

# Privacy with Health Information Technologies: What Story Do Data Breaches in US Tell Us?

Sylvestre Uwizeyemungu<sup>1</sup><sup>a</sup> and Placide Poba-Nzaou<sup>2</sup><sup>b</sup>

<sup>1</sup>*Département des Sciences Comptables, Université du Québec à Trois-Rivières (UQTR),  
3351, boul. des Forges, C.P. 500, Trois-Rivières (Québec), Canada*

<sup>2</sup>*Département d'Organisation et Ressources Humaines, ÉSG – Université du Québec à Montréal (UQAM),  
315, Ste-Catherine Est, Montréal (Québec), Canada*

**Keywords:** Health Data Breach, Health Information Technology, Privacy.


**Abstract:** Over the last decades, health policy makers have encouraged healthcare organizations to leverage health information technology (HIT) for improving the accessibility, the quality, and the efficiency of health service delivery. The adoption of HIT has contributed to the digitization of health data, which has made these data vulnerable to information technology (IT) related security breaches. Based on data published by the US Department of Health and Human Services (DHHS), we analyze the portrait of health data breaches in the USA from 2009 to 2018 in order to figure out whether there are clear patterns of breach that stand out. In addition to descriptive statistics characterizing health data breaches, this study suggests three well-separated patterns of these breaches: (1) breaches mainly related to hacking / IT incident, (2) breaches due to unauthorized access / disclosure, and (3) breaches due to theft. All these patterns of breaches have different implications regarding priorities for health IT security and privacy professionals. However, further investigations with additional data are needed to fully comprehend the phenomenon of health data breaches and their implications in terms of IT security and privacy.


## 1 INTRODUCTION

Over about the last three decades, health policy makers have encouraged healthcare organizations to intensively adopt health information technology (HIT) (Blumenthal, 2009, 2011; Rozenblum et al., 2011). The premise was that the leverage of information technology (IT) in healthcare sector will play a major role in improving the accessibility, the quality, the safety, and the efficiency of healthcare services (Daniel, 2018; McKenna, Dwyer, & Rizzo, 2018; Tubaishat, 2019; Wani & Malhotra, 2018). Incentives for the “meaningful use” (Hogan & Kissam, 2010b) of IT has led to higher rates of HIT adoption and use among targeted hospitals (Adler-Milstein & Jha, 2017; Jones & Furukawa, 2014). But, at the same time, the increasingly digitized health information has become vulnerable to IT-related security breaches, thus exposing healthcare organizations to “the mixed blessing of the digital

age” (Myers, Frieden, Bherwani, & Henning, 2008, p. 794). Data breaches in healthcare sector have become a common occurrence: the number of data breaches is on the rise, as well as the number of individuals affected (Koczkodaj, Mazurek, Strzałka, Wolny-Dominiak, & Woodbury-Smith, 2019; Liu, Musen, & Chou, 2015). Given the highly sensitive nature of health information, the potential impacts of health data breaches may be disastrous for patients, healthcare providers, and the healthcare system as a whole. Therefore, there is a pressing need to reinforce measures aiming at shielding health data from IT-enabled privacy breaches. One of the first steps in this endeavour is for researchers and practitioners to understand the nature and the patterns underlying the health data breaches. The premise here is that different data breach patterns would require security measures that may vary from one pattern to another.

In this study, we analyze the reported health data breaches affecting at least 500 individuals published

<sup>a</sup> <https://orcid.org/0000-0002-1532-8848>

<sup>b</sup> <https://orcid.org/0000-0002-7007-764X>

by the US Department of Health and Human Services (DHHS). Our first aim is to examine the main characteristics of health data breaches: what is the nature of breaches, and where are they likely to occur. Otherwise stated, we want to outline in which form, and where patient's protected health information is most vulnerable. Our second aim is to assess to what extent meaningful patterns can be drawn from US health data breaches based on their type and location.

## 2 BACKGROUND

### 2.1 The Meaningful Use of HIT

In the United States (U.S.), the federal legislation encourages the meaningful use of electronic health records (EHRs), through notably the 1996 Health Insurance Portability and Accountability Act (HIPAA) and the 2009 Health Information Technology for Economic and Clinical Health (HITECH) Act. The "meaningful use" criteria intended to encourage hospitals to achieve a significant level of IT in healthcare processes capable of spurring the improvement of quality, safety, and efficiency of patient health information.

The financial incentives for eligible healthcare services providers were intended to alleviate the cost burden associated with the EHR implementation considered as a major impediment to EHR adoption and use (Hogan & Kissam, 2010a). It seems that this legislation has attained its objective of spurring EHR adoption, at least for hospital settings. Different studies (Adler-Milstein & Jha, 2017; Walker, Mora, Demosthenidy, Menachemi, & Diana, 2016; Wolf, Harvell, & Jha, 2012) comparing HITECH-eligible hospitals with non eligible hospitals noted higher adoption rates of electronic health records (EHR) for eligible hospitals following the HITECH incentives' implementation. According to the Office of the National Coordinator for Health Information Technology, as of 2016, "over 95% of hospitals eligible for the Medicare and Medicaid EHR Incentive Program have achieved meaningful use of certified health IT" (ONC, 2018). Even though this rate indicates a high adoption rate of health IT, it should be interpreted with caution. First of all, that rate accounts for only eligible hospitals. It thus excludes a large part of the U.S. healthcare sector that comprises inpatient rehabilitation hospitals, inpatient psychiatric hospitals, long-term acute care hospitals,

and other healthcare services providers such as nursing homes, and home health agencies (Wolf et al., 2012). Secondly, the ONC statistic does not include small and independent practitioners whose HIT adoption rates remain significantly lower even when they are eligible for the meaningful use incentive program (Hsiao, Decker, Hing, & Sisk, 2012). In spite of discrepancies in HIT adoption rates and in HIT's meaningfulness use, overall, statistics show trends of increasing adoption in US hospitals (Adler-Milstein et al., 2015; Adler-Milstein et al., 2017).

The meaningful use of HIT leads to the digitization of healthcare, which has to be accompanied by IT-security measures. However, experts agree that the implementation of IT security measures is not keeping pace with the computerization of the health sector (Kruse, Frederick, Jacobson, & Monticone, 2017). Many small-and medium-sized healthcare organizations tend to not undertake significant IT security programs due to limited human and financial resources; and implementing security and privacy measures in large healthcare organizations may take a lot of time due to managerial and structural rigidity of large systems (Uwizeyemungu, Poba-Nzaou, & Cantinotti, 2019). The absence, the insufficiency, and/or the inadequacy of IT-related security and privacy practices in healthcare organizations, despite increasing HIT adoption, may explain the data breaches that are occurring.

These IT-related security and privacy practices cover a wide range of measures: access control, data storage, data anonymization, data encryption, IT security training, etc. IT security training of all human resources is particularly important, as data breach is also a human issue: the majority (58%) of recent data breaches are due to internal actors whose actions encompass both human error and misuse (Chernyshev, Zeadally, & Baig, 2019). Training helps increase awareness, promote secure behaviour, and avoid or limit human errors.

### 2.2 HIT and Privacy Breaches

The notion of "protected health information" (PHI) is key to ensuring the privacy in the context of healthcare delivery. It refers to "health data created, received, stored, or transmitted by HIPAA-covered entities and their business associates in relation to the provision of healthcare, healthcare operations and payment for healthcare services"<sup>3</sup>. The definition

<sup>3</sup> <https://www.hipaajournal.com/what-is-protected-health-information/>

goes on to precise that PHI “includes all individually identifiable health information” and gives some details of that information: “demographic data, medical histories, test results, insurance information, and other information used to identify a patient or provide healthcare services or healthcare coverage”. Although the definition provides details that reflect the scope of the information covered, it is at the same time limiting in its very essence by its insistence on “*individually identifiable* health information”. This qualification refers to the notion of “personally identifiable information” (PII) that experts argue it has become meaningless (Hiller, 2016; Milne, Pettinico, Hajjat, & Markos, 2017): the availability of powerful IT-enabled algorithms allow to accurately identify people by combining multiple seemingly harmless data and / or by tapping multiple sources. In addition, health information is a potential target for collection by an array of organizations not subject to stringent regulation or oversight (Libert, 2015). These organizations include businesses in data broker industry, retailers that would gather some health information indirectly through the customer’s purchases of health-related products, or even healthcare organizations whose activities do not place them under the HIPAA jurisdiction or under similar rules.

The nature of PHI makes it very sensitive and very attractive for marketing or criminal purposes. Different practitioners in the healthcare sector (physicians, hospitals, pharmacies, insurers) and outside (such as advertising agencies) are interested in harvesting health information for marketing purposes, and in spite of an apparent stringent legislation, “loopholes exist where consumers’ private health information is ‘up for sale’” (Levy & Royne, 2009, p. 466).

In addition to this interest for marketing purposes, health data are particularly attractive to cybercriminals. In the underground market, healthcare records are generally more valuable than other types of data including financial data (Ablon, 2018; McNeal, 2014). The attractiveness of health information on black markets is justified by the multiple types of crime that are possible with such data (Chernyshev et al., 2019): financial fraud, identity theft, access to healthcare services for non-insured patients, access to drug prescription, vindication, extortion, etc. In 2013, health data accounted for 44% of all identity thefts reported in the US, and were far costlier than breaches in retail or financial sectors (McNeal, 2014).

Consequences of health data breaches cover a wide range of effects of various gravity and may

affect all the stakeholders in the healthcare system. The most disastrous consequence is arguably the compromise of patient safety (e.g. Sametinger, Rozenblit, Lysecky, & Ott, 2015). Patients may also suffer consequences due to identity theft (then used for fraudulent activities for instance), they could be exposed to financial losses, to psychological discomforts like mental anguish or embarrassment, loss of trust, etc. Statistics from 2017 show that two thirds of patients (66%) express concerns about the safety of their medical records in the context of electronic health information exchange (ONC, 2018). According to the same statistics, these concerns lead 10% of individuals to withhold information from their healthcare providers.

For healthcare providers, they may experience damages to their reputation that would result in loss of patients’ trust, confidence, and loyalty, in addition to being subject to liabilities (damages and fines) that would put their operational capabilities under duress. For the healthcare system and public health, if health data breaches lead patients and healthcare providers to turn away from HIT usage, the advantages related to the meaningful use of HIT represent missed opportunities (Agnò & Guo, 2013). This means that the healthcare system would deliver services of sub-optimal quality, with less effectiveness, less efficiency. There are also some risks with regard to public health: in addition to patients avoiding care, the contagious disease reporting and treatment system would be affected, as well as the sharing of data for health surveillance and for research and education (Myers et al., 2008): stringent restrictions and impediments to the sharing of health data may come as a reaction to the spread of data breaches.

In the US, the same legislation that promotes the meaningful use of health IT contains provisions meant to uphold the privacy of “protected health information” (PHI). This legislation makes it mandatory to report any privacy breach that affect 500 or more persons. Data gathered following this legislation can give us valuable insights on the nature and patterns of health data breaches.

### 3 METHODS

#### 3.1 Data Source and Sample

Following the requirements of the HITECH Act, the Office of Civil Rights (OCR) of the US Department of Health and Human Services (DHHS) maintains an online database on “breaches of unsecured protected health information affecting 500 or more individuals”

(HHS Office for Civil Rights, 2018). The database recorded its first reported breaches in October 2009. We gathered these data from the OCR website and used them as our main material. As of December 31, 2018, the database reported a total of 2530 cases of breach comprised of two categories: 402 breaches reported within the preceding 24 months and that were still under investigation by the OCR, and 2128 archived cases of breach, that is all resolved breach reports and/or reports that were older than 24 months.

In order to better assess the extent of breaches in terms of the number of incidents and the number of individuals affected, we took into account data on the number of hospitals and the size of the population by state. For the data on the number of hospitals by state, we used statistics from the 2016 annual survey database of the American Hospital Association (AHA). For the population estimates we used the 2017 statistics from the US Census bureau.

In addition to providing information about the name and the type of the health organization breached and the state to which it belongs, the OCR data characterize each breach in terms of type, location (or mode), the number of individuals affected, as well as the date of breach submission. The type of breach refers to the nature of the event that caused the exposure of the protected health information, and includes 7 categories, namely 1) hacking/IT incident, 2) improper disposal, 3) loss, 4) theft, 5) unauthorized access/disclosure, 6) unknown, and 7) other. The location of breach refers to the tool from which originated the unauthorized disclosure of the PHI, and includes 8 following categories: 1) desktop computer, 2) electronic medical record, 3) email, 4) laptop, 5) network server, 6) other portable electronic device, 7) paper/films, and 8) other. There are 4 types of health organizations considered: 1) health plan (335; 13.2%), 2) healthcare clearing house (4; 0.2%), 3) healthcare provider (1828; 72.3%), and 4) business associate (363; 14.3%).

### 3.2 Data Analysis

In order to present a portrait of health data breaches, we analyze descriptive statistics of breaches. We also performed a cluster analysis using as clustering variables a combination of different types and locations of breaches. However, in order to avoid meaningless results, we excluded from clustering variables some under-specified types or locations of breaches. From the 7 categories of the type of breach, we excluded categories 6) “unknown”, and 7) “other”. From the 8 categories of the location of breach, we excluded categories 7) “paper/films” and

8 “other”. The paper/films” category was excluded as our study is about IT-related breaches. The other categories (unknown, other) were excluded considering that they do not bring about any exploitable information. Considering the data at hand, we used the two-step clustering algorithm. This clustering algorithm is applicable for large data sets, accepts continuous and categorical variables, and moreover, it can automatically suggest the optimal number of clusters (Bittmann & Gelbard, 2007). Besides, when compared to other clustering algorithms, the two-step algorithm was the top-ranked (Gelbard, Goldman, & Spiegler, 2007). The algorithm suggests a three-cluster solution, which was confirmed by a discriminant analysis.

## 4 RESULTS

### 4.1 The Overall Portrait of Breaches

Overall, we analyzed 2530 breaches reported between October 2009 and December 2018, affecting approximately 261.9 million records. As illustrated in Figure 1, the number of breaches is globally on the rise, as well as the number of individuals affected. If one excludes the incomplete year of 2009, 2,512 breaches are recorded over a 9-year period and over 194.5 million records are affected, that is an average number of 279 breaches per year, with 21.6 million individuals annually affected.

We present in Figure 2 the overall number of breaches and the overall number of individuals affected by type of breach. In some cases, a breach can combine more than one type. For example, a theft may occur following an improper disposal of a device containing sensitive healthcare information. Thus, the total number of breaches by type of breach is 2,642 instead of 2,530 breaches recorded in figure 1. From Figure 2, it appears that theft is the most frequent cause of health data breaches (905; 34.3%), closely followed by unauthorized access/disclosure (769; 29.1%), and hacking/IT incident (580; 22.0%). However, when one considers the number of individual records affected by breaches, the type of breach that has affected so far the highest number of individuals is by far the hacking/IT incident (145.6 million; 74.3%).

With regard to the location of breach, Figure 3 shows that the three most prevalent breaches occur from paper/films (608; 20.5%), network server (511; 17.3%), and laptops (430; 14.5%). The most significant breaches regarding the number of individuals affected occur from network servers (150.3 millions; 68.4%).



The distribution of breaches according to affected entities (Figure 4) shows that if most breaches occur at healthcare provider level (1825; 72.2%), breaches occurring from health plans are

the ones that involve the higher numbers of individuals affected: health plan related incidents represent 13.3% of all breaches but they represent 58.2% of individuals affected.

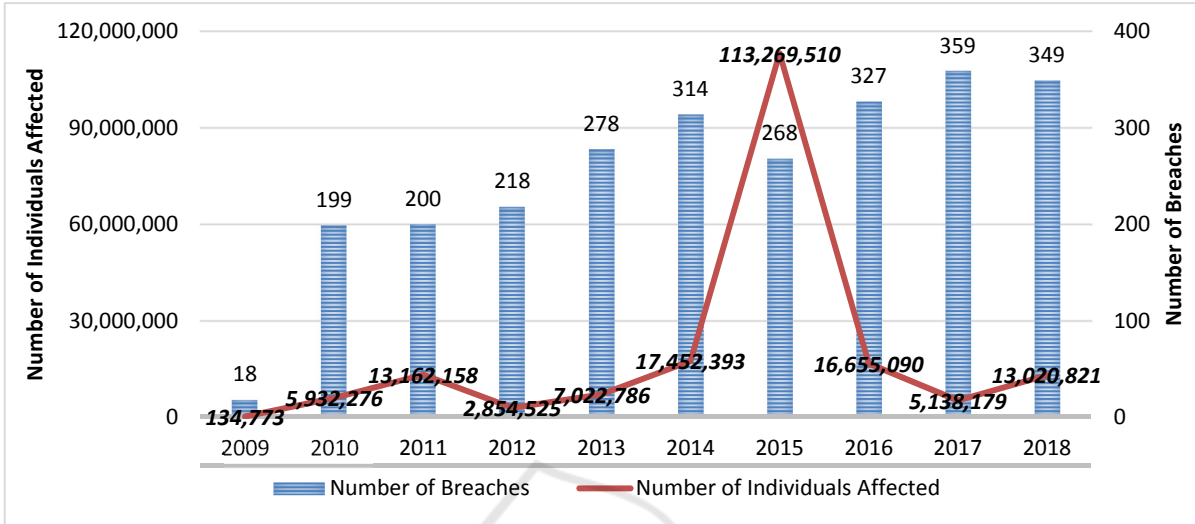


Figure 1: Evolution of Health Data Breaches (Numbers of Breaches and Individuals Affected).

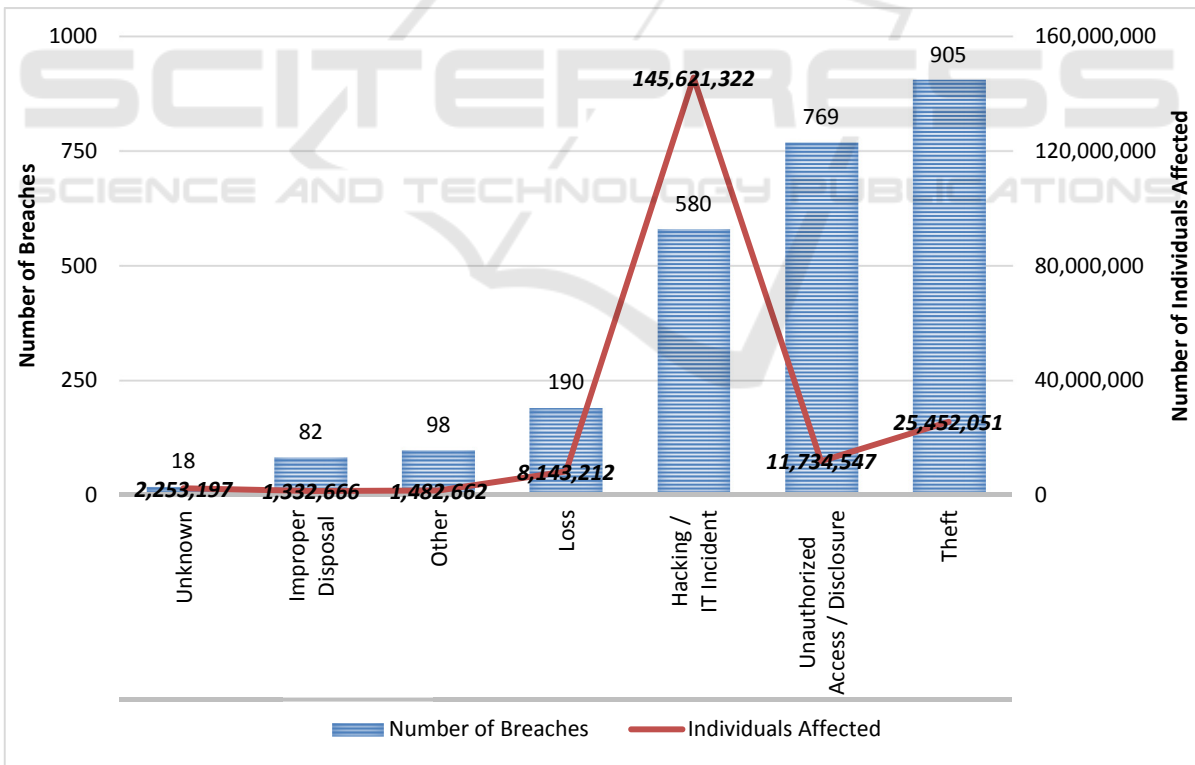


Figure 2: Number of Breaches and Number of Individuals Affected by Type of Breach.

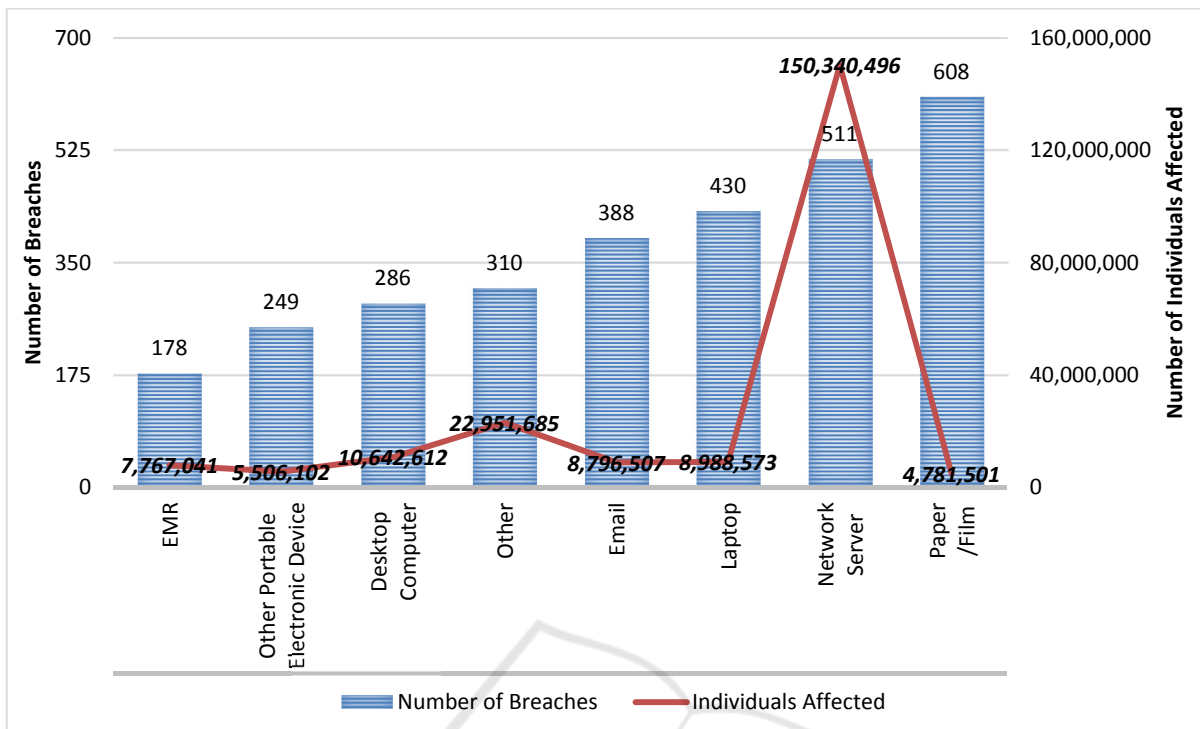


Figure 3: Number of Breaches and Number of Individuals Affected by Location of Breach.

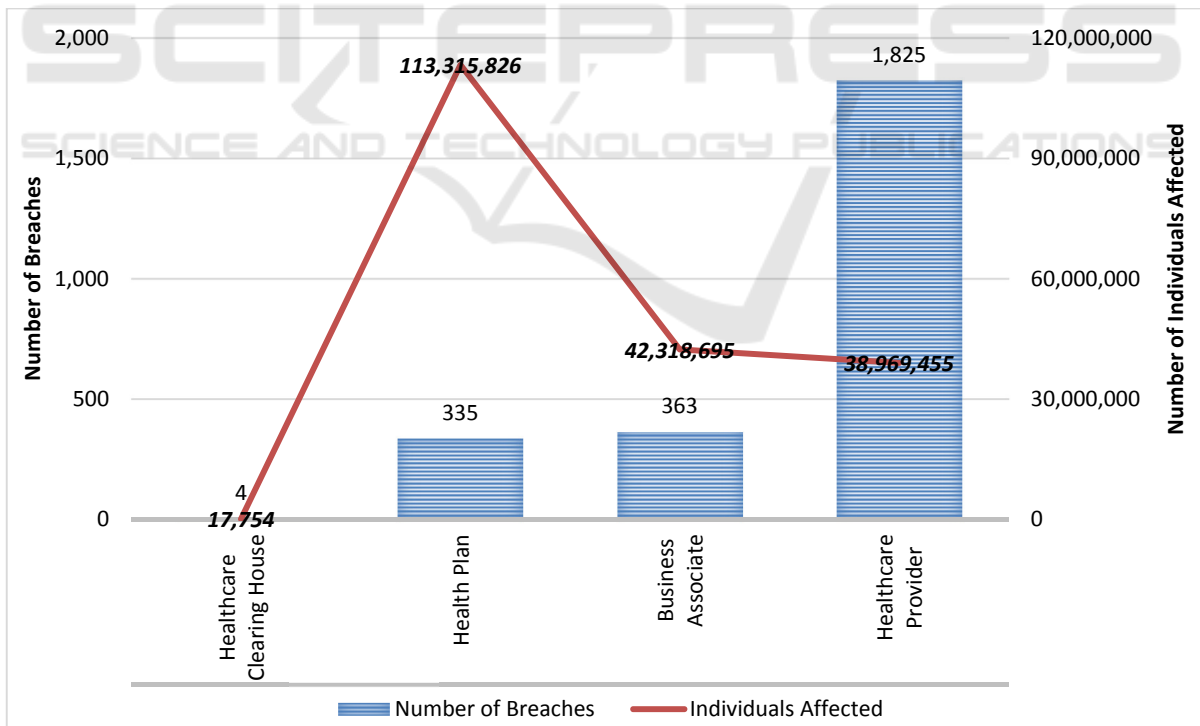


Figure 4: Number of Breaches and Number of Individuals Affected by Type of Covered Entity.

### 4.2 Most Severe Breaches

The year 2015 has set a record of over 113 millions records breached. The severest single breaches with regard to the number of individuals affected occurred in 2015 at Anthem Inc. from Indiana (78.8 millions), at Premera Blue Cross in Washington (11 millions), and at Excellus Health Plan Inc. in New York State (10 millions).

### 4.3 Breaches by State

Analyzing health data breaches by state, one notes that the states that come ahead in terms of the number of breaches are California (286), Texas (213), Florida (164), New York (140), and Illinois (125). In terms of individuals affected, the states that registered more than ten million of individuals affected are Indiana (84.7 millions), New York (17.3 millions), Washington (11.8 millions), Tennessee (11.5 millions), and California (10.1 millions).

However, it is worth noting that these numbers may depict an “unfair” portrait of each state, considering that (1) the different state healthcare systems are not of the same size, and (2) the US population is not evenly distributed across all the US states. Thus, all other things remaining equal, one would expect much more data breaches in states with larger healthcare systems, and more individuals affected in states with more people. To alleviate discrepancies stemming from the variation in the size of healthcare systems and population across states, we weighted the number of breaches by the number of registered hospitals (a proxy measure for the size of healthcare system), and we weighted the number of individuals affected by the population census (base 2017). Figure 5 presents the 15 states that come ahead in number of data breaches when one takes into account the size of each state’s healthcare system. Rhode Island comes ahead with more than one breach by registered hospital.

Figure 6 presents the 15 states that come ahead in number of individuals affected by breaches when one considers each state’s population size. In this case, the state of Indiana comes far ahead of the other states.

### 4.4 Results of Clustering Analysis

We present in Table 1 the results of the clustering procedure described in the method section. As we used dichotomous measures (1/0) to determine the type or location of each breach, the means presented in Table 1 give us at the same time the percentages of breaches for each type or location. We have highlighted characteristics that stand out (represented in more than

30% of the cases) for each cluster. The results in Table 1 are graphically depicted in Figure 7.

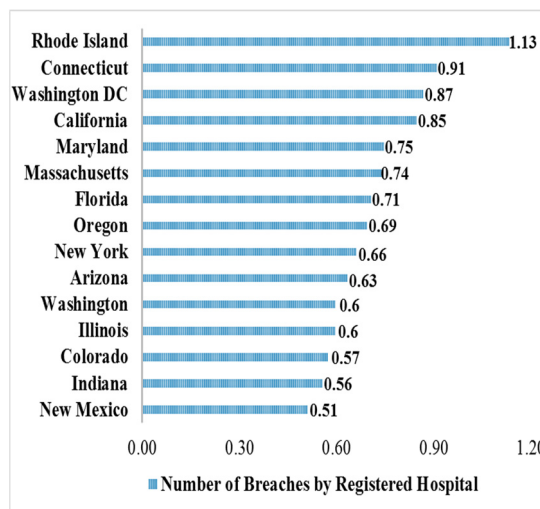


Figure 5: 15 States with Higher Number of Breaches by Registered Hospital.

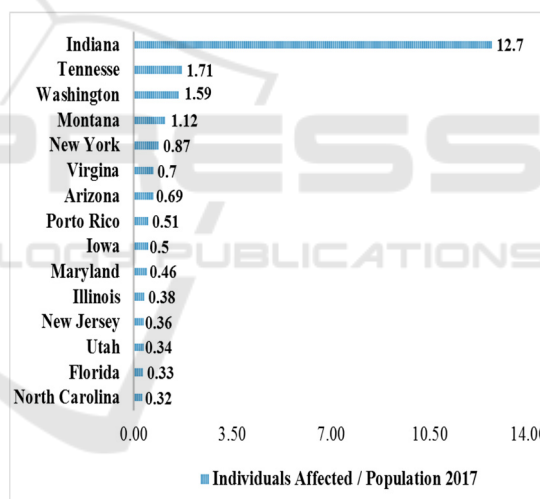


Figure 6: 15 States with Higher Number of Individuals Affected by Breaches (Weighted by Population Census).

From Table 1 and Figure 7, we can see that there are three well-separated patterns of health data breaches (3 clusters). In the first cluster (22.9% of all breaches), we find breaches that are mainly related to hacking / IT incident (99.7%), either through network server (59.2%) or e-mail (32.1%) compromise. The main characteristic of breaches in cluster 2 (29.6% of all breaches) is that they are all due to unauthorized access / disclosure (100%). The largest cluster, cluster 3 (47.5% of all breaches) groups breaches whose characteristic is that they are mainly due to theft (73.1%). It is not surprising that it is in this last cluster

that we find the relatively higher proportion of breaches of data on laptops (32.3%).

In Table 2, we breakdown the breaches in our different clusters according to the type of covered entities. Otherwise stated, we used the type of covered entity as control variable. We can infer from this figure that compared to other clusters, cluster 2 distinguishes itself by the fact that the portion of health plan entities in that cluster is statistically more present. The other two clusters have statistically comparable portions of healthcare provider entities.

## 5 IMPLICATIONS AND CONCLUSION

In this study, we present a portrait of health data breaches in the United States of America, based on public data published by the US Department of Health and Human Services (DHHS) in compliance with the requirements of the 2009 Health Information Technology for Economic and Clinical Health (HITECH) Act.

It is preoccupying to note that health data breaches are not diminishing, and that individuals affected are counted in millions every year. Given the gravity of consequences that health data breaches entail, it is important to understand their characteristics and the context in which they occur. This study contributes to this endeavour.

From a practical standpoint, our study will likely help IT security professionals in healthcare settings to prioritize their actions. Our results highlight the most prevalent events (types of breaches) that cause the exposure of the protected health information (PHI), as well as the vehicles (locations of breaches) through which the PHI is most likely to be breached. As our results show that the highest number of health data breaches is due to theft, it would be sensible to prioritize IT security practices aiming at avoiding theft in order to reduce the occurrence of data breaches. And, not surprisingly, our cluster analysis shows that theft is somehow related to laptops. Hence IT security practices aiming at securing laptops (and other portable electronic devices) and data they contain should be prioritized. As it would be naïve to expect that health data thieves will be less active in

Table 1: Patterns of Health Data Breaches in the US.

	Cluster 1 (n=579; 22.9%)	Cluster 2 (n=749; 29.6%)	Cluster 3 (n=1202; 47.5%)	Anova F
TB - Hacking / IT Incident	0.997 <sub>a</sub>	0.003 <sub>b</sub>	0.001 <sub>b</sub>	111,998.703*
TB - Loss	0.002 <sub>b</sub>	0.004 <sub>b</sub>	0.155 <sub>a</sub>	113.861*
TB - Theft	0.007 <sub>c</sub>	0.029 <sub>b</sub>	0.731 <sub>a</sub>	1,544.746*
TB - Unauthorized Access / Disclosure	0.016 <sub>b</sub>	1.000 <sub>a</sub>	0.009 <sub>b</sub>	32,963.263*
TB - Improper Disposal	0.000 <sub>b</sub>	0.003 <sub>b</sub>	0.067 <sub>a</sub>	44.035*
LB - Desktop Computer	0.150 <sub>a</sub>	0.059 <sub>b</sub>	0.129 <sub>a</sub>	16.727*
LB - EMR	0.069 <sub>b</sub>	0.131 <sub>a</sub>	0.033 <sub>c</sub>	34.463*
LB - Email	0.321 <sub>a</sub>	0.215 <sub>b</sub>	0.034 <sub>c</sub>	156.742*
LB - Laptop	0.036 <sub>b</sub>	0.028 <sub>b</sub>	0.323 <sub>a</sub>	222.846*
LB - Network Server	0.592 <sub>a</sub>	0.146 <sub>b</sub>	0.049 <sub>c</sub>	519.051*
LB - Other Portable Electronic Devices	0.007 <sub>c</sub>	0.044 <sub>b</sub>	0.176 <sub>a</sub>	86.385*

\* :  $p < 0.001$

a, b, c: Within rows, different subscripts indicate significant ( $p < 0.05$ ) pair-wise differences between means on Tamhane's T2 (post hoc) test.



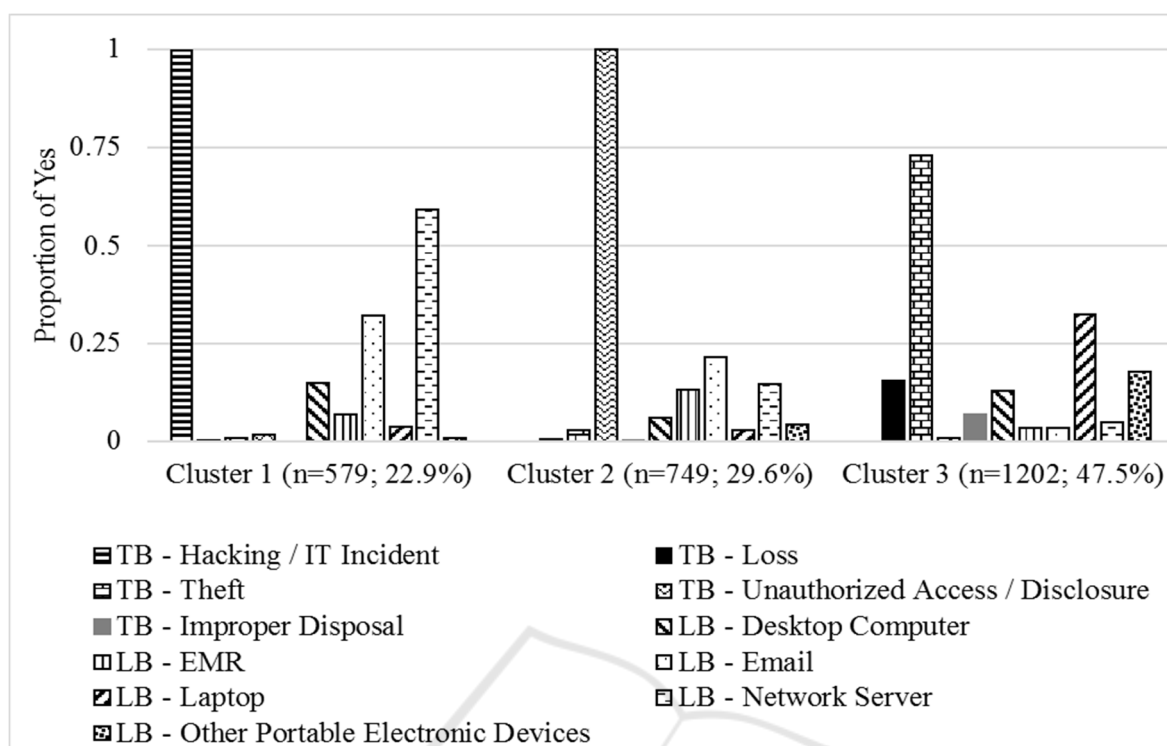


Figure 7: Graphic Representation of the Patterns of Data Breaches Resulting from the Cluster Analysis.

Table 2: Breakdown of Clusters According to the Covered Entity Type.

	Cluster 1 (n=579; 22.9%)	Cluster 2 (n=749; 29.6%)	Cluster 3 (n=1202; 47.5%)	Anova F
Business Associate	0.126	0.139	0.155	1.40
Health Plan	0.121 <sub>b</sub>	0.219 <sub>a</sub>	0.084 <sub>b</sub>	38.07*
Health-care Clearing House	0.000	0.001	0.003	0.79
Health-care Provider	0.753 <sub>a</sub>	0.641 <sub>b</sub>	0.759 <sub>a</sub>	17.96*

\*:  $p < 0.001$

a, b : Within rows, different subscripts indicate significant ( $p < 0.05$ ) pair-wise differences between means on Tamhane's T2 (post hoc) test.

the years ahead, healthcare organizations have the responsibility to be proactive, either by minimizing the opportunities available to thieves or by taking measures that would limit the damage in the event of successful theft. One way to reduce theft opportunities, for example, would be to reduce to a

strict minimum the backup of health data on portable and mobile devices, which are easy to steal. One way to limit the damage in the event of a successful theft is to encrypt all health data stored on hard drives. If healthcare organizations encrypt all sensitive records, thieves do not get anything of value from the stolen data, and the privacy of patients is preserved.

Our results also suggest that in order to reduce the number of individuals affected by health data breaches, healthcare organizations should prioritize IT security practices aiming at preventing hacking / IT incidents, notably by implementing measures related to the protection of network servers and to the elimination of unauthorized access / disclosure. The encryption of records in storage or being transmitted would limit the damage in case of hacking / IT incident. Although these measures are important for all covered entities, health plans should be prioritized as they are the main source of breaches that affect more individuals.

There are some limitations that affect the scope of our study. Obviously, the first limitation comes from our dataset itself that accounts only for breaches involving at least 500 individuals, which is the lower limit fixed by the law for mandatory reporting of the breach. Breaches under that limit go unreported, which prevent us from visualizing the whole picture of health data breaches. Another limitation is that we

did not gather data on individual healthcare organizations to analyze whether there are some characteristics of that organizations that can be correlated with the patterns of data breaches. This is an interesting research avenue that would allow one to go further in explaining our results. In the same vein, in order to better contextualize the breaches reported from different states, it would be interesting to complement data from the US DHHS with data on the states from other sources. For example, it would be worthwhile to gather data on the specificities of states with regard to health information exchange regulations, or with regard to the rates of health IT adoption; all data that would probably help explain, or at least contextualize the levels of health data breaches.

In spite of the limitations stated above, we hope that this study contributes to a better understanding of health data breaches related to the use of health IT, which is a first step to devise IT security and privacy practices to prevent data breaches from happening.

## REFERENCES

- Ablon, L. (2018). Data thieves. *The motivations of cyber threat actors and their use and monetization of stolen data*. Santa Monica, Ca: Rand Corporation.
- Adler-Milstein, J., DesRoches, C. M., Kralovec, P., Foster, G., Worzala, C., Charles, D., . . . Jha, A. K. (2015). Electronic health record adoption in US hospitals: Progress continues, but challenges persist. *Health Affairs*, *34*(12), 2174-2180.
- Adler-Milstein, J., Holmgren, A. J., Kralovec, P., Worzala, C., Searcy, T., & Patel, V. (2017). Electronic health record adoption in US hospitals: The emergence of a digital "advanced use" divide. *Journal of the American Medical Informatics Association*, *24*(6), 1142-1148.
- Adler-Milstein, J., & Jha, A. K. (2017). HITECH Act drove large gains in hospital electronic health record adoption. *Health Affairs*, *36*(8), 1416-1422.
- Agno, C. F., & Guo, K. L. (2013). Electronic health systems: Challenges faced by hospital-based providers. *The Health Care Manager*, *32*(3), 246-252.
- Bittmann, R. M., & Gelbard, R. M. (2007). Decision-making method using a visual approach for cluster analysis problems; indicative classification algorithms and grouping scope. *Expert Systems*, *24*(3), 171-187.
- Blumenthal, D. (2009). Stimulating the adoption of health information technology. *The New England Journal of Medicine*, *360*(15), 1477-1479.
- Blumenthal, D. (2011). Implementation of the federal health information technology initiative. *The New England Journal of Medicine*, *365*(25), 2426-2431.
- Chernyshev, M., Zeadally, S., & Baig, Z. (2019). Healthcare data breaches: Implications for digital forensic readiness. *Journal of Medical Systems*, *43*(1), 1-12.
- Daniel, O. U. (2018). Effects of health information technology and health information exchanges on readmissions and length of stay. *Health Policy and Technology*, *7*(3), 281-286.
- Gelbard, R., Goldman, O., & Spiegler, I. (2007). Investigating diversity of clustering methods: An empirical comparison. *Data & Knowledge Engineering*, *63*(1), 155-166.
- HHS Office for Civil Rights. (2018). Breach portal: Notice to the secretary of HHS breach of unsecured protected health information. Retrieved from [https://ocrportal.hhs.gov/ocr/breach/breach\\_report.jsf](https://ocrportal.hhs.gov/ocr/breach/breach_report.jsf)
- Hiller, J. S. (2016). Healthy predictions? Questions for data analytics in health care. *American Business Law Journal*, *53*(2), 251-314.
- Hogan, S. O., & Kissam, S. M. (2010). Measuring meaningful use Health affairs (Project Hope), April 2010, Vol.29 (4), pp.601-6. *Health Affairs*, *29*(4), 601-606.
- Hsiao, C.-J., Decker, S. L., Hing, E., & Sisk, J. E. (2012). Most physicians were eligible for federal incentives in 2011, but few had EHR systems that met meaningful-use criteria. *Health Affairs*, *31*(5), 1100-1007.
- Jones, E. B., & Furukawa, M. F. (2014). Adoption and use of electronic health records among federally qualified health centers grew substantially during 2010-12. *Health Affairs*, *33*(7), 1254-1261.
- Koczkodaj, W. W., Mazurek, M., Strzałka, D., Wolny-Dominiak, A., & Woodbury-Smith, M. (2019). Electronic health record breaches as social indicators. *Social Indicators Research*, *141*(2), 861-871.
- Kruse, C. S., Frederick, B., Jacobson, T., & Monticone, D. K. (2017). Cybersecurity in healthcare: A systematic review of modern threats and trends. *Technology and Health Care*, *25*(1), 1-10.
- Levy, M., & Royne, M. B. (2009). Up for sale: Consumer medical information. *Journal of Consumer Marketing*, *26*(7), 465-467.
- Libert, T. (2015). Privacy implications of health information seeking on the web. *Communications of the ACM*, *58*(3), 68-77.
- Liu, V., Musen, M. A., & Chou, T. (2015). Data breaches of protected health information in the United States. *Journal of the American Medical Association*, *313*(14), 1471-1473.
- McKenna, R. M., Dwyer, D., & Rizzo, J. A. (2018). Is HIT a hit? The impact of health information technology on inpatient hospital outcomes. *Applied Economics*, *50*(27), 3016-3028.
- McNeal, M. (2014). Hacking health care. *Marketing Health Services*, *34*(3), 16-21.
- Milne, G. R., Pettinico, G., Hajjat, F. M., & Markos, E. (2017). Information sensitivity typology: Mapping the degree and type of risk consumers perceive in personal data sharing. *The Journal of Consumer Affairs*, *51*(1), 133-161.
- Myers, J., Frieden, T. R., Bherwani, K. M., & Henning, K. J. (2008). Privacy and public health at risk: Public health confidentiality in the digital age. *American Journal of Public Health*, *98*(5), 793-801.

- ONC. (2018). Health IT dashboard - Quick stats. Retrieved from <https://dashboard.healthit.gov/quickstats/quickstats.php>
- Rozenblum, R., Jang, Y., Zimlichman, E., Salzberg, C., Tamblyn, M., Buckeridge, D., . . . Tamblyn, R. (2011). A qualitative study of Canada's experience with the implementation of electronic health information technology. *Canadian Medical Association Journal, 183*(5), E281-E288.
- Sametinger, J., Rozenblit, J., Lysecky, R., & Ott, P. (2015). Security challenges for medical devices. *Communications of the ACM, 58*(4), 74-82.
- Tubaishat, A. (2019). The effect of electronic health records on patient safety: A qualitative exploratory study. *Informatics for Health and Social Care, 44*(1), 79-91.
- Uwizeyemungu, S., Poba-Nzaou, P., & Cantinotti, M. (2019). European hospitals' transition toward fully electronic-based systems: Do IT security and privacy practices follow? *JMIR Medical Informatics, 7*(1), 1-16.
- Walker, D., Mora, A., Demosthenidy, M. M., Menachemi, N., & Diana, M. L. (2016). Meaningful use of EHRs among hospitals ineligible for incentives lags behind that of other hospitals, 2009-13. *Health Affairs, 35*(3), 495-501E.
- Wani, D., & Malhotra, M. (2018). Does the meaningful use of electronic health records improve patient outcomes? *Journal of Operations Management, 60*(4), 1-18.
- Wolf, L., Harvell, J., & Jha, A. K. (2012). Hospitals ineligible for federal meaningful-use incentives have dismally low rates of adoption of electronic health records. *Health Affairs, 31*(3), 505-5013.

