Confidentiality, Information Technology, and Health Care

Thomas C. Rindfleisch
Director, Center for Advanced Medical Informatics
Section on Medical Informatics
School of Medicine, Stanford University

We are well into the digital information age. Digital communications and information resources affect almost every aspect of our lives — business, finance, education, government, entertainment, ... Clinical medicine is, of course, highly information intensive, but is one of the few areas of our society in which computer access to information has had only limited success — except in selected areas such as billing and scheduling, laboratory result reporting, and diagnostic instrument systems (e.g., in radiology and cardiology). The move to widely accepted, electronic patient records (EPRs) is accelerating, however, and is inevitable because of many pressures — the desire to improve health care through timely access to information and decision-support aids; the need for simultaneous access to records by doctors, nurses, and administrators in modern, integrated provider and referral systems; meeting the needs of highly mobile patients; the push toward improved cost effectiveness based on analyses of outcomes and utilization information; the need for better support of clinical research; and the growing use of telemedicine and telecare [5].

We are, of course, motivated by the great benefits to patient care and medicine that can derive from this effort. But almost daily we hear about network computer break-ins — often close to home and in terms arousing vivid fears [4]. By putting our personal medical records on-line, might we be increasing the risk of exposing highly private and sensitive information to outsiders?

In this article we take a “systems” view of privacy and information security in health care. We will put in perspective the nature of the most urgent threats to patient information confidentiality, the new threats that almost certainly will arise because of the technologies of digital information, the kinds of countermeasures that can be effective, the places where technology can be of use and where not, the risk/cost trade-off decisions that must be made for real-world systems, the overarching policy issues that must be addressed, and impediments to the resolution of these issues.

On-Line Health Care Information

Once largely a fee-for-service cottage industry, health care has seen the steady growth of Health Maintenance Organizations (HMOs) and Integrated Delivery Systems (IDSs), and the transformation of reimbursement from an invoiced service basis to a capitation basis — under which providers receive a prenegotiated fee for each patient under their care, independent of actual services rendered. Growing fierce competition in the health care industry is resulting in regional IDSs that provide one-stop shopping for ambulatory clinic care, urgent care, and inpatient hospital care. Deloitte and Touche indicates that 24% of US hospitals now belong to an IDS and 56% of hospitals are pursuing EPRs. Outpatient clinics are also aggressive in pursuing EPRs, driven by their roles in regional IDSs and the pressures of streamlined, modern patient care.

These developments bode well for improved health care. Providers will have access to the most current information and decision-support aids for diagnosis and treatment— no matter where the point of care. Researchers and public health officials will have access to better information for their studies of disease epidemiology and treatment efficacy. And health care enterprise managers will have better information on which to base business decisions for care standards and optimization of clinical care pathways.

Need for and Nature of Medical Confidentiality

Our medical records contain much mundane information about us, such as height and weight readings, blood pressures, and notes about bouts with the flu, cuts, or broken bones. These records also may contain some of the most sensitive information about who and what we are —about topics such as fertility and abortions, emotional problems and psychiatric care, sexual behaviors, sexually transmitted diseases, HIV status, substance abuse, physical abuse problems, genetic predispositions to diseases, and so on. Access to this information must be controlled because disclosure can harm us— for example, by causing social embarrassment or prejudice, by affecting our insurability, or by limiting our ability to get and hold a job. Of course, such damage can (and does) occur no matter whether our medical records are in paper or electronic form. We have only to glance at grocery store tabloids or election year news stories to see the allure and marketability of "interesting" health information about well known people (see for example [11]).

We have a strong (but often implicit) expectation that such information will be used only in the context of providing effective care, and otherwise, will be kept secret. This expectation is based on a number of principles, beginning with the Hippocratic Oath of more than 2,000 years ago[1], and reinforced by the Code of Ethics of the American Medical Association[2] and by the federal Privacy Act of 1974[3]. Without broad confidence in medical privacy, we know there are consequences — patients

[1] "What I may see or hear in the course of the treatment or even outside of the treatment in regard to the life of men, which on no account one must spread abroad, I will keep to myself..."

[2] "A physician shall respect the rights of patients, colleagues, and of other health professionals, and shall safeguard patient confidences within the constraints of the law..."

[3] The 1974 Privacy Act specifies restrictions on federal agencies maintaining records on individuals including a right to know that identifiable information is being kept and why, and a right to review and amend/correct data.

avoid needed health care and physicians do not enter all information into patient records (or may even keep double sets of records).

Paradoxically, our medical records contain information about us that is of the utmost sensitivity, yet this information is only useful *to us* when it is shared with the medical providers and system under which we get our care. Indeed, our physicians need and expect access to our complete medical records in order to help diagnose diseases correctly, to avoid duplicative risky or expensive tests, and to design effective treatment plans that take into account many complicating factors. The desirable sharing goes beyond our personal care and includes our relationships to society as a whole — through support of medical research, public health management, and law enforcement. Thus, we must distinguish among three concepts involved in protecting health care information:

- *Privacy* — the right and desire of a person to control the disclosure of personal health information.
- *Confidentiality* — the controlled release of personal health information to a care provider or information custodian under an agreement that limits the extent and conditions under which that information may be used or released further.
- *Security* — a collection of policies, procedures, and safeguards that help maintain the integrity and availability of information systems and control access to their contents.

Threats to Confidentiality of Health Care Information

To understand the risks of disclosure of health care information and where information system technologies might be of help, we need to know how health care information is used. In 1976, Alan Westin developed a diagram [12] that shows the overall flow of information in the U.S. Health Care system (Figure 1). Other more recent studies have documented the extent of this flow as well (see for example, [6, 9, 10]). We normally think of the medical record as a tool at the point of care — the doctor's office, clinic, or hospital. It supports primary care physicians, specialists, nurses, and administrators and has contributions from the many testing and treatment services. It is a memory aid to help a team of providers manage a patient during an encounter, to provide continuity of care from encounter to encounter, and to serve as an institutional record of care rendered.

Medical records also serve a variety of functions for organizations not involved directly in care. Records are sent to insurers (government and private) to justify payment for medical services rendered, and to detect fraud. They are used for quality reviews, administrative reviews, and utilization studies to manage the business aspects of health care. And they are used for societal purposes, such as medical research, public health management, social service and welfare system management, law enforcement, screening and licensing for professions such as airline pilots, and determining life insurance eligibility.

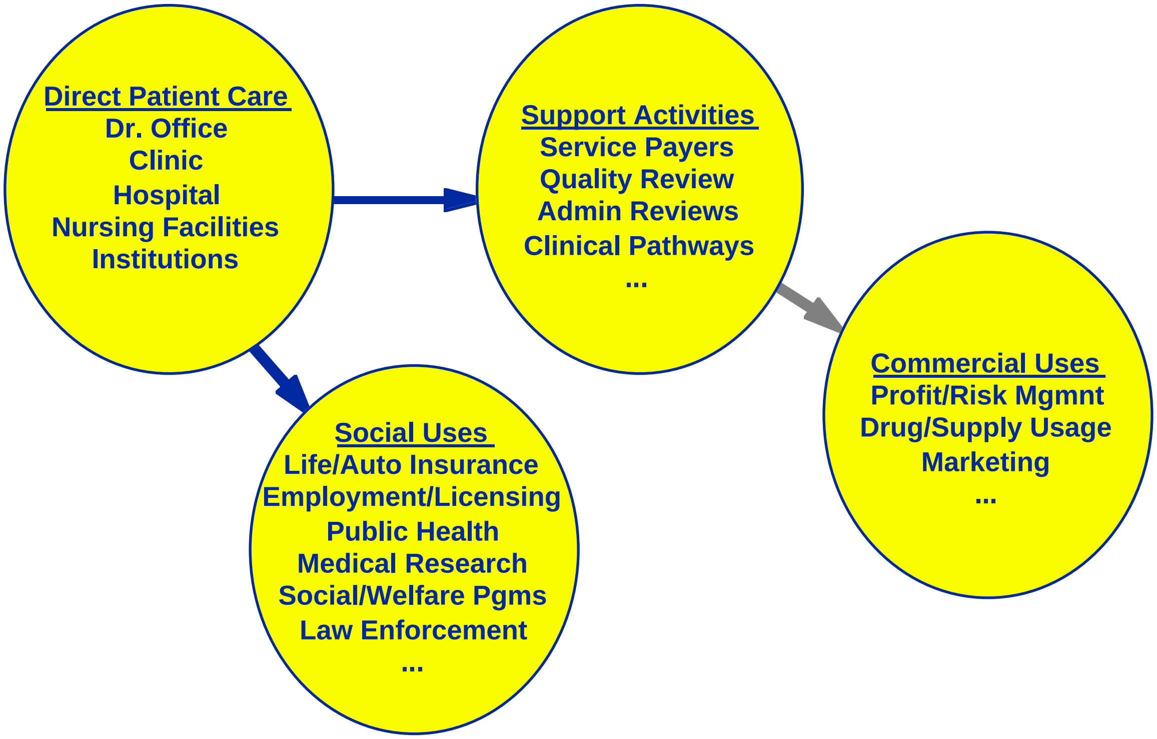

Figure 1: Flow of health care information in the US system (After Alan Westin, 1976)

Despite signing general consent forms as a requirement for obtaining health care in the US, the great majority of people (patients and physicians alike), have only a vague understanding of where health care data flow, often with little control of its use. In this complex system, risk of disclosure arises often. As in most information systems, few quantitative data exist on the nature and extent of security problems in health care institutions. There are few incentives or mechanisms to report incidents, and specific cases are most often handled quietly, unless a legal proceeding is filed. The consensus among health care CIOs is that the most important threats to patient information confidentiality are the following, in decreasing frequency of occurrence [9][4]:

1. **From inside the patient care institution**

   - *Accidental Disclosures* — medical personnel make "innocent" mistakes and cause unintentional disclosures. A conversation may be overheard between care providers in the corridor or elevator. A lab technician may notice test results for

[4] We ignore threats from environmental and system failures in this list as outside the scope of this article. Good practice to cover these kinds of failures have been in place for decades and lower cost systems and peripheral equipment, such as RAID arrays, have made redundancy and backup more convenient and cost-effective than ever.

an acquaintance. Information may be left on a computer screen where it can be seen by a passerby, or email or FAX messages may be misaddressed.

- *Insider Curiosity* — medical personnel abuse their record access privileges out of curiosity or for their own purposes. Some do so out of concern for the well being of fellow employees or family members. Some want to know about celebrities being treated. Some may be concerned about the possibility of sexually transmitted diseases in a colleague they are dating.
- *Insider Subornation* — medical personnel knowingly access information and release it to outsiders for spite, revenge, or profit. Embarrassing health information about prominent people finds its way into grocery-store tabloids or the public press with relative ease. It is said that Nicole Brown Simpson's (paper) medical records were available to the press within a week of her murder in 1994. The London Sunday Times reported in November 1995 that the contents of anyone's (electronic) medical record in Great Britain could be purchased on the street for £200.

2. **From within secondary user settings**

   Uncontrolled secondary usage — those who have access rights to patient information for a purpose in support of primary care may exploit that access for other purposes not envisioned in patient consent forms — broadly "data mining" in modern parlance (see below).

3. **Outsider intrusion into medical information systems**

   Unauthorized access —vindictive former employees, angry patients, network intruders, or others may steal information, damage systems, or disrupt operations. A recent NRC study of security practices in health care institutions found no examples of (detected) outside intruder break-ins [9]. Nevertheless, reports abound of intrusions in business, academic, and government sites on the Internet (see for example, [1, 4]). It must be considered an artifact of the fact that the US health care industry is still almost totally reliant on paper records that such intrusions have not occurred at health care sites.

The threats from insider disclosures and break-ins from outside intruders are easy to understand. But the risks to patient information confidentiality from secondary users need further explanation. We must emphasize that such secondary users as medical research, public health, governmental administration (e.g., Health Care Finance Administration and MEDICARE), and law enforcement are very carefully controlled. Institutional Review Boards closely review and control medical research activities, courts supervise law enforcement access, and the federal Privacy Act controls government use.

Secondary users of concern include insurers, pharmaceutical payers, some employers, and other players in the emerging health information services industry. While each of these users has a justified need to access patient information to carry out their function in the system, few controls are currently in place to ensure that the information is used only for the authorized purposes [9]. Some secondary users are highly conflicted. For example, self-insuring employers, under the Employment

Retirement Income Security Act (ERISA), are entitled to receive fully identified patient information for employees being covered. Such information is nominally used to help the employer/insurer make sound benefits management decisions, but it can also affect whether employees get promoted, or even whether their employment is continued.

There is a temptation to use medical record information for business purposes other than those initially authorized — for example, to manage risk in insurance underwriting, to guide marketing of medical products, or to target special market segments for non-medical "services" based on health status (e.g., presence of Alzheimer's disease). In a March 1996 consent decree filed in Minnesota and joined by 17 other states, a drug company agreed to stop using questionable marketing practices in interfering in the prescription of medications made by other companies as a by-product of seeing information to assess the allowability of drug insurance coverage[5].

The potential for (ab)use of personal genetic information is very sobering. A recent study [3] reported 206 cases of direct discrimination —employment and insurability problems — from unauthorized use of genetic-test information. These cases reflect discrimination on the basis of *future* potential for (treatable) diseases. The patients exhibited no current phenotypic evidence of disease.

We must expect the privacy threat from data mining to grow. In a June 1996 cover story, *Information Week* predicted that overall industry revenues for data warehousing and mining technologies ran at $2B in 1995 and were estimated to jump to $8.8B by 1998. At least three companies in the health information services industry that are members of the "terabyte club" — organizations with very large-scale data warehouses used to collect and analyze data for business applications. It may be argued that there is nothing wrong with using health care information to make prudent and profitable business decisions — this is merely capitalism at work. But these uses conflict deeply with the confidentiality understandings most patients have when they sign consent forms. They certainly result in patients avoiding needed treatment in sensitive areas. And they make part of our population uninsurable or place the burden of costs on a group that can least afford them. We should at the very least openly discuss and decide these policy issues at a national level.

Technologies to Help Protect Health Care Information

Unlike paper-based patient records, where access control is almost entirely manual and procedural, technological security tools are an integral part of EPR systems and offer a number of advantages. The applicable technologies come largely from

[5] Some pharmaceutical companies like Merck and Lilly have subsidiaries services (Medco and PCS respectively) that administer insurance coverage for drugs.

cryptographic and distributed systems research in computer science and, at the highest level, serve five key functions: [9]

- *Availability* — ensuring that accurate and up-to-date information is available when needed at appropriate places.
- *Accountability* — helping to ensure that health care providers are responsible for their access to and use of information, based on a documented need and right to know.
- *Perimeter Definition* — knowing and controlling the boundaries of trusted access to the information system, both physically and logically.
- *Role-Limited Access* — enabling access for personnel only to information essential to the performance of their jobs, and limiting the real or perceived temptation to access information beyond a bona fide need.
- *Comprehensibility and Control* — ensuring that record owners, data stewards, and patients can understand and have effective control over appropriate aspects of information privacy and access.

We do not have space to detail the relevant technologies in this article, but only to provide a summary listing of various interventions, their functions, and how they relate to protecting privacy[6]. As summarized in Table 1, there are three general classes of technological interventions to improve system security — deterrents, obstacles, and system management precautions. Deterrents depend upon the ethical behavior of people and provide reminders and oversight to reinforce those standards. Obstacles directly control the ability of a user to get at information, with the goal of constraining access only to information for which they have a need or right to know. System management precautions involve proactively surveying an information system to ensure that known sources of vulnerability are eliminated. It has been shown that deterrents — alerts, reminders, and education of users — are very effective in reinforcing already highly ethical behavior of the great majority of health care providers [9]. Also, audit trails are effective. If it is known that the system will record the identities, times, and circumstances of all users accessing information, and that these records are reviewed regularly, ethical users will think twice about abusing their privileges.

Technological obstacles can be equally effective. They support strong user and computer authentication, and ensure that users can access only information for which they have a bona fide need and right to know based on their identify and job function. They also can protect information against eavesdropping, ensure the integrity of information and software content, and validate the origin and content of orders and other critical transactions (digital signatures). Firewalls enforce

[6] Detailed technological descriptions are readily available from other sources — see for example [2, 7, 8].

manageable perimeters around distributed information systems, limit modes/protocols for access. Finally, rights management software, such as the IBM Cryptolope system, offer interesting future possibilities for securely delivering information. Content is segmented and encrypted, the software used to access the record is standardized and distributed from the information custodian, and users are granted access keys based on their identity and need/right to know. Obtaining the key serves as a basis for an audit trail, even perhaps across institutional boundaries.

Intervention	Function	Example Technologies
Deterrents		
Alerts and reminders	Reinforce user ethics	Vendor-specific
Audit trails	Document access/give alerts	Custom research systems
Obstacles		
Authentication	Determine who is connecting	Accounts/passwords, kerberos, tokens (e.g., SecurID), public key systems, biometric systems
Authorization	Define who can access what information	OS file and database vendor access controls, DCE access control lists
Integrity management	Ensure information content is as intended	Cryptographic checksums
Digital signatures	Validate notes and orders	Evolving standards
Encryption	Prevent eavesdropping	PGP, kerberos, DES, public key systems, secure sockets
Firewalls and network service management	Define system perimeter and control means of access	Many vendors
Rights management tools	Control information distribution and access	IBM Cryptolopes
System Management Precautions		
Software management	Guard against viruses, Trojan horses, etc.	Tripwire and controls over loading of uncertified software
System vulnerability analysis tools	Detect unintended system vulnerabilities	SATAN, crack, National Computer Security Association

Table 1: Summary of technologies applicable to information system security management.

System management precautions are crucial and include taking advantage of accumulated community experience about security vulnerabilities. Software management prevents introduction of programs like viruses, Trojan horses, or other aberrant codes. The Computer Emergency Response Team (CERT) emphasizes that many on-going network break-ins come from failures to configure systems properly and maintain them at current releases of system/service software.

The Role of Technology — Deployment, Trade-Offs, Strengths, and Limitations

The application of any administrative or technical intervention to protect information requires an explicit policy defining what is appropriate use of information and what is not. Such a policy should include as a minimum a statement of institutional philosophy and goals regarding privacy and security; a classification of information assets by type; standards for administering, controlling, and monitoring information use by type; standards for information system design, implementation, and operation; and a definition of procedures for detecting and handling abuses.

In principle, many of the technologies needed to do a prudent job of protecting medical information system security are available, if not deployed in commercial systems or in routine practice [9]. We can relate the classes of disclosure threats to available tools as shown in Table 2.

Threat	Principal Countermeasures
Insider abuse	
Accidental disclosures	Education, alerts, reminders
Insider curiosity	Education, authentication, authorization, audit trail, rights management tools (future possibility)
Insider subornation	Same as above
Secondary users	Rights management tools (future possibility)
Outsider intrusion	All available obstacles and system management precautions

Table 2: Relation of disclosure threats to security technologies.

Simple, mostly non-technical measures are appropriate to avoid accidental disclosure of confidential information or curiosity-driven disclosure. Technology, such as audit trail systems, can play an important role to curtail insider curiosity or subornation. In the future EPR technology might help by maintaining patient anonymity through use of coded patient identifiers (pseudonyms) in at least some parts of the care process.

To date, technological deterrents and obstacles play almost no role in controlling exploitation of patient information by secondary users. Once information leaves the hands of the health care provider, it is stored off-site by the secondary user and access and use controls are subject to the ethics and procedures in place by that user site. With such an unsupervised system, ethical controls fall short. In the future, in addition to tighter legal restraints, rights management software may provide a more effective way to control inappropriate secondary use.

Blocking outsider intrusions will be a major problem judging from the successes intruders now enjoy in attacking academic, business, and government systems. Special diligence is needed for health care systems to ensure state-of-the-art protections. This might include special dedicated network segments for health care enterprises and establishment of "medical CERT" and industry oversight groups to ensure high security standards.

Ultimately, security and privacy of health care information is a "people" problem. Technology can help to make sure that only health care personnel access information they have a right and need to know, and that information gets from one place to another accurately and securely. But technology can do very little to ensure that the person receiving the information will handle it according to confidentiality standards. That depends on ethics and an effective supervisory and legal structure that provides sanctions against detected misuse.

Security measures in medicine must be chosen and integrated rationally. The measures must be balanced so they protect against a realistic assessment of risks and costs. Real-world information systems will always be vulnerable. Also, threats, particularly those arising from outside the enterprise, will continue to evolve with overall technological developments in computing and networking. Finally, each security intervention must be evaluated jointly in terms of its functional benefits for protecting patient, provider, and institutional privacy and in terms of its costs. These costs include the cost of purchase and integration into the information system environment; the cost of on-going management, operations, and maintenance; the cost of user time lost to satisfy security protections; and the cost of user frustration with clumsy interfaces and procedures.

It is unthinkable that we would impose system security constraints so tight that they would prevent an emergency room doctor from accessing the record of a seriously ill, comatose patient, needed to guide treatment. Such exigent access may only be needed from special locations, but their existence means an enterprising intruder may more easily fool system access controls and break in. How should we make this trade-off?

Also, we must recognize that physicians are under growing pressure to increase productivity. They are asked to see more patients in less time while making better (or at least better justified) decisions about diagnosis and treatment. Providers can not tolerate time delays and frustrations in passing frequent record access security hurdles.

Individual technologies vary widely in terms of these cost/benefit characteristics and, as new technologies are developed and reduced to commercial practice, their characteristics change with time. System managers must choose a set of technological interventions that provide effective protection against perceived threats to system security but which overall impose acceptable costs. This choice is difficult at best and no acceptable standards of performance exist. These remain to be defined and will certainly require on-going updates of threat models; evaluations of technologies; reconsideration of integration and operation strategies; and education of management, systems staff, and users.

Summary Observations — Opportunities, Actions, and Impediments

The development and integration of electronic patient record systems into modern US health care institutions is absolutely essential and inevitable for optimal health care, medical research, public health, and the operation of modern health care enterprises. Such systems do not exist in most health care institutions today, but they have been demonstrated in a variety of academic, public, and commercial medical settings. Even paper-based medical record systems entail substantial risks of disclosure of sensitive personal health care information. The primary threats arise from various kinds of disclosures by members of the health care provider community themselves, and from uncontrolled use of information among secondary users. Longer-term threats to health information confidentiality will come from the mere fact of having information computerized and managed in network environments — from network intruders. The rapidly developing uses of data mining technologies in business and health care signal still more threats that raise significant policy and legal issues.

As we move toward the era of computerized medical record systems, we must design the systems from the start to accommodate evolving policies and security management technologies, and develop standards to integrate and administer computerized health information systems prudently. The broad and effective use of existing, but largely undeployed, technological tools in the computerization of patient-identified health-care information can help prevent exploitations of sensitive information, and/or make it clear to data owners that exploitations have happened.

Substantial public policy and legislative issues have to addressed soon in the US that will define the standards and safeguards that are to be applied to all health care information, not just that in digital form. These must address the current hodgepodge of state-based privacy laws and loop holes in current laws that allow uncontrolled access to and exploitation of patient-identified health care information in parts of a developing health information services industry. There are extremely difficult trade-offs to be made in this debate. The significant advantages of facile information access for improved medical care, enhanced research, and more cost-effective management of medical institutions have to be traded off with the privacy consequences. In cold business terms, this comes down to assessing the value of health care information, the magnitude of the risks of improper disclosure, the costs of an improper disclosure incident, and the costs of preventative measures. However, whereas financial enterprises such as banks and credit card systems can absorb the costs of abuse over the user community, without undue hardship on individuals, medical enterprises can not. Once sensitive information about an individual is exposed and the resulting damage is done to that person, the information can not be withdrawn and made secret again. Thus we must move very carefully on issues such as strong legal restraints for abuse, incorporating effective cryptographic tools for security management, and not implementing features such as universal patient identifier systems too hastily — until we have demonstrated and are confident about the strength of properly designed medical information systems.

Existing medical information systems are mostly home built, or involve collections of legacy systems that do not interoperate. For wide deployment, we will need a uniformly interoperable, vendor-supplied set of system components that incorporate highest performance security features, based on public and secret key encryption technologies. Only then will the standards develop and the costs come down to allow these tools to be integrated in effective ways. The needed technologies exist or are under development in the Internet distributed computing and commerce arenas, but they are not yet deployed in a way that allows integration into enterprise computing environments.

Contentious policy issues and potential roadblocks surround on-going federal government controls on the export of strong cryptographic technologies. These restraints, based on understandable fears for national security and law enforcement, in turn delay the willingness of US computer companies to invest in commercializing strong security tools that will be denied access to the world market — the essential transnational market for modern Internet systems.

Acknowledgments

I have benefited greatly in developing this article from my participation in the National Research Council study group on Security and Privacy in Medical Information in the National Information Infrastructure, sponsored by the National Library of Medicine from 1995-1997. Stimulating discussions with expert study group members, visits to review security systems and procedures at six modern health care enterprises, and testimony from a large number of presenters provided an invaluable resource for writing this perspective. This article is my own perspective of health care information security and does not represent any official view of the NRC or of other members of the study group.

References

[1] GAO/DISA, *Information Security: Computer Attacks at Department of Defense Pose Increasing Risks*, General Accounting Office/Defense Information Systems Agency GAO/AIMD-96-84, May 22, 1996 1996.

[2] S. Garfinkel and E. Spafford, *Practical UNIX & Internet Security*, O'Reilly & Associates, Inc., 1996.

[3] L. N. Geller, J. S. Alper, P. R. Billings, C. I. Barash, J. Beckwith, and M. Natowicz, "Individual, Family, and Societal Dimensions of Genetic Discrimination: A Case Study Analysis," *Science and Engineering Ethics*, vol. 2, 1996.

[4] M. Gembicki, *1996 Information Systems Security Survey*, WarRoom Research, LLC http://www.infowar.com/sample/survey.html-ssi, November 23, 1996 1996.

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

[6] IOM, *Health Data in the Information Age: Use, Disclosure, and Privacy*, Institute of Medicine, National Academy of Sciences, Washington, DC, 1994.

[7] R. Khanna, *Distributed Computing – Implementation and Management Strategies*, Prentice-Hall, 1994.

[8] NRC, *Computers at Risk: Safe Computing in the Information Age*, National Research Council/National Academy of Sciences, 1990.

[9] NRC, *For the Record: Protecting Electronic Health Information*, National Research Council, National Academy of Sciences, March 1997.

[10] OTA, *Protecting Privacy in Computerized Medical Information*, Office of Technology Assessment, OTA-TCT-576, US Government Printing Office, Washington, DC, 1993.

[11] J. Rothfeder, *Privacy for Sale*. New York, NY, Simon and Schuster, 1992.

[12] A. F. Westin, *Computers, Health Records, and Citizen Rights*, National Bureau of Standards, Monograph 157, US Government Printing Office, Washington, DC, 1976.

Side Box — Discussion of Background of Cryptography

Cryptography as an underlying technology

The majority of technological interventions for protecting health care information depend in some way on cryptographic methods. At present two kinds of cryptography are of potential use — symmetric or secret-key cryptography, a system in which the same key is used for encryption and decryption, and asymmetric or public-key cryptography, a system in which two different keys are used, one for encryption and one for decryption. The most common secret-key system in use today is the Data Encryption Standard (DES) developed by IBM and the National Bureau of Standards in the early 1970s and adopted as a federal standard in 1976[7]. DES uses a 56-bit key to encrypt and decrypt information based on a bit-manipulation algorithm that is well suited to rapid execution on modern computers. Because only a single key is involved, it must be shared (and therefore transported) between parties wishing to exchange information securely. Safe key transport can be a major problem.

The most common public-key system available today is the Rivest, Shamir, Adleman (RSA) system patented in 1983. RSA depends on the difficulty of factoring very large numbers and uses Euclid's algorithm from algebra to define key pairs that are used to encrypt and decrypt information by modular exponentiation. The order of key use is commutative so that if data is encoded by key 1 of a set, key 2 is used to decode the data and if data is encoded by key 2, key 1 is used to decode it. Because two keys are required, only one need be kept secret by the user to whom the key set is assigned. The other (public) key can be made generally available. If the public key is used to encrypt information, the sender can be assured that only the holder of the (other) secret key can decrypt it. Similarly, if the holder of the secret key encrypts information, someone with the public key can be sure the information came from the secret key holder — with proper certification that a public key is assigned to a given individual, this is the basis of the digital signature and related services.

Public-key systems run about 1000 times more slowly than DES systems and require keys about 10 times longer[8]. For this reason secret-key cryptography and public-key cryptography are often used together. Public-key cryptography is used for transactions in which the certified identity of the sender and/or receiver of a given message is crucial (and hence worth the computational

[7] FIPS Publication 46, 1977.

[8] W. Diffie, *The First Ten Years of Public-Key Cryptography*, vol. 76, pp. 560-577, 1988.

cost). One such application is to transfer secure DES session keys to be used in higher volume subsequent encrypted communication between entities.

Two points should be noted about cryptographic technology: (1) security tools based on cryptography are still largely undeployed *anywhere* in the public computing industry, much less in health care, and (2) cryptography does not *solve* the security problem — cryptography transforms the problem into a key management problem. Despite the ready availability of much cryptographic technology and numerous specifications for how to incorporate that technology into operational services (e.g., Privacy-Enhanced Mail, PGP, Kerberos, OSF Distributed Computing Environment, Secure Sockets, etc.), very few users of modern distributed computing systems actually take advantage of cryptographic protections. A number of universities have set up kerberos-based authentication systems based on software exported by MIT, some groups are using Zimmerman's Pretty Good Privacy system to authenticate and protect email traffic, and there is some use of Secure Sockets to protect sensitive world-wide web communications. But these are still limited and represent a very small fraction of the overall user population and traffic on intranets or the Internet.

With respect to key management, much of current day discussion about commerce systems, legally binding digital document management, strong authentication, etc. centers on the problems of secure and certified key management. The foundation of strong, public-key-based user authentication is an infrastructure system by which unforgeable certificates are issued with public keys that are trusted and that ensure that a key is associated with a bona fide real person. This certificate authority acts much like a notary public in the signing of conventional legal documents, where the notary seal certifies that the document signature was performed at an indicated time and place by an identified real person. The analog to a notary in digital authentication is a "certification authority" — some third party which signs a certificate containing the user's identity and public key. In turn, the third party's key is often signed by a higher-level certification authority. We have only a few examples of civilian key management systems today, such as Lotus Notes, campus kerberos deployments, and beginning experiments with public-key systems in Internet commerce (e.g., MasterCard/Visa). For Internet commerce, the banking system is stepping forward to attempt to provide this function. In a broader setting, it has been suggested that the federal government establish a certification authority system, perhaps administered by the United States Postal Service or the Social Security Administration -- but these are only postulated mechanisms at this point. As the scope of key management services grows, trust in the integrity of key assignments tends to diminish, and the problems of revocation in the case of key compromise become much more difficult. However this key-certification function is carried out, it is an essential part of the necessary infrastructure for public-key authentication and digital signature systems, and the economical development of commercially supported, trusted security tools based on these technologies. We have only begun to demonstrate workable, trusted systems using modern cryptographic tools.