Instances in which private information has been breached are becoming more commonplace in the United States, making the security of this type of information a significant concern. Healthcare information is particularly vulnerable, due to the sensitivity of these data and how they can be used by criminals. Demographic data, Social Security numbers, and clinical information, including medical diagnoses, are housed in both paper and electronic health records (EHRs). For these reasons, multiple attempts have been made through federal legislation to help curtail the occurrences of healthcare privacy breaches, including the 1996 Health Insurance Portability and Accountability Act (HIPAA), the Health Information Technology for Economic and Clinical Health Act of 2009, and the Omnibus Final Rule in 2013. Despite these initiatives, however, large data breaches are still occurring in US hospitals.

The adoption of EHRs among hospitals has increased rapidly over the past several years. As healthcare systems digitized to keep up, the healthcare sector was unable to adopt electronic security components at the same pace, leading to vulnerabilities in record systems. In some cases, technology purposed to assist healthcare delivery processes are now having costly difficulties. The majority of small- and medium-size health organizations do not possess the financial or personnel resources necessary for sufficient information technology (IT) and security investments. This, along with their highly valuable data, has left hospitals vulnerable to breaches of sensitive information.

Hospitals have begun to implement strategies to help prevent data breaches that most often occur through theft, loss, unauthorized access, or hacking. Strategies include the adoption of systems that include 2-factor authentication requirements to ensure that patients’ health information is only accessible to and usable by those with rights to access it. Two-factor authentication often incorporates a biometric component to verify the user’s identity, such as a fingerprint, voice recognition, or iris scan, along with a password, personal identification number, or physical verification object, such as a token or key.

ABSTRACT

OBJECTIVES: The objectives of this study were to describe the locations in hospitals where data are breached, the types of breaches that occur most often at hospitals, and hospital characteristics, including health information technology (IT) sophistication and biometric security capabilities, that may be predicting factors of large data breaches that affect 500 or more patients.

STUDY DESIGN: The Office of Civil Rights breach data from healthcare providers regarding breaches that affected 500 or more individuals from 2009 to 2016 were linked with hospital characteristics from the Health Information Management Systems Society and the American Hospital Association Health IT Supplement databases.

METHODS: Descriptive statistics were used to characterize hospitals with and without breaches, data breach type, and location/mode of data breaches in hospitals. Multivariate logistic regression analysis explored hospital characteristics that were predicting factors of a data breach affecting at least 500 patients, including area characteristics, region, health system membership, size, type, biometric security use, health IT sophistication, and ownership.

RESULTS: Of all types of healthcare providers, hospitals accounted for approximately one-third of all data breaches and hospital breaches affected the largest number of individuals. Paper and films were the most frequent location of breached data, occurring in 65 hospitals during the study period, whereas network servers were the least common location but their breaches affected the most patients overall. Adjusted multivariate results showed significant associations among data breach occurrences and some hospital characteristics, including type and size, but not others, including health IT sophistication or biometric use for security.

CONCLUSIONS: Hospitals should conduct routine audits to allow them to see their vulnerabilities before a breach occurs. Additionally, information security systems should be implemented concurrently with health information technologies. Improving access control and prioritizing patient privacy will be important steps in minimizing future breaches.

Data Breach Locations, Types, and Associated Characteristics Among US Hospitals

Meghan Hufstader Gabriel, PhD; Alice Noblin, PhD, RHIA, CCS; Ashley Rutherford, PhD, MPH; Amanda Walden, MSHSA, RHIA, CHDA; and Kendall Cortelyou-Ward, PhD
The objectives of this study were to describe the locations in hospitals where data are breached, the types of breaches that occur most often at hospitals, and hospital characteristics, including health IT sophistication and biometric security capabilities, that may be predicting factors of large data breaches that affect 500 or more patients. In spite of these health IT strategies, it is unclear what the most common types of breaches are and where patients’ health information is most vulnerable. Under federal legislation, if a healthcare privacy breach affects 500 or more patients it must be reported to the Office of Civil Rights (OCR). Then, information regarding the breach is publicly posted on the OCR data breach portal.13,17 Although several studies have examined OCR data breach information,11-13 none have specifically focused on pediatric, academic, and nonfederal acute care hospitals, which house millions of patient records.

METHODS

Data Sources

The OCR data breach portal provides an online database describing data breaches of protected health information (PHI) that affect 500 or more individuals.15,16 This portal users the option of examining breach information from 3 types of covered entities: health plans, healthcare clearing houses, and healthcare providers. As of July 2016, the OCR portal included 1085 healthcare providers that had PHI breaches affecting 500 or more individuals between October 2009 and July 2016. Of these, 185 were nonfederal acute care hospitals and 27 were Veterans Affairs (VA) hospitals. Nonfederal acute care hospital breach information was linked with the 2015 Health Information and Management Systems Society (HIMSS) analytic data file (HIMSS Analytics, unpublished data) and information from the 2015 American Hospital Association (AHA) Health IT Supplement Survey regarding the use of 2-factor authentication.19

Variables to Characterize Hospitals

Hospital data breaches of PHI that affected 500 or more individuals were characterized by: 1) type of breach and 2) location or mode of breached information. Data breach types included 6 categories: 1) hacking/IT incident, 2) improper disposal, 3) loss, 4) other/unknown, 5) theft, and 6) unauthorized access/disclosure. Data breach locations or modes included 7 categories: 1) desktop computer, 2) EHR, 3) email, 4) laptop computer, 5) network server, 6) paper/films, and 7) other location. To gain a more detailed view of which provider types were most frequently breached and had the most individuals affected, the OCR data were further categorized by “name of covered entity” into 9 health provider categories: 1) colleges/universities; 2) emergency response; 3) government; 4) group/physician practices; 5) health systems; 6) hospitals; 7) nursing homes, home/hospice care, and treatment facilities; 8) pharmacies; and 9) research facilities, laboratories, and medical supply companies.

Inclusion/Exclusion Criteria

Only nonfederal acute care hospitals, which include children’s, teaching, and public or private hospitals, were included in this study. All other health provider categories were excluded.

Variables to Characterize Data Breaches

A False Sense of Security
hospital status, such as not-for-profit, investor-owned (for-profit), and government (nonfederal). In addition, a hospital was considered to be a member of a hospital system if it belonged to an integrated healthcare delivery system. Market concentration was measured by the Herfindahl-Hirschman Index,²⁰ constructed on bed shares within systems at the HRR level.²¹

Data Analysis
Descriptive analyses to characterize provider facility, data breach type, and location/mode in hospitals were performed. Number of patients affected by data breaches was log transformed and a factorial 2-way analysis of variance (ANOVA) was conducted to examine the differences between data location/mode and type of breach and the number of patients affected by data breaches. Univariate analyses were conducted on hospital and area characteristics. To explore factors associated with hospitals having a data breach affecting 500 or more individuals, multivariate logistic regression analyses were performed using SAS Enterprise Guide (SAS Institute Inc; Cary, North Carolina). Significance was determined at the $P < .05$ level.

RESULTS
In total, 215 breaches, each affecting 500 or more individuals, occurred at 185 nonfederal acute care hospitals that reported to the OCR during the study period. Thirty hospitals had multiple breaches during that time. Twenty-four hospitals had 2 breaches, 5 hospitals had 3 breaches, and 1 hospital had 4 breaches (Table 1).

Descriptive Results
Significant differences were found between hospitals that had at least 1 breach and hospitals that did not have a breach affecting 500 or more individuals during the study period (Table 2). Bivariate descriptive statistics comparing hospitals with and without data breaches showed unadjusted differences in terms of hospital type, size, and ownership. Specifically, teaching hospitals (18% with a data breach vs 3% without a breach) and pediatric hospitals (6% with a breach vs 2% without) had higher percentages of data breaches. Larger hospitals also had a higher percentage of data breaches (26% with a data breach vs 10% without). In addition, a lower percentage of investor-owned (for-profit) hospitals (15% with a data breach vs 22% without) and other specialty hospitals (6% with a data breach vs 12% without) had at least 1 data breach. In bivariate descriptive analyses, health IT sophistication, biometric security use, health system membership, hospital region, and area characteristics were not significantly different in terms of data breach percentages.

Location of Data Breaches in Hospitals
The location of breached data and the number of individuals affected varied greatly among hospitals (Figure 1). Data breaches of paper/
films occurred most frequently (65 hospitals). Data located in “other locations” (eg, breaches not from paper/films, laptop computers, email, desktop computers, EHRs, or network servers, which were reported in 56 hospitals) and in laptops (in 51 hospitals) were the second and third most prevalent, respectively. The numbers of unsecured PHI breaches from email (in 34 hospitals) and desktop computers (in 33 hospitals) were approximately equal during the study period. EHR data were breached in 19 hospitals. Although network server breaches occurred most infrequently (in 10 hospitals), these breaches compromised the highest number of individuals (4,613,858 affected).

**Types of Data Breaches in Hospitals**

Types of data breaches and the number of individuals affected by those types of breaches varied significantly among hospitals (Figure 2). Thefts occurred most frequently (in 112 hospitals), followed by unauthorized access/disclosure (in 54 hospitals), whereas hacking/IT incidents from 27 hospitals affected the most individuals (4,685,426).

Two-way ANOVA indicated no statistically significant differences in the number of patients affected between data location/mode ($P = .455$) or type of breach ($P = .443$). There were, however, statistically significant differences between frequency of data breaches occurring from network servers and EHRs ($P = .018$) and between network servers and paper films ($P = .003$).

**Multivariate Results**

Adjusted results showed similar associations among biometric use for security and hospital characteristics, including type and size. Pediatric hospitals (odds ratio [OR], 5.1; 95% CI, 2.5-10.2) and teaching hospitals (OR, 5.0; 95% CI, 2.9-8.4) were associated with an increased risk of data breaches compared with other types of hospitals. In addition, small (OR, 0.3; 95% CI, 0.2-0.6) and medium (OR, 0.6; 95% CI, 0.3-0.9) hospitals by bed size were associated with a decreased risk for data breaches compared with larger hospitals. Hospital setting, health system membership, health IT sophistication, federal status, market concentration, and ownership status were not predicting factors of a data breach affecting 500 or more patients (Table 3).

**DISCUSSION**

Approximately one-third of all healthcare data breaches occurred in hospitals, and the most individual patients were impacted when hospitals were breached compared with other types of healthcare providers, such as doctors, nurses, and social workers. Therefore,
federal incentives to do so, the most common type of data breach in hospitals occurred with paper records and films. These paper and film breaches occurred mostly due to theft, improper disposal, and unauthorized access. However, the overall number of patients affected by these breaches was relatively small. Conversely, network and film breaches occurred mostly due to theft, improper disposal, and password reliance by utilizing fingerprint, gesture, eye, facial, and voice recognition modalities. This is further strengthened via 2-factor authentication protocols, which often combine a username and password with a physical biometric scan to grant access. As shown by the findings of this study and others, computers have served as a source of data breaches because generic usernames and passwords make them easily accessible. Hospital unit computers are easy targets because they contain patient and staff information, such as referral letters, nursing reports, patient charts, audits, handovers, and staff sick leave lists, directly on the desktop. Biometric technology is a valuable means of reducing username and password reliance by utilizing fingerprint, gesture, eye, facial, and voice recognition modalities. This is further strengthened via 2-factor authentication protocols, which often combine a username and password with a physical biometric scan to grant access. Given that the most common location of breaches in a hospital is currently paper files/films, the addition of biometric technology is not likely to impact this number. However, as the diffusion of EHR technology continues in the United States and cyber threats become more prevalent, these hard-copy breaches will presumably continue to be minimized as long as necessary security policies are upheld and security audits are practiced.

As shown by the findings of this study and others, computers have served as a source of data breaches because generic usernames and passwords make them easily accessible. Hospital unit computers are easy targets because they contain patient and staff information, such as referral letters, nursing reports, patient charts, audits, handovers, and staff sick leave lists, directly on the desktop. Biometric technology is a valuable means of reducing username and password reliance by utilizing fingerprint, gesture, eye, facial, and voice recognition modalities. This is further strengthened via 2-factor authentication protocols, which often combine a username and password with a physical biometric scan to grant access. Given that the most common location of breaches in a hospital is currently paper files/films, the addition of biometric technology is not likely to impact this number. However, as the diffusion of EHR technology continues in the United States and cyber threats become more prevalent, these hard-copy breaches will presumably continue to be minimized as long as necessary security policies are upheld and security audits are practiced.

As shown by the findings of this study and others, computers have served as a source of data breaches because generic usernames and passwords make them easily accessible. Hospital unit computers are easy targets because they contain patient and staff information, such as referral letters, nursing reports, patient charts, audits, handovers, and staff sick leave lists, directly on the desktop. Biometric technology is a valuable means of reducing username and password reliance by utilizing fingerprint, gesture, eye, facial, and voice recognition modalities. This is further strengthened via 2-factor authentication protocols, which often combine a username and password with a physical biometric scan to grant access. Given that the most common location of breaches in a hospital is currently paper files/films, the addition of biometric technology is not likely to impact this number. However, as the diffusion of EHR technology continues in the United States and cyber threats become more prevalent, these hard-copy breaches will presumably continue to be minimized as long as necessary security policies are upheld and security audits are practiced.

As shown by the findings of this study and others, computers have served as a source of data breaches because generic usernames and passwords make them easily accessible. Hospital unit computers are easy targets because they contain patient and staff information, such as referral letters, nursing reports, patient charts, audits, handovers, and staff sick leave lists, directly on the desktop. Biometric technology is a valuable means of reducing username and password reliance by utilizing fingerprint, gesture, eye, facial, and voice recognition modalities. This is further strengthened via 2-factor authentication protocols, which often combine a username and password with a physical biometric scan to grant access. Given that the most common location of breaches in a hospital is currently paper files/films, the addition of biometric technology is not likely to impact this number. However, as the diffusion of EHR technology continues in the United States and cyber threats become more prevalent, these hard-copy breaches will presumably continue to be minimized as long as necessary security policies are upheld and security audits are practiced.
Although many aspects surrounding healthcare privacy and data breaches were included in this study, not all could be accounted for. One limitation was that the OCR only had data available on breaches that affect 500 or more patients per case. Information on breaches and the at-fault facilities for all breaches affecting 499 patients or less was not accessible. In addition, the policy language around what is considered reportable in privacy breaches is vague, so these occurrences may have been under- or over-reported depending on the individual facility. The OCR data did not give specific dates and months of data breaches, only the year, and the 2016 data were not all available when we conducted the analysis. In addition, the survey questions used to measure biometric security systems only measured capability and infrastructure support and did not reflect use. Furthermore, all data used for this analysis were self-reported. Not all hospitals in the OCR database could be matched to the HIMSS and AHA analytics data files due to the possibility of facility closure, unverified city of breach, or inadequate information. In total, 15 hospitals were unable to be matched. VA hospitals were also unable to be matched because HIMSS does not monitor VA information. This study focused on hospitals due to the number of individuals affected by breaches in these types of facilities. However, other types of facilities, including physician practices, health plans, and clearinghouses, do have breaches and should be an area of focus for future research. Finally, it is important to note that most states have passed additional legislation that is more or less strict than the current federal legislation, which may or may not impact breach reporting.

CONCLUSIONS

Although there are more group/physician practices within the United States than hospitals, the overall number of individual patients treated, and who thus have data created and stored within the record system, is greater within hospitals. Routine audits required by cyber-insurance coverage may help healthcare facilities recognize, and repair, their vulnerabilities before a breach occurs. Accordingly, information security systems should be concurrently implemented alongside health information technologies. Improving access control and prioritizing patient privacy will be important steps in minimizing future breaches.

REFERENCES


A False Sense of Security

Full text and PDF at www.ajmc.com