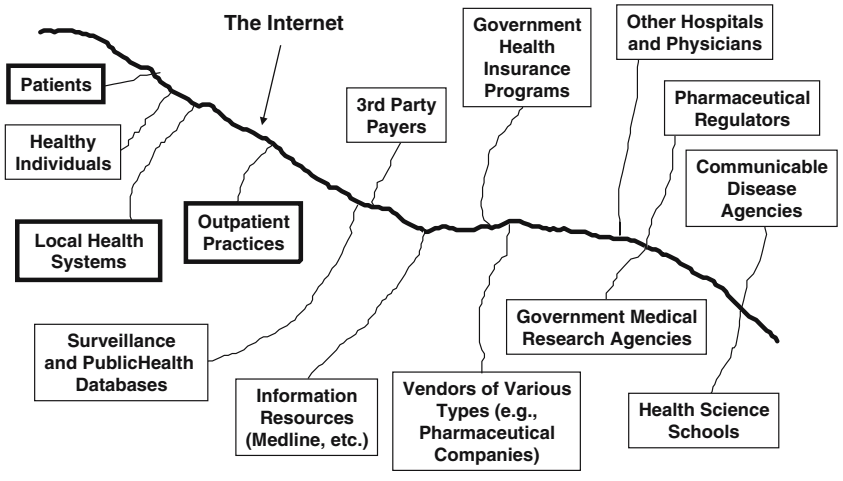

Fig. 4 The integrated interconnectivity of all the clinical systems, building on networking technology and standards for data exchange and privacy protection, creates a National Health Information Infrastructure (NHII) that supports clinical care, research, and the public health

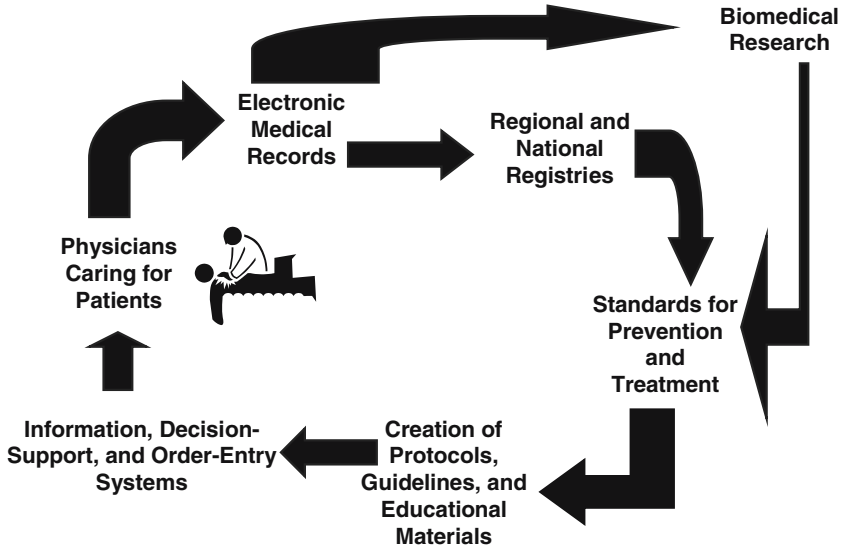

Fig. 5 The ultimate goal is to create a cycle of information flow, whereby data from distributed EMRs are automatically submitted to registries and research databases. The resulting new knowledge then can feed back to practitioners at the point of care, using a variety of computer-supported decision-support delivery mechanisms

shown in Fig. 5. Beginning at the left of the diagram, physicians caring for patients use EMRs. Information from these records will be forwarded automatically to regional and national registries as well as to research databases (if the patient is enrolled in a community-based clinical trial). The information can be used to develop standards for prevention and treatment, with major guidance from biomedical research. Researchers can draw information either directly from the medical records or from the pooled data in registries. The standards for treatment in turn will be translated into protocols, guidelines and educational materials. This new knowledge and decision-support functionality will be delivered via the NHII back to the clinicians so that the information informs the practice of medicine at the point of care, where it is integrated seamlessly with EMRs and order-entry systems.

We believe this future view is logical and ultimately feasible. The irony, however, is that it is a not new vision. The Internet has existed for over 30 years, and the diagram in Fig. 5 could have been drawn by many of us at least 15 years ago, as the Internet revolution began to expand. What, then, have been the barriers to achieving this vision?

Barriers to effective use of information technology in health care[2]

As was mentioned earlier, the barriers to successful implementation of integrated IT support for cancer care and surveillance are no longer primarily technical in nature. With the maturity of networking technology and methods for secure management of data systems in a networked world, introspection and observation reveals that the principal challenges lie in other areas.

The individual physicians who practice in this country (including both oncologists, and those who refer their patients to them) are an important element in any solution that we propose, but their ability to participate effectively is highly constrained. A variety of important issues need to be considered before formulating any incentive programs or implementation plans for health care IT in cancer.

The vast majority of health care in this country is provided by physicians in ambulatory settings, and most commonly in relatively small offices. Our view of what is needed cannot be overly skewed by the perspectives of those who practice in large, multispecialty practices or in clinics associated with academic medical centers. Although well implemented IT in a single institution can provide major quality and cost benefits for that entity, it is in the integrated penetration of health care IT throughout essentially all practice settings that our ability to carry out effective surveillance stands to gain the most. This means creating an infrastructure, both regional and national, that ties into all practice settings, and that helps individual practices to make wise decisions and investments.

Viewed from the perspective of a clinician in a small office, the issues are overwhelming in many respects. It is too easy to say that physicians are simply resistant to change or overly committed to antiquated approaches to data management. We see many examples, in fact, where clinicians have embraced new technologies rather quickly. But IT presents some special problems for

[2] An expanded discussion of these issues appears in Shortliffe, EH. Strategic action in health information technology: Why the obvious has taken so long. Health Affairs 2005;24:1222–1233.

practitioners. It is not their area of expertise, and they are uncertain how to evaluate the options that are provided to them. It is not a part of their education, and seems foreign to the major thrusts of their professional interests. System implementations are often disruptive to office operations, at least in transition, and too often physicians find that major investments have resulted in inadequate systems solutions that fail to meet expectations, integrate poorly with other systems, or are difficult to adapt to the special needs of a particular practice.

We have also seen that IT is sometimes viewed as a distraction from an organization's (or practitioner's) primary goals. Similarly, given the many other pressures on today's clinicians, and health workers' relative lack of experience with computing during their training, there can be a reluctance to learn new skills in an area that seems foreign and tangential to medical care.

It can be difficult to make the business case for investing in IT, especially at the level of individual practitioners. Clinicians tend to feel that they are being asked to make the investment in clinical computing systems, whereas the major beneficiaries of such investments are health systems, payers, and patients. This misalignment of fiscal incentives is cited often as a major barrier to widespread dissemination of IT into practice settings where, ironically, the primary data are gathered and where decision-support capabilities could be utilized with the most benefits. When physicians see clear benefits from their IT investments, and see efficiencies and cost savings as well as enhancements to information access, a major barrier to suitable investments will be overcome.

Fragmentation of the health care system in the United States has meant that decision-making regarding IT investment generally is a local decision, with local optimization in mind, leading to poor coordination and a lack of generally accepted standards. Competitive pressures among both providers and vendors have made coordination and integration evanescent goals. Health care organizations are complex social environments, and many IT users do not work for the organizations that provide the systems to them. This differs markedly from most segments of US society, where an employer can simply require employees to use the computational tools that they make available. Physicians on a hospital's staff often resist such requirements unless they see the benefits of system use.

We have already stressed the critical role of technical standards if IT is to realize its potential to link systems to achieve the integrated cycle of Fig. 5. Frequently cited as a problem, standards development continues to attract attention in both the research and vendor communities, but the political and logistical challenges remain formidable.

The growing climate for change

For those of us who have observed the barriers to IT dissemination and use in health care for several decades, the recent changes in attitude and attention to this topic are truly remarkable. What has happened recently to lead to a new momentum for creating and maintaining the NHII? What led President Bush to call for use of EMRs for all Americans within a decade? Why, after more than ten years of encouragement, has the Federal government recently created a position of the National Coordinator for Health Information Technology in the Department of Health and Human Services (the first time there has been a high-level official in HHS with a specific charge related to health care IT and its optimal use and dissemination)?

We believe there are several explanations for the recent trends. A number of forces have come together to create enthusiasm for health IT solutions. Simply stated, these forces are advances in data safety and quality, costs, and privacy. Although the health care community long has been concerned with all these issues, certain recent landmark events greatly broadened our awareness of their dependence on IT solutions:

- A decade of studies by the Institute of Medicine (IOM) and the National Research Council (NRC), dealing with issues such as computer-based patient records [3], patient data protection [4], the role of the Internet in health and health care [5], and the importance of demonstration projects to show the value of systems integration within regions and nationally [6];
- A series of three recent and influential reports from the Institute of Medicine [7–9], all of which made strong cases for the role of IT in addressing problems with medical errors and enhanced patient safety;
- Federal advisory activities, including seminal contributions in the report, *Information for health: A strategy for building the national health information infrastructure* from the Workgroup on the NHII from the National Committee on Vital and Health Statistics [10] and two important sets of recommendations, first in 2001 [11], then again in 2004 [12] from the Subcommittee on Health within the President's Information Technology Advisory Committee (PITAC);
- Employer concerns regarding the burgeoning costs of health care, leading to the creation of the Leapfrog Group and its active promotion of more

effective implementation and use of IT in health care settings;
- The privacy, security, and transaction rules that were announced by DHHS in response to the requirements of the 1996 Health Insurance Portability and Accountability Act (HIPAA) and that in many respects require informed technological solutions in order to be compliant;
- The influence of the Internet and the World Wide Web, which has greatly increased the access to health information by the public and transformed their familiarity with, and expectations of, health IT in the settings where they seek care.

The list could be much longer, and would certainly include the large number of recent reports, from a variety of public and private sources, that reiterate and refine the recommendations that have come before. Seldom has there been more consensus on the need for action and the promise that awaits us if we handle things well.

The NHII as an information and communications infrastructure

As we have described, the NHII will provide an information and communications infrastructure that can be used for a variety of public health goals, as well as in clinical care and research. Clinical data collection and analysis based on NHII methodology can assist us in preventing and responding to threats to national security, particularly bioterrorism. An NHII, therefore, will be important in supporting national security. The information and communications infrastructure we and others are promoting will establish a state-of-the-art infrastructure in a geographic region that supports communication, access, knowledge management, and decision-making. The Department of Health and Human Services recently funded proposals to establish some eight to ten sites for 5-year NHII demonstration programs. This one-time financial support from the Federal government can be a major step forward in realizing this infrastructure to support health, medicine, and public health.

Advances specified as objectives for cancer care and surveillance are profoundly dependent on the infrastructure we have described. An NHII will mean that much more detailed information will be available for analysis of individual cancer cases and populations of cancer patients. Cancer registry standards can be expanded to include a wider range of information and more timely data. Moreover, the infrastructure will facilitate more widespread use of cancer registry and surveillance data by facilitating its availability, by incorporating data from and to health care delivery and the practice of medicine, including information on cancer prevention. An improved infrastructure, coupled with analytic support, can mean improved targeting of health care resources to those places and persons in greatest need, and a greater ability to evaluate health care quality and outcomes.

Conclusion

To sum up, a 10-year vision for the future of cancer care, surveillance and information dissemination, although appealing and widely shared, requires major changes in the current culture of cancer care and data policy, an investment of capital, and extensive logistical planning. It is critical as well to take account of the competitive nature of the medical marketplace. That competition, coupled with fiscal pressures on providers and health systems, means that leadership for regional and national coordination will need to come from elsewhere, such as from Federal and state governments.

The effort is worthwhile. We are poised to achieve today what has been sought and anticipated for at least three decades. The gains for cancer, and for medicine in general, should encourage us to move forward with enthusiasm and commitment.

References

1. Shortliffe EH (1993) Doctors, patients, and computers: Will information technology dehumanize health care delivery? Proc Am Phil Soc 137(3):390–398
2. Shortliffe EH, Cimino JJ (eds) (2006) Biomedical informatics: computer applications in health care and biomedicine. Springer, New York
3. The computer-based patient record: An essential technology for health care. Institute of Medicine. National Academy Press, Washington, DC; 1991; revised 1997
4. For the record: Protecting electronic health information. National Research Council. National Academy Press, Washington, DC; 1997
5. Networking health: Prescriptions for the Internet. National Research Council. National Academy Press, Washington DC; 2000
6. Fostering rapid advances in health care: Learning from system demonstrations. Institute of Medicine. National Academy Press, Washington, DC; 2002
7. To err is human: Building a safer health system. Institute of Medicine. National Academy Press, Washington, DC; 2000

8. Crossing the quality chasm: A new health system for the 21st century. National Academy Press, Washington, DC. Institute of Medicine 2001
9. Patient safety: Achieving a new standard for care. Institute of Medicine. National Academy Press, Washington, DC; 2004
10. Information for health: A strategy for building the national health information infrastructure. Workgroup on the NHII. National Committee on Vital and Health Statistics, Washington, DC; 2001
11. Transforming health care through information technology. Panel on Transforming Health Care. President's Information Technology Advisory Committee (PITAC). Report to the President. National Coordinating Office for IT Research and Development, Washington, DC; 2001
12. Revolutionizing health care through information technology. Panel on Transforming Health Care. President's Information Technology Advisory Committee (PITAC). Report to the President. National Coordinating Office for IT Research and Development, Washington, DC; 2004