Steve Whitaker
4/10/19
VHIE

CONFIDENTIALITY, ELECTRONIC HEALTH RECORDS, AND THE CLINICIAN

STUART GRAVES

ABSTRACT The advent of electronic health records (EHRs) to improve access and enable research in the everyday clinical world has simultaneously made medical information much more vulnerable to illicit, non-beneficent uses. This wealth of identified, aggregated data has and will attract attacks by domestic governments for surveillance and protection, foreign governments for espionage and sabotage, organized crime for illegal profits, and large corporations for "legal" profits. Against these powers with almost unlimited resources no security scheme is likely to prevail, so the design of such systems should include appropriate security measures. Unlike paper records, where the person maintaining and controlling the existence of the records also controls access to them, these two functions can be separated for EHRs. By giving physical control over access to individual records to their individual owners, the aggregate is dismantled, thereby protecting the nation's identified health information from large-scale data mining or tampering. Control over the existence and integrity of all the records—yet without the ability to examine their contents—would be left with larger institutions. This article discusses the implications of all of the above for the role of the clinician in assuring confidentiality (a cornerstone of clinical practice), for research and everyday practice, and for current security designs.

FROM ANTIQUITY TO THE PRESENT the ability of clinicians to assure confidentiality has been a cornerstone of practice. Though the expectations and emphases of the various ethical codes and laws concerning confidentiality have evolved over time, it has always been the practitioner's responsibility to observe them. The use of computers for the generation and storing of individual medical records is a significant change from our current paper-based records. That

Washington County Mental Health Services, Inc., 885 South Barre Road, Barre, VT 05670.
E-mail: stuartg@wcmhs.org.

Perspectives in Biology and Medicine, volume 56, number 1 (winter 2013):105–25

change makes the security of records a technological problem generally outside the realm of physician knowledge or control.

THE PROBLEM

Researchers and vendors first began developing electronic health records (EHRs) in the late 1960s to 1980s. "Peripherals" were hardwired to "mainframes," and computing power was a limiting factor. During the 1970s and 1980s, the internet was only nascent and primarily government sponsored for research, education, and government uses. Independent commercial networks not needing to use the government's National Science Foundation Network (NSFNet) backbone did not develop until the early 1990s. The degree of connectivity that is commonplace today was not easily or commonly available then, and as a consequence it did not enter into design considerations for EHRs. The focus was on the individual record and copying into electronic form the information flow functions of a hospital or clinic that the record reflected. Consequently the storage of records was not a focus of concern, and most likely it was taken for granted that the EHRs would be possessed by and reside in record room–like places, just as paper-based records did. If one were to explicitly state the design assumption that those technologic times produced, it would be something like this: one can transpose the idea, or the logical construct, of a record room onto the organization of collections of digital records, and therefore the same principles that secure a record room will provide adequate security for collections of EHRs, and thus confidentiality for individual patients. That, this article will argue, was a faulty assumption, and it led directly to clinicians' current inability to assure their patients of confidentiality.

Paper records derive much of their security from their physical nature. The frustrations we clinicians or patient-observing clinicians have all experienced trying to find the information we need in a voluminous, precariously held together stack of paper called a chart, illustrates the point. The difficulty paper records inherently present in searching for and retrieving information, even in the care of an individual patient, protects the information in them. Because of that, security is a relatively simple matter of locking up the filing cabinet, locking up the record room, and insuring the patient has given us a signed release of information form before we copy something and send it along to another record room. We all understand how this works, and how to do it. Breaches in security occur one patient at a time, from such things as a clinician checking up on her daughter's boyfriend, or political operatives engaged in "dirty tricks." It is doubtful there has ever been a case of gangsters backing a truck up to a record room and heisting the whole lot.

Of course, it is the inherent obstinacy of paper records in divulging their secrets that has led us to yearn for what digital data seems so easily to produce: information instantly at our fingertips. The fulfillment of that desire will undoubtedly be of great medical value. Not only will we be able to abstract the information when and as it is needed by a particular patient for a particular clinical issue, but we will also be able to search for and tabulate information across EHRs. The suffering that we as clinicians observe and record seven days a week, 24 hours a day will no longer be lost to us, moldering away in a record room, but will be available to us for learning—analyzing our recorded experience. We are on the verge of inventing a new tool of clinical investigation. Much like the invention of the microscope and the x-ray machine, this new tool will enable us to "see" things that are right in front of us that we have never seen before.

It is the accessibility of the whole body of records, so potentially useful for research, that will unfortunately also make the records valuable for many other purposes. And the larger the aggregate, the greater its value will be. Whether one imagines a single record for a person from birth to death that may be accessed from any location, or imagines our usual scattering of records over time in different offices, hospitals, and clinics that are then pulled together by a health information exchange (HIE), one has imagined a virtual single record room. The searchable nature of an EHR, and the interconnections between the physical locations of EHRs, create a degree of accessibility—and therefore value of the data—that will routinely attract attackers. The problem for EHRs becomes how to reserve their bountifully accessible information only for those fingertips dedicated to its beneficent clinical use, rather than for those pursuing money and power.

THE CURRENT SOLUTION

Since the 1970s, there has been an intricate history of research, vendor competition, and acquisitions, as well as gradually evolving standards for EHRs. The standards process in the United States has culminated in the Office of the National Coordinator for Health Information Technology (the ONC), created by executive order in 2004, and then legislatively mandated by the Health Information Technology for Economic and Clinical Health (HITECH) Act of 2009. The ONC is charged with "promoting the development of a nationwide health information technology (IT) infrastructure that allows for electronic use and exchange of information," which, among other things, "ensures secure and protected patient health information." To do this, the ONC designates minimum capabilities an EHR must make available for "users" to meet their assessed security needs. The ONC is at pains to make clear that it is the purchaser/owner/user of an EHR who has the responsibility for assessing security needs, deciding which available security measures to use, and ensuring their proper implementation and function at their facilities. (However, "meaningful use" criteria set by the Center for Medicare and Medicaid [CMS] may in turn encourage one to use or not use one or the other of available security measures.) The ONC further specifies that the criteria by which one is to assess security needs are the privacy and security rules of the Health Insurance Portability and Accountability Act (HIPAA) of 1996.

The HIPAA privacy rules are meant to apply to health information in any medium, and the security rules to electronic information. Reading them makes clear that the underlying (and unaddressed) assumption of these various rules is that they are codifying business as usual to make "portability" more possible. Practitioners, clinics, hospitals, agencies, nursing homes, pharmacies, urgent care centers, ambulance companies, insurance companies, health information exchanges (HIEs), "clearinghouses," and so forth are assumed to possess and protect the patient information—and thus the pre-internet design assumptions of the 1970s prevail. The venerable medical record room is simply transposed on to its envisioned and still developing modern electronic forms.

The security rules are divided into physical, technical, and administrative categories. The physical rules have to do, as one would imagine, with securing the hardware, such as the machine room and the "work stations." Since one does not actually travel to the machine room to pick up a hard disk and read it, as one would a paper record in the record room, this set of rules, though necessary, has the least to do with controlling actual access to the information. The technical rules are meant to put in place data and network security tools that will support the permitted and not permitted "disclosure rules" of a later section. These tools are such things as monitoring and allowing or disallowing traffic based on who it is coming from and its contents ("firewalls"), logs, and identity-, role-, and location-dependent access to different categories of data in a patient's record. One can see in these the analogs to procedures used in record rooms. These rules are categorized as either "required" or "addressable," meaning they either *must* be done or *may* be done. It is worth noting that encryption, the analog of locking the record room, is considered "addressable." The administrative rules have to do with creating a security plan for routine operation and various emergencies (utilizing the kinds of tools listed in the technical section, and the actual tools in the EHR required by the ONC), implementing the plan, monitoring the plan, evaluating and updating the plan, and creating and maintaining the secure work force to do all of that. Here it is worth noting that all this implies a not inconsiderable group of administrative personnel, with not inconsiderable technical knowledge.

The "Security Standards: General Rules" section of the HIPAA security rules sets out what must be considered in developing a security plan for one's organization. Among the considerations is that one is to "protect against reasonably anticipated threats or hazards to the security or integrity of such information" (HIPAA 2006). The next section gives examples of threats that can reasonably be anticipated, and it will be seen that simply mapping the record room and its protections on to networks and computers can only be deemed acceptable security if the "reasonably anticipated threats" are in fact ignored.

The Attackers: Not Only Anticipated, But Here

Potential attackers who have and will continue to target personal data such as EHRs include: domestic governments, for surveillance and "protection"; foreign governments, for espionage and sabotage; organized crime, for illegal profits; and large corporations, for "legal" profits from data mining and related activities. These groups are attracted to any large database of identified personal information, and as our electronic medical information databases grow, so will the attraction.

An attack is one party intentionally taking information from another party against its will or without permission. These are not examples of unintentional loss of information through lapses by its owner, although the unintentional lapses are already monumental—for example, a 2011 headline in the *Boston Herald* announced that "Mass. Data Breaches Strike 5 Million" (Turner 2011)—and what one mostly reads about, though even these are surely not reported to their full extent. It is a completely different problem to admit publicly that one's security has been (and therefore probably continues to be) inadequate to protect the data that the entity has been entrusted with (who would want to trust that entity in the future?), or to reveal that one knows how to overcome another's security (and thus destroy one's own ill-gotten but treasured advantage over a competitor). It is impossible to know the extent of these problems, but what is publicly known is likely only the tip of the iceberg.

Domestic Governments

An infamous U.S. example of misuse of personal data occurred during the Second World War. Specific law preventing use of detailed census data was changed in March 1942 to allow neighborhood and individual data to be used to aid surveillance and internment of U.S. citizens of Japanese ethnicity. The law was changed back in 1947. HIPAA includes "enforcement rules," and the Office for Civil Rights (OCR) is charged with using these rules to investigate and penalize those who do not comply with the HIPAA privacy and security measures. The thought clearly is that law can protect us from misuse of databases. However, as the WWII incident demonstrates, such laws are akin to legislation prohibiting alligators from eating ducks. It is not in our human nature to follow such laws—especially if we are threatened or, like the alligator, feeling hungry.

Another more recent example is the "warrantless wiretap" controversy of post–9/11. This grew from an initial estimate of a few hundred U.S citizens being affected, to allegations of a secret National Security Agency (NSA) monitoring and data-mining room at an AT&T fiber optic hub in San Francisco, to comments that the only way to know the size of the program would be to note the number of entities granted immunity by the 2008 Foreign Intelligence Surveillance Act (FISA) Amendments Act—in other words, internet companies such as Google. William Binney, a former senior NSA mathematician who resigned in 2001, described the secret project Stellar Wind, which has 10 to 20 such fiber-optic monitoring stations about the nation, as well as satellite monitoring. "They violated the Constitution setting it up," he says bluntly (Bamford 2012). Notable in this example is the amount, and therefore the value, of the data available (20 terabytes a minute, according to Binney), the threat, the action by government without supporting law, the inability or unwillingness of private partners to resist the government, and the government's ability to make the actions legal retroactively.

Another example involves medical data from my home state of Vermont. In 2006, legislation was passed to create a registry of prescriptions for controlled substances. The registry was to be used to help physicians identify, and thus get into treatment, those suffering from substance abuse. A registry is basically a slice of data from medical records created to allow searching and sorting in the way that paper records don't, and precisely as we hope EHRs will be able to. It was clear to the legislative leaders involved in 2006 that the law would not have been passed if access had not been limited to physicians: if police had also been allowed access it would not have passed. However, in 2012, legislation supported by the governor (who supports EHRs) was introduced to allow the police access to the database for anything they deem a "bona fide investigation"—that is, access without obtaining a warrant. To date the outcome is hanging in the balance. In this example we have the following components: a medical database of identified personal health information of many individuals (though notably much smaller than that of the EHRs to come) created for a medical use; a threat (an epidemic of opiate abuse); and the ability of domestic government to access the medical data for nonmedical purposes (arrest and prosecution) at the stroke of a pen.

Or imagine the case of a cash-strapped state or HIE trying to make ends meet. States typically require that Medicaid data be sent to various departments for analysis. HIEs regularly pass through identified information from all sources or store it for future requests for forwarding. An HIE or a state can ease its financial worries by lopping a few identifiers off the information and selling it for "research," say to Blue Cross and Blue Shield. Such "de-identification" has many different implementations and therefore security levels. Rarely is consideration given to the problem of how long it would take an attacker with various levels of computing power to use the associations of data in the record to data outside the record to identify it. Basic security measures usually involve simply specifying which information data fields to write over (as opposed to de-indexing), so human readers might not know who they are reading about. But once this more-or-less anonymous data is in the hands of the "researcher," it is liable to being exposed again, depending on the security capabilities of the researcher and how conscientious the researcher is. HIEs of Texas are in exactly this position (DuBois 2010).

The HIE being developed in the state of Vermont has in its privacy and security policies provision for the sale of data: "If approved, Vermont Information

Technology Leaders (VITL), through an approved Data Subcontractor, shall prepare the de-identified protected health information (PHI) requested and shall be reimbursed for its expenses by the requesting party" (VITL 2010).

Foreign Governments

Before giving specific examples, it should be noted that any domestic surveillance program one creates is vulnerable to loss of control for surreptitious use by others—the technical equivalent of a mole. This in fact is one of the criticisms of the NSA program mentioned above, and there is a notorious known example of the Greek government deliberately building wire-tapping capabilities into their phone system, only to lose control of it to unknown parties from early 2004 until March 2005, before discovering they were the "listenees" and not the listeners.

The first example again comes from the state of Vermont. At the New England Chapter of the Healthcare Information and Management Systems Society (NEHIMSS) in January 2012, Chuck Podesta, Chief Information Officer (CIO) of Fletcher Allen Health Care (FAHC) a 560-bed academic medical center in Burlington, Vermont, reported that in March 2011 a computer virus, eventually identified with the help of outside experts as Pinkslipbot, was attacking their EHR, Epic. Podesta stated that McAfee, a company specializing in system security, rated Pinkslipbot as "low risk," but "we found out that means there is a low risk of getting it—but they might want to add that once you do get it, you're screwed" (Smith 2012). The virus simultaneously gains administrative rights, shuts down installed virus protection software, sends data back to its host computer in China, and receives variant code to continue concealing itself. In the end, 1,000 of 6,000 hosts were infected, each requiring 90 minutes to clean up; 700 laptops also had to be examined and fixed if necessary; and Podesta estimated the financial impact to be "in the six figures due to lost productivity." Note that without the discovery of this known virus infecting it, Epic would have continued operating while data was continually "sent home" to the virus's originator. Podesta stated that throughout the crisis, as they worked around the clock to isolate infected hosts, he hoped the loss would not affect over 500 medical records, so that the threshold for mandated reporting to the department of Health and Human Services (HHS) for posting on their website would not be exceeded. The report of his presentation did not mention whether or not FAHC's security plan had elected to use Epic's capability to support the Advanced Encryption Standard (AES) by encrypting the data while "at rest," the equivalent of locking the record room. That Podesta was worried about exceeding 500 compromised patient records implies that it was not. Nonetheless, Podesta is to be saluted for publicly relating his experiences to his colleagues as a cautionary tale.

In April 2010, a group of computer security researchers reported turning the tables as they watched Chinese attackers systematically access sensitive Indian Defense Ministry data (Markoff and Barboza 2010). Along the way the attackers also accessed NATO data in Afghanistan, prompting Rafal Rohozinski, a member of the Toronto research team, to observe: "It's not only that you're only secure as the weakest link in your network, but in an interconnected world, you're only as secure as the weakest link in the global chain of information."

Of course, the most dramatic publicly known example of inter-country sabotage in the computer world is the STUXNET virus (Sanger 2012). No doubt it will one day appear as a spear-thrower compared to its descendants. Though its goal was sabotage of a machine, there is no reason similar approaches could not be used to gather information, or to alter the integrity of information.

Janet Napolitano, Director of Homeland Security, was asked in an interview how many cyber attacks might have occurred during her 45-minute conversation. According to reporter Adam Levin: "Napolitano replied, 'Thousands.' And if that weren't enough by itself, her most ominous remark was delivered in almost desultory terms: 'I think we all have to be concerned about a network intrusion that shuts down part of the nation's infrastructure in such a fashion that it results in a loss of life'" (Levin 2011).

Organized Crime

On July 14, 2010, National Public Radio reported that 300 members of the 'Ndrangheta crime syndicate in Italy had been arrested. Among them was the "director of state medical services in the city of Pavia" (Poggioli 2010). Thus, as HIPAA states, one's security plan must consider how the administrative work force is to be secure.

According to Edward Powers, principal of Deloitte & Touche, these attackers "know what assets they're going after and they're going after you in a very orchestrated way." At a Wall Street Technology Association conference on managing the risks of running information systems in capital markets, he told technology and operations executives that "for these 'criminal elements,' cybercrime is now replacing drug trafficking as a primary source of revenue. The cybercriminals are getting quite specific about the data and the infrastructure they want to capture or control" (Steinert-Threlkeld 2011). I would only add that identified medical information is as good as money in the pocket.

Corporations

Corporations are also implicated in compromising personal data. For example, in 2011 it was reported that the data storage service Dropbox had changed its terms of service. According to *PCWorld*:

> Previously, the company stated in its terms of service that "all files stored on Dropbox servers are encrypted (AES-256) and are inaccessible without your account password." But, Dropbox has continued to modify the terms of service, and backpedal on exactly how secure customer data is—sometimes putting its foot in its proverbial mouth. After a few amendments, the terms have been altered such that it now reads more to the effect that Dropbox can access and

> view your encrypted data, and it might do so to share information with law enforcement if it is compelled, but that employees are prohibited from abusing that power and viewing customer data. (Bradley 2011)

Furthermore, it is nearly, if not completely, ubiquitous insurance company procedure to refuse to sell insurance unless the purchaser grants the company access to their medical information. (Medicare and Medicaid follow the same practice, hence the flow of data to state agencies.) This is despite the fact that the ONC privacy and security framework of 2008 states that "Individuals should be provided a reasonable opportunity and capability to make informed decisions about the collection, use, and disclosure of their individually identifiable health information" (ONC 2008). One supposes the ONC had in mind informed decisions about offers other than those one can't refuse.

The Consequences

Given our security design and likely attackers, our situation is something like this. It is as if we had each taken the valuable contents of our homes and deposited them, carefully indexed, in a huge warehouse somewhere in the country. We then give the warehouse owner and his employees careful instructions about who may access and use these valuables. Thoughtfully, the warehouse owner creates a special passage into the warehouse—a "portal"—by which we may access a few of or own possessions if we wish. The builder of the warehouse has equipped it with a very fancy lock (AES with 128- to 256-bit keys) that the owner may chose to use if he thinks he's in a bad neighborhood. Meanwhile, there are four different large, well-trained, well-equipped, and materiel-hungry armies in easy marching distance. Will the warehouse owner and his employees lay down their lives, or will they open the lock? Or will they discover their warehouse is riddled with secret back doors, or that the armies have really big bolt cutters (secret supercomputers that can do a brute force attack on AES), or that the lock has an equally fancy secret master key?

There are no key holders, or key and lock manufacturers, that can ultimately resist the power of a government or criminal syndicate or large, legal corporation. The defection of an employee to another company for a larger salary who takes along the access to proprietary information is a well-known "minor" example of the key holder problem. The problem with AT&T and the NSA is a more major example of the key holder problem. DigiNotar, a certificate-issuing authority (a kind of "key manufacturer," in that it authenticates to permit key distribution), was forced into bankruptcy after its infrastructure security was breached and fake secure sockets layer (SSL) certificates issued "into the wild," some of which were then used to spy on Iranian e-mail (Henderson 2011). Who would want to do that? In March 2011, RSA, another key and lock manufacturer, had its security breached:

> RSA, which is based in Bedford, Mass., posted an urgent message on its Web site on Thursday referring to an open letter from its chairman, Art Coviello. The letter acknowledged that the company had suffered from an intrusion Mr. Coviello described as an "advanced persistent threat." In recent years a number of United States companies and government agencies have been the victim of this type of attack, in which an intruder either exploits an unknown software vulnerability or in some way compromises the trust of an employee to take command of a computer or an entire network within a company. (Markoff 2011)

The *New York Times* article goes on to state: "Despite the lack of detail, several computer security specialists said the breach could pose a real threat to companies and government agencies who rely on the technology." According to computer security specialist Whitfield Diffie, who invented some of the cryptographic systems now widely used in electronic commerce, "One possibility is that a 'master key'—a large secret number used as part of the encryption algorithm—might have been stolen. The worst case would be that the intruder could produce cards that duplicate the ones supplied by RSA, making it possible to gain access to corporate networks and computer systems" (Markoff 2011). In a way, this kind of covert intrusion, which has a chance of being detected and repelled, is safer than a secret overt agreement as occurred with AT&T and the NSA. (Incidentally, RSA derives its name from Ronald Rivest, Adi Shamir, and Leonard Adelman, the mathematicians who found a way in 1977 to implement Whitfield Diffie's concept of asymmetric encryption.)

Then there is the problem of "back doors," or software manufactured so that it allows another party access unbeknownst to the software owner or user. A controversy about whether this occurred or not came about with an operating system called Open BSD. *Linux Journal* reported in 2010 that Theo de Raadt, OpenBSD founder and developer, "posted an email that claimed the Federal Bureau of Investigations (FBI) paid OpenBSD developers to leave backdoors in its IPsec network security stack. Since then early audits have found some questionable code, contributors denied any wrongdoing, and the original source reaffirmed his allegations" (Linton 2010).

The problem comes down to this: by simply imposing the "security structure" of a record room on interconnected digital data, we have created a gold mine of information. Between that gold mine and a great deal of raw power stand a few unfortunate people. It does not matter what armaments, technologically dazzling locks and monitoring ability, or salary we give to them. What matters is that the gold mine now exists, and those few (or perhaps worse, many) people directly guarding it or supplying the guardians will always be ducks receiving offers they can't refuse from the alligators.

A WORD ABOUT ENCRYPTION

Before discussing potential solutions to this dilemma, and their implications, it may be helpful to have some background about the lock and keys of the information age, encryption.

Throughout recorded history, encryption, or disguising the meaning of a message, has utilized a process and some device called a key to encrypt—or change a message into its disguised form, a ciphertext—and to change it back—deciphering or decrypting—to its understandable form, plaintext. Take, for example, the Spartan scytale from the fifth century BCE. The scytale was nothing more than a stick with several flattened edges on it. A strip of leather was spirally wound the length of the stick, and the message written on the leather along the flat edges. When the strip was unwound, the positions of the symbols in the message were changed (a transposition cipher), thereby hiding its meaning. The receiver of the leather strip would take an identical scytale, wind the strip around it, and read the message. In this example, the scytale is the key, and the winding, writing, unwinding, winding, and reading is the algorithm for encrypting and decrypting. A scheme in which the same key is used to encrypt and decrypt is now called "symmetrical encryption."

The Advanced Encryption Standard (AES) uses symmetrical encryption, and it is a direct descendent of the WWII rotor machines: ECM MarkII (U.S.), Enigma (Germany), and Coral (Japan). These, via Horst Fiestel at IBM in the 1960s, became a computer-based algorithm, Lucifer, which was redesigned by the NSA to become the Data Encryption Standard (DES). This in turn has become through a second NSA-directed project the Advanced Encryption Standard (AES), which was published by the National Institute of Standards and Technology (NIST) in 2001.

DES utilizes a key length of 56 bits, or 2^{56} possible keys. In 1977, Diffie and Hellman thought it possible to build a parallel machine of 1 million devices, each capable of one encryption or decryption per microsecond for an effective rate of 10^6 decryptions per microsecond. On average, it would take 10 hours to break DES (open the lock) with a brute force attack. They guessed the machine would cost $20 million in 1977 dollars. In 1998, the Electronic Frontier Foundation (EFF) built a special purpose machine for less than $250,000 that took less than three days to crack DES. This essentially rendered DES worthless. AES utilizes a 128-, 192-, or 256-bit key size. The same machine Diffie and Hellman imagined in 1977 would take, on average, 5.4×10^{18} years for a brute force attack on the 128-bit key.

In symmetrical encryption, the main problem (eventually because of expense) becomes how to get the key around to everybody who needs it, and yet keep it secret. One of the remarkable and major intellectual achievements of the late 20th century was the conception and realization of "asymmetric encryption." (In the United States, the public development of asymmetric encryption occurred in the mid- to late 1970s, about the time EHRs were first being designed, and likely would not have been of much note to those developers.) In asymmetric encryption, a mathematically related pair of keys is created.The key that encrypts a plaintext cannot decrypt it; the other one must be used to decrypt, or vice-versa. Most often one key of the pair is designated the private key (and kept private), and the other is designated the public key (and made freely available). Thus, to send me a secure message one would obtain my published public key, encrypt the message, and send it on down the line to me. I have the private key—which nobody else has—by which I can decrypt and understand the message. The key distribution problem has been pretty much solved. Wonderfully, the same scheme utilized in reverse is quite useful for authentication of messages as well. If I encrypt a message—say, my name and a date and time—with my private key and send it to somebody, then only the public key of the pair can decrypt the message, authenticating that I originated it.

One use of encryption is a set of protocols called Secure Sockets Layer (SSL), familiar to anyone who has used a credit card to buy a book over the internet from Amazon. Most people find it simple, quick, and straightforward to carry out such transactions. However, behind the scenes is a complex process involving the purchaser, Amazon, a certificate authority, asymmetric encryption protocols, and symmetric encryption protocols. It takes 13 pages in a textbook on cryptography and network security to explain it all (Stallings 2011). The speed and simplicity for the user of this extraordinary complexity is a point to be remembered as we discuss possible solutions in the next section. It is also important to realize that though the processes of authentication, key distribution, encryption, and decryption are being utilized to agree on an exchange of money and merchandise in the case of Amazon, they could as easily be used to agree upon an exchange of information.

POTENTIAL SOLUTIONS

Who would have thought that for all these centuries medicine has been so technologically advanced that it was using cloud computing? For convenient access when needed, most of us have long been keeping our medical data (history, physical exam, labs, images) on another entity's storage medium (a sheet of paper in my doctor's office).

Of course, over the centuries we have become a bit confused by this. On the one hand, we have gradually come to consider the data (the record) to belong to the physician; on the other, we have split what seemed like a legal and platonic hair for paper records and have considered the information to be the patient's, but the physical record the physician's. However, in our brave new world of digital data and cloud computing, such hair-splitting is not so crazy. It is possible to give the owner of the data actual physical control over its access (not just legal control, through a release of information form, over an agent who actually controls the physical process), and yet give actual physical control of the data's existence (integrity) to another entity. In more succinct and fancier language: control of access and existence may be dissociated. Going back to our warehouse metaphor, it is possible to have a warehouse in which the owners of the warehouse or intruders into the warehouse cannot know the content of the stored boxes or who the owners of the boxes are. The warehouse proprietors only know how many, and how big, the boxes are, while the owners of the boxes know which one is theirs, and are able to access its and only its contents. In a nutshell, the cloud computing problem is how to do that well. For example, the dropbox solution mentioned above did not do so well.

One idea, the "personal" health record, almost solves the problem, but in reality it only sometimes gets the access half of the problem right, and misses entirely the existence side of the problem. By giving ownership of the record to the patient, personal health records are also intended to give control of access to the individual. In fact, some marketed personal records manage to do that. Asymmetric encryption is used, and despite the fact that the information is not necessarily kept on a hard drive in the patient's possession, the company owning the hard drive cannot decrypt the information, and yet with the public key it is easy for others (a physician, a hospital) to add to the record.

The problem for personal health records arises with data integrity or existence. Knowing that if I have myself for a physician, I have a fool for a physician, I would do well to leave the information in my record well enough alone. However, personal health records leave that completely up to me. I can add, delete, or even completely destroy or fabricate to suit my own tastes. Any sensible hospital or physician caring for me would therefore keep a parallel record—so back we go to the record room problem. Additionally, digital personal health records lack the legal safeguards built up over the centuries for paper health records. The company on whose medium I am keeping the data may decide I'm not paying enough rent or that they want to start their own data-mining business and hold the records hostage, or they may go out of business and end the record's existence.

The personal health record almost succeeds in being a solution because in some implementations it uses encryption in which each person holds the secret or private key for his or her own record, and thus the physical continuity of the data is dismantled. Unlike locking the record room with one key, or encrypting all the records in one's possession with one key and algorithm, each individual's record is encrypted by a different key. Since each record alone is not worth very much for data mining, the strength of the encryption needed to dissuade an attack is much less than that needed for a large collection of records encrypted with one key and algorithm. The time an attacker would spend in decipherment of each independent record prohibits trying to decipher a significant number. It is as if we've gone down to that beyond imagination cavernous warehouse, and taken our possessions back to our own home where we have the keys to the outside and inside doors, and so control access.

Kristen Lauter and her colleagues have written a paper applying ideas about cloud computing security to medical records. The record remains just that, an "immutable" record of a person's interactions with health care, but the individual patient physically controls access to the record (Benaloh et al. 2009). According to Lauter and colleagues:

> We shall propose a design that we refer to as Patient Controlled Encryption (PCE) as a solution to secure and private storage of patients' medical records. PCE allows the patient to selectively share records among doctors and healthcare providers. The design of the system is based on a hierarchical encryption system. The patient's record is partitioned into a hierarchical structure, each portion of which is encrypted with a corresponding key. The patient is required to store a root secret key, from which a tree of subkeys is derived. The patient can selectively distribute subkeys for decryption of various portions of the record. The patient can also generate and distribute trapdoors for selectively searching portions of the record. Our design prevents unauthorized access to patients' medical data by data storage providers, healthcare providers, pharmaceutical companies, insurance companies, or others who have not been given the appropriate decryption keys. (p. 2)

Sadly, to the best of my knowledge, this system has not been implemented for field trials, but it is there waiting, ready to be tried out. There are other cloud computing ideas and even products. Although they are clearly not for medical records, one can see that current industrial solutions are groping their way in that direction, and could form the nidus of an approach to medical records. Take, for example, a product recently described in the trade journal *PCWorld*, which is based on the concept of the safe deposit box with two keys, one for the customer and the other for the banker, dubbed the Porticor Virtual Key Management Service:

> Just like the safe deposit box, the customer can't decrypt the data without the key held by Porticor, and Porticor can't decrypt the data without the master key held by the customer. In practice, the customer actually has one key per project, which is usually an application. Porticor has thousands of keys, one for each file or disk belonging to that project. Still, the keys must pair up in order to provide access to the encrypted data. Beyond the keys being split between the customer and Porticor, the unique part of the solution is the keys themselves are encrypted by the customer's master key, which only the customer holds and knows. As a result, Porticor holds project keys but the vendor can't read them because they are encrypted. By encrypting the "banker" keys with the customer master key, Porticor gives the customer complete mitigation of end data protection. (Musthaler 2012)

The overall effect of these techniques is to restore the security of a system in which the unauthorized intrusion into records requires a one-by-one process as it does now for paper records. If one of the four "attackers" wants your particular record for some reason, they can get it, just as they can now with paper records, but there can be no data mining for non-beneficent use of the culled information. With such a system in place, a clinician could reasonably assure patients of the confidentiality of their medical records.

IMPLICATIONS FOR PRACTICE

Trying to imagine what life might look like if the "record room" is made physically discontinuous by giving each patient control over access to his or her own record is a good way to induce panic attacks, whether one is a clinician or a system designer. This section is meant to provide some small steps of desensitization to that image by lingering with it long enough to think through some of the design implications and by sketching out approaches for a few of the commonly raised questions. Of course, the real proof or disproof of these ideas can only come through research, actual implementations of systems, and field trials. It is hoped, though, that by persevering through the initial shock of turning on its head something so seemingly commonsense as a record room, one can experience the opening up of unexpected promising avenues. By viewing the record from the vantage point of the patient rather than that of the physician, things can fall into place.

Physicians seem to live interminably hassled lives. Partly from an accretion of historical accidents, and partly from misguided approaches to utilization control, much energy and time must be spent in activities that lend little meaning to one's own or the patient's life. Consequently, even a hint at making something simple more complicated can brush a bruise, and protest vigorously follows. Appearing to ask a physician to acquire permission from each patient to peer into a record is more than a hint of complication. But although it is politically expedient to avoid this protest, it is not wise. Any number of things that look complicated in description, particularly for those that are foreign to our sensibilities, turn out in practice not to be so. The description of the inner workings of SSL is complex, but its use is simple, and it is an example of how two parties could agree to exchange information. This agreement could happen prior to face-to-face meetings, could be done by office staff, or could involve more than one clinician, office, or consultant. The fact that all this will be, by our current sensibilities, complicated under the hood does not necessarily mean it will be complicated to drive. Further, the ability to assure patients that their information will remain confidential, and that they can decide what information to give and to whom they will give it, puts us back in the familiar, tried-and-true territory of establishing rapport and trust. This is not a small prize to re-win.

The relief and excitement generated among patients and their clinicians simply at the prospect of having the same medical record if they cross the street from office to office or hospital to hospital is a wonderful promise of EHRs, but there is another prize also of untold value: research based in the real world. The exclusion criteria of randomized double-blind studies divorce them from everyday practice. The money required to stage them often makes them—surprise—the pawns of the moneyed, and putting them out to bid to offshore "research" companies can make them of dubious quality. Studies of the larger-scale "system" issues are usually done using billing data from Medicaid or Medicare, and perhaps indemnity insurance companies. This approach is akin to trying to understand the ecology and health of a forest by studying the few items from it that are of market value and found in the local village. A question that arises then is if we adopt cloud computing solutions in which individual records are individually encrypted each with a different set of keys, each held by a different person, are we not, in our zeal to make data mining impossible, making research impossible as well? The answer is not at all—but to do research, we will need to anonymize a person's data as it is created and move it into some pooled collection. How to make this a good enough one-way street to anonymity, and who will "own" the anonymous data, is beyond the scope of this article. But such a project is necessary for our own well-being; is not beyond our capabilities; is better than hospitals considering the data proprietary, and freely examining identified data believing all is well if the published results are made anonymous; and is far, far better than an HIE clearing out a few fields and selling the data.

Invariably the question of emergencies comes up. If individuals control access to their records, what do we do if they emergently can't participate? Assuming there is a record to be had, there are innumerable ways to solve this, and it is really a question for field trials. One possible solution, for example, is that some subset of the data in a person's record could be tagged as useful in an emergency, and then some biometric of the person could be used as the data to generate the key (or key pair) by which that subset is encrypted, and by which the key can be generated for decryption as well. More than one biometric could be used to guard against illness or injury altering or destroying the tissue basis for the biometric. In a sense, the person himself or herself becomes the equivalent of the Spartan scytale.

Variants of the above are key loss, and some form of chronic incapacity, such as dementia or intellectual disability or being a child. Since this is just one record we are worried about, and since most of us are small enough fish that the data has little value to anybody else, we can solve this as we do for our house: an extra key in the garden or given to one's partner, trusted friend, parent or guardian, advanced directive agent, or doctor's office. If one hasn't taken those precautions and actually does get locked out of the house, there is always, for a fee, the locksmith—and the fancier the lock, the fancier the fee. Since we are mostly attempting to discourage data mining—large-scale plunder of our homes—then the individual encryption and decryption schemes for persons could be within the reach of more-or-less everyday computers for a brute force run through of the possible keys.

There are many more questions of this sort one could work through, but they can mostly be resolved with a little patience and ingenuity. There is, however, one concern of HIEs—who must potentially ship records from all over the country—whose solution turns out to be particularly interesting and surprising: that is, the problem of a "unique identifier." HIEs are pulling out their hair trying to make sure the right record gets pulled or sent merrily off to Timbuktu. Thus, the need arises for a unique number so a technician can distinguish between the 500 Joe Millers in the country and instruct the computer to send the correct record. The truth is, though, that most of us most of the time know who we are, and where we live. I know which Joe Miller I am, and therefore where I live, without needing to know where the other 499 Joe Millers live. There is no need for me to give myself a unique identifier to know myself, and therefore if I am my own "record librarian," then there is no need to do so for my record either. I can arrange to have my record forwarded and accessed as need be without resort to numbers (of course the various symbolic names I use would eventually devolve into binary numbers). And even if I somehow manage to mistakenly send my record where I don't want it to be, I am more likely to forgive myself than I would be a technician, and if asymmetric encryption were used, it might be in the wrong place, but it would still have its meaning disguised.

The issue of a unique identifier does come up for our pooled, anonymous, research data, but it is much less terrifyingly solved there, than it is out here in identity theft land. (For our individual, identified and encrypted medical record one could truthfully say, as for a house, that its physical location and associated address is the unique identifier, but that number, should somebody else acquire it, has no other use, and leads to nothing but disguised meaning—a locked house.)

The Current Solution Revisited

The ONC has among its advisory committees one called the Privacy and Security Tiger Team. On August 19, 2010, it presented some recommendations and core values based on specific questions posed to it by the ONC (Elmore 2010). Among these were:

> 2.1: Trust Framework for Exchange Among Providers for Treatment
>
> The responsibility for maintaining the privacy and security of a patient's record rests with the patient's providers, who may delegate functions such as issuing digital credentials or verifying provider identity, as long as such delegation maintains this trust.
>
> 3.4: Consent Implementation Guidance
>
> Based on our core values, the person who has the direct, treating relationship

> with the individual, in most cases the patient's provider, holds the trust relationship and is responsible for educating the patients about how information is shared and with whom.

Core Values

> The relationship between the patient and his or her health care provider is the foundation for trust in health information exchange, particularly with respect to protecting the confidentiality of personal health information.
>
> As key agents of trust for patients, providers are responsible for maintaining the privacy and security of their patients' records.

These are tall orders for clinicians. The interlocking laws, rules, regulations, and policies found in the privacy and security rules of HIPAA, the EHR design criteria and policies of the ONC, the investigatory and protective functions of OCR, and the sticks and carrots of the CMS are simply overwhelming and disorienting compared to the function of a paper record room. The ONC, OCR, CMS, and HIPAA are like mirrors lining the inside of a circle, at whose center the clinician stands. "You da reality!" they say enthusiastically, even as the clinician him- or herself searches for an anchor in reality. "You ensure the confidentiality, and expend your patient's trust to assure him or her of that confidentiality," the mirrors instruct. However surreal, this is an understandable intuitive step on the ONC's part. Prior to the information age, a vow of maintaining confidentiality by professionals was the solution to the cloud computing problem: people have come to trust clinicians to keep their confidences. But such a vow is not the solution today, and the sad reality with the EHRs of today is that the clinician has little to nothing to do with assuring the confidentiality, and likely does not understand the technical mechanisms, laws, and administrative structures being used that purport to ensure security of the information. Clinicians today cannot assure the patient's confidentiality with anywhere near the certainty one could with a paper record in one's office or the local hospital.

Among the ONC's initiatives regarding privacy and security is the Strategic Healthcare IT Advanced Research Projects on Security (SHARPS), being carried out at three major universities and two major hospitals. On the home page of the SHARPS website, it states the research projects are "aimed at reducing security and privacy barriers to the effective use of health information technology." This is a discouraging way to frame the problem, seemingly assuming that privacy and security is adequate at present but is somehow getting in the way. It is hard (for me) to determine from the site what specific research is ongoing, but there is a fairly long list of publications/papers generated by their work. None of these seem to contemplate a patient having direct physical control over access to their record. One fears it has been judged that such an approach is not necessary for security and would inevitably increase security and privacy barriers to "effective use of health information technology." If this is the assumption by the government agency funding research on EHRs, and writing the criteria for acceptable EHRs, then the probability is high that a major catastrophe stemming from loss of control over protected health information looms down the road.

CONCLUSION

The confluence of EHRs first developing in the absence of the internet; encryption standardization (DES) and breakthroughs (asymmetric) not occurring until the late 1970s; a focus on individual EHR design; the belief that such design could be carried out simply by transposing paper record functions into EHRs—in particular, the unexamined assumption that a record room would be replicated for EHRs; the significant countrywide investments in EHRs, with designs elaborating on such assumptions; the sudden political will to reap the potential benefits of EHRs; and the late arrival of the ONC, which in its mission to move ahead seems unwilling to step back and reconsider the stresses EHRs built to its specifications will have to withstand—all of these factors have placed the confidentiality and beneficent use of the nation's medical information at serious risk. The current security standards and plans are so much tissue paper standing between the data being amassed and the forces, irresistible as a flooding river, that are moving toward it.

Further, individual clinicians, despite the vows of their profession, are completely marginalized in the actual maintaining of confidences. How many day-to-day practitioners are involved in EHR design, or, when it comes to confidentiality, know much more than the sales talk of "robust security" that has been "hardened"? What is a block cipher with a 128-bit key anyway, and "asymmetric what"? Few individual practices can sort through the various vendors and consider the options. Most rely on advice from state and federal programs that are necessarily there to help promulgate EHRs, or they hire consultants. Larger organizations (such as hospitals) form committees to select their "product," and they then hire compliance officers and chief information officers, each with large staffs who need to insure the correct technical and administrative functions of EHRs and people, and who also need to detect, contain, and repel any security lapses or attacks. CEOs are not used to the "record room" and its personnel being a major part of their budget, so the departments of information officers are chronically underfunded. And all these people need to understand what HIPAA, the ONC, the CMS, and the OCR require of them, keep the requisite documentation to prove their compliance, and receive necessary recompense. This complex architecture of technology and people is held together by contracts—not by the loyalties that come from personal contact, and professional vows—and it is very vulnerable to entropy, ineptitude, coercion, and infiltration. Patients will continue to have direct personal contact with clinicians, but the two remain isolated from the levers on the machines and the people directing the flow of information. They say what they think they would like to happen, and hope for the best. This is particularly tragic because we can and have thought of better ways to do things.

It is hard to imagine that a reading of history and an awareness of current events would not compel us to explore these ideas. We must make every effort to prepare for the excesses of domestic government, attacks from other nations, and the corrosive inroads of organized crime and commercialism upon the social use of privacy to promote individual and common good.

References

Bamford, J. 2012. The NSA is building the country's biggest spy center (watch what you say). *Wired Mag*, April 1. http://www.wired.com/threatlevel/2012/03/ff_nsadatacenter//.

Benaloh, J., et al. 2009. Patient controlled encryption: Ensuring privacy of electronic medical records. *ACM Cloud Computing Security Workshop*, Nov. 13. http://research.microsoft.com/en-us/um/people/horvitz/ccsw_2009_benaloh_chase_horvitz_lauter.pdf.

Bradley, T. 2011. Dropbox drops the ball on data security. *PCWorld*, May 16. http://www.pcworld.com/article/228018/dropbox_drops_the_ball_on_data_security.html.

Dubois, S. 2010. Electronic medical records: Great, but not very private. *Fortune*, Oct. 6. http://money.cnn.com/2010/10/06/technology/electronic_medical_records_safety.fortune/.

Elmore, R. 2010. Tiger Team presents MU Stage 1 privacy and security recommendations. *HealthcareTechnologyNews*, Aug. 19. http://news.avancehealth.com/2010/08/tiger-team-presents-mu-stage-1-privacy.html.

Health and Human Services (HHS). 2006. HIPAA Administrative Simplification, Regulation Text, 45 CFR, 164.306 a. 2. Feb16, 2006. http://www.hhs.gov/ocr/privacy/hipaa/administrative/privacyrule/adminsimpregtext.pdf.

Health and Human Services (HHS). 2010. Health information technology: Initial set of standards, implementation specifications, and certification criteria for electronic health record technology; interim final rule, Jan. 13, 2010. *Federal Register* 75(8). http://www.gpo.gov:80/fdsys/pkg/FR-2010-01-13/html/E9-31216.htm.

Henderson, N. 2011. Hack drives Dutch certificate company DigiNotar to bankruptcy. *Web Host Industry Rev*, Sept. 20. http://www.thewhir.com/web-hosting-news/hack-drives-dutch-certificate-authority-diginotar-to-bankruptcy.

Levin, A. 2011. The next Osama Bin Laden already has your Social Security number. *Credit.com*, Nov. 17. http://blog.credit.com/2011/11/the-next-osama-bin-laden-already-has-your-social-security-number/.

Linton, S. 2010. Allegations of OpenBSD backdoor may be true, updated. *Linux J*, Dec. 22. http://www.linuxjournal.com/content/allegations-openbsd-backdoors-may-be-true.

Markoff, J. 2011. SecurID company suffers a breach of data security. *NY Times*, March 17. http://www.nytimes.com/2011/03/18/technology/18secure.html?_r=1&src=busln.

Markoff, J., and D. Barboza. 2012. Researchers trace data theft to intruders in China. *NY*

Times, April 5. http://www.nytimes.com/2010/04/06/science/06cyber.html?_r=2&hpw&partner=TOPIXNEWS.

Musthaler, B. 2012. New key technology simplifies data encryption in the cloud. *PC World*, March 10. http://www.pcworld.com/article/251623/new_key_technology_simplifies_data_encryption_in_the_cloud.html.

Office of the National Coordinator for Health Information Technology (ONC). 2008. Nationwide privacy and security framework for electronic exchange of individual identifiable health information, Dec. 18. http://healthit.hhs.gov/portal/server.pt?open=512&objID=1173&parentname=CommunityPage&parentid=34&mode=2&in_hi_userid=10732&c.

Poggioli, S. 2010. Italian police arrest hundreds in anti-Mafia raids. *NPR*, July 14. http://www.npr.org/templates/story/story.php?storyId=128506439.

Privacy and Security Tiger Team. Federal Advisory Committee of the ONC. 2010. Draft letter to the Chair, HIT Policy Committee, Aug. 12, 2010, in Aug. 16 minutes. http://healthit.hhs.gov/portal/server.pt%3Fopen%3D512%26mode%3D2%26objID%3D2833%26PageID%3D19477.

Sanger, D. E. 2012. Obama order sped up cyberattacks against Iran. *NY Times*, June 1. http://www.nytimes.com/2012/06/01/world/middleeast/obama-ordered-wave-of-cyberattacks-against-iran.html?pagewanted=all&_r=0.

Strategic Health Care IT Advanced Research Projects on Security (SHARPS). 2012. Home page. http://sharps.org/.

Smith, L. B. 2012. Case study: Computer virus containment at Fletcher Allen Health Care. *HIT Community*, Jan.. https://www.thehitcommunity.org/2012/01/case-study-a-virus-attack-at-fletcher-allen-health-care-led-to-best-practices-for-hit-security/.

Stallings, W. 2011. Secure socket layer and transport layer security. In *Cryptography and network security: Principles and practice*. 5th ed., 489–501. Boston: Prentice Hall.

Steinert-Threlkeld, T. 2011. Organized crime's new drug, web attacks on financial firms. *Insurance Networking News*, March 3. http://www.insurancenetworking.com/news/insurance_technology_risk_data_security_mobile_social_media_hackers-27284-1.html.

Turner, G. 2011, Mass. data breaches strike 5 million. *Boston Herald*, July 3. http://www.bostonherald.com/business/technology/general/view/2011_0703mass_data_breaches_strike_5_million/srvc=home&position=5.

Vermont Information Technology Leaders (VITL). 2010. Policy on secondary use of PHI, 11/22/2010. http://www.vitl.net/sites/default/files/documents/HIE/SEC006%20Secondary%20Use.pdf.