POLICY

A cybersecurity primer for translational research

Eric D. Perakslis[1,2]* and Martin Stanley[3]

Virtually all health care organizations have had at least one data breach since 2012. Most of the largest data breaches and Health Care Information Privacy and Accountability Act fines could have been prevented by the simplest of strategies. Each researcher must clearly understand his or her responsibilities and liability.

Modern technologies enable medical research with data storage and computation on scales not previously envisioned by research institutions. Computing and data science are ubiquitous, and collaboration is global and takes place in real time. The scientific need and appetite for these advances are ravenous, yet there are daily reminders that substantial risk accompanies the benefits. At the heart of these risks is the rapidly growing prevalence of criminal cyber attacks on health care systems used to store and manage patient data, which have risen 100% since 2010 (*1*). In fact, the cyber threat has become so clear as to warrant multiple new federal initiatives, including a Comprehensive National Cyber Security Initiative as well as several more targeted executive orders to combat what is now widely considered a true threat to our national security (*2*).

The prevalence and impact of these threats are reflected in the reporting of health information data breaches to the U.S. Department of Health and Human Services (HHS) Office of Civil Rights database; the list was most recently updated with the highly publicized Anthem data breach, which has jeopardized potentially 78.8 million people with identity theft and exposure of their personal information. Of health care organizations surveyed in a 2014 Ponemon Institute study, 90% have had at least one data breach since 2012, and many are also reporting that the rapid adoption of new technologies such as cloud services, mobile devices, and health care information exchanges is introducing new and concerning vulnerabilities. The rapid advance of technology has left all electronic devices potentially vulnerable to compromise, including medical devices, telephone and video systems, and security devices themselves (*3*).

Here, we describe the underlying causes of some of the largest health care data breaches of the past several years and provide practical advice on how future data breaches could be prevented (Table 1).

HEALTH DATA AT HIGH RISK

Cyber threats, such as those that challenge the integrity of research environments and the consequences of working with personally identifiable information (PII) and personal health information (PHI), must be considered when planning research studies. When the physical risks to patients are considered along with the legal liability, regulatory liability, and costs of remediation and damages, the health care delivery setting contains extremely high-risk data (*4*). Typically, patients are asked to explicitly agree to risks about their PII and PHI when they agree to the risks of a clinical study. However, waivers and acknowledgments do not greatly reduce the liability of those conducting the studies. Compliance and security are not the same thing. The most commonly understood risks to study data are covered by the Health Care Information Privacy and Accountability Act (HIPAA) and can occur when HIPAA security or privacy rules are not properly implemented. In addition to the massive potential fines levied by the HHS Office of Civil Rights, other liabilities include substantial reputational risk, civil litigation, and possible theft of precious intellectual property.

Until recently, one of the more surprising aspects of data loss has been the lack of involvement of computer hacking and intrusion via the Internet. Many of the largest HIPAA data breaches reported to the HHS database were caused by basic failures: lost or stolen laptops that were not equipped with encryption, improper disposal of microfiche, and computer programming errors (*5*). It is even more important to understand the asymmetric nature of data loss. In the case of the seventh largest data breach ever reported to the HHS database as of this writing, more than 4 million people were affected by the theft of only four laptops from Advocate Medical Group. The 10 largest HIPAA data breaches reported to the HHS database as of December 2015 are shown in Table 2.

In addition to basic vulnerabilities, the same types of malicious threats that have been seen in retail and banking to the integrity, security, and resilience of financial account data are present with research data and health care data, as recently evidenced in the Anthem breach (*6*). Beyond the theft of PHI and intellectual property, there exists the threat of disruption and "hacktivism" by motivated parties that wish to protest or stop clinical care and practices. Such an event occurred at Children's Hospital in Boston in April 2014, which greatly hindered the daily operations of the hospital (*7*). Along with the risk of legal and fiduciary liability incurred by data loss and theft, the risk of system disruption and destruction must be considered. A prolific virus introduced onto the network of a research laboratory can easily destroy data and equipment and affect laboratory operations for weeks or months or, in the worst case, permanently, unless data are properly managed via fully redundant backup and recovery capabilities.

RISK MANAGEMENT

Despite these risks, research requires real-time collaboration with data that must be accessed for use, shared, and properly protected. The rise in the prevalence and importance of patient-reported outcomes via initiatives, such as the Patient-Centered Outcomes Research Institute (PCORI), and the almost endless opportunities for consortia and data sharing are all positive for patients who are waiting for new therapies (*8*–*10*). Fortunately, cyber risk can be greatly reduced across the research enterprise through a basic understanding of regulatory compliance, security principles, and the roles of procedural and technical defenses.

The data and systems used within the research environment must be well understood if researchers, research subjects, and research progress are to be protected against cyber threats. Most biomedical research occurs within universities, academic medical centers, and small and large private-sector laboratories. Although these environments are highly diverse and complex, they do have many aspects in common that can serve as a basis for cyber protection across the research landscape.

[1]Takeda Pharmaceuticals International, R&D Informatics, Cambridge, MA 02139, USA. [2]Harvard Medical School, Center for Biomedical Informatics and Countway Library of Medicine, Instructor in Pediatrics, Boston, MA 02115, USA. [3]Department of Homeland Security, Federal Network Resilience, U.S. Department of Homeland Security, Office of Cybersecurity and Communications, Arlington, VA 22203, USA.
*Corresponding author. E-mail: eric.perakslis@takeda.com

COMPLIANCE VERSUS SECURITY

For translational researchers, HIPAA likely is the most familiar form of regulatory compliance. Proper records management and retention policies are also compulsory, as are human subjects protections and myriad financial regulations that are based on whether an organization is private, public, or nongovernmental (*11*). Detailing the complex landscape of compliance efforts is well beyond the scope of this writing. The best practice for any researcher is to study and understand the regulatory situation of a proposed effort and ensure that he or she has a clear and correct view of how the liability may be split between a researcher and his or her institution. Most scientists in government and academia underestimate their own personal liability and overestimate the liability of their institution. In fact, existing case law suggests that patients should be able to bring researcher malpractice suits and that institutional review board (IRB) approval is only a partial defense against the liabilities and damages that a researcher may face if found to not have used a suitable standard and duty of care (*12*). A brief history of HIPAA legislation and the evolution of information security standards are presented in Fig. 1. Ideally, a primary investigator has a working knowledge of this complex alphabet soup of regulations, guidance, and standards when they are applicable.

First, compliance does not equal security, and the differences and relationships between them are easily misunderstood. Security is the application of protections and management of risk posed by cyber threats. Compliance is typically a top-down mandate based on federal guidelines or law, whereas security is often managed bottom-up and is decentralized in most organizations (Fig. 1). Compliance processes typically revolve around documentation, whereas security processes are embedded within the technology life cycle as systems are acquired, used, and discarded. Regulations and standards are typically updated and assessed on an annual basis, whereas the landscape of security threats and necessary protections changes so rapidly that security controls often must be updated daily, and even hourly. Security and compliance officers often report to different organizations, and their levels of accountability may be unclear. Similarly, both are best managed in a data-driven and risk-based approach, but this can be difficult if a compliance-driven culture is already established and is exclusively focusing the security resources on compliance efforts. Last, in complex research organizations, scientists frequently assume that security and compliance are someone else's job and are often overdocumented and undertested.

Last, compliance can actually be a competitive advantage for research institutions when it comes to federal grants and industry collaborations. With increasing federal requirements for research grants—such as the ability of a research institution to ensure that their technology infrastructure can comply with Federal Information Security Management Act (FISMA) standards—organizations that can demonstrate high levels of compliance will have greater opportunities for funding and data-centric collaborations. One example of this is the Coordinating Center grant for the NIH Undiagnosed Diseases Network at Harvard Medical School (HMS). The successful implementation of this program, which involves the sharing of sensitive data across multiple research institutions, required that HMS implement a FISMA-compliant solution. Organizations that have poor compliance histories will be at a disadvantage despite the merits of their research.

Table 1. Six steps that will improve the cybersecurity posture of any organization.

• Categorize and classify your systems and data according to risk of downtime, data loss, data destruction, and data theft
• Grant access to systems and data only to those who truly "need to know."
• Work with your institutional security officer to select, implement, train, and routinely test appropriate procedural and technical controls.
• Assess the effectiveness of all controls via a third-party testing or audit.
• Ensure that the security controls are monitored on a regular basis.
• Have a clearly articulated incident-response plan and trained users.

Table 2. Ten of the largest HIPAA data breaches reported to the HHS database as of December 2015.

Organization	Method/breach	Date	Number of records	Location of breached information
Anthem	Hacking/IT incident	March 2015	78,800,000	Network server
Premera Blue Cross	Hacking/IT incident	March 2015	11,000,000	Network server
Excellus Health Plan	Hacking/IT incident	September 2015	10,000,000	Network Server
Science Applications International Corporation (SAIC)	Loss	November 2011	4,900,000	Other
University of California, Los Angeles Health	Hacking/IT incident	July 2015	4,500,000	Network Server
Community Health Systems Professional Services Corporation	Theft	August 2014	4,500,000	Network Server
Advocate Medical Group	Theft	August 2013	4,029,530	Desktop computer
Medical Informatics Engineering	Hacking/IT incident	July 2015	3,900,000	Electronic Medical Record Network Server
Xerox State Healthcare	Unauthorized access/disclosure	September 2014	2,000,000	Desktop computer or other device
IBM	Unknown	April 2011	1,900,000	Other

ENSURING SECURITY

There are qualitative and quantitative assessment methodologies that represent cyber risk in dollar values as well as the potential impact on an organization or mission. These methodologies are well documented in the National Institute of Standards and Technology (NIST) Risk Management Framework and the NIST Cybersecurity Framework (summarized in supplementary materials) and provide model approaches for assessing cyber risk and determining a budget for protecting IT systems and data (*13*, *14*). Effective risk management requires that business owners, such as scientific

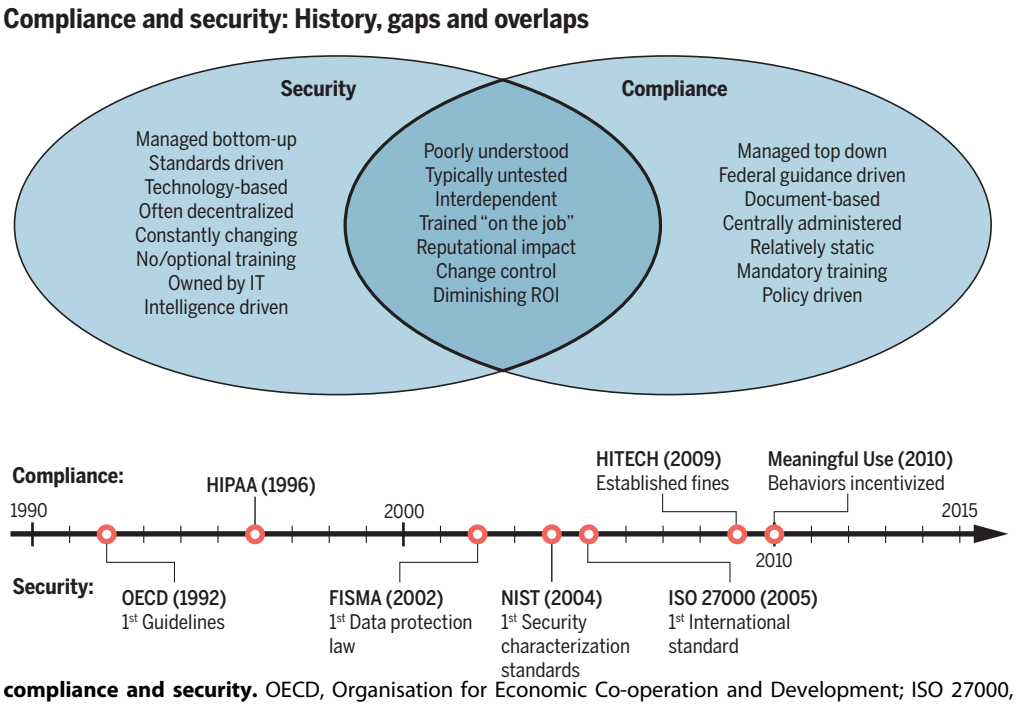

Fig. 1. Comparing compliance and security. OECD, Organisation for Economic Co-operation and Development; ISO 27000, International Organization for Standardization information on security standards; HITECH (2009), Health Information Technology for Economic and Clinical Health Act.

researchers, remain involved in all phases of the risk management process because they intuitively understand what is most important to them and can most effectively direct what information must be protected and to what extent.

All data are not equal, and the necessary first step to determining where to focus cybersecurity efforts is knowing which data and systems are sensitive and most essential to an organization's mission. This knowledge then leads directly to the second step, which is to ensure that only users with a genuine need—one that supports the institute's mission—are granted access to sensitive data. In the case of collaborative translational medical research in which highly specific phenotypic traits and molecular profiling information must be shared and discussed, researchers must take due care to deidentify and share, using proper encryption, only the minimum amount of PII and/or PHI required to conduct the study. In addition, although one size does not fit all, there are basic risk-based protections that form the cornerstone of good cybersecurity (*15*). Implementing basic cybersecurity protections virtually mitigates the most common cyber vulnerabilities, such as a lost laptop or phone, and affords the same advantages as securing one's home with a system that is superior to one's neighbors' systems: Intruders will often opt for an easier break in.

Researchers should not count on others to implement these critical basic protections; instead, they should be well versed in their organization's security and privacy policies as well as the important security contacts at their institutions, such as the chief information security officer (CISO), who can help researchers to understand and implement protections. The CISO is essential for the protection of data and of biomedical research operations (*16*), and if an organization lacks an internal CISO, the role should be contracted out. Data protection depends on a well-functioning cooperative and collaborative partnership among scientists, clinicians, computer scientists, and security officers.

First, virtually all research data are input, manipulated, and accessed via some form of device that represents an "endpoint" to the network. Laptops, cell phones, tablets, desktop computers, and even medical devices are all types of end points and must be highly protected. Failures in end point security are the most common causes of data loss and theft, and most are completely avoidable. Single passwords are ineffective once a device falls into malicious hands (*17*). Researchers should rely on the organizational CISO to provide a federated identity management solution that ensures that users are securely authenticated for access to any and all devices and systems; at a minimum, all IT systems must support two-factor authentication for systems that use sensitive data (*15*). Further, all end points should be adequately encrypted so that the device becomes useless to any unauthorized user. Some institutions have implemented such security measures, but many research institutions lag behind. The real challenge here is that the technology and policy infrastructure in use at most institutions was put into place years ago, long before many of the current threats existed, and it is impossible to fix everything quickly and simultaneously. The online guide at the University of Washington (UW Medicine) provides an excellent example of how a basic but comprehensive cybersecurity program can be used effectively to secure data and be integrated in a complex research and clinical environment (*18*). There also are many commercial encryption tools and services available. The best approach is to work with the information security office at your institution to select the tools that will be most effective in your particular technology environment.

The second line of defense in research data protection is the computing network to which these end points connect. Network firewall and antivirus technologies are limited because they are only capable of detecting and protecting against threats that they have seen before. There is great debate about the utility of protection via firewalls and sole reliance on a

strong network perimeter; however, these tools serve well as part of a systematic application of security controls commonly referred to as "defense in depth." Many technology vendors continue to build better mousetraps and to insist that they are impenetrable. Others take the opposite view, that networks cannot be completely secure and that a mix of approaches is best (*19*, *20*). Our view is that network and other basic security controls such as antivirus software serve important purposes and contribute to a security baseline but require continual updating through subscription maintenance and daily updates to ensure their efficacy against newly identified threats.

In addition to the general security tips discussed here, the following are critical cybersecurity protections for PII/PHI and should be considered before beginning any new clinical research effort. These protections are typically available through the institution's CISO office or as part of commercially available IT services.

(i) Protect computers and data with antivirus software and encryption so as to ensure that known threats are quickly identified and contained automatically and that data are secure if computers are lost or stolen, respectively.

(ii) Require the use of one-time passwords or two-factor authentication or, preferably, a federated identity management solution so as to ensure only authorized users are able to access computers and data (*15*).

(iii) Restrict access to sensitive data and systems to personnel who require access or establish time limits for personnel to access data and systems consistent with their needs. Often, a researcher will need access to particular sensitive information while completing a particular study, and their access to that information should expire when they complete the effort.

(iv) Ensure that personnel engage in recurring cybersecurity training that covers institutional policies and practical aspects of cybersecurity such as "don't click the link"; recognizing suspicious e-mails and attachments; sharing, transmitting, and storing sensitive information; and how to report cyber incidents and data breaches.

Similarly, best practices for responding to cyber incidents are consistent with the due-care standard, which is a legal construct in which negligence is tested against what a reasonable person would do in a given situation. Many organizations have experienced some kind of cyber incident, and what distinguishes an effective from an ineffective response is making sure that the appropriate measures that would protect an injured party are taken in an expedited and practical manner. These include ensuring that affected parties are notified, correcting the vulnerability in accordance with the severity of the breach, and performing accurate reporting of the breach, as required by federal law, regulations, or other industry standards.

In the event of a known or suspected data breach, five important steps are considered to be the minimum response (*21*). First, find the point of intrusion and immediately patch the hole. Second, engage the organizational incident-response team or assemble one in partnership with IT if none exists. Third, test whatever fixes and remediation are proposed and implemented. Fourth, resolve any related issues or risks that may have led to the breach. And fifth, contact appropriate external parties such as law enforcement or outside experts to validate the remediation and recommended next steps.

Health care and proprietary research data and systems are highly attractive targets for criminals because of the personal information and intellectual property they contain. Such systems carry substantial personal, legal, and regulatory risks for researchers and their institutions, but they can and must be protected.

SUPPLEMENTARY MATERIALS

www.sciencetranslationalmedicine.org/cgi/content/full/8/322/322ps2/DC1
Table S1. NIST framework categories and definitions.

REFERENCES AND NOTES

1. The Ponemon Institute, Fourth annual benchmark study on patient privacy and data security (March 2014); available at www.ponemon.org/library/fourth-annual-benchmark-study-on-patient-privacy-data-security.
2. www.whitehouse.gov, The Comprehensive National Cybersecurity Initiative, (January 2015); available at www.whitehouse.gov/sites/default/files/cybersecurity.pdf.
3. B. Filkins, Health care cyberthreat report: Widespread compromises detected, compliance nightmare on horizon. *SANS Institute* (February 2014); available at www.sans.org/reading-room/whitepapers/analyst/health-care-cyberthreat-report-widespread-compromises-detected-compliance-nightmare-horizon-34735.
4. E. D. Perakslis, Cybersecurity in health care. *N. Engl. J. Med.* **371**, 395–397 (2014).
5. D. Vogel, Top 10 HIPAA breaches of 2013. *DataPipe* (28 January 2015); available at www.datapipe.com/blog/2015/01/28/top-10-hipaa-data-breaches-of-2014.
6. Anthem Facts, (13 February 2015); www.anthemfacts.com.
7. D. J. Nigrin, When 'hacktivists' target your hospital. *N. Engl. J. Med.* **371**, 393–395 (2014).
8. L. Frank, E. Basch, J. V. Selby, Patient-centered outcomes research institute, the PCORI perspective on patient-centered outcomes research. *JAMA* **312**, 1513–1514 (2014).
9. H. G. Eichler, F. Pétavy, F. Pignatti, G. Rasi, Access to patient-level trial data—A boon to drug developers. *N. Engl. J. Med.* **369**, 1577–1579 (2013).
10. R. Jagsi, A. Chiang, B. N. Polite, B. C. Medeiros, K. McNiff, A. P. Abernathy, R. Zon, P. J. Loeher Sr., Qualitative analysis of practicing oncologists' attitudes and experiences regarding collection of patient-reported outcomes. *J. Oncol. Pract.* **9**, 290–297 (2013).
11. OCR HIPAA Privacy, Research [45 CFR 164.501, 164.508, 164.512(i)], (2003).
12. R. L. Jansson, Researcher liability for negligence in human subject research: Informed consent and researcher malpractice actions. *Wash. Law Rev.* **78**, 229–263 (2003).
13. NIST special publication 800-37, Revision 1, 93 pages (February 2010).
14. NIST Framework for Improving Critical Infrastructure Cybersecurity, version 1.0 (February 2014).
15. J. Stewart, in *Best Practices in Computer Network Defense Incident Detection and Response*, vol. 35 of NATO Science for Peace and Security Series–D, M. Hathaway, Ed. (IOS Press, 2014), pp. 30–33.
16. C. Burgess, Why the role of the CISO is vital in every company. *Security Intelligence* (October 2014). https://securityintelligence.com/why-the-role-of-the-ciso-is-vital-in-every-company.
17. L. Div, Three reasons why endpoints cannot remain a security blind spot. *Forbes online* (26 September 2014). www.forbes.com/sites/frontline/2014/09/26/three-reasons why-endpoints-cannot-remain-a-security-blind-spot.
18. UW Medicine Information Technology Services Web page; https://security.uwmedicine.org/guidance/technical/encryption/mobiledevice_encryption/default.asp.
19. B. Schneier, Securing medical research: A cybersecurity point of view. *Science* **336**, 1527–1529 (2012).
20. T. Scully, The cyber threat, trophy information and the fortress mentality. *J. Bus. Contin. Emer. Plan.* **5**, 195–207 (2011).
21. J. Brandon, 5 steps to take when a data breach hits. *CIO Online* (October 2014); www.cio.com/article/2692972/data-breach/5-steps-to-take-when-a-data-breach-hits.html.

Author note: The opinions, conclusions, recommendations, or other matters should be considered as the authors' and do not necessarily reflect the position of the United States or the U.S. Department of Homeland Security.

10.1126/scitranslmed.aaa4493

Citation: E. D. Perakslis, M. Stanley, A cybersecurity primer for translational research. *Sci. Transl. Med.* **8**, 322ps2 (2016).

Science Translational Medicine

A cybersecurity primer for translational research

Eric D. Perakslis and Martin Stanley

Sci Transl Med **8**, 322ps2322ps2.
DOI: 10.1126/scitranslmed.aaa4493