Health IT, hacking, and cybersecurity: national trends in data breaches of protected health information

Jay G. Ronquillo, J. Erik Winterholler, Kamil Cwikla, Raphael Szymanski, and Christopher Levy

Western Michigan University Homer Stryker M.D. School of Medicine, Kalamazoo, Michigan, USA

Corresponding Author: Jay G. Ronquillo, MD, MPH, MMSc, MEng, Western Michigan University Homer Stryker M.D. School of Medicine, 1000 Oakland Drive, Kalamazoo, MI 49008, USA (jay.ronquillo@med.wmich.edu)

Received 30 December 2017; Revised 22 April 2018; Editorial Decision 21 May 2018; Accepted 31 May 2018

ABSTRACT

Objective: The rapid adoption of health information technology (IT) coupled with growing reports of ransomware, and hacking has made cybersecurity a priority in health care. This study leverages federal data in order to better understand current cybersecurity threats in the context of health IT.

Materials and Methods: Retrospective observational study of all available reported data breaches in the United States from 2013 to 2017, downloaded from a publicly available federal regulatory database.

Results: There were 1512 data breaches affecting 154,415,257 patient records from a heterogeneous distribution of covered entities ($P<.001$). There were 128 electronic medical record-related breaches of 4,867,920 patient records, while 363 hacking incidents affected 130,702,378 records.

Discussion and Conclusion: Despite making up less than 25% of all breaches, hacking was responsible for nearly 85% of all affected patient records. As medicine becomes increasingly interconnected and informatics-driven, significant improvements to cybersecurity must be made so our health IT infrastructure is simultaneously effective, safe, and secure.

Key words: cybersecurity, ransomware, electronic health records, clinical informatics, hacking

OBJECTIVE

The rapid adoption of health information technology (IT) coupled with growing reports of ransomware, and hacking has made cybersecurity a priority in health care. This study leverages federal data in order to better understand current cybersecurity threats in the context of health IT.

BACKGROUND AND SIGNIFICANCE

The Health Information Technology for Economic and Clinical Health (HITECH) Act of 2009 increased electronic medical record (EMR) adoption across the country, making health IT security a growing concern for health care organizations. Indeed, cybersecurity has become a priority in health care because of the highly sensitive nature of patient information, the prevalence of large databases of diverse health data, and the development of highly interconnected medical and health information technologies. With growing reports of ransomware and other deliberate hacking activities initiated by unauthorized users, there is a need for data-driven approaches to better understand the extent of these threats to the health care system.

The HITECH Act requires that health care organizations publicly report all breaches of protected health information involving more than 500 patients to the U.S. Department of Health and Human Services (HHS) Office for Civil Rights (OCR). To our knowledge, a formal analysis of cybersecurity breaches nationwide with a focus on EMR and hacking-related incidents has not been performed. This study leverages publicly available regulatory data in order to better understand the current state of cybersecurity and to identify opportunities to make health IT safer and more secure.
Table 1. Characteristics of reported data breaches of protected health information by covered entity, 2013–2017

| Characteristic                        | 2013  | 2014  | 2015  | 2016  | 2017  |
|--------------------------------------|-------|-------|-------|-------|-------|
| Reported data breaches, n (%)*       |       |       |       |       |       |
| Health plan                          | 17 (6.1) | 38 (12.1) | 62 (23.0) | 51 (15.6) | 47 (14.5) |
| Health care provider                 | 184 (66.2) | 180 (57.3) | 194 (72.1) | 256 (78.3) | 259 (79.9) |
| Other                                | 77 (27.7) | 96 (30.6) | 13 (4.8) | 20 (6.1) | 18 (5.6) |
| Patient records affected, n (%)      |       |       |       |       |       |
| Health plan                          | 88549 (1.3) | 2135600 (16.8) | 102919905 (90.9) | 880455 (5.3) | 330728 (6.8) |
| Health care provider                 | 5773597 (83.2) | 2051214 (16.2) | 6392806 (5.6) | 12213969 (73.3) | 4328916 (88.9) |
| Other                                | 1078487 (15.5) | 8496026 (67.0) | 3954663 (3.5) | 3560666 (21.4) | 209876 (4.3) |

*P < .001.

RESULTS

General characteristics of data breaches

Over the 5-year study period, there were 1512 reported data breaches of protected health information affecting 154 415 257 patient records (Table 1). This included 215 (14.2%) health plan breaches, 1073 (71.0%) breaches of a hospital or health care provider, and 224 (14.8%) breaches involving other (170 business associate, 2 health care clearing house, 52 unknown). Furthermore, there were 106 355 237 (68.9%) patient records affected due to health plan breaches, 30 760 502 (19.9%) records affected by hospital or health care provider breaches, and 17 299 518 (11.2%) patient records affected from other (17 131 045 business associate; 6504 health care clearing house; 161 969 unknown). State-level trends in breach characteristics were calculated and visualized after adjusting by population estimates from the U.S. Census Bureau (Figure 1) and by physician totals provided by the AHRQ Compendium of U.S. Health Systems (Figure 2).

Hacking incidents and EMR-related breaches of protected health information

There were 128 breaches involving EMRs which affected a total of 4 867 920 patient records, in contrast to 1384 non-EMR-related breaches involving 149 547 337 patient records. In addition, there were a total of 363 breaches classified as hacking incidents which involved 130 702 378 patient records, in contrast to 1149 breaches that were not hacking-related but impacted 23 712 879 records. The breakdown of hacking-related data breaches by media type, as well as EMR breaches according to breach category, are shown in Table 2. By media type, network server incidents made up more than 50% of all hacking-related breaches of protected health information, affecting more than 90% of hacking-related patient records and more than 75% of breached patient records overall. Furthermore, the majority (87%) of patient records affected by EMR-related breaches were due to hacking incidents (Table 2).

DISCUSSION

Health IT cybersecurity is a serious issue, with millions of individuals across the country already affected by cybersecurity breaches each year. Hacking incidents are becoming a very real concern because cyberattacks can lock down health systems and directly hinder patient care. Despite making up less than 25% of all breaches, hacking was responsible for nearly 85% of all affected patient records over the last 5 years, highlighting the broad potential reach of this kind of breach. The 2017 WannaCry ransomware attack on Britain’s National Health Service, which forced numerous clinics to...
close their doors and refuse patient care, further highlights just how devastating a single cyberattack can be simply because of outdated hospital IT systems and infrastructure. Because ransomware is a relatively new and complex phenomenon in the published literature, it will be important to make data, information, and solutions addressing ransomware easily accessible to the public and to informatics researchers. Covered entities play an important role in the scale of breaches, with most reported data breaches occurring in the hospital and health care provider settings, while the largest volume of affected patient records comes from the health plan setting, consistent with the local/regional nature of most hospitals and the national scope of most health plans (eg, the 2015 Anthem breach of 78 million records), respectively. These breaches can have a serious financial impact on organizations, with some cost estimates as high as $355 per breached patient record. Furthermore, in addition to disrupting physician and other health care provider workflows, EMR breaches may encourage patients to withhold critical information from their providers out of privacy concerns, further impacting the quality of care patients receive. As diverse software companies (eg, Google, Apple, Microsoft, Amazon) known for handling large amounts of user data enter the health care space, it will be useful to see how the growing number of business associates and related covered entities further alter the national landscape of cybersecurity breaches.

Recently implemented health IT policies provide further insight into the current and future risks of health IT cybersecurity breaches. The HITECH Act helped drive rapid adoption of EMRs nationwide, but existing certification criteria focused heavily on Meaningful Use functionality without equal attention to patient privacy and security, usability, or interoperability. The Office of the National Coordinator for Health IT recently described its plans to (1) reduce random surveillance of certified products and (2) replace more than half of formal health IT certification test procedures with simple vendor self-declaration raising concerns from both the medical and informatics communities. Finally, laws such as the 21st Century Cures Act passed in 2016 will catalyze new cancer and precision medicine initiatives, but also formally remove EMRs and other health IT from U.S. Food and Drug Administration regulatory oversight, creating the need for new ways to monitor the safety, effectiveness, and security of health IT. As new policies affecting health IT and cybersecurity continue to be implemented, the informatics community must provide evidence-based, data-driven expertise, and leadership to support these initiatives and protect patients.
physician populations, reflects the fragmented state of cybersecurity measures currently in place.19,20 The scientific community has recognized that health care organizations are seriously unprepared for cyber threats, due to the strong focus on health IT adoption without similar attention to security, leading to systems that are both highly interconnected and vulnerable to cyberattacks.10,11 Indeed, cybersecurity issues easily become serious patient safety concerns: the health IT infrastructure provides a robust medium for problems to propagate throughout large populations, while many technologies (e.g., drug infusion devices) contain software flaws that could be exploited by malicious users to directly impact clinical care.18,21 Serious consequences range from patient identity theft to inhibited patient care (via EMR ransomware) to direct patient harm (via unauthorized control of devices).22 While organizations may already have disaster mitigation procedures for unplanned system downtime, specific resources must also be devoted to preventing cybersecurity threats from posing a risk to patient safety.

Effective solutions to reduce cybersecurity vulnerabilities are broad and will require engagement at multiple levels. At the individual level, physicians and other health care providers must follow good data and cyber hygiene, which includes creating, updating, and protecting strong passwords as well as remaining vigilant against cybersecurity threats like email phishing.6,22 Similarly, at the organizational level, institutions must have sufficient IT resources to implement robust user authentication, strong data encryption, and frequent software updates (for technologies ranging from EMRs to operating systems).3,19,23 Finally, as big data, machine learning, and artificial intelligence are integrated into the practice of medicine, policies must be revised to help build the informatics infrastructure necessary to safely and effectively monitor, detect, and eliminate cyber threats.20,24 In the context of the current regulatory landscape, our findings highlight a critical need to strengthen key administrative, physical, and technical safeguards for health IT cybersecurity.4,14

Our study had several limitations. First, the number of breaches and affected records were likely underestimated, since only breaches that affected more than 500 individuals were reported to the HHS database. Second, while we quantified the number of breaches and patient records affected, there were insufficient details on each breach to determine the underlying factors (e.g., human, technical, and policy) that would explain yearly breach fluctuations in greater depth. Finally, while our analysis addressed several important cov-ered entities, future studies should also focus on the characteristics of business associates and related organizations as this data becomes available. Despite these limitations, the HHS breach dataset currently represents the most comprehensive database of reported breaches and provides a valuable snapshot of important trends in breach characteristics at a national level.

CONCLUSION

Cybersecurity breaches of protected health information are a serious problem that currently affects millions of patients nationwide. Hacking-related breaches, particularly involving health IT and ransomware, are a rapidly growing concern. As we work towards an interconnected learning health system focused on personalized and precision medicine, it will be crucial that our future informatics infrastructure is designed to be simultaneously effective, safe, and secure.

CONTRIBUTORS

All authors included in the manuscript provided substantial contribution to (1) conception and design, acquisition of data, or analysis and interpretation of data, (2) drafting the article or revising it critically for important intellectual content, and (3) final approval of the completed manuscript. J.G.R. had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; or decision to submit the manuscript for publication.

Conflict of interest statement. None declared.

ACKNOWLEDGEMENTS

The authors would like to thank Bob Brown, PhD and Jon Walsh, MD, MPH; Co-Chairs of the Biomedical Informatics Program at the Western Michigan University Homer Stryker M.D. School of Medicine; as well as Mr. Joe Costello for their valuable contributions to the study.

SUPPLEMENTARY MATERIAL

Dataset available in the Dryad Digital Repository: http://dx.doi.org/10.5061/dryad.24275c6.

REFERENCES

1. HIMSS North America. 2016 HIMSS Cybersecurity Survey. HIMSS. 2016. http://www.himss.org/library/2016-himss-cybersecurity-survey. Accessed August 17, 2017.
2. Nigrin DJ. Cybersecurity in health care. N Engl J Med 2014; 371 (5): 393–6.
3. Cohen IG, Hoffman S, Adashi EY. Your money or your patient’s life? Ransomware and electronic health records. Ann Intern Med 2017; 167 (8): 587–8.
4. Krause CS, Smith B, Vanderlinden H, Nealand A. Security techniques for the electronic health records. J Med Syst 2017; 41 (8): 127.
5. Liu V, Muser MA, Chou T. Data breaches of protected health information in the United States. Jama 2015; 313 (14): 1471.
6. Bai G, Jiang J, Flasher R. Hospital risk of data breaches. JAMA Intern Med 2017; 177 (6): 878–80.
7. U.S. Department of Health and Human Services Office for Civil Rights. Breach Portal: Notice to the Secretary of HHS Breach of Unsecured Pro-

Table 2. Breakdown of hacking and EMR-related breaches

| Breach characteristics | Breaches, n (%) | Records affected, n (%) |
|------------------------|----------------|------------------------|
| Portable electronic device or laptop | 1 (0.3) | 1911 (0.0) |
| Desktop, email, or EMR | 106 (29.2) | 1984418 (1.5) |
| Network server | 192 (52.9) | 119590428 (91.5) |
| Multiple types | 48 (13.2) | 8822024 (6.7) |
| Other/unknown | 16 (4.4) | 303597 (0.2) |
| Hacking-related breaches by breach category | | |
| Theft | 21 (16.4) | 146496 (3.0) |
| Unauthorized access or disclosure | 66 (51.6) | 377088 (7.7) |
| Hacking or IT incident | 34 (26.6) | 4240218 (87.1) |
| Multiple categories | 5 (3.9) | 102614 (2.1) |
| Other/unknown | 2 (1.6) | 1504 (0.0) |

Abbreviations: EMR: electronic medical record; IT: information technology.
1. Schoenberger A. The Global Health Informatics Data Protection Revolution. 2017. https://www.who.int/gho/health_info/hipaa/en/. Accessed December 14, 2017.
2. Health Sector- Specific Data Breach Report. 2017. https://ocrportal.hhs.gov/ocr/breach/ breach_report.jsf. Accessed December 14, 2017.
3. Blankenship B. HIPAA Breach Report: 2017. Data Breach Digest. 2017. https://www.databreachdigest.com/2017/11/22/hipaa-breaches-in-2017/. Accessed December 31, 2017.
4. GAO. Health Care Security Breaches and Healthcare Organizations. 2017. https://www.gao.gov/products/GAO-18-227. Accessed December 31, 2017.
5. GNU. PSPP [computer program]. 2017. https://www.gnu.org/software/pspp/. Accessed December 31, 2017.
6. Clarke R, Youngstein T. Cyberattack on Britain’s National Health Service—a wake-up call for modern medicine. *N Engl J Med* 2017; 377 (5): 409–11.
7. Kruse CS, Frederick B, Jacobson T, Monticone DK. Cybersecurity in healthcare: a systematic review of modern threats and trends. *Technol Health Care* 2017; 25 (1): 1–10.
8. Sittig DF, Singh H. A socio-technical approach to preventing, mitigating, and recovering from ransomware attacks. *Appl Clin Inform* 2016; 07 (02): 624–32.
9. Ponemon Institute LLC. 2016 Cost of Data Breach Study: Global Analysis. 2016. https://www-01.ibm.com/common/ssi/cgi-bin/ssialias?htmlfid=SEL03094WWEN. Accessed January 20, 2018.
10. Campos-Castillo C, Anthony DL. The double-edged sword of electronic health records: implications for patient disclosure. *J Am Med Inform Assoc* 2015; 22 (e1): e130–40.
11. Ratwani RM, Benda NC, Hettinger AZ, Fairbanks RJ. Electronic health record vendor adherence to usability certification requirements and testing standards. *Jama* 2015; 314 (10): 1070.
12. Anthony ES, Posnack S. Certification Program Updates to Support Efficiency & Reduce Burden. HealthITBuzz. 2017. https://www.healthit.gov/buzz-blog/healthit-certification/certification-program-updates-support-efficiency-reduce-burden/. Accessed January 31, 2018.
13. Sweeney E. ONC scales back EHR certification process requirements, catching industry groups by surprise. FierceHealthcare. 2017. https://www.fiercehealthcare.com/regulatory/onc-ehr-certification-health-it-himss-amia-chime-interoperability. Accessed February 1, 2018.
14. Ronquillo JG, Zuckerman DM. Software-related recalls of health information technology and other medical devices: implications for FDA regulation of digital health. *Milbank Q* 2017; 95 (3): 535–53.
15. Walker BJ, Koppel R. For Healthcare Cybersecurity the Whole is Weaker Than the Sum of the Parts. 2016. http://thehealthcareblog.com/blog/2016/09/23/for-healthcare-cybersecurity-the-whole-is-weaker-than-the-sum-of-the-parts/. Accessed February 1, 2018.
16. Blumenthal D, McGraw D. Keeping personal health information safe: the importance of good data hygiene. *Jama* 2015; 313 (14): 1424.
17. Sittig DF, Singh H. Electronic health records and national patient-safety goals. *N Engl J Med* 2012; 367 (19): 1854–60.
18. Jarrett MP. Cybersecurity—a serious patient care concern. *Jama* 2017; 318 (14): 1319–20.
19. Koppel R, Thimbleby H. Lessons From the 100 Nation Ransomware Attack. 2017. http://thehealthcareblog.com/blog/2017/05/14/lessons-from-the-100-nation-ransomware-attack/. Accessed February 1, 2018.
20. Char DS, Shah NH, Magnus D. Implementing machine learning in health care—addressing ethical challenges. *N Engl J Med* 2018; 378: 981–3.