Archives of Surgical Research | Original Investigation

Recent Healthcare Information Breaches and Their Lessons

Haroon Elahi, Oana Geman

IMPORTANCE Modern health facilities are continually reporting data breaches, which raises the need to understand underlying factors and design mitigation strategies accordingly.

OBJECTIVE The objective of this research is to analyze real healthcare data breaches, identify underlying factors, and report the lessons that may help in preventing the future breaches.

DESIGN This study is based on exploratory analysis of electronic health record breaches reported to the U.S. Department of Health and Human Services, Office for Civil Rights during the past twenty-four months (December 2018 - December 2020).

DATA SOURCES This study uses the data of breaches reported within the last twenty-four months, and currently under investigation by the Office for Civil Rights and provided, obtained from the breach portal of the U.S. Department of Health and Human Services, Office for Civil Rights. These breaches affected more than 42 million individuals.

METHODS: We analyze 698 breach cases affecting more than 42 million individuals to identify underlying attack patterns, trends, outliers, and unexpected results, and major factors leading to these breaches in recently reported electronic health record breach cases.

RESULTS The frequency of data breaches reported during the past twenty-four months shows an increasing trend; their overall impact size is consistent, with a few exceptions. The most significant data breaches involved business associates, healthcare providers, health plans, and healthcare clearinghouses, with healthcare providers and business associates impacting about 83% of the affected individuals' privacy. However, most breaches occurred at smaller entities. Hacking of emails and network servers are the most common breach types, followed by unauthorized access, theft, improper disposal of records and devices, and loss.

CONCLUSION The nature and size of the incidents suggest paying particular attention to human factors, small-sized healthcare entities, business associates, and continuous revision of related security standards and frameworks.

KEYWORDS: Security attacks, Data breach, Electronic health records, Human factors

HOW TO CITE Elahi H, Geman O. Recent Healthcare Information Breaches and Their Lessons. *Archives of Surgical Research*. 2020;1(4):17-23. https://doi.org/10.48111/2020.04.05

Original Investigation

Author Affiliations: Author affiliations are listed at the end of this article.
Corresponding Author: Oena Geman, Department of Health and Human Development, Faculty of Electrical Engineering and Computer Sciences, Stefan cel Mare University of Suceava, Romania. Email: oana.geman@usm.ro

https://doi.org/10.48111/2020.04.05

Computing-based health information technology (HIT) systems and electronic health records (EHRs) target to improve the overall healthcare system by enabling efficient data sharing among different healthcare system stakeholders [1]. The confidentiality and accuracy of the EHRs and HIT systems' security and reliability are prerequisites for building users' trust in these systems, particularly patients whose information is collected, processed, stored, and transmitted by these systems. Since this information includes sensitive data that, if exposed, can have severe implications for the data subjects and the health service providers, new privacy and security issues are arising [2].
For example, on the one hand, easy access to this information by different HIT stakeholders improves the healthcare delivery system. Still, the exposure of this information can affect the patients' insurance, career, or relationships. Various malicious parties can try to access and use this data for illegal gains and purposes, e.g., for blackmailing healthcare providers for money.

Due to healthcare's central role in society, any disruption due to the unavailability of EHRs or discontinuity of HIT services can be considered a worst-case scenario to topple a society. Therefore, in recent years, a rise in the number of cyber-attacks against healthcare systems has been observed. Recently, the president of the International Committee of the Red Cross warned about the increasing frequency of malicious attacks against hospitals and other critical public infrastructural facilities [3]. “If hospitals cannot provide life-saving treatment in the middle of a health crisis or an armed conflict, whole communities will suffer,” he explained in a meeting of the United Nation’s Security Council.

Considering the sensitivity of the data processed by HIT systems, organizations like the National Institute of Standards and Technology (NIST) and the European Commission have issued special guidelines for managing electronic health records [4–6]. Likewise, dedicated security frameworks have been introduced for installing and operating HIT systems and exchanging EHRs [7].
Consequently, we see that special attention has been paid in recent literature to investigate the privacy and security issues of HIT systems and relevant infrastructures [8–11]. These researches identify the security requirements for setting up HIT systems, the nature of privacy and security issues in these systems, and internal and external attack vectors for pre-emptive security.

However, modern health facilities, equipped with sophisticated medical devices and computing systems recording and efficiently exchanging patient data and various health information with different stakeholders to deliver superior services, are continually reporting data breaches. This introduces the need to identify underlying factors contributing to these breaches for effective security. One of the approaches is to use automated risk detection models and designing and implementing different controls to mitigate these risks. But recent research proposes that many of the automated risk detection models designed for HIT systems and related infrastructural facilities are faulty [12].
Furthermore, due to the continually evolving nature of cyber threats, we need to look at privacy, data protection, and security in a completely new, fresh way and adapt our activities to the afresh cyber reality [13]. The objective of this research is to analyze real healthcare data breaches, identify underlying factors, and report the lessons that can help in mitigation.

RELATED WORK

The Healthcare system comprises of different players, including but not limited to sickness funds, hospitals, laboratories, etc., who need to communicate health data for treatment and other purposes [4]. A modern HIT system may process clinical information, handle telemedicine, or offer personalized care services or remote patient monitoring services. It can include teleconsultation and teleradiology. It integrates health information networks, distributed EHR systems, e-prescriptions, e-referral, etc. The nature of the processed data, the distributed nature of infrastructure facilities, and the variety of actors introduce unique privacy and security concerns [14]. Many researchers have investigated these issues and their implications.

Shoniregun, Dube, and Mtenzi [15] proposed that the increased privacy and security concerns among the general public led to the development of different legal frameworks for information protection in HIT systems. They proposed that laws and standards stipulating IT security adoption and sanctions for non-compliance were common features in securing HIT systems. They further proposed that effective compliance could only be achieved by putting different security and privacy controls in place.

Wuytz [2] focused on data privacy issues in HIT systems. They proposed that an analysis of the privacy issues emerging from integrating EHR, PHR (public health record), and community data should be the first step towards implementing HIT systems. They proposed a taxonomy for classifying health data into different categories for better access management. They suggested that access control in HIT systems could be achieved by defining different access levels and corresponding access rights.

Zeadally et al. [10] conducted a study to explore the underlying possibilities of security and privacy risks for HIT systems, discussed security attacks reported during the first six months of 2016 for these systems, and proposed different solutions to mitigate these attacks. They also identified future challenges for achieving end-to-end security and privacy in HIT systems and associated these challenges with integrating various emerging technologies. However, they focused only on deliberate attacks intended to compromise information security and privacy and further narrowed down their study to only three specific domains: body area, communication infrastructure, and service.

Mcleod and Dolezel [16] modeled different exposure, security, and organizational factors to determine their associations with healthcare data breaches. They found that increased connectivity of healthcare facilities meant increased exposure and higher chances of data breaches. They found that somehow laboratory barcoding was related to an increase in the data breach chances. Establishing an association with business associates with vulnerable computing systems was also recognized as a factor that could lead to data breaches. They also discovered that healthcare facilities with complex structures were at great risk of experiencing data breaches. Finally, they associated the spending on HIT systems' security with data breaches: the lower the spending, the higher the chances of a breach.
Keshta and Odeh [8] performed a review of literature to identify the privacy and security concerns in HIT systems and to examine the solutions that could address the identified concerns. They found that recent research identified physical, technical and administrative factors as major causes of potential security and privacy threats in HIT systems.

Niazkhani et al. [17] conducted a review of recent original studies assessing barriers to EHR adoption/use in chronic care. They discovered that these systems privacy and security concerns were significant barriers to the adoption of HIT systems.

Hoffman et al.[13] identified six types of vulnerabilities for the security and privacy of a computing system. They proposed there could be vulnerabilities lying in legislation gaps, human factors, organizational structures, processes, technical implementations, and physical protection of the system.

Despite that these past studies make valuable contributions, due to the continually evolving nature of cyber threats, we need to look at privacy, data protection, and security in a completely new, fresh way and adapt our activities to the afresh cyber reality [13]. This research aims to analyze real healthcare data breaches, identify underlying factors, and report the lessons that can help in mitigation.

METHODS

We analyzed data consisting of 698 cases of real data breaches reported to the United States Department of Health & Human Services. We obtained the data for this study from the official breach portal of the United States Department of Health & Human Services. The data consisted of all under investigation breaches reported within the last 24 months beginning from 17th of December 2018 till 14th of December 2020.

DATA DESCRIPTION

The analyzed data consisted of 698 cases of data breaches reported by four different types of covered entities. Each case contained information on the name of the covered entity, covered entity type, the number of individuals affected, breach submission date, type of breach, location of breached information, and business associate presence. The covered entities included business associates, health plans, healthcare clearinghouses, and healthcare providers. These entities were all located in different parts of the United States of America.

The reported breaches included hacking/IT incident, improper disposal, theft, loss, and unauthorized access/disclosure. These breaches' locations included desktop computers, laptops, network servers, emails, electronic health records, paper films, other portable devices, and other.

PROCEDURE

An exploratory data analysis was performed. Exploratory data analysis aims at classifying behaviors within a given area of research, identifying potentially important variables, and identifying relationships between those variables and the behaviors [18]. We downloaded the data in Microsoft Excel format. We explored and analyzed data using SPSS and Microsoft Excel tools. Certain data pre-processing was needed, particularly in the case of the location of the breached information.

In many cases, more than one location was involved. We used Microsoft Excel functions to sort individual locations for finding the role of individual locations of the data breaches. We used Microsoft Excel tools like built-in functions, pivot tables and graph builder to generate tables and graphs. SPSS was used to generate descriptive statistics for making sense of the data.

RESULTS

Figure 1 shows an overall increase in the number of reported breaches over the past 24 months. However, a sharp decrease in the number of attack frequencies can be observed in October,

November, and December of 2020. The figure shows that the lowest numbers of breach reports were made in December 2018, and the highest numbers of breaches were reported in September 2020.

Figure 2 shows the month-wise distribution of the number of affected forty-two million individuals. It can be seen that July 2019 was the most damaging month when more than 12 million individuals were hit by data breaches involving their medical data.

Entity type	Number
Business Associate	79
Health Plan	57
Healthcare Clearing House	2
Healthcare Provider	560
Total	698

Table 1: Breakup of the reported breaches according to the types of reporting covered entities

Entity Type	Number of Affected Individuals
Business Associate	15,927,846 (38%)
Health Plan	5,473,369 (13%)
Healthcare Clearing House	1,611,070 (4%)
Healthcare Provider	18,990,233 (45%)

Table 2: Distribution of number of affected individuals according to covered entity types

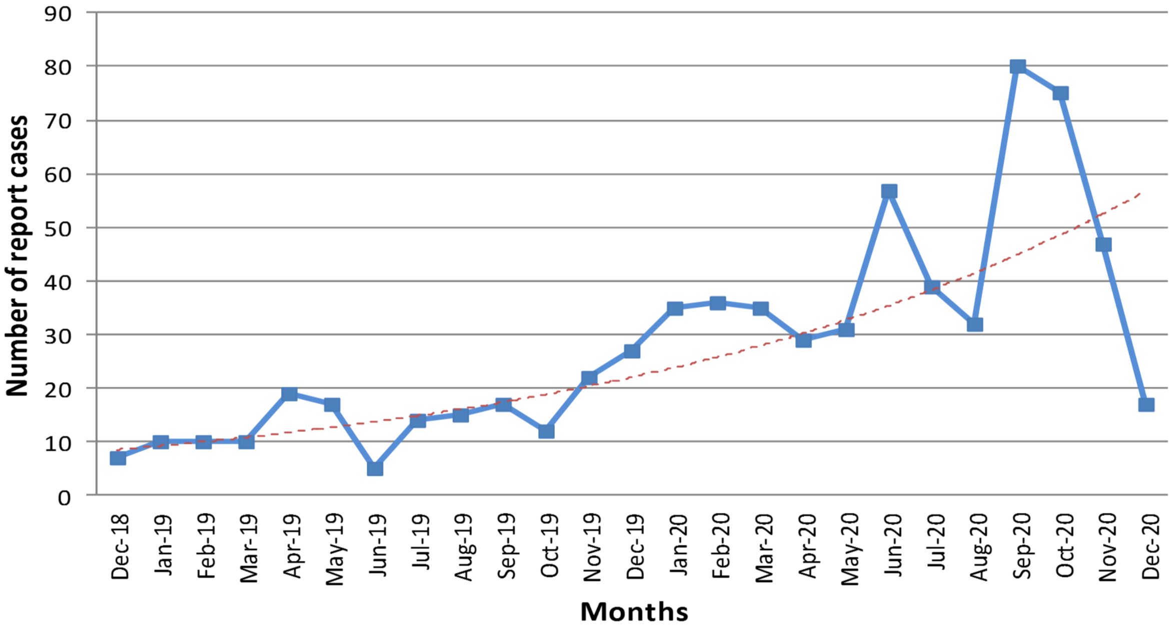

Figure 1: There is an overall increase in the number of reported data breach cases in the past 24 months.

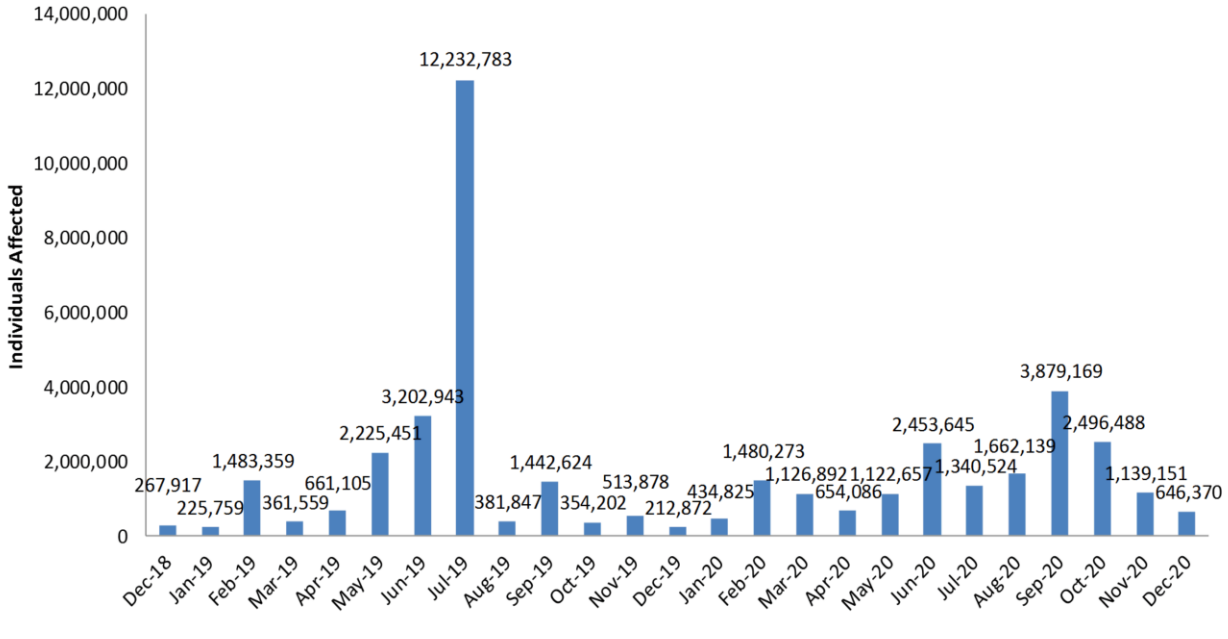

Figure 2: Month-wise distribution of the number of affected individuals

Table 1 shows that among the 698 cases, healthcare providers reported about 80% of the breaches. However, looking at the number of individuals affected by breaches, as shown in Table 2, although the number of breaches reported by the business associates was significantly lower than those reported by the healthcare providers, these breaches affected a large number of individuals. Table 3 shows the further breakup of the impact of breaches on individuals according to incident types.

Entities and Incidents	Sum of Individuals Affected
Business Associate	15927846
Hacking/IT Incident	15774406
Loss	6723
Theft	23394
Unauthorized Access/Disclosure	123323
Health Plan	5473369
Hacking/IT Incident	4627773
Theft	656020
Unauthorized Access/Disclosure	189576
Healthcare Clearing House	1611070
Hacking/IT Incident	45732
Unauthorized Access/Disclosure	1565338
Healthcare Provider	18990233
Hacking/IT Incident	17092365
Improper Disposal	571535
Loss	172766
Theft	135041
Unauthorized Access/Disclosure	1018526
Grand Total	42002518

Table 3: Breach types affecting different entities and the numbers of affected individuals

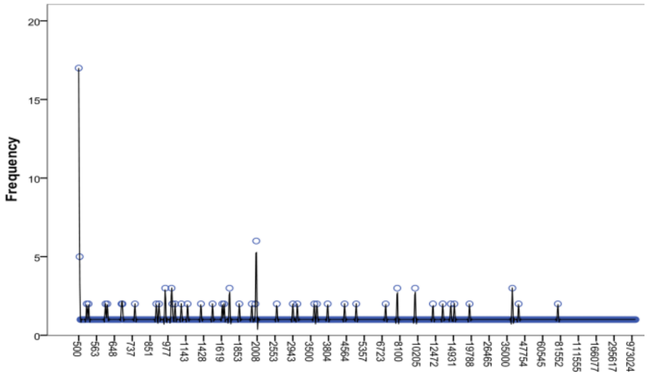

Figure 3: Frequency distribution of breaches according to number of affected individuals

Table 3 shows a total of forty-two million individuals were affected by the breaches analyzed in this study. It presents the distribution of the number of individuals affected across the four types of covered entities.

Figure 2 shows the frequency distribution of breaches according to the number of affected individuals. It is clear from the figure that most of the breaches affected less than 5000 individuals each. Notably, the frequency of breaches affecting 500 individuals is very high.

Tables 4 shows the distributions of the number of affected individuals in cases where business associates are present or absent. It is learned that although a little less than half of the cases involved business associates.

Row Labels	Sum of Individuals Affected
Business Associate	15927846
Yes	15927846
Health Plan	5473369
No	4634925
Yes	838444
Healthcare Clearing House	1611070
No	1611070
Healthcare Provider	18990233
No	12873960
Yes	6116273
Grand Total	42002518

Table 4: Distribution of affected individuals in the presence or absence of a business associate

Incident Locations	Hacking	Improper Disposal	Loss	Theft	Unauthorized Access
Email	250	0	0	0	39
Desktop Computer	22	1	0	8	3
Laptop	4	1	0	15	5
Network Server	250	0	1	2	16
Electronic Medical Record	15	0	1	1	28
Paper/Films	0	12	3	17	38
Portable Devices	1	0	7	8	4
Other	26	0	2	2	11

Table 5: Breach locations and corresponding incidents

More than half of affected individuals come from cases where there were not business associates present. However, looking at the total number of business associates in Table 1, it can be observed that despite making up only 11% of the total covered entities that reported the breaches, the total number of individuals affected due to these breaches is very high.

Table 5 lists the locations where the breaches occurred, the types of the incident leading to breaches, and the respective incidents. The major incidents leading to data breaches were hacking and unauthorized access and mostly involved emails and network servers. Hacking-based breaches made about 71% of the overall reported incidents. Among these 91% hacking incidents involved emails and network servers. Desktop and laptop computers and other portable devices like smartphones or USB storage devices can also become a means of breaches.

Data shows that desktops and laptops were almost equally affected by the hacking, theft, and unauthorized access. However, in the case of laptops, theft was a major reason for the potential data breach. Surprisingly, paper/films became a source of data breaches in 70 cases and were subject to unauthorized access, theft, improper disposal, and loss. In forty-three cases, other factors were also involved. They can include hacking and unauthorized access.

DISCUSSION

This paper analyzed 698 cases of healthcare data breaches reported to the U.S. Department of Health and Human Services, Office for Civil Rights during the past twenty-four months (December 2018 - December 2020) with an aim to identify underlying factors and report the lessons that can help in mitigation. The analysis showed an increasing trend in the number of reported breaches during the past 24 months. In the presence of a large number of security and privacy guidelines for the healthcare IT systems [4,7,9,15], it is an alarming situation.

Previous research shows that privacy and security concerns of the healthcare data led to the development of security standards and regulatory and security frameworks for these systems [15]. These frameworks need to be reviewed, keeping in view the emerging threats. More efforts should be put into compliance audits to ensure that entities handling the healthcare data are sticking to regulatory guidelines and standards.

The analysis showed that healthcare providers were most frequent in reporting breaches among different stakeholders of the healthcare systems. However, despite being a less frequent target, healthcare system business associates experienced data breaches whose impact size was comparable to those of the far larger numbers of healthcare providers. Previous research has associated these entities with the vulnerabilities of HIT systems[16]. We assume that these business associates can have associations with multiple healthcare providers and, therefore, bigger databases with a larger number of electronic health records and patient personal data. Dedicated security standards and guidelines should be developed for these entities.

It was also discovered that the majority of the breaches involved data of less than 5000 individuals. It can be that smaller entities pay less attention to their security infrastructures due to limited budgets and become an easy target of hackers. Or due to less investment in training their employees in cybersecurity, these entities fail to develop an organizational culture that ensures following the security best practices [16]. Effective security frameworks with low budgetary requirements need to be developed for small size entities. Mandatory online security training can be one effective solution to improve security awareness and security skills of HIT systems users.

The major incidents leading to data breaches were hacking and unauthorized access and mostly involved emails and network servers. While exploiting emails depicts a lack of following security practices among email users, professionals manage network servers. Past research shows that network security hacks mainly result from IT infrastructure mismanagement [19]. However, the role cloud service providers and vulnerabilities in the underlying platforms should also be determined. The specifics of the HIT systems should not be ignored in this process [20].

Previously, different models have been proposed to address the issues resulting from unauthorized access [2]. Integration of new technologies such as cloud computing-based infrastructures redefines the systems' trust boundaries, and therefore, new models are needed to manage access. Similarly, to avoid the theft of equipment carrying healthcare data, bring your own device (BYOD) culture should be discouraged in the cases where a user handles the data. Likewise, mobile access to these systems should be restricted as these devices can be easily stolen or lost. Improper disposal of health records is one of the factors leading to data breaches in healthcare systems. The data shows that the number of incidents involving improper disposal of records is low and can be controlled through effective disposal policies such as mandatory in-house record disposal.

Finally, previous research has focused on technical and organizational aspects of security issues in HIT systems [8,10,16]. The nature of incidents like hacking, unauthorized access, improper disposal, and theft requires investigating the human factors' role. Extensive field studies need to be conducted in this regard. In this regard, our findings are in line with those of [13].

LIMITATIONS

This study is based on the data comprising healthcare data breaches that occurred and reported in the United States. However, similar technical infrastructures and used to support healthcare facilities across the globe. It is also evident that most of the incidents are driven by the skill and awareness of the healthcare systems' users/operators. The disparity among the levels of these skills in different countries is known [21] and should be considered.

CONCLUSION

In this paper, we performed an exploratory analysis of 698 cases of data breaches reported to the U.S. Department of Health and Human Services, Office for Civil Rights during the past twenty-four months (December 2018 - December 2020). We found an increase in the frequency of incidents involving healthcare data breaches. We understand that there is a need to continually reviewing the existing security frameworks, developing dedicated security standards and performing stricter security audits for business associates, developing security frameworks keeping in view the capabilities of small-sized entities, and mandatory online training for the users of HIT systems. Furthermore, with the involvement of technologies like cloud computing, special access control measures need to be developed. We observed a prominent role of the human factors in the data breaches. Extensive studies are required to identify the underlying factors and develop procedural and usability improvements.

ARTICLE INFORMATION Accepted for Publication: December 11, 2020
Published Online: December 30, 2020.
https://doi.org/10.48111/2020.04.05

Author Affiliations: Haroon Elahi is with Department of Computer Science, Guangzhou University, China. Oana Geman is with the Department of Health and Human Development, Faculty of Electrical Engineering and Computer Sciences, Stefan cel Mare University of Suceava, Romania.

Financial Support and Sponsorship: Nil.

Conflicts of Interest: There are no conflicts of interest

REFERENCES

1. Al-Zinati M, Almasri T, Alsmirat M, Jararweh Y. Enabling multiple health security threats detection using mobile edge computing. Simul Model Pract Theory. 2020;101(July 2019):101957. doi:10.1016/j.simpat.2019.101957

2. Wuyts K, Verhenneman G, Scandariato R, Joosen W, Dumortier J. What electronic health records don't know just yet. A privacy analysis for patient communities and health records interaction. Health Technol (Berl). 2012;2(3):159-183. doi:10.1007/s12553-012-0026-3

3. Lederer EM. Red Cross chief : cyber attacks increasing on hospitals. 2020:1-1. https://apnews.com/article/3676505e1426752f90b87e490eca42bb.

4. Callens S. The EU legal framework on e-health. In: Mossialos E, Permanand G, Baeten R, Hervey TK, eds. Health Systems Governance in Europe. Cambridge: Cambridge University Press; 2010:561-588. doi:10.1017/CBO9780511750496.014

5. The European Commission. COMMISSION RECOMMENDATION (EU) 2019/243 of 6 February 2019 on a European Electronic Health Record exchange format. Off J Eur Union. 2019;4(March 2011):18-27. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv:OJ.L_.2019.039.01.0018.01.ENG.

6. Lowry SZ, Ramaiah M, Taylor S, et al. Technical Evaluation, Testing, and Validation of the Usability of Electronic Health Records: Empirically Based Use Cases for Validating Safety-Enhanced Usability and Guidelines for Standardization. Gaithersburg, MD; 2015. doi:10.6028/NIST.IR.7804-1

7. Scholl M, Stine K, Lin K, Steinberg D. NISTIP 7497 "Security Architecture Design Process for Health Information Exchanges (HIEs).”; 2020.

8. Keshta I, Odeh A. Security and privacy of electronic health records: Concerns and challenges. Egypt Informatics J. 2020;(xxxx). doi:10.1016/j.eij.2020.07.003

9. Alanazi HO, Zaidan AA, Zaidan BB, Kiah MLM, Al-Bakri SH. Meeting the Security Requirements of Electronic Medical Records in the ERA of High-Speed Computing. J Med Syst. 2015;39(1):9-12. doi:10.1007/s10916-014-0165-3

10. Zeadally S, Isaac JT, Baig Z. Security Attacks and Solutions in Electronic Health (E-health) Systems. J Med Syst. 2016;40(12):263. doi:10.1007/s10916-016-0597-z

11. Brady K, Sriram RD, Lide B, Roberts K. Testing the Nation's Healthcare Information Infrastructure: NIST Perspective. Computer (Long Beach Calif). 2012;45(11):50-57. doi:10.1109/MC.2012.290

12. Montibeller G, Patel P, del Rio Vilas VJ. A critical analysis of multi-criteria models for the prioritisation of health threats. Eur J Oper Res. 2020;281(1):87-99. doi:10.1016/j.ejor.2019.08.018

13. Hoffmann R, Napiórkowski J, Protasowicki T, Stanik J. Risk based approach in scope of cybersecurity threats and requirements. Procedia Manuf. 2020;44(2019):655-662. doi:10.1016/j.promfg.2020.02.243

14. Shoniregun CA, Dube K, Mtenzi F. Introduction to e-Healthcare Information Security. In: ; 2010:1-27. doi:10.1007/978-0-387-84919-5_1

15. Shoniregun CA, Dube K, Mtenzi F. Laws and Standards for Secure e-Healthcare Information. In: Electronic Healthcare Information Security, Advances in Information Security. ; 2010:59-100. doi:10.1007/978-0-387-84919-5_3

16. McLeod A, Dolezel D. Cyber-analytics: Modeling factors associated with healthcare data breaches. Decis Support Syst. 2018;108:57-68. doi:10.1016/j.dss.2018.02.007

17. Niazkhani Z, Toni E, Cheshmekaboodi M, Georgiou A, Pirnejad H. Barriers to patient, provider, and caregiver adoption and use of electronic personal health records in chronic care: a systematic review. BMC Med Inform Decis Mak. 2020;20(1):153. doi:10.1186/s12911-020-01159-1

18. Bordens KS, Abbott BB. Research Design and Methods: A Process Approach. Ninth Edit. McGraw-Hill Education; 2014.

19. Liu Y with a chance of breach: F cyber security incidents, Sarabi A, Zhang J, et al. Cloudy with a chance of breach: Forecasting cyber security incidents. Proc 24th USENIX Secur Symp. 2015:1009-1024.

20. Bahtiyar Ş, Çağlayan MU. Trust assessment of security for e-health systems. Electron Commer Res Appl. 2014;13(3):164-177. doi:10.1016/j.elerap.2013.10.003

21. The European Commission. Europeans' Attitudes towards Cyber Security. Brussels, Belgium; 2020. doi:10.2837/672023