SECONDARY USE OF GENETIC INFORMATION IN RESEARCH BENEFITS, PROBLEMS, AND POTENTIAL APPROACHES

By

Carolyn Petersen

A CAPSTONE

Presented to the Department of Medical Informatics & Clinical Epidemiology
and the Oregon Health & Science University
School of Medicine
in partial fulfillment of
the requirements for the degree of

Master of Biomedical Informatics

May 2009

School of Medicine

Oregon Health & Science University

CERTIFICATE OF APPROVAL

This is to certify that the Master's Capstone Project of

Carolyn Petersen

"Secondary Use of Genetic Information in Research: Benefits, Problems, and Potential Approaches"

Has been approved

David A. Dorr, M.D., M.S.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

I would like to thank Dr. David Dorr for his guidance as I worked through this multi-phase project. His patience enabled me to more thoughtfully assess the changing legislative environment and its implications for genetic privacy and the practice of medical informatics. Dr. William Hersh stimulated my interest in informatics through his excellent presentation of the AMIA 10x10 course and the provision of extracurricular learning opportunities. Andrea Ilg and Diane Doctor supported my success as a distance student through constant attention to and assistance with the many details inherent in pursuing an advanced degree.

I would also like to thank Mr. Brian Kaihoi, a Mayo Clinic colleague who introduced me to the field of medical informatics. His encouragement and support during my years in the program helped me meet the challenges associated with development of technical skills and knowledge.

ABSTRACT

Scientific and technological developments have made it possible to test human DNA for the presence of abnormalities associated with some inherited disorders. This genetic information may become part of an individual's medical record, creating the potential for discrimination in employment, insurance purchase, and other areas. The Genetic Information Nondiscrimination Act and Health Insurance Portability and Accountability Act have provisions intended to protect individuals from genetic discrimination, but health care industry and societal influences threaten individuals' genetic privacy. At the same time, health care reform and the movement to identify the most effective medical treatments have resulted in increased interest in secondary use of individuals' medical information, including genetic information. Through a literature review, this paper examines how data privacy, data security, and electronic health records have influenced the current understanding of genetic privacy. It explores evolving forces – social media, patient-driven research, personal health records, online marketing of genetic tests, and data ownership and stewardship issues – that are changing the way patients interact with the health care system. After assessing consumers' expectations and the impact of legislation on genetic privacy, this paper identifies systems- and process-based approaches to genetic data management and secondary use of genetic data.

INTRODUCTION

Advances in laboratory techniques and bioinformatics are making it possible for researchers and health care organizations to collect and store greater amounts of genetic data about patients and clinical research participants for longer than ever before. Such techniques as chain termination[1] and high throughput sequencing[2] allow researchers to determine the heritable genetic make-up of human beings and, in some cases, determine whether an individual carries the gene for a particular condition. Information obtained through gene sequencing can be stored nearly indefinitely for future use. In addition to providing information about whether an individual carries the gene for a condition, gene sequencing can be used in predictive testing, such as to predict the likelihood that a fetus will have cystic fibrosis.[3] It can also be used to select an approach to treatment, as with the use of CYP enzyme testing for antidepressant selection.[4]

At the same time, improvements in computer hardware and medical information systems (e.g., electronic health records) are making it possible to store and transfer large quantities of medical data efficiently. Sophisticated databases and high-capacity storage devices facilitate sharing of identifiable medical information among multiple physicians, researchers, and institutions. When individuals consent to the collection and storage of medical and/or genetic information, they typically have no way to determine what

happens to the information derived from the sample(s) provided. In particular, they cannot determine how information about them is used, whether researchers followed confidentiality preservation procedures, and whether their right of genetic privacy has been protected adequately.

These and other questions take on greater importance with the increasing interest in and, on occasion, need for secondary data use. Though medical data have been subject to limited secondary uses in the United States for decades, e.g. in certain public health activities, technology and medical practice standards have constrained widespread development of secondary data uses. The marriage of biotechnology and complex health information systems makes possible many more large-scale secondary data use applications than were previously feasible, but questions about privacy and confidentiality of personally identifiable genetic information – particularly that given without consent for secondary uses – remain.

Using a literature review, this project will 1) review existing federal laws relevant to privacy, security, and genetic nondiscrimination; 2) examine existing issues relevant to genetic privacy and the electronic health research environment, including privacy concerns and requirements, security, electronic health records, and use of data in research; and 3) explore evolving issues relevant to genetic privacy, including social media, patient-driven research, personal health records, online marketing of genetic tests, and data ownership and control issues. The final chapter presents some approaches to genetic privacy protection that permit secondary use of personal health information data for research.

The first chapter of this document defines two concepts (genetic privacy and secondary data use) and two pieces of legislation (the Genetic Information Nondiscrimination Act of 2008 (GINA) and the Health Insurance Portability and Accountability Act of 1996 (HIPAA)) central to thinking about the potential affects of secondary data use on genetic information. Without assurance that researchers will preserve confidentiality and protect privacy, consumers and patients will be unwilling to support secondary uses of genetic data due to fear of difficulty obtaining employment, insurance, and financial services.

Genetic Privacy

Anyone who has interacted with a medical provider, financial institution, insurance company, public utility, or other entity that collects personal information has been exposed to the concept of privacy. Stated broadly, privacy is the notion that all individuals have identifying characteristics or information about them that they may wish to make available selectively, and that individuals have a right to decide who may access to their personal information. The idea of privacy has become ubiquitous in American society; by law, Americans annually receive a privacy policy or similar document from each entity that holds personal information about them. They also sign a statement authorizing release of personal information at the office of every medical care provider with whom they have a relationship, which makes it possible for providers to share medical information with other providers, bill insurers, and comply with government health reporting regulations. These events and related activities provide the basis for a

general understanding of medical privacy, privacy that relates specifically to episodic and ongoing medical care.

Genetic privacy is a related but separate concept. Though one's DNA may be analyzed in the diagnosis or treatment of illness, genetic testing and any resulting genetic profile have come to be viewed as more than just medical information. Genetic information tells a story of not only the past, but also the future. DNA functions as a "probabilistic medical record," a property that differentiates genetic privacy from medical privacy.[5] Medical records exist as historical information, but genetic "records" – DNA – exist as both historical medical record and encoded prediction based upon known probabilities.

- Additional factors support the separation of medical and genetic privacy. According to medical ethicist Lawrence Gostin, genetic data have been viewed as different from other types of data for reasons that go beyond genetic information's predictive capability, including:
- Genetic profiles yield a far greater amount of information about a person than do medical or other types of records;
- Genetic data can reveal secrets about an individual, including those that could not be discovered through other means;
- Genetic data provide a means by which to identify individuals with near-absolute certainty;
- Genetic information can be stored and used in applications unimaginable at the time of sample collection; and

- Genetic information can be used to analyze specific populations and make predictions about their personal characteristics in ways that have never been possible with other technologies.[6]

In effect, genetic information and the research possibilities it offers far exceed the limits of medicine's previous experience with medical information. Genetic information answers not only questions asked by patients – but also questions patients have not thought to ask and or may not wish to consider.

Compared to medical privacy, genetic privacy takes on additional dimensions in that specific privacy concerns differ before and after a genetic condition becomes apparent.[7] Before an individual shows signs of a condition, genetic testing is a source of possible harm, should test results indicating predisposition to the condition be shared against an individual's wishes. After a genetic condition becomes observable or has been diagnosed, genetic information related to that condition (but not other genetic conditions to which the individual is predisposed) is treated the same as non-genetic medical information. The primary privacy concerns shift to the individual's family members, who themselves may experience harm as a result of their genetic relationship to the diagnosed individual.

Given the relatively recent availability of DNA testing, it is difficult to predict how genetic privacy will evolve. Advances in technology allow researchers today to perform analyses that were unimaginable when genetic testing first was practiced. Health law attorney Gostin noted that, as a result, patient consent given years ago for tissue storage and future testing can no longer be considered "informed" because donors could not have foreseen – and thus agreed to – the analyses to which tissue samples are subject now.[6]

The evolution of more sophisticated genetic testing methods also may affect the genetic privacy expected by individuals entering into research agreements as research subjects. Patients once donated tissue samples to researchers funded by the U.S. National Institutes of Health, charitable foundations, or other not-for-profit institutions, but now may be asked to consent to unrestricted use of tissue samples by private, for-profit companies. Such agreements may lack the genetic privacy rights contained in not-for-profit entities' agreements and expose research subjects to unanticipated and unwanted use of personal information.[8]

The development and increasingly widespread implementation of electronic health records (EHRs) poses other implications for genetic privacy. On the one hand, rare conditions that would likely not be recognized in an emergency situation would be documented in the patient's EHR, facilitating appropriate care. On the other hand, the increased information sharing associated with health information networks increases the risk that improperly disclosed personal health information will reach those not intended to see it.[9] Less integration and organization of health information, i.e. through the use of paper records, provides a measure of protection against unintentional or unapproved data sharing.

Some question whether it is possible to achieve true genetic privacy. Regulatory measures authored to prevent unauthorized disclosure of genetic information are necessary but insufficient to protect genetic privacy.[10] Attention to the situations in which data are acquired by third parties (e.g., for life insurance underwriting, selection of a cancer treatment, or mortgage application review) will do more to protect genetic privacy than regulatory initiatives detailing the information that may be disclosed and the

procedures for doing so. Such an approach would shift the focus to the potential harms to be experienced by the individual whose data is shared, rather than to achievement of a procedural specification.[10]

Genetic Information Nondiscrimination Act of 2008

Congress approved GINA[11] in April 2008 after 13 years of efforts to pass federal antidiscrimination legislation.[12] When the Human Genome Project began in 1995, researchers and ethicists involved with the project recognized that the ability to determine the genes an individual carried could place some individuals at a disadvantage when trying to purchase health insurance, get a job, secure a mortgage, or in other situations. Such forms of bias had been reported after individuals have been diagnosed with genetic or other serious illnesses,[13] and consumer advocates feared that knowledge of an individual's genetic profile could facilitate prospective discrimination.

Concerns about the potential for genetic discrimination proved accurate. In 2001, the U.S. Equal Employment Opportunity Commission (EEOC) filed suit against Burlington Northern Santa Fe Railroad for secretly testing employees for Chromosome 17 deletion, which some believe can predict development of carpal tunnel syndrome.[14] Because the United States had no law protecting individuals against genetic discrimination, EEOC filed the action under the Americans with Disabilities Act of 1990 (ADA).[15] The ADA protects individuals with symptomatic genetic disabilities, but not those who show no symptoms.[16] The ADA also does not prevent employers from requiring genetic testing of employees and people who have been given a conditional job offer. EEOC might not

have won its action in court under the ADA, but Burlington Northern close not to fight EEOC's petition.

GINA prohibits genetic discrimination against individuals by health insurance companies and employers.[17] It applies to health insurance, but not to long-term care insurance, disability insurance, or life insurance.[18] It also fails to cover the period prior to employment, permitting employers to require consent to access applicants' complete medical record before hiring.[13] GINA requires that the EEOC issue regulations related to genetic discrimination in employment by May 21, 2009,[19] and these regulations go into effect on November 21, 2009.[11] As written, GINA will not prevent all instances of discrimination based on a person's DNA, but it does address two issues that, historically, have particularly concerned ethicists and consumer advocates.

As with all other laws, GINA's effect on individuals and society depends upon the way it is implemented. Though health care providers and others who have access to personal health information have incentive to keep confidential patient information private, the risk of breach of confidentiality and resulting discrimination cannot be entirely eliminated.[20]

Health Insurance Portability and Accountability Act of 1996

Background

Several federal and state statutes affect the concept of privacy as it relates to medical information, including genetic information. HIPAA is the best-known federal mandate addressing consumer privacy with regard to health care. HIPAA's Privacy Rule was published December 28, 2000, after Congress failed to enact privacy statutes as mandated

by HIPAA, and revised to its final form in March 2002. Those subject to HIPAA and associated administrative rules – "covered entities," in the parlance of the U.S. Department of Health and Human Services – include health care providers such as physicians, nurses, and pharmacists; medical clinics, hospitals, nursing homes, and pharmacies; health insurance companies, managed care organizations, and other health plans; health care clearinghouses; and health care-related government agencies such as Medicare.[21] Any individual or organization that transmits health information as a part of transactions such as claims filing and health service pre-authorization is considered a covered entity; third-party organizations that act on behalf of providers (business associates) also are covered entities.[22]

Health records may contain large quantities of information. The Privacy Rule covers all individually identifiable health information relating to a person's past, present, or future physical or mental health; the health services given to a person; and past, present, or future payment for health services given to an individual that can be used to identify the person or that may reasonably be thought to identify the person. Demographic information and common identifying items such as a birthday or Social Security Number also are considered to be protected health information. The Privacy Rule defines information meeting any of these criteria as protected health information (PHI) except information held by a covered entity as employment records rather than patient health records.

In the course of being evaluated and treated, identifiable information about a patient often must be used by multiple care providers. To facilitate necessary flow of information among providers, patients sign a consent form permitting release of the information to

providers and associated business partners before receiving care. The HIPAA Privacy Rule includes detailed descriptions of when and how PHI may properly be shared in various situations.

The Privacy Rule has become a foundational principle in efforts nationwide to provide consumers with a standardized set of rights related to health record access and confidentiality. The Privacy Rule also is intended to facilitate the flow of health information in support of quality health care and promotion of the patient's well-being while protecting the public's health and well being.[22] When implemented correctly, HIPAA is intended to protect patients from inappropriate disclosure of their personally identifiable information while facilitating provision of appropriate, high-quality care. As a provision of GINA, HIPAA protection will be extended to genetic information in July 2009.

Although HIPAA places stringent requirements on how PHI may be used, it does not limit the disclosure or use of information that has been de-identified. The regulation defines de-identified information as information that does not identify a specific individual or provide a reasonable basis for identifying an individual.[22] Personally identifiable information can be de-identified through two processes, examination by a statistician or removal of certain identifying pieces of information (e.g., name, Social Security number) from a record. Although human tissue contains genetic information that can be used to identify individuals, the Privacy Rule does not treat blood and tissue samples as individually identifiable personal information.[23]

HIPAA and Health Research

HIPAA's Privacy Rule affects not only patient care, but also health care research. "Research" includes "systematic investigation, including research development, testing, and evaluation, designed to develop or contribute to generalizable knowledge."[22] Under this definition, creation of databases or repositories for research, as well as use or disclosure of information from such databases and repositories, qualify as research.[23]

Although health care researchers use PHI, which is subject to the Privacy Rule, they are not classified as covered entities unless they provide patient care or perform any of the activities subject to the Privacy Rule.[24] They may also be affected if they obtain data from covered entities. Covered entities that provide data for research must adhere to the minimum necessary standard; providing no more information than is needed to conduct the research.[24]

As for non-research uses, the Privacy Rule permits researchers to use PHI with written consent from the individual or in some circumstances in which authorization has not been obtained. Such circumstances include: [24]

- When a data set has been de-identified
- When a limited data set is used and the covered entity has entered into a data use agreement with the researcher stipulating acceptable data uses and privacy protections to be implemented
- When an Institutional Review Board (IRB) or other privacy review board waives or alters the requirement for individual authorization in accordance with the waiver/alteration criteria set out in the Privacy Rule
- When the research involves the PHI of deceased persons

- When a researcher wishes to review PHI while developing research protocols
- When the researcher obtained an authorization, informed consent, or IRB waiver for use of PHI prior to implementation of the Privacy Rule on April 14, 2003

In a limited data set, everything required for de-identification except elements of dates and ages and "other unique, identifying characteristics" must be removed.[24]

HIPAA-Related Challenges to Health Research

The requirements of the Privacy Rule have created additional complications for investigators conducting clinical and health care research. Reported challenges include:

- Difficulty obtaining authorization to use PHI because researchers frequently do not have contact with those whose PHI they seek[24]
- Difficulty accessing research data from other institutions, whose data protection and use policies and procedures may differ from the researcher's institution[24]
- Confusion over which regulations (e.g., HIPAA Privacy Rule, human subjects protection regulations) apply to specific research projects and, thus, how the needed research data must be handled and protected[24]
- Inhibition of cancer survivorship research and information sharing among cancer patients, their physicians, and researchers[25]

As a result of these and other research challenges, various health care-related organizations have begun calling for changes to HIPAA and the broader issue of data protection. For example, the President's Cancer Panel called for re-evaluation of HIPAA

provisions in its 2003–2004[25] and 2005–2006[26] annual reports, citing the need to conduct increased research in the areas of cancer survivorship.

Perhaps the strongest call for review of the Privacy Rule came in a 2009 Institute of Medicine report that recommended development of a new approach to protecting privacy.[27] The new approach should be written to facilitate uniform application to all health research. If a new approach cannot be developed, the report noted, the Department of Health and Human Services should revise the Privacy Rule to necessitate fewer variations in its application, facilitate more effective data use, and eliminate provisions that fail to enhance privacy.

Secondary Data Use

After collection, information can be put to multiple uses. When a physician orders a blood test for a pregnant patient to screen her fetus for cystic fibrosis (CF), the mother assumes that her blood will be tested for factors associated with the presence of CF. Testing for CF, thus, is the intended use – the primary use – of the blood sample provided. The blood sample can also be used for other purposes, such as testing to determine if she had adequate stored iron in her blood. Ferritin testing, in this case, is a secondary use. With the development of more genetic tests, it has become possible to use the mother's blood sample for other purposes – additional secondary uses – that may not benefit her or her child. In secondary data use of this kind, where the benefit is not immediate or direct, the risks of gathering this new information outside of the original testing agreement must be weighed against potential benefits.

In the white paper "Toward a National Framework for the Secondary Use of Health Data," American Medical Informatics Association leaders defined secondary data use as "non-direct care use of personal health information including but not limited to analysis, research, quality/safety measurement, public health, payment, provider certification or accreditation, and marketing and other business including strictly commercial activities."[28] This definition is broad enough to cover not only genetics-related secondary data uses, but also research, commercial, and other uses. Examples of secondary data uses to which genetic information could be subject include:

- Recruitment of subjects for a diagnostic or therapeutic clinical trial
- Identification of households likely to be interested in a medical device or durable medical equipment
- Reporting of disease incidence in support of public health surveillance regulations
- Medication selection and dosage determination
- Insurance application review and decision making
- Marketing of disease-specific medications or services

Clearly, some of these secondary data uses are likely to benefit an individual to a degree that the individual regards the reduced genetic privacy as a reasonable cost. Other uses, such as product marketing or insurance rating, are less likely to be regarded positively by individuals who have provided a tissue sample for some other purpose. The remainder of this paper will examine existing issues related to genetic information, personal health information, and the electronic health care environment; emerging issues related to genetic information and personal health information in the era of electronic

health records and social networks; and opportunities, challenges, and suggestions for the future with regard to secondary data use and genetic privacy.

ESTABLISHED ISSUES RELATED TO SECONDARY DATA USE

Privacy and confidentiality are among the issues of greatest concern to consumers. When a patient is treated by a physician or receives care in a hospital, personal details about the individual's life are compiled and stored, not just for the duration of the care episode but potentially for decades or a lifetime. Development and widespread deployment of electronic health records will make possible the rapid sharing of information across providers and health care systems, facilitating clinical care and research initiatives that previously were difficult, if not impossible, to undertake. At the same time, the potential impact of unintentional disclosure or knowing violation of confidentiality has increased due to the ease and speed with which information can be shared.

An important related issue is data security, the systems and technologies that together support protection of privacy and promotion of confidentiality within the health care system and others with whom data are shared. For most individuals, the technical details of such systems lie outside the realm of day-to-day activity, but the expectation that health care providers and organizations will protect data remains. A survey performed for the Institute of Medicine in September 2007, however, suggests that Americans do not fully trust health care professionals to protect their information. In response to the statement, "I generally trust my health care providers – doctors and hospitals – to protect the privacy and confidentiality of my personal medical records and health information,"

30 percent agreed strongly, 54 percent agreed somewhat, 12 percent disagreed somewhat, and 5 percent disagreed strongly.[29] If the 17 percent disagreement rate does not distress, the fact that slightly more than half only somewhat trust professionals should.

This chapter examines several established issues that are relevant to the secondary use of genetic data. The issues described in this chapter – privacy and privacy-related concerns, technical requirements supporting privacy, security and security-promoting technologies, electronic health records and their implementation, and the impact of privacy and confidentiality requirements on data use for research – are termed "established" because they have been studied for some years and are well-understood relative to evolving issues, such as the role of social media in the sharing of personal health information sharing. Experienced informaticians certainly can enumerate other pertinent established issues; those discussed in this chapter were selected because they affect most, if not all, fields within medical informatics and bioinformatics.

Privacy

Among the constellations of concepts comprising health information technology, few have achieved the awareness – or the notoriety – of privacy. Whether informed of the latest technological advances or just the most recent privacy breach, almost everyone has an opinion. Whereas some advocate for more stringent consent processes, others argue that privacy protection must go beyond consent forms to a system that balances the risks and benefits of a proposed data use and does not require consent for every use.[30] This section reviews key privacy concerns, notes some of the most well-known privacy

breaches, and summarizes work that is being done to develop technology requirements for maintenance of privacy.

Privacy Concerns

"Privacy" has come to be used as a shorthand reference for many ideas and issues. Similarly, privacy-related issues span a broad range of concerns. Some privacy issues, such as the use of stored blood samples, directly intersect the science of genomics, while others are indirectly related. Some of the most prominent privacy concerns relevant today and into the future include:

- Use of stored blood samples – The Citizens Council of Health Care has filed suit against the state of Minnesota over the storage of blood samples taken from infants at birth to screen for such treatable genetic conditions as phenylketonuria.[31] The consumer group charges that the state has used more than 52,000 dried blood spots for research.
- Medical identity theft – Medical identity theft involves the theft of personal and insurance information about an individual so that another individual can obtain care or submit medical claims in someone else's name. One example of medical theft occurred at the Cleveland Clinic, where a desk attendant sold patient information to a cousin, who submitted false bills to Medicare.[32] Once genetic information has been stolen, others can sell it for profit or use it to gain access to medical care that might be unavailable otherwise, such as for enrollment in a clinical study. The incidence of medical fraud has increased during the past decade, and was reported to affect 1.8 percent of patients in

2005.[33] In 2006, 3 percent of reports made to the Federal Trade Commission for misuse of consumer information contained in existing accounts involved medical insurance accounts, and 0.4 percent of complaints related to new accounts involved medical accounts.[34]

- Undisclosed data sales – Although privacy policies typically specify how personally identifiable information will be handled within the context of health care provision, the policies may not clearly communicate what may happen to the data should the patient fail to submit payment as expected. When hospitals turn past-due accounts over to collection agencies, the agencies may auction off the debt over the Internet, sharing confidential information – and potentially, genetic information – with buyers in the process.[35]
- Proliferation of data fusion centers – The number of data fusion centers – operations that aggregate data about individuals from multiple sources – increased significantly as part of post-9/11 efforts to reduce the likelihood of further terrorist activity. Such centers operate under state laws, which has resulted in minimal oversight of the firms and little transparency or accountability to the public.[36] Consumers typically are unaware that aggregated files that may contain personal health and genetic information have been created about them.
- Willful disclosure via electronic media – Like consumers, health care professionals use electronic and social media, such as email lists, Wikis, and blogs. Although the Privacy Rule extends to all uses of personally identifiable

> health information gathered by medical professionals, disclosure of protected information occurs. A content analysis of 271 medical blogs published by health care providers found that 42.1 percent contained descriptions of individual patients, and that 16.6 percent provided enough information for patients to identify themselves or their doctors. Three blogs displayed photographs in which patients were recognizable.[37] Electronic media can function as a distribution route for genetic information.

These examples highlight the need for vigilance against privacy violations from a wide variety of sources. Secondary data use can lead to unauthorized and potentially negative use of data.

Privacy Breaches

Consumers may not understand the legal definition or privacy or the specific implications of the HIPAA Privacy Rule, but they are aware that privacy breaches occur. In a June 2008 Harris poll of approximately 2,400 American adults, 69 percent reported hearing about a loss or theft of personal health data, and 4 percent said they think their personal health information or that of a family member has been lost or stolen.[38] Even health information professionals are divided; in an online survey conducted by the Healthcare Information and Management Systems Society in May 2008, 54 percent thought HIPPA privacy and security rules are strong enough and 34 percent said the rules are not.[39]

Determining how many privacy breaches have occurred is difficult, but the popular media provide a broad, if not comprehensive, view of what can happen to data intended to remain confidential. Some recent privacy breaches include:

- In August 2008, NIH removed access to two databases containing genetic information from the NIH Web site.[40] NIH had made the databases available publicly to facilitate research, and stored data had been masked, summarized, and aggregated in the belief that these steps were sufficient to block identification of individuals. The event showed that assumptions protecting privacy may be more difficult than previously believed.[41]
- WellPoint, the health insurer with the largest number of covered lives in the United States, in April 2008 acknowledged that confidential information about 130,000 members had been placed on the Internet during the previous year.[42]
- On March 17, 2008, the University of Miami announced that a container of back-up tapes containing patient information had been stolen from a university vendor's vehicle.[43]
- The University of California San Francisco inadvertently placed online the records of 6,000 patients, which were publicly accessible from July 1 to October 9, 2007.[44] The records were made available online when UCSF shared patient information with a vendor hired to mine the records for fundraising information. The breach occurred in direct violation of the University of California's HIPAA Privacy Rule Implementation Guidelines.[45]
- During 2006 and 2007, UCLA Medical Center administrative specialist Lawanda Jackson accessed the medical records of Farrah Fawcett, Maria

Shriver, and approximately 60 other celebrities, then selling confidential information to media outlets.[46]

- In November 2006, licensed practical nurse Andrea Smith accessed the record of a Northeast Arkansas Clinic patient and gave the information to her husband, who called the patient and threatened to disclose the information during a legal proceeding.[47]
- As these examples indicate, privacy breach-related data exposure results from many sources, and all types of personal health information can be affected. In some cases, inappropriate human activity was at fault; in others, unforeseen consequences resulting from the use of technology were the cause. Even when regulations intended to protect confidential patient information were in place, privacy breaches occurred. If institutions that collect personal health information are unable to protect patients' confidentiality when collecting information for primary uses such as clinical care, patients are less likely to consent to secondary use of their information. Regardless of cause, the future management and use of personally identifiable health information depends in part upon addressing privacy concerns highlighted by the cases, as well as other actions.

Technical Requirements Supporting Privacy

Numerous health information products have been developed and are available for purchase and commercial use, but vendors have pursued proprietary strategies to achieve compliance with HIPAA and other statutory mandates. As a result, the industry has yet to

adopt common data management and sharing standards. As the industry matures and greater numbers of physician practices and health care institutions implement electronic health records, the need for industry standards will become even more acute. Privacy requirements can increase consumer acceptance of and support for secondary data use by conferring credibility. Identification of tangible product functionality and institution processes that promote privacy preservation engender confidence that researchers can, in fact, protect privacy. At present, the Health Information Technology Standards Panel and Health Level Seven are working to clarify requirements and standards for privacy protection in data management and sharing.

Health Information Technology Standards Panel. The Health IT Standards Panel (HITSP) is one group working to develop standards for interoperable IT systems. Several organizations, including the Healthcare Information Management Systems Society and the American National Standards Institute, jointly sponsor HITSP.[48] The panel brings together industry and government with the goal of developing standards for applications used in health care environments. The U.S. Department of Health and Human Services (HHS) funds the panel's work, which includes developing standards for a national data exchange network.

HITSP is developing specifications for numerous health care applications, including electronic health record lab results, medication management, personalized medicine, biosurveillance, quality reporting, and patient-physician secure messaging, among others.[49] After announcement of the American Reinvestment and Recovery Act

(ARRA), the organization moved to provide “lightweight interoperability specifications” to aid vendors in applying for ARRA funds.[50]

To date, HITSP has approved and released 10 complete sets of standards, including standards for patient-provider secure messaging, consumer access to clinical information via networks and media, medication management, personalized medicine, immunizations and response management, consultations and transfers of care, remote monitoring, public health reporting, and emergency responder systems. The standards are written to ensure that systems support privacy mandates, and HHS has mandated that federal health systems implement the consumer clinical data accession standard.[51] However, HITSP has no authority to compel private-sector health care systems, providers, or vendors to build applications to its standards.

Health Level Seven. Health Level Seven (HL7) is an international group of health information specialists working together to develop standards for administrative and clinical data exchange and integration. Accredited as an American National Standards Institute standards-developing organization, HL7 was founded in 1987, and its members serve voluntarily. HL7’s most well-known work to date is a messaging standard for health data exchange.[52] HL7 differs from other standards initiatives in that it seeks to develop interface standards that could be applied across health care institutions, industry market sectors, and countries.

Security

Health care professionals invariably are acquainted with HIPAA's privacy protection provisions because much of their workflow and patient care processes have been created to promote privacy or maintain confidentiality. Providers often are less familiar with HIPAA's data security provisions, which have come to be known as the Security Rule, in part because the Security Rule applies only to electronic health information.[53] The Security Rule addresses implementation of the security standards contained within HIPAA, establishing the level of security to be provided rather than specifying the technology or tactics to be used to achieve security. More frequent accession of PHI from locations outside the clinic or hospital, greater quantities of information contained in electronic records, and electronic sharing of data with business partners (e.g., third-party payers) have resulted in increasing data security risks. The Security Rule is intended to protect patients and providers by defining data management outcomes that, at least in theory, support privacy, among other objectives.

Because the Security Rule is considered to be an administrative requirement, the Centers for Medicare and Medicaid Services oversees its implementation. The Security Rule requires that covered entities must 1) implement appropriate administrative, technical, and physical safeguards to protect the privacy of protected health information; 2) protect such health information from intentional or unintentional use or disclosure; and 3) limit incidental uses and disclosures related to permitted or required uses or disclosures.[53] With a few exceptions, the Security Rule preempts state laws. Security safeguards include multiple standards that, together, form required or technically feasible implementation specifications. If a covered entity determines that a specification is

reasonable and appropriate to its circumstances it must implement the specification; if it determines that a specification is unreasonable and/or inappropriate, it must document how it reached the decision and implement equivalent specification(s). Covered entities should develop an environment that emphasizes risk assessment, risk reduction, and continuous improvement.[54]

Security-Promoting Technologies

The Security Rule permits great flexibility in how covered entities achieve data security. Health information systems vendors and academic organizations have developed numerous technological solutions that support Security Rule specifications. Some examples of security-promoting technologies include:

- Commutative encryption involves the serial encryption of all data pieces in each covered entity's data set until all covered entities have encrypted the data. The data sets are joined on key attributes that have been encrypted by all covered entities, and entities cannot obtain individually identifiable personal health information from the joined data.[55]
- The National Cancer Institute's cancer Biomedical Informatics Grid (caBIG) developed the Grid Authentication and Authorization with Reliably Distributed Services (GAARDS) infrastructure to address challenges with user account authentication across multiple institutions, trust fabric management for credential provisioning across multiple institutions and multiple levels of system access, and access control and enforcement.[56]

- Implementation of database design that removes personal health information from exported data sets and permits only authenticated users to view and perform pre-specified tasks on only certain data types involving the minimum personal information necessary to complete a task reduces the security risk associated with clinical research databases.[57]
- These security-promoting technologies must continuously evolve to address the threats to discovery of electronic records. Proponents of EHRs often note that while paper records are easily available on racks in physicians' offices, electronic records are safely locked behind passwords and firewalls. Skeptics counter that while EHRs may not be immediately accessible, their format makes it easier for would-be thieves to both gather en mass and share information that should remain private. Reports of data thefts, such as a recent theft of 8 million pharmacy records in State of Virginia pharmacy records, suggest that skeptics may have a point.[58]

Even when physicians make a point of staying aware of data security threats and updating their health information systems, some threats slip through the cracks. In a study of 16 practices in Ontario, Canada that use the Internet, physicians generally believed that they complied with privacy legislation and that data security was adequate. A survey of equipment and security-promoting measures in the offices, however, revealed that none of the physicians managed their computers' firewalls and only 40 percent of computers had had a virus scan in the previous month.[59] Spam detection software, spyware, and other products that enhance security were implemented even less frequently, despite physicians' beliefs that their systems met federal and provincial security mandates.

In 2008, the Healthcare Information and Management Systems Society (HIMSS) commissioned a report evaluating the security of patient data and the potential for health care and identity fraud. The hospitals and clinics surveyed tended to focus on HIPAA compliance rather than fraud risk reduction, and few had plans for dealing with attempts to obtain patient data for fraudulent use.[60] When data breaches occurred, health care organizations typically discovered that the exposure was greater than initially believed.

Survey respondents also indicated a lack of understanding of the Sarbanes-Oxley Act, which likely has led to storage of greater amounts of patient data than are required by law. The personally identifiable data most frequently compromised in security breaches included patient names, followed by care-related information such as diagnoses; patient addresses; treatment information; Social Security numbers; and insurance information.

Internal personnel accounted for the breach in about one-quarter of cases that were serious enough to require patient notification. Inadvertent breaches were reported to be more common than planned malicious activity, so a majority of organizations prefer staff education as the primary strategy for preventing data security breaches. Organizations were less likely to be concerned about, or to have developed, strategies for addressing theft or access with malicious intent by employees.

Particularly critical of current data security practices is a report by the HHS' Office of the Inspector General. It noted that the Security Rule was subject to minimal enforcement by the Centers for Medicare and Medicaid Services. In an October 27, 2008 memo to CMS, Inspector General Daniel Levinson noted, "Ongoing Office of Inspector General audits of various hospitals nationwide indicate that CMS needs to become more proactive in overseeing and enforcing implementation of the HIPAA Security Rule by focusing on

compliance reviews. Preliminary results of these audits show numerous, significant vulnerabilities in the systems and controls intended to protect ePHI at covered entities. These vulnerabilities place the confidentiality and integrity of ePHI at high risk."[61] Prior to the OIG audits, CMS enforced the Security Rule by responding to complaints of security failures. Following release of the OIG report, CMS agreed to develop policies and procedures for proactive compliance reviews of covered entities, a key OIG recommendation.

Electronic Health Records

Electronic health records have been promoted as both the solution to many of the most challenging problems within the American health care system and as the probable source of many new problems. Some argue that even with minimal implementation, the ability to identify drug interactions, retrieve test results, and remind patients of needed care already has saved millions of dollars and reduced negative health outcomes.[62] Others suggest that though they may facilitate improvements in patient care, the costs of systems, the difficulties in implementing them, and the privacy issues related to data storage and transfer render EHRs impractical for all physicians and health care organizations to implement.[63,64] Privacy within the clinical care setting is not the focus of this document, but electronic health records are an important source of research data, particularly when secondary use is undertaken, so EHRs merit consideration.

Consumer Perception of EHRs and Privacy

Consumer perceptions about electronic health records are mixed. On the one hand, consumers have expressed interest in products that can help them improve their health, and there is a perception that EHRs are one such tool. On the other hand, among individuals who have heard about medical record theft or loss, seven percent believe that their medical records or those of a family member have been lost or stolen.[65] In addition, three times as many consumers (47 percent) believe that EHRs are lost or stolen most often as believe paper records are stolen or lost most (16 percent). Paek previously reported that patients' previous roles and life experiences influence their perception of risk related to EHR and computerized physician order entry use within hospitals,[66] so the skepticism reflected in the Harris poll is not surprising.

Despite the ambivalence about the security and privacy of EHRs, consumer interest remains. Nine percent of respondents to the Deloitte 2009 Survey of Health Care Consumers reported having an electronic personal health record (PHR), an increase from the 8 percent reported the previous year.[67] More telling, perhaps, is that 42 percent of respondents would like an EHR that is connected to their physician's office, and 57 percent would like to be able to access their medical records and perform other health care-related functions (e.g., schedule appointments, pay bills) electronically. Survey respondents have concerns about privacy – 38 reported being "very concerned" about privacy and security of personal health information – but those issues haven't reduced the interest in electronic health record use.

Benefits of EHRs for Research

As a quality management tool, EHRs facilitate data abstraction and research in ways that would be impossible or extremely time-consuming or costly when done using paper records. Search engines such as the Electronic Medical Record Search Engine make it possible to query patient records for specific terms or data, facilitating data abstraction.[68] This capability could greatly reduce the time now required to search paper records for health conditions or terms associated with genetic conditions and genetic tests.

Some examples of secondary data use in quality research include:

- Automated abstraction of EHR clinical notes to develop a surrogate measure for patient's health-related quality of life[69]
- Integration of patient questionnaires into EHRs[70]
- Patient survival prediction for individualized care planning[71]
- Drug efficacy research on cardiovascular outcomes[72]
- Utilization review and clinical implementation of diabetes care guidelines[73]
- Review of text-based clinical notes to identify patients with heart failure[74]

Secondary data uses involving genetic data also exist. For example, researchers can assess the efficacy of genetic testing for medication selection by analyzing the medical records of patients who underwent testing.

Implementation of EHRs

Although EHRs have been in use by a small number of providers for decades, and have received a great deal of attention in recent months, their implementation in the United

States is still fairly limited. In a national survey of electronic record use in ambulatory care settings, 13 percent of the 2,578 responding physicians reported using a basic EHR, and 4 percent indicated use of a more comprehensive system.[75]

Physicians report using EHRs because they believe the systems enhance clinical decision making, facilitate communication with other physicians and pharmacies filling patients' prescriptions, and reduce medical errors.[75] However, the features likely to provide the greatest improvements in patient care tend to be those that are most advanced (e.g., clinical decision support, reporting and analytical tools) are also the features physicians used least frequently in a study of physician use patterns in a large Western health system.[76] The failure to implement these beneficial functions is an important consideration when weighing EHR benefits versus the risks to privacy.

Privacy, Confidentiality, and Data Use for Research

In their most basic form, EHRs perform largely the same functions as paper charts, albeit in ways that facilitate integration with other clinical care tools, such as computerized order entry. The digital format changes the way providers and patients access patients' information and, depending on the environment in which EHRs are used, may offer functionality not feasible with paper charts, such as rapid remote access to medical information. However, EHRs fundamentally remain medical records, and are subject to many of the same limitations and issues as paper charts. This is true even when they are used as a source of data for research.

Privacy issues related to personal health information are fairly intuitive. Few people, if any, benefit from inadvertent or malicious disclosure of their personal health

information, and the potential harm that may result from such disclosures is clear. Some patients may be interested in drug marketing information sent to them as a result of data mining of their health record, but others will see this effort as misuse of personal information. Consumers are concerned about privacy regardless of whether data use is primary or secondary.

With regard to medical record use for research purposes, however, the issue is less black and white. Though some people will never be comfortable with use of their medical information for research, others may be willing to share information about themselves if there are potential benefits for their care or the care of others. A recent study of consumer attitudes toward psychiatric genetic testing and research found a positive attitude regarding research, as had been noted in previous studies.[77] However, respondents became more concerned about negative effects of testing and research when treatment for identified conditions could not be guaranteed.

Despite differences in EHRs, data sharing networks and systems, and state laws, existing data standards do support medical data integration and analysis. HL7's work, as well as that of others such as the Clinical Data Interchange Standards Consortium, may facilitate the sharing of research data. Kush and colleagues identify eight initiatives in development that address health data exchange, and others (as noted above and elsewhere) also exist.[78] Data sharing and integration challenge the conduct of health research today, but in the future will likely cease to be obstacles.

Some of the challenges of using patient data, particularly when consent for secondary uses has not been acquired, are apparent from the previous discussion of privacy. For some types of research, such as quality improvement research, privacy may be less of an

issue than for other areas because data sharing among institutions or singling out individuals with unique traits are not required. For example, researchers at Johns Hopkins University implemented a quality-improvement project to determine whether use of a procedure checklist for catheter insertion in intensive care units would reduce the number of infections.[79] In the first 3 months of the trial the infection rate dropped from 2.7 infections per 1,000 insertions to 0 per 1,000 insertions, a rate that persisted through 18 months. When the researchers published the work, the U.S. Office for Human Research Protections (OHRP) launched an investigation into the work because the university's institutional review board had designated the study exempt from federal regulations. After investigating, OHRP determined that Johns Hopkins acted improperly in treating the study as exempt from institutional review board review and informed consent requirements.[80]

Although the Hopkins infection-reduction case raises more issues than just those related to secondary data use, it illustrates a central challenge inherent in research involving secondary uses of personal health information: the difficulty in balancing patients' privacy needs and expectations with improvements in care that may benefit a large number of patients or population. Researchers intended to see if they could reduce infections and, in so doing, caused no reported harm to the patients in their care. Yet research conventions demanded that they obtain permission not only for use of the protocol, but also for use of the data in outcomes analysis. Note British Heart Foundation epidemiologists Davies and Collins, "Clearly, research should conform to good practice, but it remains appropriate to consider whether over-interpretation of data protection legislation represents another real, albeit difficult to quantify, risk to the public."[81]

They argue that rather than applying rules, researchers should determine "the likely effect of alternative approaches to protecting personal data on the potential health gains from the research."

Functionally, the informed consent process represents a kind of opt-in mechanism in that patients need to actively agree to participate. The absence of required informed consent is not equivalent to an opt-out mechanism in that without the informed consent process, patients generally would not know of their own participation in the trial or analysis. In thinking about how to protect patient confidentiality while forwarding a research agenda with implications for the larger population, it may be useful to think about approaches in terms of opt-in/opt-out.

For research efforts that require relatively small amounts of personal health information and offer the possibility of major benefits to the entire population (e.g., influenza prevalence monitoring), it is likely that most people would consent to participate. For situations in which decision makers can determine with a high degree of certainty what potential research subjects would choose, and when subjects' decisions generally would favor the common good, proceeding as though consent was an opt-out process may be reasonable.[82] Given that data re-use is easier and less costly than data re-collection, perhaps with a second informed consent process, acting without an opt-in might be appropriate, at least in some circumstances.

The arguments against the opt-in approach to consent go beyond cost and convenience. Proponents of dropping the opt-in approach argue that in a majority of cases, the adverse effects of obtaining consent are greater than the potential benefits to patients.[83] Permitting patients to exclude themselves from research introduces bias in

results, skews prevalence studies, and masks the effects of interventions, among other effects.

Also of concern is the possibility that the expected benefits of confidentiality may not be real.[83] The requirement that patients opt in to receiving marketing materials or having parts of their health record analyzed may result in their never being asked, thereby depriving them of an opportunity to obtain a benefit.

The population(s) to which patients belong is another consideration in thinking about the opt-in and non-consent approaches to data use. Depending on the type of data being sought and the planned use(s), the advantages and disadvantages of opting in or declining to participate may be different for different populations.[83] When DNA analysis is the intended use, patients who can afford treatment for conditions identified through testing may regard the research approach differently than those who cannot afford treatment.

EVOLVING ISSUES RELATED TO SECONDARY DATA USE

Many of the issues described in the previous chapter involve systems and processes that are controlled by health care providers, organizations, and payers. Even if consumers can envision the technical features that would make an electronic health record more efficient or easier for a physician to use, there is little, if anything, they can do to implement their ideas. This chapter examines several trends that consumers do influence, in some cases quite significantly.

Web developers build Internet-based applications such as Facebook, but consumers drive their growth and dictate how applications evolve, voting with their feet and their pocketbooks. In contrast to the clinical setting, when consumers engage social media they voluntarily decide how much personal information, including personal health information, to share. Although consumers report concerns about the privacy of their medical information, they demonstrate a remarkable willingness to make personal information available to others, including strangers, outside the clinical environment. In a 2008 Deloitte survey 37 percent of consumers reported an interest in online tools that can help them assess, monitor, and manage their health, and 57 percent want secure Web applications that allow them to schedule appointments, manage health information, and help them adopt a more healthful lifestyle.

These use rates suggest that at least some consumers are more than ready to trust people they have not met and commercial entities whose behavior is not easily verified. Development of trust is critical to the establishment of patient-provider relationships, electronic health care-related transactions such as laboratory test results retrieval, and general e-commerce. Abuse of trust engenders behaviors – unwillingness to participate in future electronic initiatives, actions taken to preserve privacy that create other undesirable consequences, and others – that have implications going beyond the flawed transaction. Although consumers are aware of information privacy and security challenges and take steps they believe will protect privacy, they also expect online tools to keep safe their private information. This chapter explores the intersection of social media, personal health records, online genetic testing, data ownership issues, and how these tools may affect consumers' trust in health information systems in general.

Social Media and Online Behavior

The terms "social media" and "social networking" have become cultural buzzwords for electronic tools that facilitate active or passive communication between individuals and groups. In December 2008, 35 percent of American adults had an online profile, compared to just 8 percent in 2005.[84] A June 2008 survey found that 40 percent of users connected with others about a health concern over the Internet or a mobile phone, and 35 percent used social media tools to obtain health information.[85] About a quarter of the survey respondents reported creating health-related content via one-to-one or social media tools.

Within health care, social media have precipitated a shift in how consumers and patients relate to the system. Social media shift the locus of control to the patient and change the relationships among patients and care providers.[86] The tools enhance the interactions between physicians and their patients, even when they replace face-to-face office visits. Although some physicians have expressed concern that the Internet will weaken the physician-patient relationship, nearly half (46 percent) of patients reported that Internet use has a positive effect on the patient-physician relationship.[87] Use of the Internet, including development of trust-based relationships using social media tools, prompted greater interest in their health, greater adherence to physician instructions, and improved diet.

At the same time, social media shift the communication paradigm from one-to-one to many-to-many. Where patients once learned about a condition via a consultation with a physician, and thus needed many such consultations to obtain a spectrum of opinions, they now can gain many perspectives just by joining a disease-specific electronic list or subscribing to an email newsletter. They also can obtain benefit from others' ill-gained experience, such as stories of misdiagnosis.[88] Positive experiences occurring via social media stimulate and reinforce trust in others in the online environment.

Support Groups and Behavior Change Resources

Online support groups are among the oldest and highest visibility initiatives of all social media forms in use on the Internet. Electronic bulletin boards such as Psycho-Babble[89] offered users an asynchronous way to communicate with others. Social media applications such as Facebook[90] and Twitter[91] allow users to communicate in real

time, permitting a greater sense of immediacy and intimacy. Such spontaneity comes at a price, however, in that whereas bulletin board messages can be taken down following a change of heart, Twitter messages sent and received cannot be unviewed.

Some Web-based support groups may operate in conjunction with a research or patient advocacy association. The Association of Cancer Online Resources (ACOR) hosts 150+ private and public email lists and informational resources links, most related to specific forms of cancer or to cancer issues.[92] Launched to facilitate development of cancer patient groups, ACOR has evolved into a hub that connects communities of cancer patients and caregivers with others who have related concerns but likely would not meet outside the virtual environment. Members agree not to share others' comments outside the list, but compliance cannot be monitored, and expulsion from a list is the only penalty for breaching the promise of confidentiality.

Some consumers join social networks as an expression of something they aspire to accomplish, rather than as a function of who they are. Communities centered on support for weight loss, smoking cessation, and other health practices have become readily available on the Internet, and the sites' longevity suggests that consumers have found something of value. In a year-long study of utilization of a weight loss Web site, use of social support features such as chats was the best predictor of weight loss maintenance after a six-month weight loss period.[93] Features involving objective feedback, such as progress charts and calculators, encourage users to submit personal information.

Virtual Worlds

In the popular game/simulation Second Life, participants create avatars that live in a virtual world, with the game player deciding how closely the avatar parallels the player's life. Second Life has drawn the interest of the health care industry, which now is using Second Life to extend users' real-world experience. Over time, health insurer Cigna will build out virtual seminars, interactives, games, and other features to support behavior change and healthier living.[94] Second Life also is being used as a health behavior research site. The Texas Obesity Research Center (TORC) at the University of Houston has implemented a program in which participants earn lindens (the Second Life currency) for having their Second Life avatars (virtual personas) exercise or try healthy foods.[95,96] To participate, users must provide personal health information, though the applicability of HIPAA, GINA, and other privacy regulations in cyberspace is undetermined.

Patient-Driven Research

With consumers taking greater ownership of their clinical care, it was only a matter of time before they empowered themselves to take charge of the testing of developmental therapies. The most prominent example of patient-driven research is PatientsLikeMe, an online community frequented by patients with "life-altering diseases." As with bulletin boards and e-lists, patients share personal health information, but in a structured manner that permits quantitative outcome analysis.[97] Other patients with the condition use the published analyses to make decisions about their own therapy and then post information about their progress, further growing the sample size and adding texture to the discussion.

PatientsLikeMe has developed communities for people with amyotrophic lateral sclerosis (ALS), multiple sclerosis, Parkinson's disease, and other neurological conditions; fibromyalgia; HIV infection and AIDS; and mood disorders.

To potential participants, PatientsLikeMe's analytical approach appears (and may be) objective, and its rigorous data collection process resembles the protocol-based approach used by academic medical center and government institute researchers. As the data collections grow, they may become suitable for uses unrelated to the treatment analyses for which they were established. However, the applicability of federal and state privacy statutes is unclear, despite the appearance of privacy protection, so users' privacy may not be protected.

Personal Health Records

The requirement that users have a particular condition limits participation in patient-driven initiatives such as PatientsLikeMe. However, there are no similar requirements for establishment of a personal health record, which in essence is an electronic health record created by a consumer for personal use or sharing with health care providers. The American Recovery and Reinvestment Act includes incentives for health care professionals to implement electronic health records. Some consumers have already taken the initiative by creating online health records of their own known as personal health records (PHR). According to a Deloitte survey, last year 9 percent of consumers had a PHR, up from 8 percent in 2007, and 42 percent report wanting a PHR.[67] PHRs were developed so that consumers could store all relevant information about themselves and their family in one central location accessible via the Web.

Some attributes of the PHR include control by the individual who created the record; portability for access anytime, anywhere; and assumed privacy and security of the record.[98] A number of PHRs are being developed, including one by the search engine Google (Google Health PHR)[99] and one by Microsoft (HealthVault).[100] Proponents note that PHRs may improve health care efficiency,[101] such as by making data available to multiple insurers for claims processing or by documenting information not typically recorded in an EHR.

Although PHRs have benefits, there are risks associated with their use. Online health storage of data is not covered under HIPAA, and promises about privacy made by PHR vendors may lack legal force, potentially leaving users unprotected.[102] If the company that developed a PHR sells it to another company, the new owner may not be bound to follow the privacy policy set out by the previous owner, creating the possibility that consumers' information will be subject to unintended secondary use, perhaps without their knowledge or consent.

Online Marketing of Genetic Tests

Use of both social media and personal health records is increasing, and this trend seems unlikely to reverse direction. A significant number of Americans make personal health information available via social media applications, which typically provide minimal privacy protection. Consumers are gaining awareness of the benefits of personal health records, which offer greater privacy protection but have not been widely adopted yet. As people become more comfortable sharing personal health information over the Internet and more aware of the effect of genetics on health, they become more likely to consider

obtaining genetic testing over the Internet. This, in turn, creates further opportunities for unintended disclosure of an individual's genetic information.

Within the health care industry, direct-to-consumer advertising (DTCA) refers to the promotion of products available by prescription.[103] Consumers have become all too familiar with DTCA of prescription drugs for such conditions as erectile dysfunction, hay fever, high blood cholesterol, and chemotherapy-related fatigue. However, consumers are less familiar with DTCA for genetic tests. A survey of residents of Oregon, Utah, and Michigan conducted by the National Office of Public Health Genomics indicated awareness of direct-to-consumer tests ranging from 24.4 percent (Oregon) to 7.6 percent (Michigan), and less than 1 percent had taken such a test.[104]

Genetic testing is among the latest health-related technologies to be offered to consumers via the Internet. In August of 2008, the state of California licensed two companies, 23andMe and Navigenics, to provide online DTC genetic testing in the state.[105] The following month 23andMe announced that it had reduced the cost of gene mapping from $999 to $399, citing a desire to generate business and expand its database of genetic profiles, which the company would like to make available for research.[106]

Though the U.S. Food and Drug Administration (FDA) has published several documents intended to provide guidance about acceptable marketing practices,[107] consumer advocacy groups continue to express concerns about the accuracy and appropriateness of DTCA.[108] Initial DTCA initiatives focused on pharmaceutical products, but over time manufacturers of genetic tests and other diagnostic procedures began marketing their services directly to consumers. A 2003 survey of family physicians' attitudes suggested that physicians recognize that DTCA can have both

positive and negative effects on the physician-patient relationship and on the quality of care.[109] Among patients, the evidence is mixed. Concerns about breast cancer did not increase significantly among women at higher risk for breast and ovarian cancers after exposure to ads for genetic tests, though interest in being tested rose among women with all levels of risk.[110] However, another study indicated that women who are exposed to information about the risks of genetic testing for the *BRCA* breast cancer gene feel more negatively about testing and report less interest in being tested.[111]

Nonstandard Process

Generally, tests can be ordered by an individual, though individuals may be required to talk with a genetic counselor or family physician. Some companies include a telephone counseling session with a genetic counselor with the service, whereas others charge the consumer an additional fee to speak with a counselor. Testing services typically advise consumers to discuss the results of the test(s) with their physician, though most also provide information to help the consumer and physician interpret the test results.[112]

Tests promoted on the Internet address a range of concerns, such as determining the risk of developing a specific condition, establishing identity, confirming paternity, banking DNA, or analyzing nutritional status. The cost of genetic tests varies widely; in 2003, a home-use DNA banking home cost $14.95, while mitochondrial DNA maternity testing cost $3,200.[113] Five years later, a basic gene scan cost about $100, and $350,000 would buy complete sequencing of an individual's genome with medical counseling.[114]

Genetic tests are classified as those for diagnosis (tests for single genes that have been associated with specific conditions); risk assessment (tests performed to determine the likelihood of developing a particular condition); or enhancement (tests that provide information related to health or nutritional status).[115] In an evaluation of 24 companies, 13 promoted diagnostic tests, 13 promoted tests for risk assessment, and 10 companies offered genetic tests for enhancement (some companies offered multiple tests, with different tests in different categories). About half the companies required consumers to engage with their personal physician in some way, such as by having the physician collect the tissue sample for submission to the company or by sending the test results to the physician for discussion with the patient. An earlier study reported that 4 of 14 sites marketing genetic testing services directly to consumers required that test results be sent to a physician,[113] though it is unclear whether the physician had to be one with whom the ordering consumer had an established relationship. Both investigators noted that companies providing the least clinically valid tests, those for health enhancement, were least likely to require physician involvement.

The American College of Medical Genetics recommends in its policy statement on DTCA for genetic tests that a knowledgeable professional be involved in ordering genetic tests and interpreting their results.[116] However, in many cases researchers and physicians do not really understand what a genetic test result means for a patient's health.[112] Genetic tests look at genotypes that influence the likelihood that an individual will develop a condition (rather than at the genes themselves), and can't take into account lifestyle factors that may have much more significant impact.

Potential Problems

The clinical utility of tests marketed over the Internet to consumers remains a primary concern for opponents of DTCA for genetic testing services.[117] A test may be able to identify the gene it is purported to identify, but will be meaningful to consumers only if the gene is implicated in development of the condition of interest. Similarly, tests that aren't sensitive and specific enough to reliably identify the gene of interest also don't provide useful information, even if the gene has been shown to affect development of disease.[118] The possibility that tests may lack clinical utility is high because FDA does not enforce regulations related to laboratory genetic testing services, though it is involved in monitoring tests sold at retail.[114]

The American College of Medical Genetics affirms the importance of administration of genetic tests by a trained medical geneticist and interpretation of test results with a trained genetic counselor for consumers undertaking testing,[116] but this recommendation is frequently not met when consumers purchase genetic testing services online. For purposes of oversight, genetic tests are not different from other types of laboratory tests, and as with other tests, quality improvement processes that ensure consistent delivery to appropriate patients are needed.[119]

DTCA opponents also raise the potential for privacy breaches. Brick-and-mortar lab facilities, as well as physicians' offices, are subject to the privacy mandates embodied in HIPAA, which offers consumers a measure of protection. Web-based genetic testing services are not subject to HIPAA, leaving consumers open to the possibility that their personal and genetic information may be used in unanticipated and unintended ways.[120] Though facilities may advertise that they are HIPAA-compliant, consumers

have little opportunity to ascertain that a laboratory's claim is true when purchasing services over the Internet.[112]

Data Control, Ownership, and Stewardship

The establishment of social media, personal health records, and online genetic testing services opens up new possibilities for consumers to create and share personal health information outside traditional health care settings. As providers and health care institutions adopt EHRs, the opportunity for information sharing will further increase, with patients being able to download test results and other information into their PHRs. The research possibilities conferred by the shift to an electronic health care environment, though promising, depend upon the availability of data. Without the ability to access patients' health records for secondary data uses, the obstacles that slow research in today's paper chart-based world will hamper EHR-based research. Unless access to data can be assured, the technological advances currently underway and future work to be funded by ARRA will not result in the maximum possible improvement in clinical care or the optimal patient outcomes. Changes in the control and ownership of the information in patient heath records are crucial to the future of EHR-based research.

State laws influence the ownership of patient information contained in medical records, so there is some variation in consumers' rights to their data. Information contained within electronic records lacks the physical property of paper records, and thus, issues of ownership are less clear. No statute provides defines specifically the rights of patients and providers with regard to information about a patient that is stored and maintained over time.

Functionally, *data ownership* is practiced by the entity that holds data and decides who may see it and how it is to be used.[121] Historically, physicians and other providers who created paper charts owned those records and were responsible for managing the information contained in them and maintaining patients' privacy as mandated by law. Digital records, however, are different in that there is no physical medium to own or manage. Patients have rights related to the use of the information.

As the electronic health care environment evolves, this perspective may be changing. In a survey conducted by the Healthcare Information and Management Systems Society, 92 percent of health IT workers said patients should own the data they place into PHRs and 1 percent said the PHR manufacturer should own the data.[122]

Data stewardship refers to the management of data that (presumably) belongs to others.[121] It involves the development of a trusting relationship between patients and providers in which patients come to trust that providers will proactively manage data to avoid breaches of privacy or other negative effects, and thus allow providers to use data for other purposes.

The American Health Information Management Association (AHIMA) has taken the position that control over personal health information must rest with the person to whom the information relates, which in most cases will be the patient.[121] The patient's right of control should be limited, however, creating a balance of power with providers and facilitating the performance of some activities involving secondary uses of data. AHIMA has published a list of 15 actions it believes will lead to responsible handling of health information. The suggested actions emphasize development of enforceable policies

defining patient-centered approaches to data use, procedures for handling data requests and investigating potential misuse, and proactive engagement with patients.

The National Committee on Vital and Health Statistics (NCVHS) supported the strategy of limited control of data by patients in a February 2008 letter to the secretary of HHS.[123] The recommendation was made following four years of deliberation regarding the best way to manage data in the Nationwide Health Information Network (NHIN), a national data exchange. NCVHS recommended that HHS permit patients limited control over the disclosure of sensitive health information for treatment purposes, taking into consideration patients' concerns about privacy and confidentiality, how best to engender trust and encourage participation in NHIN, and promote integrity within the health care system. NCVHS proposed to achieve these ends through sequestration of certain categories of sensitive patient information (e.g., psychiatric treatment information, history of domestic abuse) that patients could choose to not share.

The approach suggested by AHIMA and NCVHS has benefits for both patients and clinicians/researchers. However, it addresses data control rather than data ownership. As EHRs are implemented throughout the United States, the approach taken will need to address data ownership as well as data control.

APPROACHES TO SECONDARY USE OF GENETIC DATA

In the preceding chapters, this report has defined concepts relevant to use of genetic information for health care quality research, described some key differences between personal medical information and genetic data, and explored some of the most prominent challenges associated with management and use of genetic data. This chapter briefly assesses the effect of legislation on genetic privacy and identifies systems-based and process strategies that will support quality research involving patient-specific genetic data.

Consumer Perceptions and Expectations

The way consumers and patients perceive electronic health records, personal information and privacy issues, and evolving events such as the incorporation of social media into daily life affect the way they interact with the health care system and whether they will allow data about them to be used. Consumer surveys and studies indicate that although Americans have privacy-related concerns, they also see opportunities and are open to change. About one-third of Americans are very concerned about privacy,[67] though consumers generally trust their health care providers and hospitals to protect the privacy and confidentiality of their medical records,[29] though a small number believe their personal health information or that of a family member has been lost or stolen.[38],

Americans are divided about whether HIPAA privacy and security rules are adequate; 54 percent believe that they are, and 34 percent believe they are not.[39] Confidence in EHRs is not high, with three times more Americans believing that electronic records are more easily stolen than paper records.[38] Consumers regard psychiatric genetic testing positively when treatment is available for identified conditions or conditions to which they are predisposed, but are less enthusiastic when there is no treatment for a condition.[77] Thirty-nine percent have a general interest in screening and genetic testing.[67]

Americans also report clear expectations related to health care and the electronic health care environment. Fifty-seven percent want to use access their medical records and perform basic health care-related functions (e.g., appointment scheduling) online, and almost as many (42 percent) want an EHR that is connected to their physician's office.[67] About a third of American adults report being interested in online tools that can help them assess, monitor and manage their health,[67] about the same number have used social media tools to access health information.[85] In 2008, 35 percent of American adults reported having an online profile of themselves, suggesting a willingness to share personal information under at least some circumstances.[84] Although the popular media regularly reports on privacy breaches and general events involving criminal activity or fraud related to use of the Internet, consumers remain interested in using the Internet to facilitate improved health and the provision of health care services. This interest suggest that consumers may be willing, rather than reluctant, participants in health care research, including quality research involving secondary use of personal data.

Impact of Legislation

Several federal laws influence the management and use of medical information, including genetic information. Beyond those laws lay dozens more state laws, which lie outside the scope of this project. At present, it is difficult to predict how the forthcoming regulations for the Genetic Information Nondiscrimination Act and the health information technology-related provisions of the American Recovery and Reinvestment Act will affect the evolution of genetic information management and use within medical records. GINA, certainly, will play a significant, if not major, role.

GINA specifies a limited number of situations in which individuals' genetic profiles may not be taken into account. In the near future, HHS, Department of Labor, Department of the Treasury, and EEOC will publish regulations describing impermissible behavior, establishing procedures for investigating possible violations, and defining penalties.[124] Though the measure goes further toward protecting Americans than any previous federal legislation, GINA will function as a remedy after genetic discrimination has been proven rather than as a preventive protection against discrimination. Filing a complaint and obtaining a decision that a violation has occurred will take time and will not restore individuals' previous level of genetic privacy. In effect, GINA offers the possibility of victory without the promise of remedy.

GINA's effectiveness also is limited by logistical considerations. Although GINA's authors intended to protect individuals against discrimination in the health insurance market, true protection against genetic discrimination is unlikely, if not impossible, in an insurance market based on actuarial underwriting.[125] The United States heath

insurance market is based on the premise that neither the insurer nor the insured have information about the insured's future health status not known to both. The availability of genetic testing services online makes it possible for consumers to obtain information relevant to their future health outside the claims/health records review process. Thus, health insurers have financial incentive to seek and use genetic information about insured persons and insurance policy applicants despite legislation that discrimination based on this information.

Given that GINA cannot restore the genetic privacy of individuals who experience privacy breaches, and that the implementation of electronic health records in the United States is imminent, development of personal information protection strategies and tools that reduce the likelihood of unintended sharing of personal information is appropriate. This chapter will present some data management strategies and tools that may mitigate the unintended and potentially harmful consequences associated with inclusion of individuals' genetic information in their medical records.

Approaches to Genetic Data Management and Use in EHRs

The surveys and studies of consumer beliefs and behavior suggest that individuals may be open to use of personal information, including genetic data, for secondary purposes such as quality research. However, consumers also are aware of previous privacy breaches that exposed patients to potential harm, and have expressed concern over inadequate protection of their information. Thus, if researchers wish to obtain or retain access to personally identifiable health information, it is in the health care industry's best interest to continuously improve the management of privacy-preserving and security-

sustaining technology. There are many ways to achieve this objective; this section focuses on systems-based and health care process-based approaches because medical informaticians can exercise some control in these areas. Changes to federal and state legislation and in consumer behavior might promote additional improvements in privacy and security maintenance, but informaticians can influence those areas only indirectly, so they will not be considered here.

Systems-Based Approaches

Data Masking. Members of the American Health Information Community's Personalized Health Care Work Group have advocated data masking – control of access – as a strategy for consumers to control access to certain types of information in their EHR, including genetic information.[126] The authors do not believe that genetic information should be treated differently from other types of medical information, but frame this approach as a feasible strategy should policymakers adopt data masking for other types of health information considered sensitive (e.g., psychiatric history, illegal drug use). Consumers may appreciate the opportunity to control access to their genetic information, and masking could prevent the sharing of genetic information with other providers, such as occurs when a patent signs a consent form permitting his or her physician to send a medical record to other providers or to insurers. However, consumers also may experience lower quality of care if the masked information is relevant to treatment. If masking of genetic data is offered as an option, consumers should undergo an informed consent process similar to those used in clinical trials so they understand the benefits and risks of masking genetic information before making a decision.

Sequestration. Data sequestration, as recommended by NCVHS, would prevent providers from seeing any patient information in categories selected by the patient. Patients could give individual providers the option of reviewing sequestered types of information, which would permit some flexibility for patients receiving care from many physicians. For example, the patient could permit the primary care physician and psychiatrist to access cytochrome P450 enzyme test results to facilitate antidepressant prescription while restricting the podiatrist's access to the test results. EHRs could be designed to indicate categories of information that had been sequestered, allowing a provider to know when genetic information that may be relevant to clinical care exists and inquire about it. If the provider has a specific concern, he or she could share that concern with the patient, who then could permit access to the information, if desired. EHRs could be designed to permit access by any provider during emergency care without patient consent and resequestration following emergency treatment. Technical approaches to data sequestration already are under development in Canada, as well as in the Netherlands and England, so this approach may have the added advantage of timely availability.[123]

Improved Data De-identification Tools. Researchers, including those involved in health care quality research, use data de-identification techniques to reduce the likelihood that individual patients will be identified among data aggregations. However, re-identification is often possible, and is sometimes intended, depending upon the data set and the purpose of the research. Development and use of data de-identification methods that make data re-identification more difficult may reduce patients' concerns about permitting secondary

data uses and thereby facilitate health care quality research. For example, the technique k-unlinkability allows administrators to specify the necessary degree of anonymity needed to prevent re-identification of DNA records using office visit location data, which typically is included in the patient record.[127] The IdentiFamily tool can link de-identified records of a family to named persons in public genealogy records, thereby permitting clinicians to consider pedigree anonymity before disclosure. In one trial, IdentiFamily indicated that approximately 70 percent of the population can be identified using public records, suggesting that many individuals' data could be de-identified when genetic testing and genetic data inclusion in EHRs and PHRs become commonplace.[128]

Process-Based Approaches

Personally Controlled Health Record Model. In the personally controlled health care record (PCHR) model, patients are able to merge medical information from multiple sources into a single record and then authorize access to the record or parts of the record by individual providers, health plans, family members, or automated health-related applications (e.g., pharmacy refill reminders).[129] Patients also can sign up to be alerted to clinical trials, disease management programs, support groups, or other initiatives of interest, and can perform related functions such as signing up for a class. The technologies used in a PCHR are similar to those now used in PHRs and EHRs; this approach differs in that the medical record may be owned by a party other than the health care providers.

Formalized Provider Accreditation. Consumers' development of trust is a critical requirement for both electronic systems and those who use such systems, whether they be electronic health records, social media, or other applications. Robust systems containing multiple barriers to privacy breaches such as passwords, encryption, and de-identification are only as sound as the individuals who use them. A formal confidentiality accreditation program certifying the staff's competence in protecting privacy may encourage consumers to share personal information they wish to remain confidential, particularly in research settings.[130] Given that virtually all academic medical centers and research institutions already require staff to complete privacy, mutual respect, and similar trainings prior to the start of employment, implementation of an accreditation program may not require significant resources beyond those already committed to such training. Health care institutions will need to communicate the scope and depth of accreditation efforts to consumers, which likely will require additional resources.

Genetic Test Vendor Accreditation. At present, the sale of genetic tests online is not regulated by FDA or other governmental body, and vendors of such tests are not required to complete and specific training related to privacy and confidentiality. Depending upon the terms of the contract, vendors also may not be subject to HIPAA and other privacy regulations. At present, consumers may purchase genetic testing services and place the results in PHRs unaware that the company performing the test(s) may not adhere to the same privacy practices as the PHR developer or the physician providing other information placed into the PHR. Vendors may strengthen consumers' trust by ensuring that employees complete accreditation programs like those used by health care providers,

adhering to consumer-friendly privacy policies, and clearly communicating data storage and handling policies to consumers.

Summary

The evolution of sophisticated DNA testing techniques has made possible the collection of personal identifiable health information that previously was unavailable. Because genetic data are both descriptive of current health status and predictive of future health status, the protection of genetics-related information – genetic privacy – researchers should treat genetic data with particular attention to privacy and confidentiality. Social media, patient-driven research, personal health records, and direct-to-consumer genetic testing promote consumer behavior that may result in unintentional loss of genetic privacy. Because consumers often cannot foresee the privacy-related implications of their choices, researchers and clinicians who collect health information have a responsibility to develop systems and processes that preserve patients' genetic privacy.

The way forward for integrating genetic data into electronic health records and personal health records will require thoughtful regulation and effective enforcement of GINA, HIPAA, and other laws relevant to genetic privacy; provision of compelling benefits for creating PHRs (EHRs won't be optional under ARRA); and implementation of secure, functional EHRs and related systems. When these objectives have been achieved, clinicians and health care quality researchers can begin working toward the level of consumer trust necessary to conduct meaningful research benefiting both patients and providers.

REFERENCES

[1] National Human Genome Research Institute. A brief guide to genomics. National Institutes of Health, 2008.

[2] Wikipedia. DNA sequencing. 2008 [cited 2008 Mar 23]; Available from: http://en.wikipedia.org/wiki/DNA_sequencing

[3] National Library of Medicine Genetics Home Reference. Cystic fibrosis. 2008 [cited 2008 Mar 23]; Available from: http://ghr.nlm.nih.gov/condition=cysticfibrosis

[4] Black JL, O'Kane DJ, Mrazek DA. The impact of CYP allelic variation on antidepressant metabolism: a review. Expert Opin Drug Metab Toxicol. 2007 Feb;3(1):21-31.

[5] Roche PA, Annas GJ. Protecting genetic privacy. Nat Rev Genet. May 2001;2(5):392-6.

[6] Gostin LO. Genetic privacy. J Law Med Ethics. 1995 Winter 1995;23(4):320-30.

[7] Troy ES. The genetic privacy act: an analysis of privacy and research concerns. J Law Med Ethics. 1997;25(4):256-72.

[8] Roche PA, Annas GJ. DNA testing, banking, and genetic privacy. N Engl J Med. 2006;355(6):545-6.

[9] Rothstein MA. Tougher laws needed to protect your genetic privacy. Scientific American. 2008 Aug 19.

[10] Rothstein MA. Genetic privacy and confidentiality: why they are so hard to protect. J Law Med Ethics. 1998;26(3):198-204.

[11] Genetic Information Nondiscrimination Act of 2008. United States of America, 2008.

[12] National Human Genome Research Institute. Genetic Information Nondiscrimination Act: 2007-2008. National Institutes of Health, 2009.

[13] Kass NE, Medley AM, Natowicz MR, Hull SC, Faden RR, Plantinga L, et al. Access to health insurance: experiences and attitudes of those with genetic versus non-genetic medical conditions. Am J Med Genet A. 2007 2007 Apr 1;143(7):707-17.

[14] Equal Employment Opportunity Commission. EEOC petitions court to ban genetic testing of railroad workers in first EEOC case challenging genetic testing under Americans with Disabilities Act. Equal Employment Opportunity Commission, 2001.

[15] National Human Genome Research Institute. Cases of genetic discrimination. National Institutes of Health, 2008.

[16] National Human Genome Research Institute. Existing federal anti-discrimination laws and how they apply to genetics. National Institutes of Health, 2008.

[17] National Human Genome Research Institute. Genetic Information Nondiscrimination Act of 2008 (fact sheet). National Institutes of Health, 2008.

[18] Freedman L. Privacy of genetic information -- issues for employers to consider. 9/18/2008 [cited 9/18/2008]; Available from: http://www.kiplinger.com/printstory.php?pid=14630

[19] Equal Employment Opportunity Commission. EEOC seeks public comment on proposed regulations implementing Genetic Information Non-Discrimination Act. Equal Employment Opportunity Commission, 2009.

[20] Wilfond B. The genetic information nondiscrimination act: fear factor or fantasy island? Hastings Cent Rep. 2008 Nov-Dec;38(6):11-2.

[21] Department of Health & Human Services. Your health information privacy rights. U.S. Department of Health & Human Services, 2008.

[22] Office for Civil Rights. Summary of the HIPAA privacy rule. U.S. Department of Health & Human Services, 2003.

[23] Department of Health & Human Services. Research repositories, databases, and the HIPAA privacy rule. U.S. Department of Health & Human Services, 2004.

[24] Department of Health & Human Services. Health services research and the HIPAA privacy rule. U.S. Department of Health & Human Services, 2005.

[25] Reuben S. Living beyond cancer: finding a new balance. Washington, DC: National Cancer Institute. May 2004.

[26] Reuben S. Assessing progress, advancing change, 2005-2006. Rockville, MD: National Cancer Institute. May 2006. Report No.: P078.

[27] Nass SJ, Levit LA, Gostin LO, (eds.). Beyond the HIPAA privacy rule: enhancing privacy, improving health through research. Washington, DC: National Academies Press, 2009.

[28] Safran C, Bloomrosen M, Hammond WE, Labkoff S, Markel-Fox S, Tang PC, et al. Toward a national framework for the secondary use of health data: an American Medical Informatics Association white paper. J Am Med Inform Assoc. 2007;14(1):1-9.

[29] Westin AF. How the public views privacy and health research. Washington, DC: Institute of Medicine, November 2007.

[30] Center for Democracy & Technology. Rethinking the role of consent in protecting health information privacy. Washington, DC: Center for Democracy & Technology, January 26, 2009.

[31] Medical News Today. Nine families sue state of Minnesota -- allege violations of state genetic privacy law in newborn screening. *Medical News Today,*.

[32] Appleby J. Identity thieves prey on patients' medical records. USA Today. May 6, 2008.

[33] Federal Trade Commission. Identity theft victim complaint data -- figures and trends January 1-December 31, 2005. In: Federal Trade Commission, ed.

[34] Synovate. Federal Trade Commission 2006 identity theft survey report. McLean, VA: Synovate November 2007.

[35] Rubenstein S. Hospitals put patients' debt up for auction. Wall Street Journal. June 3, 2008.

[36] O'Harrow Jr. R. Centers tap into personal databases. Washington Post. April 2, 2008.

[37] Lagu T, Kaufman EJ, Asch DA, Armstrong K. Content of Weblogs written by health professionals. J Gen Intern Med. 2008 October 2008;23(10):1642-6.

[38] Harris Interactive. Millions believe personal medical information has been lost or stolen. Rochester, NY.

[39] California Healthcare Foundation. Do you think HIPAA privacy and security rules are strong enough? *iHealthBeat* 2008.

[40] Felch J. DNA databases blocked from the public. Los Angeles Times. August 29, 2008.

[41] Couzin J. Whole-genome data not anonymous, challenging assumptions. Science. 2008 2008 Sep 5;321(5894):1278.

[42] Krauskopf L. WellPoint probing data breach for 130,000 members. Reuters. April 9, 2008.

[43] University of Miami. Announcement from the University of Miami. Miami, FL.

[44] Fernandez E. 6,000 UCSF patients' data got put online. San Francisco Chronicle. May 2, 2008.

[45] University of California. University of California Implementation of HIPAA Privacy Rule.

[46] Ornstein C. Celebrity-snooping ex-UCLA Medical Center staffer is indicted. *Los Angeles Times*. Los Angeles, CA.

[47] Park C. Nurse pleads guilty to privacy violation. Arkansas Democrat-Gazette. April 17, 2008.

[48] American National Standards Institute. Healthcare information technology standards panel. [cited March 28, 2009]; Available from: http://www.ansi.org/standards_activities/standards_boards_panels/hisb/hitsp.aspx?menuid=3

[49] Health Information Technology Standards Panel. Welcome to www.HITSP.org. [cited March 28, 2009]; Available from: http://www.hitsp.org/

[50] Manos D. 'HITSP is not on vacation,' chair John Halamka says. Healthcare IT News. April 8, 2009.

[51] Health Information Technology Standards Panel. HITSP and consumers. [cited; Available from: http://www.hitsp.org/consumer.aspx

[52] Health Level Seven. What is HL7? [cited April 12, 2009]; Available from: http://www.hl7.org/

[53] Centers for Medicare and Medicaid Services. Security 101 for covered entities. In: U. S. Department of Health & Human Services, ed.: U. S. Department of Health & Human Services.

[54] Centers for Medicare and Medicaid Services. HIPAA security guidance. In: U.S. Department of Health & Human Services, ed.: U.S. Department of Health & Human Services.

[55] Adam N, White T, Shafiq B, Vaidya J, He X. Privacy preserving integration of health care data. AMIA Annu Symp Proc 2007; Chicago, IL. p. 1-5.

[56] Langella S, Osler S, Hastings S, Siebenlist F, Phillips J, Ervin D, et al. The cancer biomedical informatics grid (caBIG™) security infrastructure. AMIA Annu Symp Proc 2007; Chicago, IL. p. 433-7.

[57] Fearn PA, Lafferty HJ, Bauer MJ, Kattan M. A clinical research database solution for HIPAA privacy and security requirements. MedInfo 2004; San Francisco, CA.

[58] Lenzer J. Hackers demand $10m for eight million medical records they are holding "hostage". BMJ. 2009 May 12;338(b1917).

[59] Keshavjee K, Pairaudreau N, Bhanji A. Physician office readiness for managing internet security threats. AMIA Annu Symp Proc 2006 November 11-15, 2006; Washington, DC. p. 981.

[60] HIMSS Analytics. 2008 HIMSS Analytics report: security of patient data. April 2008.

[61] Office of Inspector General. Nationwide review of the Centers for Medicare & Medicaid Services health insurance portability and accountability act of 1996 oversight. In: Office of Inspector General, ed.: U.S. Department of Health & Human Services 2008.

[62] McGirt E. $19 billion for what? Fast Company. May 2009:69.

[63] Arvantes J. FP congressional testimony addresses hidden costs of electronic health records. *AAFP News Now*.

[64] California Healthcare Foundation. Federal health IT efforts spark concerns about EHR costs, privacy. *iHealthBeat* 2009.

[65] Harris Interactive. The Harris poll #74: Millions believe personal medical information has been lost or stolen. Harris Interactive, 2008.

[66] Paek HM, Swiatek-Kelley J, O'Connell R, Brandt C. Qualitative study of patients' perceptions of safety and risk related to electronic health records in a hospital. AMIA Annu Symp Proc 2006 November 11-15, 2006; Washington, DC. p. 1054.

[67] Deloitte Center for Health Statistics. 2009 survey of health care consumers: key findings, strategic implications. Washington, DC: Deloitte Center for Health Statistics, March 2009.

[68] Hanauer DA. EMERSE: the electronic medical record search engine. AMIA Annu Symp Proc 2006; Washington, DC. p. 941.

[69] Pakhomov S, Shah N, Hanson P, Balasubramaniam S, Smith SA. Automatic quality of life prediction using electronic medical records. AMIA Annu Symp Proc 2008 2008 Nov 8-12; Washington, DC. p. 545-9.

[70] Moore BJ, Gaehde S, Curtis C. Architecture choices and challenges of integrating electronic patient questionnaires into the electronic medical record to support patient-centered care. AMIA Annu Symp Proc 2008 2008 Nov 8-12; Washington, DC. p. 490-4.

[71] Le Duff F, Muntean C, Cuggia M, Mabo P. Predicting survival causes after out of hospital cardiac arrest using data mining method. MedInfo 2004; San Francisco, CA. p. 1256-9.

[72] Tannen RL, Weiner MG, Xie D. Use of primary care electronic medical record database in drug efficacy research on cardiovascular outcomes: comparison of database and randomised controlled trial findings. BMJ. 2009 Jan 27;338(271):b81-9.

[73] Brokel J, Delaney C. Secondary data analysis of primary care practice utilization of evidence-based diabetes guidelines and patient outcomes. *MedInfo 2004*. San Francisco, CA:1537.

[74] Pakhomov S, Weston SA, Jacobsen SJ, Chute CG, Meverden R, Roger VL. Electronic medical records for clinical research: application to the identification of heart failure. Am J Manag Care. June 2007;13(6):281-8.

[75] DesRoches CM, Campbell EG, Rao SR, Donelan K, Ferris TG, Jha A, et al. Electronic health records in ambulatory care -- a national survey of physicians. N Engl J Med. 2008 Jul 3;359(1):50-60.

[76] Wilcox A, Bowes WA, Thornton SN, Narus SP. Physician use of outpatient electronic health records to improve care. AMIA Annu Symp Proc 2008 November 8-12, 2008; Washington, DC. p. 809-13.

[77] Laegsgaard MM, Kristensen AS, Mors O. Potential consumers' attitudes toward psychiatric genetic research and testing and factors influencing their intentions to test. Genet Test Mol Biomarkers. 2009 February 2009;13(1):57-65.

[78] Kush RD, Helton E, Rockhold FW, Hardison CD. Electronic health records, medical research, and the tower of babel. N Engl J Med. 2008 2008 Apr 17;358(16):1738-40.

[79] Pronovost P, Needham D, Berenholtz S, Sinopoli D, Chu D, Cosgrove S, et al. An intervention to decrease catheter-related bloodstream infections in the ICU. N Engl J Med. 2006 2006 Dec 28;355(26):2725-32.

[80] Borrer KC. Human Research Protections Under Federalwide Assurances FWA-5752, FWA-287, and FWA-3834. In: Office for Human Research Protections, ed. 2007.

[81] Davies C, Collins R. Balancing potential risks and benefits of using confidential data. BMJ. 2006 Aug 12;333(7563):349-51.

[82] Singleton P, Wadsworth M. Consent for the use of personal medical data in research. BMJ. 2006 Jul 29;333(7561):255-8.

[83] Hewison J, Haines A. Overcoming barriers to recruitment in health research. BMJ. 2006 Aug 5;333(7562):300-2.

[84] Lenhart A. Adults and social network Websites. Washington, DC: Pew Internet & American Life Project January 14, 2008.

[85] California Healthcare Foundation. Do online users create and consume health content using one-to-one and social media? *iHealthBeat* 2008.

[86] Hawn C. Take two aspirin and tweet me in the morning: how Twitter, Facebook, and other social media are reshaping health care. Health Aff. Mar-Apr 2009;28(2):361-8.

[87] Iverson SA, Howard KB, Penney BK. Impact of internet use on health-related behaviors and the patient-physician relationship: a survey-based study and review. J Am Osteopath Assoc. 2008 Dec;108(12):699-711.

[88] California Healthcare Foundation. Social networking site lets patients share stories of misdiagnosis. *iHealthBeat* 2008.

[89] Psycho-Babble. [Web site] [cited April 21, 2009]; Available from: http://www.dr-bob.org/babble/

[90] Facebook. [Web site] [cited April 17, 2009]; Available from: http://www.facebook.com/

[91] Twitter. [Web site] [cited April 17, 2009]; Available from: http://twitter.com/

[92] Association of Cancer Online Resources. Association of Cancer Online Resources. [Web site] [cited April 21, 2009]; Available from: http://www.acor.org/

[93] Krukowski RA, Harvey-Berino J, Ashikaga T, Thomas CS, Micco N. Internet-based weight control: the relationship between web features and weight loss. Telemed J E Health. 2008 Oct;14(8):775-82.

[94] California Healthcare Foundation. Cigna launches virtual health community to promote healthy lives. *iHealthBeat* 2008.

[95] Texas Obesity Research Center. TORC's international health challenge in Second Life. [Web page] [cited April 22, 2009]; Available from: http://grants.hhp.coe.uh.edu/obesity/sl.htm

[96] Colliver V. Fat people get online chance to lose weight. San Francisco Chronicle. August 8, 2008.

[97] Frost JH, Massagli MP, Wicks P, Heywood J. How the social web supports patient experimentation with a new therapy: the demand for patient-controlled and patient-centered informatics. AMIA Annu Symp Proc 2007 November 8-12, 2008; Washington, DC. p. 217-21.

[98] Ball MJ, Coslin MY. Whose record is it, anyway? Consumers bank on health. HIMSS 2006.

[99] Tanne JH. Fears over security as Google launches free electronic health records service for patients. BMJ. 2008 May 31;336:1207.

[100] Yee CM. Mayo opens private medical-data website. Minneapolis Star-Tribune. 2009 Apr 20.

[101] Anderson HJ. A major payer pushes PHRs. Health Data Management 2009 Apr 1.

[102] Steinbrook R. Personally controlled online health data -- the next big thing in medical care? N Engl J Med. 2008 Apr 17;358(16):1653-6.

[103] National Library of Medicine. What is direct-to-consumer genetic testing? In: National Library of Medicine, ed. 2006.

[104] Goddard KAB, Duquette D, Zlot A, Johnson J, Annis-Emeott A, Lee PW, et al. Public awareness and use of direct-to-consumer genetic tests: results from 3 state population-based surveys, 2006. Am J Public Health. 2009 Mar;99(3):442-5.

[105] Pollack A. California licenses 2 companies to offer gene services. New York Times. 2008 Aug 20.

[106] Kaiser Family Foundation. Personal DNA scanning service 23andMe drops price to $399. *Kaiser Daily Health Policy Report* 2008 Sep 10.

[107] Division of Drug Marketing A, and Communications,. Laws, regulations, guidances, and enforcement actions. In: Food and Drug Administration, ed.: Department of Health & Human Services,.

[108] Public Citizen. Direct-to-consumer drug ads not beneficial; federal government should do more to educate patients, Public Citizen tells senators. 2005.

[109] Murray E, Lo B, Pollack L, Donelan K, Lee K. Direct-to-consumer advertising: physicians' views of its effects on quality of care and the doctor-patient relationship. J Am Board Fam Pract. 2003 Nov-Dec;16(6):513-24.

[110] Lowery JT, Byers T, Axell L, Ku L, Jacobellis J. The impact of direct-to-consumer marketing of cancer genetic testing on women according to their genetic risk. Genet Med. 2008 Dec;10(12):888-94.

[111] Gray SW, O'Grady C, Karp L, Smith D, Schwartz JS, Hornik RC, et al. Risk information exposure and direct-to-consumer genetic testing for BRCA mutations among women with a personal or family history of breast or ovarian cancer. Cancer Epidemiol Biomarkers Prev. 2009 Apr;18(4):OF1-9.

[112] Annals of Neurology. Genome scans get personal with online consumer services. Ann Neurol. 2008 Feb;63(2):A15-7.

[113] Gollust SE, Wilfond BS, Hull SC. Direct-to-consumer sales of genetic services on the Internet Genet Med. 2003 Jul-Aug;5(4):332-7.

[114] Schmidt C. Regulators weigh risks of consumer genetic tests. Nat Biotechnol. 2008 Feb;26(2):145-6.

[115] Geransar R, Einsiedel E. Evaluating online direct-to-consumer marketing of genetic tests: informed choices or buyers beware? Genet Testing. 2008;12(1):13-23.

[116] American College of Medical Genetics. ACMG statement on direct-to-consumer genetic testing. 2008.

[117] Hunter DJ, Khoury MJ, Drazen JM. Letting the gnome out of the bottle -- will we get our wish? N Engl J Med. 2008 Jan 10;358(2):105-7.

[118] National Human Genome Research Institute. Promoting safe and effective genetic testing in the United States. In: National Institutes of Health, ed.: National Institutes of Health,.

[119] Secretary's Advisory Committee on Genetics H, and Society,. U.S. system of oversight of genetic testing: a response to the charge of the secretary of health and human services. In: Department of Health & Human Services, ed.: Department of Health & Human Services.

[120] Wolfberg AJ. Genes on the web -- direct-to-consumer marketing of genetic testing. N Engl J Med. 2006 Aug 10;355(6):543-5.

[121] American Health Information Management Association. Health data access, use, and control. J AHIMA. 2007 May;78(5):63-6.

[122] Health Information and Management Systems Society. Personal health records: Health Information and Management Systems Society 2008 May

[123] Cohn SP. Re: Individual control of sensitive health information accessible via the Nationwide Health Information Network for purposes of treatment. In: National Committee on Vital and Health Statistics, ed.: Department of Health & Human Resources, 2008.

[124] Department of Health & Human Services. “GINA” -- The Genetic Information Nondiscrimination Act of 2008 Information for Researchers and Health Care Professionals. In: Department of Health & Human Services, ed. 2009.

[125] Rothstein MA. Is GINA worth the wait? J Law Med Ethics. 2008 Spring;36(1):174-8.

[126] McGuire AL, Fisher R, Cusenza P, Hudson K, Rothstein MA, McGraw D, et al. Confidentiality, privacy, and security of genetic and genomic test information in electronic health records: points to consider. Genet Med. 2008 Jul;10(7):495-9.

[127] Malin B. A computational model to protect patient data from location-based re-identification. Artif Intell Med. 2007 July;40(3):223-39.

[128] Malin B. Re-identification of familial database records. AMIA 2006 Annual Symposium 2006 Nov 11-15; Washington, DC. p. 524-8.

[129] Mandl KD, Kohane IS. Tectonic shifts in the health information economy. N Engl J Med. 2008 Apr 17;358(16):1732-7.

[130] Kalra D, Gertz R, Singleton P, Inskip HM. Confidentiality of personal health information used for research. BMJ. 2006 Jul 22;333(7560):196-8.