2021

Impact of Information Breaches on Health Care Records

Anton Antony Arockiasamy
Walden University

Walden University

College of Management and Technology

This is to certify that the doctoral dissertation by

Anton Antony Arockiasamy

has been found to be complete and satisfactory in all respects,
and that any and all revisions required by
the review committee have been made.

Review Committee
Dr. Aridaman Jain, Committee Chairperson, Management Faculty
Dr. Robert Levasseur, Committee Member, Management Faculty
Dr. Jean Gordon, University Reviewer, Management Faculty

Chief Academic Officer and Provost
Sue Subocz, Ph.D.

Walden University
2021

Abstract

Impact of Information Breaches on Health Care Records

by

Anton Antony Arockiasamy

Mini-MBA, Rutgers University, 2012

MS, Birla Institute of Technology and Science, 2011

MBA, Annamalai University, 2009

BE, Anna University, 2007

Dissertation Submitted in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Philosophy

Management

Walden University

February 2021

Abstract

Although there were almost 3.5 million reported information breaches of health care data in the first quarter of 2019, health care providers do not know the extent of digital and nondigital breaches of patient medical records. The purpose of this quantitative, comparative study was to identify the difference between the individual patient records affected by digital versus nondigital breaches for three types of health care entities in the United States, health care providers, health care plans, and health care clearinghouses. Allman's privacy regulation theory, the National Institute of Standards and Technology Privacy Framework, and ecological systems theory comprised the theoretical framework. The focus of the research questions was on the difference between digital and nondigital breaches for each of the health care entities. The study data consisted of 2,601 digital and nondigital breach reports for the three healthcare entities for the years 2010 to 2018 retrieved from the public database of HIPAA breach and violations maintained by the U.S. Department of Health and Human Services. Significant *t* tests of the hypotheses for each health care entity indicated that more breaches occurred digitally than nondigitally, and that health plan provider breaches resulted in a greater number of individuals impacted per incident than breaches of healthcare providers or healthcare clearinghouses. The implication for positive social change is that the study findings may help health care entities make better decisions about how to allocate scarce information security resources to lower health care costs by reducing the breaches of health care records.

Impact of Information Breaches on Health Care Records

by

Anton Antony Arockiasamy

Mini-MBA, Rutgers University, 2012

MS, Birla Institute of Technology and Science, 2011

MBA, Annamalai University, 2009

BE, Anna University, 2007

Dissertation Submitted in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Philosophy

Management

Walden University

February 2021

Dedication

I dedicate this research study to my wife, Abini Cross, for her love, unwavering support, and patience during my long journey. I also dedicate this to my daughter, Amy Anton. My doctoral journey started during the same year when Amy was born. We both grew together and shared beautiful memories during this long journey. Amy has been a motivational force for me to complete my doctorate.

I also would like to dedicate this study to my brother, Frank Duff, who has suffered from hydrocephalus all his life and always reminds me of the importance of healthcare in our community. I also would like to thank my parents, Antoni and Anuncia, who emphasized the importance of education since my childhood.

I also dedicate this study to my mentor, Ramon Collante; without his support, for which I will forever be thankful, this journey would not have been possible. Above all, I am thankful that God has showered his blessings and strength on me during my doctoral journey and all my life.

Acknowledgments

I would like to acknowledge Dr. Aridaman K. Jain, my chair, for his continued support and motivation during this whole process. I would also like to thank Dr. Robert Levasseur, my second committee member, for his guidance in shaping this dissertation to be aligned with the required elements. I also would like to thank my first, second committee member, Dr. Thomas Spencer, who helped during the initial phases of this journey until his retirement. I would like to acknowledge Robert James Babin Sr., Chief Information Security Office at Saint Peter's Healthcare System, and Dr. Capecomorin Pitchumoni for their inspiration in initiating this study based on the HIPAA security breaches.

I would like to extend my thanks to the Information Technology leadership members, William Rears, Dr. Jordan Tannenbaum, and Frank DiSanzo, from Saint Peter's Healthcare System for their support during this doctoral journey. In addition, I would like to thank the Human Resources department at Saint Peter's Healthcare System, especially Pam Teufel and Susan Ballestero, for facilitating the scholarship, which greatly helped during this long journey.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction to the Study

Cyberattackers widely target financial and health care entities because they are critical infrastructures. In 2019, there were 94 reported information breaches in the U.S. healthcare system from January to March, affecting a total of 3,486,735 individuals (U.S. Department of Health and Human Services, n.d.).The general problem is that providing adequate information security within the health care sector is both challenging and costly. For health care entities, providing information security increases the cost of health care for the U.S. public (Toé, 2013). Toé (2013) examined the explicit costs of sensitive information security breaches and found that breaches involving financial information, medical protected records, social security numbers, names, and addresses have an impact on the explicit costs associated with health care. The U.S. government has regulated the information security protocols for the health care organizations under the umbrella of Health Insurance Portability and Accountability Act (HIPAA) regulations since the enactment of HIPAA in 1996 to protect patient health information in this electronic era (Pekala, 2017).

The objective of this research study was to estimate the difference between digital and nondigital breaches of individual patient records for each of the three types of health care entities in the United States. The identification of the type of breach that occurs the most and the most targeted entity determined where health care organizations should focus their efforts and resources to decrease the number of breaches of individual patient records. Current literature provides analyses of which type of breach costs the health care industry the mostas well as information on breaches by entity type and the numbers of

breached individual patient records (U.S Department of Health and Human Services, n.d.). The goal of this study was to examine these sets of data to determine any potential differences between digital and nondigital breaches.

This chapter includes the following key sections: (a) background of the study, (b) problem statement, (c) purpose of the study, (d) research questions and hypotheses, (e) theoretical foundation, (f) nature of the study, (g) definitions, (h) assumptions, (i) scope and delimitations, (j), limitations, and (k) significance of the study. A summary of the key points of the study and an overview of the succeeding chapters conclude the chapter.

Background of the Study

Phishing refers to the practice of obtaining computer credentials from authorized users through manipulation and deceit (Wright et al., 2016). In this instance of phishing, the hospital's information technology (IT) department claiming the user needs to update their computer credentials through fraudulent emails sent to users pretending to be legitimate senders. While most computer users have heard of and are watchful for potential phishing schemes, the ever-evolving technology of scams means that phishing schemes are getting more elaborate and convincing, tricking users into giving away sensitive information that can grant a hacker access to the hospital's private records (Wright et al., 2016). While the health care organizations must report breaches of protected health care information, the actual number of breaches is unknown because it is challenging to catch the vast majority of attacks (Wright et al., 2016). Between 2014 and 2016, there were 10 reported breaches of health care information associated with phishing schemes (Wright et al., 2016). However, security consultants estimate that

hospitals routinely undergo several phishing attacks each week. The success rate of these attacks is difficult to estimate because users who fall for the scheme are unlikely to realize it (Martin et al., 2017).

Hackers steal and sell personal identifying information (e.g., social security numbers, Medicare numbers, and dates of birth) from healthcare organizations in online black markets or to criminal networks that use the information to commit financial fraud (Wright et al., 2016). The criminal entities highly prize the information stolen from health care organizations due to the completeness of the information. The information often sells for many times the cost of a stolen credit card number (Wright et al., 2016). While some estimates place the cost of an identity stolen from a health care organization at $10 per identity, other estimates are as high as several hundred dollars an identity (Wright et al., 2016). The sheer number of identities housed in hospital databases makes health care organizations an attractive target for hackers.

Hackers looking to illegally access health care databases may have schemes other than identifying possibilities for theft in mind. Once they gain access to the database, hackers may use falsified or stolen credentials to change direct payroll deposits to their bank accounts, thereby stealing employee wages (Martin et al., 2017). Similarly, attackers could use the credentials to forge prescriptions or steal clinical data for blackmail. Once hackers gain access to a database, it can be challenging to remove them. In 2016, Hollywood Presbyterian Hospital experienced extended downtimes due to malware installed by hackers (Wright et al., 2016). Eventually, the hospital paid the attackers a ransom of $17,000 to remove the malware from their hospital system (Wright

et al., 2016). The ransom received was unlikely to be the hacker's biggest prize because they had extended access to the hospital's database, including payroll information and personally identifiable information for employees and patients.

The HIPAA established safeguards to protect sensitive information from cybercriminals, including requiring unique user identification numbers, emergency access procedures, automatic logoff functions, encryption, and decryption (Kruse et al., 2017). However, despite these protections, there are still ways to illegally access systems, and criminals try to be ahead of the security protocols. Health care organizations face the challenge of not only maintaining their systems but continuously improving them to keep up with more and more advanced hacking methods. In recent years, health care organizations have increased spending to improve hospital integration; however, they have not spent the same amount of resources improving their data security integration (Kruse et al., 2017). According to Kruse et al. (2017), there may be several reasons for this. Updating software can be a time-intensive process, so organizations may struggle to find the downtime necessary to make updates. Improved security may also come with increased roadblocks for users. Two-factor authentication is a method of improving security with logging on to the system, but it takes longer for the user, so organizations that try to implement it face pushback from busy staff (Kruse et al., 2017). Like network integration, security improvements are expensive (Kruse et al., 2017). However, unlike network integration, security improvements are not likely to produce a positive effect on the user (Martin et al., 2017). Improving integration means users can get the data they need faster and easier. The same is not applicable for data security; therefore, the

incentive to improve security does not exist as it does for network integration (Martin et al., 2017).

While many researchers have discussed the rising prevalence of cyberattacks on health care organizations and trends relating to these attacks (Kruse et al., 2017; Martin et al., 2017; Pekala, 2017; Toé, 2013; Wright et al., 2016), there is scarce extant literature on the difference in the number of affected individuals between the types of information security breaches and types of healthcare entities in the U.S. healthcare system. This critical gap means that IT professionals, health care professionals, and health care consumers are unaware of their organization or data being at an increased risk for theft or hacking. Addressing this gap in the literature would allow organizations to be more aware of their risk and encourage them to take the steps necessary to protect the data.

Problem Statement

Even with the advent and implementation of stricter laws to prevent cyberattacks, the number of breaches increases every year, due primarily to the increasing adoption of digital infrastructure and the rise of software solutions to aid in operational tasks within the health care sector (Gomillion, 2017). From the end-user to the sensitive core health care storage infrastructure, there are several layers, such as firewalls, encryptions, and other cybersecurity measures (Shahri et al., 2012). Users remain the weakest link in information security because most users lack the awareness of the risk involved (Shahri et al., 2012). With such high reliance on the digitalization of patient records, IT is evolving to be one of the fastest-growing trends in the U.S. healthcare system (Nimkar, 2016). Although the majority of breaches occur because of human error (Lineberry,

2007), there are other ways to lower information security breaches in the health care industry. Digitalizing and securing patient data is expensive (Berwick & Gaines, 2018). As cybercriminals utilize the internet more frequently and in more varied ways, the need for more sophisticated defense countermeasures becomes increasingly apparent (Langer, 2017).

The three health care entities in the United States (i.e., health care providers, health plan providers, and health care clearinghouses) hold a large amount of digital patient data. There is a gap in the literature regarding the differences between nondigital and digital breaches of patient records within health care entities. Researchers have not studied the extent of digital and nondigital breaches of patient medical records in the past, leading to the specific management problem of health care providers not knowing the extent of digital and nondigital breaches of patient medical records. To combat cybercrime in health care organizations, identifying the type of breach that occurs the most frequently and the most targeted type of entity is necessary to determine the optimal information security resource allocation to decrease the number of breaches of individual patient records. The objective of this research study was to estimate the difference between digital and nondigital breaches of individual patient records and compare individual patient record breaches for each of the three types of health care entities in the United States. Identifying the type of breach that occurs the most and the most targeted entity can be used to determine where healthcare organizations should focus their efforts to decrease the number of breaches of individual patient records. Current literature provides analyses of which type of breach costs the health care industry the most and

information on breaches by entity type and the numbers of breached individual patient records (U.S Department of Health and Human Services, n.d.). The goal of this study was to examine these sets of data to determine any potential differences.

Purpose of the Study

The purpose of this quantitative study was to determine if there is a significant difference between digital and nondigital breaches of individual patient records for each of the three types of health care entities in the United States. The examination of digital and nondigital breaches amongst the three health care entities is essential to both reducing the number of data breaches and ensuring proper allocation of resources to achieve that end. The independent variables were the types of information security breaches and health care entities, while the dependent variable was the number of breached individual patient records. To examine the difference between variables, I used statistical analysis of group means to estimate the differences in individual patient records affected between digital and nondigital breaches of health data in the three types of health care entities.

Research Questions and Hypotheses

The theoretical framework guided the formation of the research questions. I developed the following research questions to aid in the examination of the impact of digital and nondigital security breaches on individual patient records nondigitalfor each of the three types of health care entities in the United States. The research questions and associated hypotheses were as follows:

RQ1: Is there a significant difference between the average number of individual patient records affected by digital breaches and nondigital breaches for health care providers?

$H_0$1: There is no significant difference between the average number of individual patient records affected by digital breaches and nondigital breaches for health care providers.

H_a1. There is a significant difference between the average number of individual patient records affected by digital breaches and nondigital breaches for health care providers.

RQ2: Is there a significant difference between the average number of individual patient records affected by digital breaches and nondigital breaches for health plan providers?

$H_0$2: There is no significant difference between the average number of individual patient records affected by digital breaches and nondigital breaches for health plan providers.

H_a2: There is a significant difference between the average number of individual patient records affected by digital breaches and nondigital breaches for health plan providers.

RQ3: Is there a significant difference between the average number of individual patient records affected by digital breaches and nondigital breaches for health care clearinghouses?

$H_0$3: There is no significant difference between the average number of individual patient records affected by digital breaches and nondigital breaches for health care clearinghouses.

H_a3: There is a significant difference between the average number of individual patient records affected by digital breaches and nondigital breaches for health care clearinghouses.

Theoretical Foundation

The theoretical framework for this quantitative, comparative study consisted of two theories: Allman's privacy regulation theory (and the associated National Institute of Standards and Technology (NIST) Privacy Framework) and the ecological systems theory. In the privacy regulation theory, Allman (2018) posited that the goal of privacy regulation is to achieve the most favorable level of privacy. Privacy is a nonmonotonic function, meaning that there can be too much privacy or too little privacy (Margulis, 2003). Health care information systems require health care entities to protect patient data without decreasing the ease of access for authorized users and risking security breaches. Allman's position is that privacy has two levels: individual and group. HIPAA requires the protection of an individual's data; yet, the protection of private information is a group process (CITE). The underlying assumption of the NIST Privacy Framework is that if an information system's security plan also includes privacy protections, the resilience of the system will provide the resilience of privacy (Hiller & Russell, 2017). The NIST Framework includes the position of privacy as a fundamental part of resilient cybersecurity, which should work towards maintaining the privacy of citizens as

mandated by HIPAA. The NIST Framework is useful as it can help determine what types of data breaches and in which health care entities breaches may occur prior to actual data breaches. The findings of this study may prove to be a valuable resource to the information security professionals in the health care entities and be used to bring about a positive social change by increasing information security protocols and reducing the cost of information security.

In addition to the NIST Privacy Framework, I used the ecological systems theory as part of the theoretical framework for this quantitative, comparative study. Although the NIST Security Framework gives context regarding how to employ cybersecurity and how it relates to addressing cybersecurity threats, it does not adequately explain the connection between cybersecurity and the health care system. The ecological systems theory can be used to create the context for why cybersecurity criminals commit crimes, how these cyber breaches affect targeted individuals, and in what ways the employees within the health care systems in the United States can be more diligent in protecting information.

Urie Bronfenbrenner developed an ecological systems theory, or human ecology theory, in 1979 to explain better the variety of environmental factors that influence human behavior. Bronfenbrenner (1992) believed a great deal of human behavior can be attributed to various interactions between individuals and their respective ecological systems, which includes socialization between different spheres of influence. Additionally, Bronfenbrenner posited that interactions interrelate within the ecological system and can be both conscious and unconscious in nature. As such, Bronfenbrenner

recognized five main environmental systems that influence the entirety of human behavior: the individual, the microsystem, the exosystem, the mesosystem, and the macrosystem. Within Bronfenbrenner's ecological construct, all levels interact, especially with levels that are sequential. It is possible to examine and contextualize human behaviors by constructing the ecological spheres of influence on human interaction (Bronfenbrenner, 1992).

Nature of the Study

In this study, I used a quantitative approach to examine group means to determine if there was a difference between digital and nondigital breaches of health care information stored by U.S. health care entities. Specifically, I determined whether there was a significance between the number of individual patient records affected per digital breach and per nondigital breach for each of the three different types of health care entities. This allowed me to conclude whether there were significant differences between digital and nondigital breaches within health care entities based on the number of breached patient records for each type of breach.

Definitions

BYOD management: A data security process where employees who use a particular device frequently at home decide to use the same in their organization as well (Technopedia, 2019).

Data loss prevention systems: A set of tools and processes used to ensure that sensitive data are not lost, misused, or accessed by unauthorized users (Digital Guardian, 2019).

Encryption: The method by which plaintext or any other type of data converted from a readable form to an encoded version that can decode only by another entity if they have access to a decryption key (SearchSecurity, 2019).

Endpoint and malware protection: An approach to detecting malicious network activity and protecting computer networks, including servers, desktops, and mobile devices, from intrusions and malware attacks (SearchSecurity, 2019)

Health care entities: Organizations, including health care providers like hospitals and clinics, health plan providers like insurance agencies, and health care clearinghouses, also known as billing agencies (National Practitioner Data Bank, n.d,).

Information breach: The theft of either physical or virtual data. A confirmed incident in which sensitive, confidential, or otherwise protected data are exposed in an unauthorized fashion (SearchSecurity, 2019).

Information security: The processes and methodologies that protect the print, electronic, or any other form of confidential, private, and sensitive information or data from unauthorized access, use, misuse, disclosure, destruction, modification, or disruption (SANS, 2019)

Intrusion prevention systems: A system that monitors a network for malicious activities, such as security threats or policy violations (Technopedia, 2019).

OSI Layer Firewall 7: Layer 7 of a firewall sorts traffic according to which application or application service the traffic is trying to reach and what the specific contents of that traffic are. Rather than simply blocking all traffic on a certain port, Layer

7 allows some traffic through while blocking traffic that may contain a threat (DigitalGuardian, 2019)

Web proxy: A method for hiding an IP address from the websites an individual visits (SANS, 2019).

Assumptions

The primary assumption in this study was that the data gathered from the respective government agencies would be accurate. I sought assistance from the data managers or administrators of the government agencies that keep the data to ensure that I determined the correct number of affected individuals for each of the different information security breaches across different health care entities.

Scope and Delimitations

The scope of this study included individuals affected by information breaches across different U.S. health care entities. The participants, selected through purposeful sampling, consisted of a cross-representation of different individuals having different educational backgrounds, ages, and social statuses. I located records for the individuals in this study in the HIPAA breach database available for use from 2016 to 2018. The analysis involved estimation of comparative differences in the number of affected individuals by digital versus nondigital information security breaches across different healthcare entities by the comparative research design.

Limitations

There were two limitations to the study. First, I gathered data for the variables from databases of government agencies. Use of such a method might have limited the

insights produced from the analysis because the data might not have reflected the general population. Second, the use of a nonprobability sampling procedure, such as purposive sampling, reduces the possibility of generalizing the results to a larger population.

Significance of the Study

Significance to Theory

This study may enable information security professionals in the health care industry to identify the critical areas in healthcare information security that require more attention to avoid information breaches. Over the past 2 decades, the cost of health care cost in the United States has proliferated. In an annual report, the Health Care Cost Institute (2019) reported that healthcare costs rose 3.9% per year on average between 2013 and 2017. This figure is higher than the rise of the gross domestic product, which averaged a 3.1% increase between 2013 and 2017. Within the same period, the use of health care services declined by 0.2%, indicating that there has been an inverse relationship between the use of health care services and the cost of health care. One of the significant contributors to the increase in health care costs is due to the adaption of IT and the digitalization of patient data (Langer, 2017). In the overall IT budget, management allocates a significant portion towards information security (Langer, 2017).

Significance to Practice

The significance of this study to practice is two-fold. The goal was to further explore the importance of information security in the U.S. health care system. Though information security has been a growing interest for almost a decade now, only a few research studies on the topic existed in the health care field. The health care system holds

an enormous amount of digital information about patients, doctors, and health care providers; therefore, information security is of paramount priority.

Significance to Social Change

The goal of this study was to provide a better understanding of how different types of information breaches impact the number of individual patient records affected per breach. The findings may help to provide health care entities with information to incentivize them to invest in information security to safeguard health care processes and patients and lessen the opportunity for information breaches. Furthermore, the results of this research could provide positive social change by decreasing the cost of information security for the health care entity, which, in turn, would lower the cost of health care for the patients.

Summary and Transition

The specific problem addressed in this study was identifying the most vulnerable entity to data breaches and the type of breach that results in the most individuals affected to determine where the most resources are necessary to decrease the number of breaches and individuals affected. As such, the purpose of this quantitative, comparative study was to determine the difference in the number of affected individuals between the types of information security breaches for the three types of U.S. healthcare entities. In this study, I used a comparative design to analyze data from the HIPAA breach database, which contains information on the type of data breach, the number of affected individuals, and the type of health care entity.

This study consists of four more chapters. Chapter 2 is a review of the pertinent literature relating to information security in the health care industry. In the literature review, I will provide a summary of the previous findings as well as previous researchers' recommendations for implementation strategies and future research. In Chapter 3, I will present an overview of the research methodology for the study. The research design, population and sampling, methods of data collection and data analysis, ethical considerations, and the validity and reliability of the instruments will be discussed in the chapter. In Chapter 4, I will elaborate on the results of the data analysis to answer the research questions. Finally, Chapter 5 includes a discussion of the results and my recommendations to the health care industry regarding improving information security.

Chapter 2: Literature Review

Introduction

The purpose of this quantitative, comparative study was to determine whether there was a significant difference in the number of individual patient records affected in a digital breach compared to a nondigital breach of individual patient records for each of the three types of U.S. healthcare entities. The examination of digital and nondigital breaches across the three healthcare entities is important to both reducing the number of data breaches and ensuring proper allocation of resources to achieve that end. To examine the difference between variables, I used the comparative design to analyze information on the type of breach, the number of affected individuals, and the type of health care entity.

To aid in the accomplishment of the goals of this research project, I completed an exhaustive review of the literature within this chapter to more comprehensively understand previous research relevant to the topic of interest: security breaches within the U.S. health care system. Topics included in this chapter are seminal sources on the conceptual and theoretical framework employed in this study, the composition of the U.S. healthcare system, how the U.S. healthcare system operates, types of security breaches that can affect the U.S. healthcare system, and how all of the aforementioned factors interrelate.

Literature Search Strategy

I searched for relevant literature by means of an extensive online search of the following databases accessed through the Walden University Library: Communication & Mass Media Complete, Web of Science, PsycARTICLES, PsycINFO, and

PsycCRITIQUES, PsycEXTRA, ERIC, Center for Disease Control, ResearchGate, SAGE Journals, and Google Scholar. The search of relevant topics was completed with established parameters to return results applicable to this research project. The parameters set focused on peer-reviewed articles published between 2015 and 2019. Search results yielded approximately 320 articles and other appropriate forms of literature; however, I did not use all returned results for the purposes of this literature review. Instead, only literature that met all inclusion criteria were employed. The adherence to inclusion criteria ensured that all literature included in this project were current, defined as within the context of this research study, and published within the past 5 years. Any sources of literature outside of that time were defined as classic sources for relevant topics needed for the appropriate explanation of theory or concept. Additionally, I used gray literature, which is reputable non-peer-reviewed information, in this literature review for emphasis or clarity but did not use gray literature as conclusive sources of information or to dictate topics of interest in this chapter.

Key search terms used included: *NIST privacy framework, ecological systems theory, United States healthcare, healthcare, security and healthcare, healthcare security breach, consequences of healthcare breaches,* and *American healthcare*, all in the title. Additionally, I made online searches using the terms of *implications of security breach, cybercrime + American healthcare,* and *patient information and healthcare* to ensure the comprehensive nature of literature search.

After I provide an exhaustive review of relevant topics, this chapter ends with a summary of the literature and an introduction to the subsequent chapter.

Theoretical Foundation

The conceptual and theoretical framework of a research project is important because it often aids in creating the context in which the establishment and results of research questions are applied in practice. Often, it is common to consider more than one theory within the framework of research to ensure the research topic is more comprehensively explained (CITE). The theoretical framework for this research study comprise of two theories: the NIST Privacy Framework and ecological system theory.

The NIST Privacy Framework

The NIST developed a comprehensive cybersecurity framework for a variety of private sector organizations in which these entities can better protect the information, especially sensitive or confidential data (Shen, 2014). The original version of this privacy framework, developed in 2014, is known as Version 1.0 (Shen, 2014). Version 1.0 was one of the first cybersecurity frameworks that could allow organizations within the private sector to be proactive in anticipating security risk (Shen, 2014). Risk management is often defined by cybersecurity experts as the continuous array of the process that identifies, assesses, and understands cybersecurity risks (Esser, 2018). Further cybersecurity and risk management include the framework necessary to reduce the impact of possible cybersecurity breaches and the overall prevalence of risk (Esser, 2018). Cybersecurity and risk management become increasingly important in the modern economy and a variety of other services because cybersecurity breaches can cause great financial loss, damage to an organization’s reputation, or cause violations of privacy for both employees and shareholders (Esser, 2018)

An updated version of the NIST Privacy Framework was released in 2017, known as Version 1.1, and added a variety of upgrades to Version 1.0 that allows organizations in the private sector to include guidance on how to perform self-assessments on possible risks regarding security and self-assessments of the risk factors on supply chains (Shen, 2014). The latter feature is especially important because supply chain security risks can also compromise sensitive data on private organizations that interact with one another (Shackelford et al., 2015). Version 1.1 also increased information on mitigating risk management and how to more appropriately interact with supply chain stakeholders (Shackelford et al., 2015). As such, Version 1.1 of the NIST Privacy Framework provides a comprehensive security framework for many of the organizations of the private sector (Shackelford et al., 2015). The NIST Privacy Framework is an integration of three separate but interrelated facets: the Core, the Profiles, and the Implementation of Tiers (Shen, 2014).

The Core

The Core refers to an exhaustive body of privacy protection activity that permits the prioritization of possible security risks and activities that may jeopardize the privacy of data collected by organizations (Shen, 2014). When utilized appropriately, the Core allows for various ways of individual risks to be assessed and addressed to protect the organization from other related individual security or privacy risks (Shackelford et al., 2015). The Core facet of the NIST Privacy Framework comprises five simultaneously addressed functions: identify, protect, control, inform, and respond (Shen, 2014).

The first function of the Core of the NIST Privacy Framework is the ability to identify ways that organizations may become at risk for security or privacy breaches (Miron & Muita, 2014). This identification function includes a wide array of different forms of data collected by organizations for everyday operations as well as information important to development, research, or consumer trends (Miron & Muita, 2014). The first form of data that needs to be identified and assessed includes all the assets of the organization, which is known as asset management (AM; Miron & Muita, 2014). AM includes all data collected from the actual systems, devices, or technology present within an organization that allows for businesses to achieve functionality (Shen, 2014). When AM occurs, each of the aforementioned types of data are identified and then subsequently ranked by both importance to the company and how likely the data are to be compromised (Miron & Muita, 2014).

The next type of data examined by the NIST Privacy Framework is the data related to the respective business environment (BE; Shen, 2014). The BE includes all data concerning the purpose, functionality, and associated objectives of an organization (Shen, 2014). Further, the BE includes the integration of data from stakeholders and their associated activities within an organization. BE is often assessed by the NIST Privacy Framework to ensure the cybersecurity and recommended risk management protocols are in alignment with the organization (Esser, 2018).

Another form of data identified and prioritized within the Core of NIST Privacy Framework is referred to as governance (Shen, 2014). Governance includes the implementation of data related to the policies and operations procedures within the

organization (Shen, 2014). This includes the data included in regulatory compliance on legal, environmental, and operational policies (Esser, 2018). In this way, the NIST Privacy Framework can assess risk related to these types of data and prioritize them respective to organizational importance (Esser, 2018).

Risk assessment (RA) is also a part of the Core component of the NIST Privacy Framework (Esser, 2018). RA includes any form of cybersecurity risk than can be associated with the mission or overall functionality of an organization (Esser, 2018). Additionally, the execution of RA is to ensure that any possible cybersecurity breach or privacy violation does not extensively damage the reputation or public image of a company (Esser, 2018). As security breaches can be catastrophic to the organization or associated personnel, RA is a vital component of proactive cybersecurity risk assessment (Shen, 2014).

The final assessment completed by the Core component of the NIST Privacy Framework is called risk management strategy (RMS; Esser, 2018). RMS includes the identification of an organization's priorities regarding both functionality and potential cybersecurity risk (Esser, 2018). Additionally, RMS includes the identification of an organization's limitations or constraints to obtaining functionality and mitigating cybersecurity and privacy breaches for both the organization and associated stakeholders (Scofield, 2016). To accomplish and implement effective RMSs, tolerance levels for an organization's level of perceived risk are ascertained and then used to as part of the RMS (Scofield, 2016). Finally, within the RMS, assumptions about perceived security risks are

established relative to the type of organization and integrated into the Core facet of the NIST Privacy Framework (Scofield, 2016).

The second function of the NIST Privacy Framework is to protect vital parts of an organization's infrastructure (Shen, 2014). In order to protect an organization, there is a simultaneous implementation of a variety of interrelated steps that comprehensively protect the data, assets, and personnel of the organization (Esser, 2018). First, the NIST Privacy Framework establishes a variety of control measures that restrict access to sensitive data, known as access control (Esser, 2018). The establishment of access control through the physical restriction of inappropriate personnel to places in which sensitive material is housed or through the implementation of restriction protocols ensure that unauthorized personnel does not have access to sensitive data stored within the databases of the organization (Shen, 2014). Online restriction protocols can prohibit access to sensitive data by not only restricting personnel to the databases but also restricting access to sensitive processes and devices utilized to gather, store, or analyze data (Scofield, 2016).

Another facet is the categorization of awareness and training to protect function (Esser, 2018). Within this context, awareness and training refers to the training of personnel associated with the organization or the organization's stakeholders regarding what types of behaviors can lead to an increased risk of cybersecurity risk (Shen, 2014). Additionally, this type of training can facilitate awareness of what types of data are targeted in cybersecurity breaches or security risks (Scofield, 2016). In this way, organizations and stakeholders can train personnel to perform duties regarding proper

cybersecurity and integrate more comprehensive security-based initiatives within formal training sessions (Scofield, 2016).

Another aspect of the protective function of the NIST Security Framework includes the development and installation of data security measures within all portions of the company's operational protocols (Shen, 2014). This type of data protection includes a variety of measures that can encrypt or encode data within organizational databases (Shen, 2014). When data are encrypted or encoded, the data relevant to the company can be better protected. Furthermore, in this way, the integrity and availability of sensitive information are largely inaccessible, even if a cyberdata breach occurs (Esser, 2018).

Information protection processes and procedures (PRIP) is another protective measure within the NIST Security Framework (Esser, 2018). PRIP refers to the development and implementation of various security policies needed by an organization (Esser, 2018). Security-based polices can include workplace rules on the purpose of security measures, the scope of security risks, and the commitment of the organization to ensure security measures are followed (Shen, 2014). Moreover, PRIP can include workplace procedures regarding the responsibilities of each employee to ensure cybersecurity is maintained and consequences for failure to adhere to policy (Scofield, 2016). PRIP is a vital component of the protection factor of NIST privacy Framework to ensure that all personnel within an organization are in agreement on security measures and that all procedures are cohesive and implemented the same way by all employees and stakeholders (Scofield, 2016).

Maintenance (MA) of all security features instituted within a workplace culture is imperative for all security measures to operate appropriately (Shen, 2014). As such, MA is the next component of the protection factor of the NIST privacy Framework (Shen, 2014). MA within this framework refers to the repairs needed for control systems or software utilized to encrypt, encode, house, or analyze the data within an organization (Esser, 2018). Often this type of MA requires professional cybersecurity professionals that employ within or externally to the organization (Esser, 2018). Often companies that handle large amounts of sensitive data employ full-time cybersecurity to staff to ensure protective measures are always working properly and avoid cybersecurity breaches (Esser, 2018).

The last facet of the protection factor within the NIST privacy Framework is the implementation of protective technology (PT) (Esser, 2018). PT includes a variety of security measures that has the main objective of properly monitoring security features and ensure that developed security features are performing accurately (Shen, 2014). Additionally, PT can include cybersystems needed to respond to repair security measures should they fail, quickly before a cyber breach occurs (Shen, 2014).

The third factor of the NIST privacy Framework is detection (Esser, 2018). Within this context, detection refers to the examination and identification of different possible security threats and cybersecurity breaches (Esser, 2018). Detection for cybersecurity threats and security breaches focuses on searching for Anomalies and Events (AE). Often, employees responsible for cybersecurity will scan virtual activity for any events that appear abnormal or out of context (Walser, 2018). Abnormal events are

largely subjective and unique to the respective organization; however, they are defined by the online presence by unauthorized or unrecognized participants or within the discovery of pervasive online activity (Walser, 2018). The pervasive online activity includes unusual behavior that is largely unnecessary when proper access is established (Walser, 2018). Pervasive online activity may include multiple log-in attempts to restricted data, trying to gain access to data outside normal access methods, or trying to access sensitive data anonymously (Walser, 2018). If the case of identification of abnormal or pervasive online activity, then cybersecurity professionals can isolate the incident and hopefully avoid cybersecurity breaches (Esser, 2018).

The next detection method utilized by the NIST Security Framework is continuous monitoring (Shen, 2014). As the aforementioned AE locate with the training and employment of cybersecurity experts, an organization must employ continuous monitoring to ensure that if cybersecurity becomes jeopardized, the identification occurs quickly and in real-time (Walser, 2018). CM included the examination of all sensitive data within established intervals to ensure that all data can be monitored effectively (Shen, 2014). Additionally, CM implementation occurred to also monitor the efficacy of established security measures and the effectiveness of responses to cybersecurity breaches (Walser, 2018).

The last facet of detection methods utilized within the NIST Security Framework is known as detection processes (Esser, 2018). DP includes the identification of various processes or protocols that are maintained by the organization or associated stakeholders (Esser, 2018). DP also includes the establishment of novel modes of detection for

cybersecurity breaches and the development of new ways to ensure that AE's can be addressed adequately and in a timely manner (Shen, 2014).

The next facet of the Core function of the NIST privacy Framework is known as respond (Shen, 2014). This facet involves the development and execution of protocols needed when a cybersecurity breach is detected (Esser, 2018). The first part of the respond is response planning (Esser, 2018). Response planning within this context refers to the need for any processes or initiatives in case of a cybersecurity breach (Shen, 2014). Within RP, the organization and associated stakeholders create a comprehensive response protocol if a cybersecurity breach should occur to ensure that there can be timely response and little ambiguity (Walser, 2018).

To ensure that an RP is comprehensive and situationally appropriate, organizations and their stakeholders establish a variety of communications (Esser, 2018). The creation of communications ensures that verbiage and syntax describing cybersecurity risks are universally understood (Walser, 2018). In this way, when a cyber breach occurs, response times can be reduced, and the confusion largely mitigated (Esser, 2018). Further, communications are often shared with external law enforcement agencies to guarantee that if law enforcement services are needed, they understand the gravity of the situation and how best to mitigate cyber breaches (Shen, 2014).

After COs for a company are established and distributed to appropriate personnel, and analysis (AN) is often conducted by an organization to ensure that response to cybersecurity breaches were adequate (Esser, 2018). Often the AN is conducted by an outside cybersecurity firm that specializes in testing response protocols (Keller, 2017). In

this way, the AN can be more comprehensive and unbiased than the completion by cybersecurity specialists associated with the organization or stakeholders (Keller, 2017).

After the AN is complete, then cybersecurity experts can make recommendations on how to best address issues that arise during AN (Keller, 2017). Often these issues include inadequacy within the response itself or shortcomings in identifying the initial cybersecurity breach (Esser, 2018). When the identification issues occur, then mitigation (MI) can occur (Keller, 2017). MI often include the activities or policies needed to contain the consequences of a cybersecurity breach or keep the cybersecurity breach from spreading into other data while occurring (Shen, 2014). Finally, MI can include the policies or activities needed to erase traces of the incident and repair the breach within the future (Keller, 2017). Subsequently, Improvements made to the cybersecurity framework of an organization (Keller, 2017). Improvements can include novel ways to make the establishment, execution, and analyses of cybersecurity protocols for a given organization or associated stakeholders (Esser, 2018).

The last part of the Core Function is known as Recover (Esser, 2018). Recover includes the development and execution of any protocols or activities needed to mitigate the scope and effects of cybersecurity attacks (Keller, 2017). Within this type of recovery includes the restoration of all functionality of all data-related programs within the organization or associated stakeholders (Esser, 2018). Within the recovery step, it is imperative that the restoration of functionality is expedient and comprehensive. Moreover, the documentation of all necessary steps, and if these issues are novel, they

often contribute to the communications and plans associated with security within the organization (Esser, 2018)

The aforementioned list of components represents the components that create the Core function of the NIST privacy Framework. It is within the Core that the needs of an organization are identified and examined (Shen, 2014).

Further, the Core allows the organization and associated stakeholders to establish priorities regarding cybersecurity needs and initiatives (Shen, 2014). Finally, the core establishes a set language for individuals that are employed within cybersecurity and law enforcement to ensure the reduction of confusion and ambiguity if a cyber breach should occur (Esser, 2018). Once the Core needs are established and prioritized, information ascertained from the Core in order to make a profile of the organization (Keller, 2017).

The Profiles. Profiles created by the information contained within the Core of the organization (Keller, 2017). The utilization of profiles occurs in two main ways regarding cybersecurity (Esser, 2018). First, the use of the profile of an organization as the framework of a company's current cybersecurity protocols (Esser, 2018). This includes all initiatives, training, and safety measures that are implemented within an organization and by associated stakeholders (Esser, 2018). Further, the profile of a company also includes the framework needed to meet cybersecurity needs in the future, including the development of new cybersecurity protection (Shen, 2014).

The other function of the profile of an organization is to make connections between facets, which aids in the comprehensive nature of the assessment of risk and the distribution of resources needed to combat cyber breaches (Esser, 2018). Organizations

can have many profiles depending on the type of data the company obtains and utilizes within an operation (Shen, 2014). Additionally, profiles on an organization may affect the way in which an organization interacts within the third factor of the NIST Security Framework, the Implementation of Tiers (Esser, 2018).

Implementation of Tiers

The Implementation of Tiers is the last major portion of the NIST privacy Framework (Shen, 2014). Tiers refer to the level or scope that the implementation of the cybersecurity measures within an organization (Shen, 2014). The greater the tier, the more sophisticated or comprehensive the cybersecurity measure (Esser, 2018). Often the designation of Tiers is given to various measures in order to convey importance between various levels of cybersecurity experts and operational management (Esser, 2018). Within the NIST privacy Framework, there are four tiers to denote differences in sophistication and integration (Shen, 2014). The lowest tier within this system is Tier I (Shen, 2014). Tier I refer to cybersecurity measures that are largely transient in nature and only partially integrated within the cybersecurity framework (Esser, 2018). Often, Tier I measures are largely unofficial within the organizational operation and are often informal and receive little or no cybersecurity coordination (Esser, 2018).

Tier II is the next inclusive level. Within Tier II, there exists some level of awareness of potential cybersecurity threats (Esser, 2018). As such, there is a small amount of coordination and resources shared between cybersecurity employees and management to ensure implementation, and however, if the implementation fails,

cybersecurity risk is low (Keller, 2017). Moreover, if the access of data is by the cybersecurity breach, the type of data is often not sensitive in nature (Keller, 2017).

Tier III refers to cybersecurity protection measures that are primarily implemented in the company and with associated stakeholders (Esser, 2018). Often, Tier III security measures are formal parts of security procedure and, as such, are repeated regularly (Esser, 2018). Tier III security measures are revised and updated as needed, and a good amount of resources spent on Tier III security (Keller, 2017). If the breach occurred to the data protected by Tier III protocols, sensitive data becomes exposed, and often breaches within this tier can be problematic across multiple types of data (Esser, 2018).

The last tier, Tier IV, is the most sophisticated security measures (Esser, 2018). Security measures within Tier IV have large amounts of resources devoted to the implementation and appropriate integration with the system in lower tiers (Shen, 2014). Security measures within Tier IV are often largely cohesive within other security protocols and are common practice within the organization and with stakeholders (Shen, 2014). Tier IV cybersecurity often occur in response to prior breaches and, as such, are largely unique to respective organizations (Esser, 2018). Massive consequences arise if the breach occurred to the data protected by Tier IV cybersecurity of the organization (Esser, 2018).

Within the NIST privacy Framework, practices and identification of security issues are largely based on effective predictive indicators and experience (Esser, 2018). Often the development of security measures ascertained through the NIST privacy

Framework is implemented throughout organizations and within the security measures of associated stakeholders (Shen, 2014). Currently, the identification of Core security issues, the development of Profiles, and the implementation of the Tiers are largely voluntary (Shen, 2014). However, the NIST privacy Framework is one of the most comprehensive cybersecurity initiatives in existence (Esser, 2018). As such, the NIST privacy Framework is an adequate framework in which to study the topic of this study, security breaches in the healthcare system.

Ecological Systems Theory

In addition to the NIST privacy Framework, the exploration of Ecological Systems Theory (EST) helps to establish the systems theory. Although the NIST privacy Framework gives context to how cybersecurity is employed and how it relates to addressing cybersecurity threats, the NIST privacy Framework does not adequately explain the connection between cybersecurity and the healthcare system. Additionally, the utilization of the EST can give context to why cybersecurity crimes occur and why and in which ways the healthcare system within the United States can be more diligent in protecting information.

Ecological systems theory, or human ecology theory, was established by Urie Bronfenbrenner in 1979 to explain better the variety of environmental factors that influence human behavior. Bronfenbrenner (1992) believed that much of human behavior could attribute to various interactions between an individual and their ecological system, which includes socialization between different spheres of influence. Additionally, Bronfenbrenner posited that interactions within the ecological system are often

interrelated and can be both conscious and unconscious. As such, Bronfenbrenner recognized five central environmental systems that influence the entirety of human behavior: the individual, the microsystem, the exosystem, the mesosystem, and the macrosystem. Within Bronfenbrenner's ecological construct, all levels interact, especially with levels that are positioned prior and subsequently (Bronfenbrenner, 1992). With the construction of the ecological spheres of influence on human interaction, human behaviors can be examined and contextualized (Bronfenbrenner, 1992).

Bronfenbrenner (1992) identified the smallest sphere of influence as the individual. It is within this level, and intrinsic interactions create the majority of influence upon an individual (Bronfenbrenner, 1992). Interactions within the individual level include factors such as sex, age, quality of health, ethnicity, and socioeconomic station. Interactions between the aforementioned factors can explain the way in which a person interprets and reacts within their respective environment (Hertler et al., 2018). Additionally, the many factors within the individual level of behavioral influence may be the primary motivators for many behaviors as these factors create perspective (Hertler et al., 2018). An individual's perspective concerning the environment in which they often live dictates which actions and beliefs consider as normative (Hertler et al., 2018). Within the context of this project, Bronfenbrenner's individual level becomes vital for understanding why some people commit cybercrime, as behaviors that are considered risky to much of the American population, may be perceived as normative to some individuals. Further, the individual level can explain why some individuals feel

compelled to protect the personal information of others' from cybercrime and properly protect sensitive data relevant to organizations from cyber breaches.

The next most proximal level to the explanation of human behavior within the EST is the microsystem (Bronfenbrenner, 1992). Factors that influence human interaction within the microsystem includes any thoughts, emotions, or actions that are possessed by people that are close to the individual (Bronfenbrenner, 1992). The microsystem includes interaction with peers, family, schools, or religious affiliations and community members (Bronfenbrenner, 1992). The influence of the microsystem is conscious primarily to the individual within the interactions (Hertler et al., 2018). The beliefs and actions of others in the microsystem can directly impact the behaviors, actions, and disposition of the individual and often connect to deep-rooted behaviors or beliefs within an individual's value system.

Additionally, the individual can influence the microsystem, as the microsystem comprises of individuals that are important to the individual. In this way, the microsystem is reciprocal in nature (Hertler et al., 2018). In the microsystem, much of the influence on an individual's behavior that does not comprise of intrinsic factors are explained as members of the microsystem often project their own beliefs about normative and non-normative behavior within this sphere of influence (Bronfenbrenner, 1992). This level helps to understand how the collaborative behaviors of individuals within the healthcare system interact with one another and interact with patients to address the patients' needs.

The level subsequent to the microsystem refers to as the mesosystem (Bronfenbrenner, 1992). It is within the mesosystem where various microsystems can interact. This collaboration between microsystems can often influence the behavior of the individual (Bronfenbrenner, 1992). The mesosystem is also where the influence of consciousness ends, and subconscious influence begins (Hertler et al., 2018). According to Bronfenbrenner (1992), although an individual is aware of the interactions between themselves and the mesosystem, individuals are largely unaware of how the mesosystem work together with various microsystems to influence behavior. The level of the mesosystem is another level that becomes influential to this project as the mesosystem is largely representative of interactions between the patients and healthcare systems, which may lead to unintended consequences. For example, a patient can interact with a healthcare provider but may not be aware of how their information storage procedure by the overall healthcare system or how that may put them at risk of a cyber breach was to occur.

The exosystem is the next most inclusive layer within Bronfenbrenner's EST. In the exosystem, it influenze the individual, and their respective behaviors; however, the individual plays no direct role within this sphere of influence (Bronfenbrenner, 1992). The exosystem may include interactions between entities like healthcare providers and insurance companies or healthcare providers and legislative bodies (Bronfenbrenner, 1992). In this way, the resultant actions of the interactions affect the individual and future interactions; however, the individual is absent from the collaboration (Bronfenbrenner,

1992). Moreover, although the interactions can directly influence the individual, much of the influence of the exosystem is an unconscious influence (Hertler et al., 2018).

The exosystem gives way to the macrosystem, the next level in Bronfenbrenner's ecological systems theory (Bronfenbrenner, 1992). Within the macrosystem occurs all influences on behaviors that facilitate governance within the ecosystem of an individual. The macrosystem includes legislation, religious doctrine, and ethnic customs that shape and influence the behavior of an individual (Bronfenbrenner, 1992). Bronfenbrenner believes much of the influence of the macrosystem to unconscious influence as these factors are omnipresent and have been acting upon an individual since childhood. In this way, the mesosystem also influences the smaller levels of not only the individual but also of all the smaller levels of every individual that inhabit similar locales (Hertler et al., 2018).

The last and most inclusive level of influence within the ecological systems theory is the chronosystem. The chronosystem includes the time in which an individual lives. The chronosystem, according to Bronfenbrenner (1992), gives context to all the aforementioned levels of influence on human behavior (Bronfenbrenner, 1992). In addition to time, the chronosystem also encompasses all types of transition events that are important for development and may influence behavior (Bronfenbrenner, 1992) For example, the death of a loved one or the onset of sickness can influence the behavior of an individual (Hertler et al., 2018).

The ecological systems theory was included as part of the theoretical framework of this project as this theory can more comprehensively explain the influence on human

behavior. The theory of ecological systems use to infer that human behavior is not part of an isolated event but a part of an interrelated system of influence (Bronfenbrenner, 1992). As such, the ecological systems theory utilize to not only explain why people commit cybercrimes but also many of the behaviors of the healthcare system of the United States. The ecological systems theory can aid in understanding the behaviors of individuals associated with the healthcare system, including how they protect patient information, why they protect patient information, and how these protective behaviors influence cybercrime.

Review of the Literature

In addition to the conceptual and theoretical framework of the project, a comprehensive review of literature relevant to research topics occur. As the topic of this research project is increased healthcare costs associated with cyber breaches within the United States, topics covered within this section include an overview of the U.S. healthcare system, an overview of how cyber breaches occur, and how costs of healthcare increase in relation to prevention measures against cybercrime.

Healthcare Systems Within the United States

Healthcare Composition

Within the United States, much of the healthcare system privatization, meaning that they majority stake by the private entities (Barr, 2016). Conversely, there exist medical facilities that are owned and operated by varying facets of the government (Barr, 2016). Although medical practices are similar regarding medical practice in both private and governmental healthcare facilities, often the private healthcare providers are more

convenient to the patient (Singh, 2015). In most privately owned healthcare facilities, a patient can make an appointment, and accommodations are far less crowded in comparison to public health care entities (Barr, 2016). Access to the facilities varies primarily by the type of health insurance possessed by the patient (Singh, 2015). Private insurance often offers a broader array of healthcare options than does governmental insurance programs, such as Medicaid or Medicare (Barr, 2016).

In addition to private or governmentally run or private healthcare facilities, there exist healthcare options known as free clinics (Singh, 2015). Often free clinics are considered as part of social networks for community residents that do not possess the resources to obtain insurance in other ways, like employment or through government programs (Barr, 2016). Free clinics often are subsidized by private or corporate donations to continue operations (VanderWielen & Ozcan, 2015). Although free clinics are often widely used within communities, free clinics often only offer a limited number of healthcare services (VanderWielen & Ozcan, 2015). Most free clinics are able to provide services to patients based on short-term or situational needs such as acute sickness, some sexually transmitted diseases (STDs), and some limited dental or vision care (VanderWielen & Ozcan, 2015).

Conversely, free clinics are often not equipped to offer long-term care or care for chronic diseases such as cancer (VanderWielen & Ozcan, 2015). To determine which services the free clinic offers, often, the needs of the community dictate the type of services offered (VanderWielen & Ozcan, 2015). For example, if the community

experiences a high incidence of STD free clinics within that area often offer more comprehensive services for STD treatment and prevention.

As healthcare within the United States comprise of interrelated networks of privately owned and publicly operated facilities, often the costs and access to different forms of healthcare are dictated by resources available to those in need of services (Papanicolas et al., 2018). Often the costs associated with services can vary greatly between community, service type, and severity of the condition (Papanicolas et al., 2018). As such, understanding healthcare costs is important for the overall comprehension of the healthcare system within the United States.

Healthcare Costs

The healthcare system within the United States consists of many interrelated private and public entities that offer an array of services and goods that influence the health of the American people (Dietz et al., 2016). According to the World Health Organization (WHO), within the United States, approximately 65% of healthcare is subsidized by the government, including Medicare and Medicaid systems (WHO, 2016). Medicare is a governmental program that often grants health insurance coverage to Americans older than 65 years of age or younger persons with disabilities (WHO, 2016). Each year an estimated 55 million Americans utilize Medicare to meet healthcare costs. Similarly, Medicaid is a governmentally funded program that grants healthcare subsidies to Americans living in poverty, which as of 2017 was approximately 70 million Americans (WHO, 2016) The remaining costs of healthcare is assumed by U.S. persons as often respective healthcare is a result of employment benefits (Papanicolas et al., 2018).

Although the U.S. government agencies largely subsidize the price of healthcare, the cost of healthcare within the United States can become quite costly (Papanicolas et al., 2018). Currently, within the United States, it is estimated that the cost of healthcare is almost 4,000,000,000,000 dollars annually, an almost 6500% increase than 50 years ago (WHO, 2016). These costs estimates indicate that healthcare costs within the United States comprise approximately 20% of the overall gross domestic product (WHO, 2016).

This increase in health care costs results in an estimated cost of healthcare of $10,000 per year per adult, which includes a variety of both proactive healthcare measures and treatments for various illnesses and injuries (Papanicolas et al., 2018). As such, the average American adult will contribute approximately 15% of all earnings to healthcare costs (Papanicolas et al., 2018). The contribution ratio of earnings to the healthcare cost increases continuously; however, if health ailments are severe or chronic in nature (Papanicolas et al., 2018).

There are two causal factors that contribute to the inexorable rise in healthcare costs within U.S., lifestyle habits, and operational costs. (Dietz et al., 2016). The first causal factor lifestyle habits refer to the way in which the average American lives today (Dietz et al. , 2016). When compared to U.S. adults 50 years ago, the rates of disease such as diabetes, obesity, and heart disease are much more prevalent, as are other diseases attributed to unhealthy habits (Dietz et al., 2016). This is due mostly to change in lifestyle, as many adults are now more sedentary and eat less healthy diets overall (Kim & Basu, 2016). As these diseases and many others are largely preventable with proper diet and exercise, adhering to proactive measures can greatly reduce the prevalence of

these diseases; however, once they are acquired, healthcare costs can be great to mitigate symptoms (Kim & Basu, 2016).

According to research, healthy American adults only contribute to about 5% of overall healthcare costs within the United States (Dietz et al., 2016). This indicates that with proper healthy lifestyle habits reduce the overall the price of healthcare. Moreover, 50% of Americans that identify as unhealthy are responsible for the other 95% of healthcare costs annually (Dietz et al., 2016). As the population of the United States becomes less healthy, the cost of healthcare increases dramatically (Kim & Basu, 2016).

The second reason the costs of healthcare costs have risen drastically within the past few decades is operational costs (Yeganeh, 2019). As the U.S. healthcare system relies heavily on insurance costs and government initiatives such as Medicare and Medicaid, expansion of services to accommodate changing healthcare needs largely increased the cost of healthcare (Yeganeh, 2019). Additionally, research associated with the development of novel treatments and more efficient technology also raises the prices of healthcare (Yeganeh, 2019). Finally, within the past few decades, the need to digitize all data collected within the healthcare systems has been expensive and have increased the cost of healthcare (Yeganeh, 2019)

The need to digitize data facilitated by the increased need and ability to share patient files more quickly between healthcare providers and share billing information with insurance companies to decrease payment times (Bhavnani et al., 2016). To digitized data, many healthcare entities needed to code and upload a variety of data; including patient files, billing information, administrative files, employee information

and procedural protocols (Bhavnani et al., 2016) This feat required the hiring of a large amount of staff and cybersecurity professionals to complete (Bhavnani et al., 2016). Additionally, now that all information is digitized, staff must be maintained to organize, upkeep, and monitor all digital data (Bhavnani et al., 2016).

The cost of healthcare has been rising considerably within the last few decades, and the average U.S. adult is often unable to compensate for increased expenditure on healthcare options. As such, many U.S. adults suffer a variety of consequences that can be contributed directly to inadequate coverage and associated issues obtaining treatment or medication.

Consequences of Healthcare Expense

As the high costs of U.S. healthcare persist, many people are unable to gain reliable access to healthcare. According to the WHO, approximately 28 million Americans (25%) of all adults within the United States were uninsured in 2016 (WHO, 2016). As many adults lack healthcare coverage, the number of preventative deaths within the United States expect to increase within the next few years (WHO, 2016). Woolhandler and Himmelstein (2017) examined Americans that lacked adequate health insurance coverage and were able to demonstrate a connection between inadequate healthcare coverage and preventative mortality rates. Woolhandler and Himmelstein found that it is possible to avoid almost 50,000 deaths on an annual basis if Americans had access to healthcare options. Moreover, the risk of mortality is higher in adults than in lack of health insurance than within similar cohorts that possess health insurance options (Woolhandler & Himmelstein, 2017).

In addition to preventative mortality, the cost of healthcare within U.S. largely contributes to personal debt. According to Scott et al. (2018), the costs of treating injury or sickness can lead to high amounts of debt to American adults. Further, the likelihood of oppressive debt increases if the condition is chronic or rare as often treatments are more expensive (Scott et al., 2018). Increased personal debt is most common within older Americans than within younger cohorts (Banegas et al., 2016). Debt is often more common with older individuals as they often lived on fixed incomes and possess an overall higher prevalence of health issues (Banegas et al., 2016).

As such, almost 45% of individuals over the age of 65 must declare bankruptcy to try and mitigate their respective health care debt (Banegas et al., 2016). The incidence of personal bankruptcy claims is similar to the overall population (Scott et al., 2018). When examined, the debt accrued through healthcare costs can account for almost 50% of all personal bankruptcies within the United States (Scott et al., 2018). As such, the debt attributed to treating disease and injury is the highest of all developed nations globally (WHO, 2016).

As the various organization and additional stakeholders comprise the American healthcare systems, the cost and associated access to services can vary greatly (Woolhandler & Himmelstein, 2017). Although the health risks of not gaining access to appropriate medical care are significant, there is another risk to the American public concerning the healthcare system (Woolhandler & Himmelstein, 2017). As the amount of sensitive data contained within the databases of the healthcare system is great, the possible consequences of inappropriate or unauthorized access can be catastrophic

(Abouelmehdi et al., 2018). Files containing confidential medical information, sociodemographic data, and billing information are accessible within the cyber framework of almost all healthcare systems (Abouelmehdi et al., 2018). As such, ensuring these files are adequately protected is essential, and also the implementation of cybersecurity.

Cyber Security

Cybersecurity is an array of behaviors and actions of information technology (IT) experts, including the protection of data from theft, corruption, and interruption (Kimani et al., 2019). As reliance on technology increases throughout the country for completion of everyday personal and professional tasks, the amount of cyber information must be protected increases annually (Kimani et al., 2019). Information protected by cybersecurity measures can include any information accessed by the internet, personnel files, payment information, identification tools, sociodemographic data, as well as personal or private information (Kimani et al., 2019). As the breadth and depth of information that requires protection with cybersecurity increases, it should use dynamic information protection (Kimani et al., 2019). Cybersecurity tools must be useful in identifying a variety of threats and also addresses cyber breaches quickly and comprehensively.

Sensitive data may be vulnerable to cyber breaches in a few different approaches by cybercriminals. First, the programs that utilize to store sensitive files may include design flaws that allow cybercriminals to access sensitive data more easily (Kimani et al., 2019). Design flaws found in both the design of the data itself or within the framework in

which it contained (Kimani et al., 2019). Three of the most common types of cyber breaches include backdoor, denial-of-service, and direct-access cyber breaches (Kimani et al., 2019.

Common Types of Cyber Breaches

The first type of vulnerability refers to as a backdoor security breach (Tuptuk & Hailes, 2018). Within a backdoor security breach, a cybercriminal is able to bypass installed authentication or access restriction protocols (Tuptuk & Hailes, 2018). Subsequently, the cybercriminal can access sensitive data in much the same way as authorized users (Tuptuk & Hailes, 2018). Often backdoor security breaches are indicative of poor design within the data storage program (Tuptuk & Hailes, 2018). As such, the flaw in the design of the program often passes through until a backdoor security breach occurs (Tuptuk & Hailes, 2018).

Another type of cybersecurity breach includes a cyber breach called a denial-of-service attack that largely interrupts the accessibility and functionality of data or data storage (Adat et al., 2018) A denial-of-service attack can work in two ways. First, the cybercriminal may be able to corrupt the login function on sensitive data that would force the intended user to enter the "wrong" password until their login information fails (Adat et al., 2018). Similarly, cybercriminals may be able to overload the login capability of data storage platforms to disrupt the login capabilities of all authorized users (Adat et al., 2018).

The second type of denial-of-service cyber-breach attacks the sensitive data by bombarding security measures that protect the data by a collection of "ghost" accounts

(Mallela & Jonnalagadda, 2018). Due to this, it becomes incredibly difficult for IT professionals and cybersecurity experts to adequately defend the sensitive information from the array of simultaneous attacks (Mallela & Jonnalagadda, 2018). Attacks from "ghost" accounts can continue until access goes through, as security measures will inevitably fail after enough attacks have occurred (Mallela & Jonnalagadda, 2018).

The third type of cyber-data breach is called a direct-access attack. This type of attack occurs when a cybercriminal gains physical access to a computer or other device that contains sensitive information (Duffany, 2018). When this occurs, cybercriminals can physically modify the device in such a way that securing sensitive data is easier (Duffany, 2018). This type of attack may include copying software or access credentials (Duffany, 2018). Further cybercriminals can add spy-ware or other software that allows the cybercriminal to access data from a remote location (Duffany, 2018). Often when direct-access attacks occur, they can infect the rest of the devices that utilize the identical platform and can then access increased amounts or levels of sensitive data (Duffany, 2018). As cyber breaches can occur in varied ways and affect different types of fail-safes and security measures, it is imperative to have a comprehensive cybersecurity protocol to effectively thwart the efforts of cybercriminals (Mishra et al., 2018). Much of this cybersecurity measure involves protocols that integrate well into workplace culture (Mishra et al., 2018). In this way, sensitive data can be better protected and ensure that data that is sensitive to the organization, consumers, and associated stakeholders remain guarded.

Integration of Cybersecurity

The more integrated cybersecurity measures are into an organization's workplace culture, often, the more effective they are (Li et al., 2019). To ensure that a cybersecurity protocol is well integrated, it is necessary to follow five basic steps. First, a pre-evaluation must occur to bring awareness of possible security breaches to employees. In this way, employees become more aware of possible issues with cybersecurity and allow employees to feel vested within the protection of sensitive data relevant to their organization (Li et al., 2019). Often, when employees feel more integrated with the cybersecurity process, they are more likely to adhere to cybersecurity initiatives (Li et al., 2019). Also, within the pre-evaluation, the cybersecurity measures of the organization need to be evaluated to ensure that no obvious faults or shortcomings are observed (Irons et al., 2016). The next step includes the implementation and initiation of strategic planning sessions that allow the employees to become more aware of current cybersecurity protocols and measures within the future (Irons et al., 2016). Within these sessions, goals, and deadlines are established to ensure proper integration of cybersecurity measures (Irons et al., 2016). After the strategic planning stage, the initiation of the integration of cybersecurity measures occurs (Irons et al., 2016). After installation and initiation, a post-evaluation process can take place to ensure the cybersecurity measures are working correctly and quickly address the problems (Irons et al., 2016).

Effects of Cyber Breaches

A completed cyber breach can negatively impact an organization in a variety of ways (Makridis & Dean, 2018). The first way an organization is affected is through theft. Theft of data can compromise the personal and financial information of employees, consumers, and stakeholders within an organization (Makridis & Dean, 2018). Further, it can be costly for an organization to recover or repay stolen funds. The stolen intellectual property and ideas implemented elsewhere, causing a loss of revenue for the organization (Makridis & Dean, 2018). When this occurs, an organization may lose rights to the stolen intellectual property, creating setbacks within research and development (Makridis & Dean, 2018).

The reputation, or image, of an organization, can also be damaged after a security breach (Chen & Jai, 2019). If sensitive data is stolen or distributed by cyber criminals, an organization can be held accountable within public opinion (Chen & Jai, 2019). Often, consumers blame the organization for not adequately safeguarding sensitive information (Chen & Jai, 2019). When this occurs, the services or goods provided by the organization may not be used by future customers and lose the business they already possess (Biener et al., 2015). This may create a large loss of revenue and trouble with trustworthiness long term (Biener et al., 2015).

As the U.S. healthcare system is so multifaceted, the organizations affiliated amass large amounts of sensitive data that could be harmful if accessed or distributed (Martin et al., 2017). It is imperative that such a diverse system be protected from cyber breaches to ensure that sensitive data is adequately protected (Martin et al., 2017). When

cybersecurity measures fail, the healthcare system can suffer a variety of breaches that affect different individuals and different types of data (Martin et al., 2017).

Summary and Conclusions

The purpose of this quantitative, comparative study was to determine if there is a significant difference between digital and non digital breaches of individual patient records and to determine if there is a significant difference between the number of individual patient records breaches for each of the three types of healthcare entities in the United States. The examination of digital and non digital breaches amongst the three healthcare entities is important to both reducing the number of data breaches and ensuring proper allocation of resources to achieve that end. To aid in a more comprehensive understanding of how cybersecurity breaches affect individuals within the healthcare system of the United States, this chapter contains a literature review. The conceptual framework selected for this project was the NIST privacy Framework, which explains how cybersecurity breaches occur and in which ways organizations can protect themselves and their associated stakeholders. Similarly, the theoretical framework for this project is the ecological systems theory, which helps to give context between behaviors within the healthcare system and in relation to cybercrimes. This study includes an overview of the U.S healthcare system, common types of cyber breaches, and a discussion of how cyber breaches can affect organizations.

In the next chapter, I describe the methodological approach. The framework of Chapter 3 includes the methods for sampling, data collection, and subsequent data analysis in order to address the research questions associated with this research project.

Chapter 3: Research Method

The purpose of this quantitative, comparative study was to determine if there is a significant difference in the number of individual patient records affected between digital and nondigital breaches of individual patient records for each of the three types of U.S. health care entities. The examination of digital and nondigital breaches across the three types of health care entities is important to both reduce the number of data breaches and ensure proper allocation of resources to achieve that end. This chapter includes the following key sections (a) research design and rationale, (b) methodology, and (c) threats to validity. The chapter concludes with a summary of the key points presented and an overview of the succeeding chapters.

Research Design and Rationale

In order to address the purpose of the study, I developed the following three main research questions:

RQ1: Is there a significant difference between the average number of individual patient records affected by digital breaches and nondigital breaches for health care providers?

RQ2: Is there a significant difference between the number of individual patient records affected per digital breach than per nondigital breach for health plan providers?

RQ3: Is there a significant difference between the average number of individual patient records affected by digital breaches and nondigital breaches for health care clearinghouses?

In this study, I used a quantitative method with a comparative design. Quantitative methods are used to measure variables or data numerically and objectively and make use of statistical techniques to analyze the underlying difference or between and among variables or differences between groups based on some variable (Mustafa, 2011). Quantitative methods are used to deduce insights from numerically measured and statistically tested data in the hope of generalizing the findings to a larger population (Allwood, 2012). Therefore, I used a quantitative methodology to enable the determination of differences in the number of affected individuals between the types of information security breaches (i.e., digital and nondigital) for each of the three types of health care entities in the U.S. health care system (i.e., health care providers, health plan providers, and health care clearinghouses).

I used a comparative research design in this study. A comparative research design is used to attempt to determine the differences that already exist between or among groups of individuals (Gall et al., 2010). This study involved the analysis of individual patient records affected for two types of breaches and three types of health care entities, which made a comparative design appropriate.

A comparative design has been used to help to advance knowledge and contribute to the body of literature in other similar studies. Rice et al. (2018) examined the gender differences that existed in privacy concerns individuals had about unmanned aerial systems (colloquially called drones). Like the current study, Rice et al. used quantitative data to examine the differences that existed between two or more groups. While Rice et al. examined the difference in the Likert score for drone mistrust in men and women, in

the current study I examined the differences in the number of individuals affected by data breaches by type of breach or type of organization breached.

Başaran and Hama (2018) used a comparative design to compare and contrast faculty members' views towards cloud computing adoption in higher education. They used descriptive statistics and an independent *t* test to demonstrate that regional differences existed in the adoption of cloud computing. Similar to the current study, Başaran and Hamma examined quantitative data to ascertain differences across groups. While the current study differs from both Rice et al.'s (2018) and Başaran and Hamma's studies in that it examined the number of affected individuals rather than Likert scores, the methodology employed was similar in all three studies.

Methodology

This study involved the use of historical data available in the public database of HIPAA breach and violations (breach portal) maintained by the U.S. Department of Health and Human Services (HHS). I began the data collection upon receiving approval from the Walden University Institutional Review Board.

Population

The archival section of the breach portal contains a total of 2,441 data breach reports from 2009 to February 14, 2018, which is 24 months prior to the present day. Reports less than 24 months old are housed in a separate section of the portal because they are considered still under investigation. The total population of closed data breach reports at the time of this study was 2,441 reports.

Sampling and Sampling Procedures

This study involved the use of reports beginning in 2010 to the most recent report available in the archival section of the breach portal, which was about 2,601 reports. I did not consider newer data because the breaches were still under investigation and the reports were subject to change. Older data would not work due to the rapid pace of technological advancement. Data older than 2015 might not be relevant to the modern day due to improvements in technology and data security. I used 100% of the total population within the stated timeframes in the current study, so sampling was unnecessary.

Archival Data

HHS's breach portal is a publicly available database that can be accessed and downloaded online. No permission was needed to access the data because it is freely available to anyone who wants to use it. HHS investigates the reports submitted by individual entities and posts the results of their findings in the data breach portal. The portal includes the following information: name of the breached entity, state the entity resides in, entity type, individuals affected, breach submission date, type of breach, location of breached information, if a business associate is present, and a text description of the incident. For this study, I obtained data from the HHS breach portal in the form of a Microsoft Excel document and truncated the unnecessary data (i.e., name of the entity, date of the breach, state the entity resides in, and if a business associate is present).

Operationalization of Variables

I used statistical pacakage for the social sciences (SPSS) to analyze the study data to address the primary research questions in this study. To test the hypotheses for each of

the research questions, I ran a two-sample *t* test for the difference in the means of two independent variables. If the *t* test indicated that there was no statistical difference between the two means, I failed to reject the null hypothesis. If there was a statistically significant difference in the two means, I rejected the null hypothesis in favor of the alternative hypothesis.

Threats to Validity

In this section, I consider threats to validity that may have affected the current study. These threats include those related to external validity, internal validity, and construct validity as well as ethical considerations. I addressed these issues before data collection and throughout the study process.

External Validity

I conducted this study in the United States using data from the HHS breach portal. While data breaches in the United States may be similar in nature to data breaches in other countries, it is safer to assume that the findings presented in the current study relate only to the United States. Because data from the data breach portal are from all 50 states, the findings of this study should apply to all states within the United States. However, while many cybercrimes are federal offenses, states also have corresponding or supplementary state laws related to cybercrime (Jarrett et al., 2009). This may mean that some states are more or less susceptible to cybercrime. I have not taken this into account in the current study.

Internal Validity

While I used the entire breach data set available from the HHS breach portal, there is still a possibility of selection bias. The data contained in the breached database only includes incidents that affected 500 or more individuals. It is possible and likely that there are many more breaches that were either ignored or that affected less than 500 individuals. Therefore, the study results only apply to breaches that are both eventually caught and affect more than 500 individuals.

Additionally, in a few instances, I needed to make a decision whether to include or exclude a breach report. A breach report exclusion occurred when I could not determine if the breach is the result of a digital or paper breach. In these cases, personal bias may have affected the choices I made despite my efforts to avoid such bias.

In this study, I did not consider the date of the breach in the analysis of the results. It is possible that as technology shifts, types of data breaches may become more or less relevant. Therefore, this study is most applicable to the present day and may not remain applicable as technology and general practices change.

As this research involved the use of an open-access database, I had no control over how the data were collected or reported. While the database was from a reliable source (i.e., the HHS), I was not able to account for any errors or bias that may have occurred while the data were being collected or reported.

Construct Validity

I used a *t* test to evaluate the study hypotheses for each research question. In order to correctly use a *t* test, the data must be random and normally distributed (Frost, 2019). Where this was not the case, I identified the most appropriate statistical test to use.

Ethical Procedures

I did not use any human subjects in this study, meaning that an extensive institutional review board review was not necessary. Furthermore, in this study I only used information that was freely available online; therefore, the need for the confidentiality of participants did not apply to this research. I did not need to ensure the informed consent of participants because the study used archival data. However, despite not needing to protect the privacy of participants due to the public nature of the data, I did not publish the name of any company that experienced a data breach as identified in the breach portal. The names of the organizations were not pertinent to the study, so while the types of organizations that experienced a data breach were an integral part of the study, the name or identifying information of any particular company or organization was not included.

Summary

The purpose of this quantitative, comparative study was to determine if there is a significant difference in the number of individual patient records affected between digital and nondigital breaches of individual patient records for each of the three types of U.S. health care entities in the United States. nondigitalnondigitalnondigital

In this study, I used a quantitative method with a comparative design. Quantitative methods are used to measure variables or data numerically and objectively and make use of statistical techniques to analyze the underlying difference between groups based on some variable (Mustafa, 2011). Quantitative methods are also used to deduce insights from numerically measured and statistically tested data in the hope of generalizing the

findings to a larger population (Allwood, 2012). Therefore, I used a quantitative methodology to enable the determination of differences in the number of affected individuals between the types of information security breaches (i.e., digital and nondigital) for each of the three types of health care entities in the U.S. health care system (i.e., health care providers, health plan providers, and health care clearinghouses).

The major threat to validity in this study was that the data contained in the breached database only included incidents that affected 500 or more individuals. It is possible and likely that there are many more breaches that either occurred and were ignored or that affected less than 500 individuals. Therefore, the study results only apply to breaches that are both eventually reported and affect more than 500 individuals. They are not generalizable to smaller breaches.

Chapter 4: Results

The purpose of this quantitative study was to determine if there is a significant difference between digital and nondigital breaches of individual patient records for each of the three types of U.S. health care entities. The examination of digital and nondigital breaches amongst the three health care entities is essential to both reduce the number of data breaches and ensure proper allocation of resources to achieve that end. The independent variables were the types of information security breaches and health care entities, while the dependent variable was the number of breached individual patient records. To examine the difference between variables, I used statistical analysis of group means to estimate the differences in individual patient records affected between digital and nondigital breaches of health data in the three types of health care entities.

I developed the following research questions and corresponding hypotheses to aid in the examination of the impact of digital and nondigital security breaches on individual patient records nondigitalfor each of the three types of U.S. health care entities:

RQ1: Is there a significant difference between the average number of individual patient records affected by digital breaches and nondigital breaches for health care providers?

$H_0$1: There is no significant difference between the average number of individual patient records affected by digital breaches and nondigital breaches for health care providers.

H_a1: There is a significant difference between the average number of individual patient records affected by digital breaches and nondigital breaches for health care providers.

RQ2: Is there a significant difference between the average number of individual patient records affected by digital breaches and nondigital breaches for health plan providers?

$H_0 2$: There is no significant difference between the average number of individual patient records affected by digital breaches and nondigital breaches for health plan providers.

$H_a 2$: There is a significant difference between the average number of individual patient records affected by digital breaches and nondigital breaches for health plan providers.

RQ3: Is there a significant difference between the average number of individual patient records affected by digital breaches and nondigital breaches for health care clearinghouses?

$H_0 3$: There is no significant difference between the average number of individual patient records affected by digital breaches and nondigital breaches for health care clearinghouses.

$H_a 3$: There is a significant difference between the average number of individual patient records affected by digital breaches and nondigital breaches for health care clearinghouses.

This chapter contains a restatement of the data collection procedures. In this chapter, I describe the samples and the results of the statistical tests conducted in this study. This chapter also includes the results of the hypothesis testing to address the research questions posed in the study. This chapter ends with a summary of the key results of the study.

Data Collection

I collected data for this study from the historical data available in the public database of HIPAA breach and violations (breach portal) maintained by the HHS. HHS's breach portal is a publicly available database that can be accessed and downloaded online. No permission was necessary to access the data because it is freely available to anyone who wants to use it. HHS investigates the reports submitted by individual entities and posts the results of their findings in the data breach portal. The portal includes the following information: name of the breached entity, state the entity resides in, entity type, individuals affected, breach submission date, type of breach, location of breached information, if a business associate is present, and a text description of the incident. For this study, I obtained data from the HHS breach portal in the form of a Microsoft Excel document and truncated the unnecessary data (i.e., name of the entity, date of the breach, state the entity resides in, and if a business associate is present).

This study included reports beginning in 2010 to the most recent report available in the archival section of the breach portal, a total of 2,601 reports. The data did not include newer data because the breaches are still under investigation and the reports are thus subject to change. The data did not include data older than 2010 because that data

may not be relevant to the modern day due to improvements in technology and data security. I used 100% of the total population within the stated timeframes.

I imported the data from the database to Microsoft Excel and recoded it into three separate sheets for health care providers, health plan providers, and health care clearing houses. I imported the resulting data into SPSS Version 25.0 and recoded them to numerically represent categorical variables. The type of breach was recoded as 2 for nondigital and 1 for digital breach. The individuals affected variable was considered a continuous variable.

Study Results

I gathered a total of 2,601 cases from 2010 to 2020 in database and included them in the study. The covered entity types have three categories: health care providers, health plan providers, and health care clearinghouses. The majority of the cases (n = 1,876, 72.13%) come under health care providers, followed by healthcare clearinghouses (n = 376, 14.45%). Finally, there are 349 cases (13.42%) covered by health plan providers.

Data on the types of the breach were also collected. The types of breach categories include nondigital and digital breaches. Digital breaches include hacking/IT incidents, while nondigital breaches include loss, theft, and improper disposal. Unauthorized disclosure breaches get reviewed manually to determine if the breach is digital or nondigital. Unauthorized disclosures on paper/film locations were considered as nondigital, while unauthorized disclosures on email, electronic medical records, network servers, and desktop computers were considered as digital breaches.

The results of frequencies and percentages showed that 69.59% of the cases were digital (n = 1,810) while 39.1% of the cases were nondigital (n = 791).

I performed the descriptive statistics of individuals affected by the breach. Because the individuals affected variable was continuous in nature, I used measures of central tendencies, such as the mean, standard deviation, and range values, to present the data. The minimum number of affected was 500 individuals, while the maximum number of affected was 78,800,000 individuals. The mean number of affected is 74,323 individuals (SD = 1,593,587).

For the first research question, I performed two samples t-test (i.e., individual samples t-test) analyses. The first set of hypotheses considered the type of breach as the independent variable, while the individuals affected in the health care providers was the dependent variable. The type of breach was recoded into dummy variables for digital and nondigital.

Levene's test results for equality of variances has a significance of 0.002, based on which the hypothesis of equal variances was rejected in favor of the alternative hypothesis of unequal variances.

Table 1 shows the descriptive statistics of the original number of individuals affected in the health care provider entity. The results showed that the number of individuals affected is higher for digital types of the breach (M = 27,893.63, SD =

208,563.29) as compared to nondigital types of the breach (*M* = 8,917.79, *SD* = 57,863.41).

Table 1

Descriptive Statistics of the Raw Number of Individuals Affected Based on the Type of Breach

Type of Breach	*N*	*M*	*SD*	Lower 90% CI	Upper 90% CI
Digital	1,398	27,893.63	208,563.29	18,712	37,075
Nondigital	478	8,917.79	57,863.41	4,556	13,280

Based on this, I conducted an independent samples *t* test by considering the equal variances are not assumed. The results in Table 2 show that there is a significant difference between the average number of individuals affected in digital and nondigital types of breaches ($t = 3.073$, $p = .002$). Therefore, there is sufficient evidence to reject the null hypothesis, which stated that there is no significant difference between the average number of individual patient records affected by digital breaches and nondigital breaches for health care providers.

Table 2

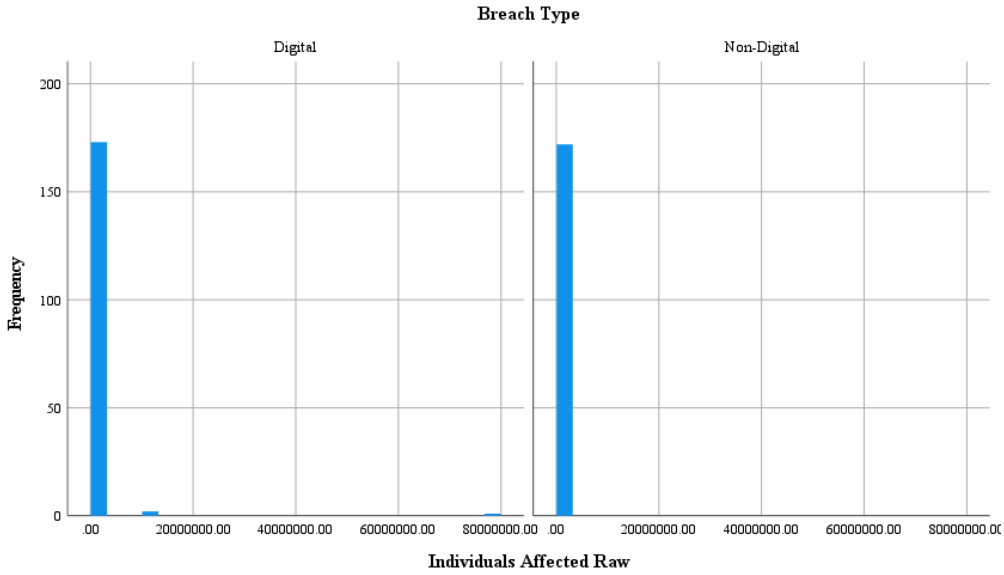
Independent Samples t-*Test Results of the Number of Individuals Affected Based on the Type of Breach for Raw Data*

	t test for Equality of Means						
						90% Confidence Interval of the Difference	
	t	*df*	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Individuals affected	3.073	1,825.79	0.002	18,974.84	6,174.10	8,815.20	29,136.48

A histogram of the data helped to identify the skewness of the data. As shown in Figure 1, it is evident that the data are very highly skewed.

Figure 1

Histogram of Raw Data of Health Care Providers



Because the raw data are highly skewed with a very long tail, I decided to trim the upper tail by removing the top 10% of the individual values. The results of the Levene's test for the equality of the variances of the raw data. The level of significance of Levene's test is 0.000, based on which the hypothesis of equal variances was rejected. Table 3 shows the descriptive statistics of the original number of individuals affected in the health care provider entity. The results showed that the number of individuals affected is higher for digital types of the breach (M = 4,094.74, SD = 4,590.11) as compared to nondigital types of the breach (M = 2,435.91, SD = 3,053.12).

Table 3

Descriptive Statistics of the 10% Trimmed Raw Number of Individuals Affected Based on the Type of Breach

Type of Breach	*N*	*M*	*SD*	Lower 90% CI	Upper 90% CI
Digital	1,230	4,094.74	4,590.11	3,879	4,310
Nondigital	458	2,435.91	3,053.12	2,201	2,671

Based on the results, as shown in Table 4, I conducted an independent samples *t* test by considering that equal variances are not assumed. The result showed that there is a significant difference between the average number of individuals affected in digital and nondigital types of breaches ($t = 8.568$, $p = .000$). This also shows that there is sufficient evidence to reject the null hypothesis, which stated that there is no significant difference between the average number of individual patient records affected for digital breach and nondigital breach for health care providers. In Table 9, the 90% confidence interval for digital and nondigital are shown.

Table 4

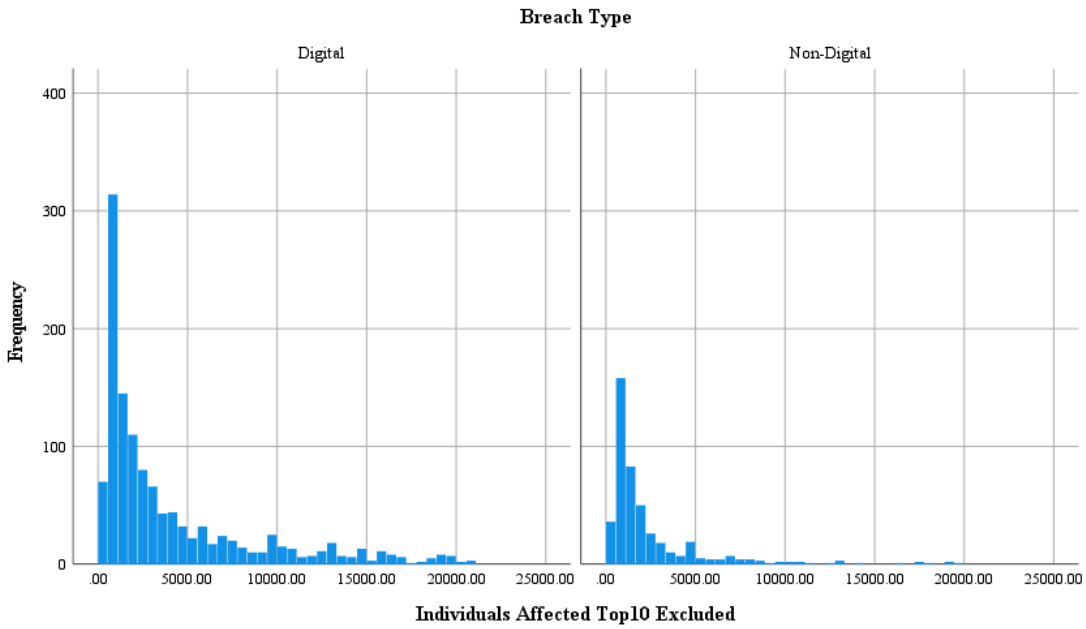
Independent Samples t*-Test Results of the Number of Individuals Affected Based on the Type of Breach for 10% Trimmed Raw Data with 90% Confidence Interval*

	t-test for Equality of Means						
						90% Confidence Interval of the Difference	
	t	*df*	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Individuals affected	8.568	1,226.82	0.000	1,658.83	193.60	1,279	2,038.66

While comparing this value with the raw data, the value of the 10% trimmed value is less skewed. I analyazed a histogram of the data to see the skewness of the data. As shown in Figure 2, the skewness is much smaller with the top 10% excluded in comparison to the original raw data.

Figure 2

Histogram When Top 10% of the Values are Excluded



In order to improve this further, the loge of raw data excluding the top 10% of the values is considered. Since the trimmed data are still highly skewed, I decided to consider loge transformation.

As shown in Table 5, the loge of the raw data excluding top 10% was analyzed for the equality of the variances. Levene's test for equality of variances had a significance of 0.002, based on which the null hypothesis of equal variances was rejected.

Table 5

Independent Samples t*-Test Results of the Number of Individuals Affected Based on the Type of Breach for Loge of 10% Trimmed Raw Data With 90% Confidence Interval*

	Levene's Test for Equality of Variances	t-test for Equality of Means

	F	Sig.	*t*	*df*	Sig. (2-tailed)	Mean Difference	Std. Error Difference	90% Confidence Interval of the Difference	
								Lower	Upper
Individuals affected	45.310	.000	8.204	984.114	0.000	0.41310	0.05035	0.31429	0.51191

Based on this, as shown in Table 6, the independent samples *t*-test was conducted by considering the equal variances are not assumed. The result showed that there is a significant difference between the average number of individuals affected in digital and nondigital types of breaches ($t = 8.204$, $p = .000$). This shows that there is sufficient evidence to reject the null hypothesis, which stated that there is no significant difference between the average number of individual patient records affected for digital breach and nondigital breach for healthcare providers.

Table 6

Descriptive Statistics of the Loge of 10% trimmed Raw Number of Individuals Affected Based on the Type of Breach

Type of Breach	*N*	*M*	*SD*	Lower 90% CI	Upper 90% CI
Digital	1,230	7.7622	1.05	7.7130	7.8115
Nondigital	458	7.3491	0.87	7.2821	7.4155

Table 6 shows the descriptive statistics of the original number of individuals affected in the healthcare provider entity. The results showed that for the loge scale, the number of individuals affected is higher for digital types of the breach ($M = 7.7622$, $SD = 1.05$) as compared to nondigital types of the breach ($M = 7.3491$, $SD = 0.87$). To bring this back to the original scale, the exponential of the mean values of digital and nondigital

breaches in the healthcare entity was obtainted and the results are displayed in Table 13. The results of the original highly skewed raw data for digital types of the breach (90% lower CI = 18712, 90% upper CI = 37075) as compared to nondigital types of the breach (90% lower CI = 4556, 90% upper CI = 13280). Whereas after the exponential of the loge of the 10% trimmed data for digital types of the breach (90% lower CI = 2237, 90% upper CI = 2469) as compared to nondigital types of the breach (90% lower CI = 1454, 90% upper CI = 1663).

Table 7

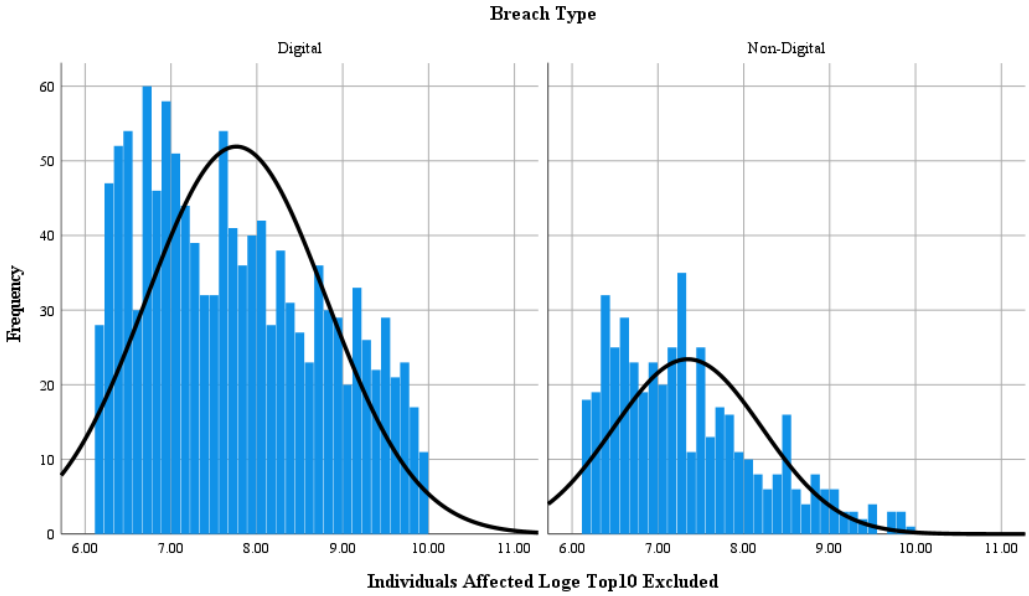
Descriptive Statistics of the Exponential of Loge of 10% Trimmed Raw Number of Individuals Affected Based on the Type of Breach

Type of Breach		*N*	*M*	Lower 90% CI	Upper 90% CI
Individuals affected	Digital	1,230	2,350.07	2237.06	2,468.79
	Nondigital	458	1,554.8	1454.03	1,662.54

In Table 7, the 90% confidence interval for digital and nondigital are shown. shown in Figure 3, there is skewness in the histogram with the loge of the top 10% excluded data. However, the skewness is lower on loge scale than the original scale.

Figure 3

Histogram of Loge of Top 10% Excluded Data



The last set of confidence intervals shown in Table 13 are the shortest confidence intervals, as they are based on the loge transformation which yields the best histograms shown in Figure 3. This graph still shows some skewness in the two histograms, but it is much smaller than those based on the raw data. In addition, the normal probability curve shows the typical patten for both the breaches. For the digital breaches, the average frequency is around 50, whereas the average frequency is at 25 for the nondigital breaches as shown in Figure 3.

In Table 8, the summary is shown with the difference based on the raw data, top 10% trimmed data, and the transformed data. These values come from the Tables 1, 3, and 6. It is evident that the trimming of the top 10% data and the transformation helped to make the data less skewed, as shown in Figure 3.

Table 8

Summary Table for Healthcare Providers

Type of Breach		Digital	Nondigital
Raw Data	Mean	27,894	8,918
	90% Lower Limit	18,712	4,556
	90% Upper Limit	37,075	13,280
	Width of 90% CI	18,363	8,724
10% Trimmed Data	Mean	4,095	2,436
	90% Lower Limit	3,879	2,201
	90% Upper Limit	4,310	2,671
	Width of 90% CI	431	470
Transformed Data	Mean	2,350	1,555
	90% Lower Limit	2,236	1,454
	90% Upper Limit	2,469	1,663
	Width of 90% CI	233	209

As shown in Table 8, the width of the 90% confidence interval has narrowed considerably by transforming the data. As the significance of the tests shown in Tables 1, 3, and 6 is 0.000, we can reject the null hypothesis that there is no significant difference between the average number of individual patient records affected for digital breaches and nondigital breaches for healthcare providers.

For the second research question, two samples *t*-test (individual samples *t*-test) analyses were performed. The second set of hypotheses considered the type of breach as the independent variable while the individuals affected in the health plan providers as the dependent variable. The type of breach was recoded into dummy variables for digital and nondigital.

The results of Levene's test for equality of variances has a significance of 0.016, based on which the hypothesis of equal variances was rejected in favor of the alternative hypothesis of unequal variances.

Table 9 shows the descriptive statistics of the original number of individuals affected in the healthcare provider entity. The results showed that the number of individuals affected is higher for digital types of the breach (M = 616,066.07, SD = 6032878.13) as compared to nondigital types of the breach (M = 24,935.24, SD = 132296.62).

Table 9

Descriptive Statistics of the Raw Number of Individuals Affected Based on the Type of Breach

Type of Breach	*N*	*M*	*SD*	Lower 90% CI	Upper 90% CI
Digital	176	616,066.07	6,032,878.13	1,135,904	1,368,036
Nondigital	172	24,935.24	132,296.62	8,252	41,618

Based on this, an independent samples *t*-test was conducted by considering the equal variances are not assumed. The results in Table 10 showed that there is a significant difference between the average number of individuals affected in digital and nondigital types of breaches (t = 1.30, p = .195). Therefore, there is no sufficient evidence to reject the null hypothesis, which stated that there is significant difference between the average number of individual patient records affected for digital breach and nondigital breach for health plan providers.

Table 10

Independent Samples t-*Test Result for the Number of Individuals Affected Based on the Type of Breach for Raw Data*

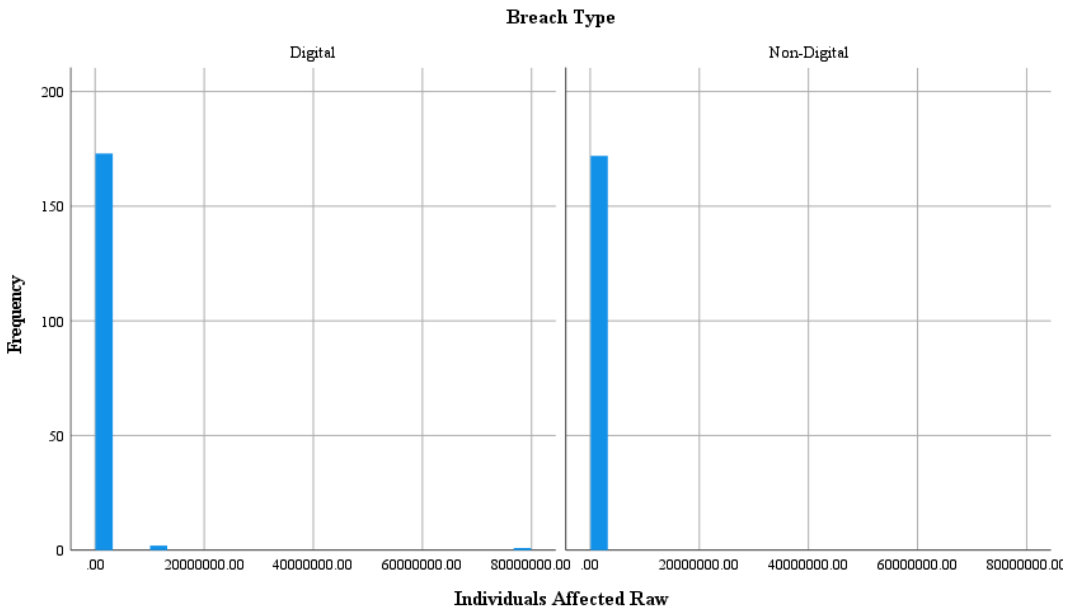
	t-test for Equality of Means

	t	*df*	Sig. (2-tailed)	Mean Difference	Std. Error Difference	90% Confidence Interval of the Difference	
						Lower	Upper
Individuals affected	1.30	175.17	0.195	591,130.82	454,857.17	-161,020	1,343,282

The histogram of the data helped to identify the skewness of the data. As shown in Figure 4, it is evident that the data are very highly skewed.

Figure 4

Histogram of Raw Data Health Plan Providers



Because the data are highly skewed with a very long tail, I decided to trim the upper tail by removing the top 10% of the individual values. The results of the Levene's test for the equality of the variances of the raw data. The signidviance of Levene's test is 0.000, based on which the hypothesis of equal variances was rejected.

Table 11 shows the descriptive statistics of the original number of individuals affected in the health plan provider entity. The results showed that the number of individuals affected is higher for digital types of the breach (M = 6458.78, SD = 8766.68) as compared to nondigital types of the breach (M = 4001.19, SD = 5386.52).

Table 11

Descriptive Statistics of the 10% Trimmed Raw Number of Individuals Affected Based on the Type of Breach

Type of Breach	*N*	*M*	*SD*	Lower 90% CI	Upper 90% CI
Digital	152	6,458.78	8,766.68	5,282	7,636
Nondigital	161	4,001.19	5,386.52	3,299	4,704

Based on this, as shown in Table 12, the independent samples t-test was conducted by considering that equal variances are not assumed. The result showed that there is a significant difference between the average number of individuals affected in digital and nondigital types of breaches ($t = 2.968$, $p = .003$). This shows that there is sufficient evidence to reject the null hypothesis, which stated that there is no significant difference between the average number of individual patient records affected for digital breach and nondigital breach for health plan providers. In Table 9, the 90% confidence interval for digital and nondigital are shown.

Table 12

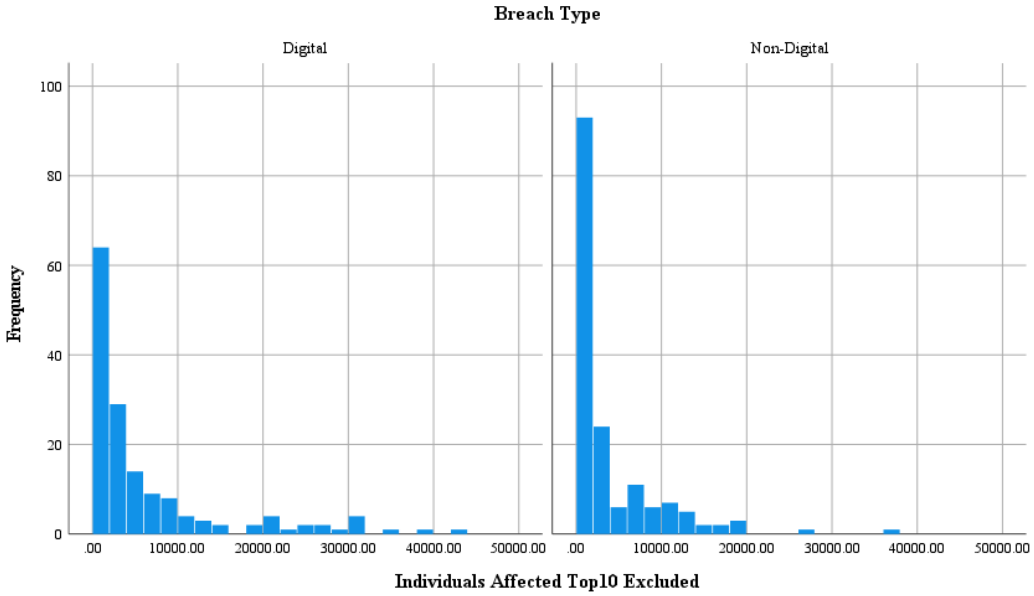
Independent Samples t-*Test Result for the Number of Individuals Affected Based on the Type of Breach for 10% Trimmed Raw Data With 90% Confidence Interval*

	t-test for Equality of Means						
						90% Confidence Interval of the Difference	
	t	*df*	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Individualsaffected	2.968	248.079	0.003	2,457.60	828.15	1,090	3,825

While comparing this value with the raw data, the value of the 10% trimmed value is less skewed. The histogram of the data was analyzed to see the skewness of the data. As shown in Figure 5, the skewness is much better with the top 10% excluded in comparison to the original raw data.

Figure 5

Histogram of Top 10% Excluded Data



In order to improve this further, the loge of raw data excluding the top 10% of the values is considered. Since the trimmed data are still highly skewed, I decided to consider loge transformation.

As shown in Table 13, the loge of the raw data excluding top 10% was analyzed for the equality of the variances. Levene's test for equality of variances had a significance of 0.003, based on which the null hypothesis of equal variances was rejected.

Table 13

Independent Samples t*-Test Result for the Number of Individuals Affected Based on the Type of Breach for Loge of 10% Trimmed Raw Data With 90% Confidence Interval*

	Levene's Test for Equality of Variances		*t*-test for Equality of Means						
								90% Confidence Interval of the Difference	
	F	Sig.	*t*	*df*	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper

Individuals affected	1.252	.264	2.979	311	0.003	0.39061	0.13114	0.17425	0.60696

Based on this, as shown in Table 14, the independent samples *t*-test was conducted by considering the equal variances are not assumed. The result showed that there is a significant difference between the average number of individuals affected in digital and nondigital types of breaches ($t = 2.979$, $p = 0.000$). This shows that there is sufficient evidence to reject the null hypothesis, which stated that there is no significant difference between the average number of individual patient records affected for digital breach and nondigital breach for healthcare providers.

Table 14

Descriptive Statistics of the Loge of 10% Trimmed Raw Number of Individuals Affected Based on the Type of Breach

Type of Breach	*N*	*M*	*SD*	Lower 90% CI	Upper 90% CI
Digital	152	8.0242	1.215	7.8611	8.1873
Nondigital	161	7.6336	1.105	7.4895	7.7777

Table 14 shows the descriptive statistics of the original number of individuals affected in the healthcare provider entity. The results showed that for the loge scale the number of individuals affected is higher for digital types of the breach ($M = 8.0242$, $SD = 1.215$) as compared to nondigital types of the breach ($M = 7.6336$, $SD = 1.105$). To bring this back to the original scale, the exponential of the mean values of digital and nondigital breaches in the health plan provider entity was obtained and the results are displayed in Table 24. The results of the original highly skewed raw data for digital types of the breach (90% lower CI = -1135904, 90% upper CI = 1368036) as compared to nondigital

types of the breach (90% lower CI = 8,252, 90% upper CI = 41,518). Whereas after the exponential of the loge of the 10% trimmed data for digital types of the breach (90% lower CI = 2592, 90% upper CI = 3605) as compared to nondigital types of the breach (90% lower CI = 1790, 90% upper CI = 2392).

Table 15

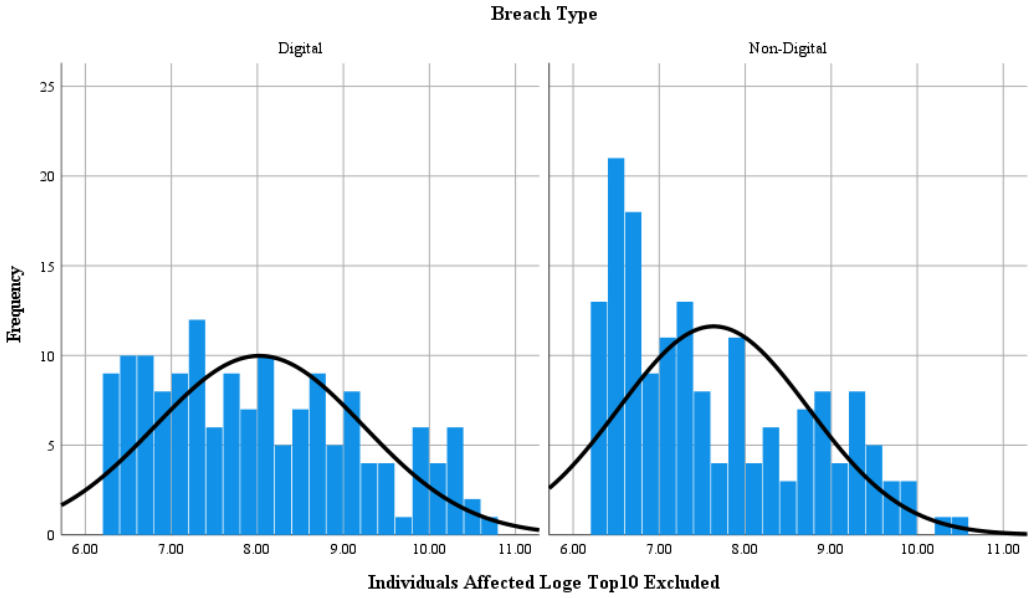
Descriptive Statistics of the Exponential of Loge of 10% Trimmed Raw Number of Individuals Affected Based on the Type of Breach

Type of Breach	N	Mean	Lower 90% CI	Upper 90% CI
Digital	152	3053.98	2592	3605
Nondigital	161	2066.48	1790	2392

In Table 15, the 90% confidence interval for digital and nondigital are shown. As shown in Figure 6, there is skewness in the histogram with the loge of the top 10% excluded data. However, the skewness is lower on loge scale than the original scale.

Figure 6

Histogram of Loge of Top 10% Excluded Data



The last set of confidence intervals shown in Table 15 are the shortest confidence intervals, as they are based on the loge transformation which yields the best histograms shown in Figure 6. This graph still shows some skewness in the two histograms, but it is much smaller than those based on the raw data. In addition, the normal probability curve shows the typical pattern for both the breaches. For the digital breaches, the average frequency is around 10, whereas the average frequency is at 12 for the nondigital breaches as shown in Figure 6.

In Table 16, the summary is shown with the difference based on the raw data, top 10% trimmed data, and the transformed data. These values come from the Tables 9, 11, and 14. It is evident that the trimming of the top 10% data and the transformation helped to make the data less skewed, as shown in Figure 6.

Table 16

Summary Table for Health Plan Providers

Type of Breach		Digital	Nondigital
Raw Data	Mean	616,066.07	24,935.24
	90% Lower Limit	-1,135,904	8,252
	90% Upper Limit	1,368,036	41,618
	Width of 90% CI	2,503,940	33,366
10% Trimmed Data	Mean	6,458.78	4,001.19
	90% Lower Limit	5,282	3,299
	90% Upper Limit	7,636	4,704
	Width of 90% CI	2,354	1,405
Transformed Data	Mean	3,053.98	2,066.48
	90% Lower Limit	2,592	1,790
	90% Upper Limit	3,605	2,392
	Width of 90% CI	13	602

As shown in Table 16, the width of the 90% confidence interval has narrowed considerably by transforming the data. As the significance of the tests shown in Tables 9, 11, and 14 is 0.003, I reject the null hypothesis that there is no significant difference between the average number of individual patient records affected for digital breaches and nondigital breaches for health plan providers.

For the third research question, two samples *t*-test (individual samples *t*-test) analyses were performed. The first set of hypotheses considered the type of breach as the independent variable while the individuals affected in the healthcare clearinghouses as the dependent variable. The type of breach was recoded into dummy variables for digital and nondigital.

Levene's test results for equality of variances has a significance of 0.05, based on which the hypothesis of equal variances was considered in favor of the alternative hypothesis of equal variances.

Table 17 shows the descriptive statistics of the original number of individuals affected in the healthcare provider entity. The results showed that the number of individuals affected is higher for digital types of the breach (M = 27893.63, SD = 208563.29) as compared to nondigital types of the breach (M = 8917.79, SD = 57863.41).

Table 17

Descriptive Statistics of the Raw Number of Individuals Affected Based on the Type of Breach

Type of Breach	*N*	*M*	*SD*	Lower 90% CI	Upper 90% CI
Digital	235	122204.93	586984.84	58972	185438
Nondigital	141	60611.07	441603.22	-968	122190

Based on this, an independent samples *t*-test was conducted by considering the equal variances are not assumed. The results in Table 28 showed that there is a significant difference between the average number of individuals affected in digital and nondigital types of breaches (t = 1.154, p = .249). Therefore, there is no sufficient evidence to reject the null hypothesis, which stated that there is significant difference between the average number of individual patient records affected for digital breach and nondigital breach for healthcare clearinghouses.

Table 18

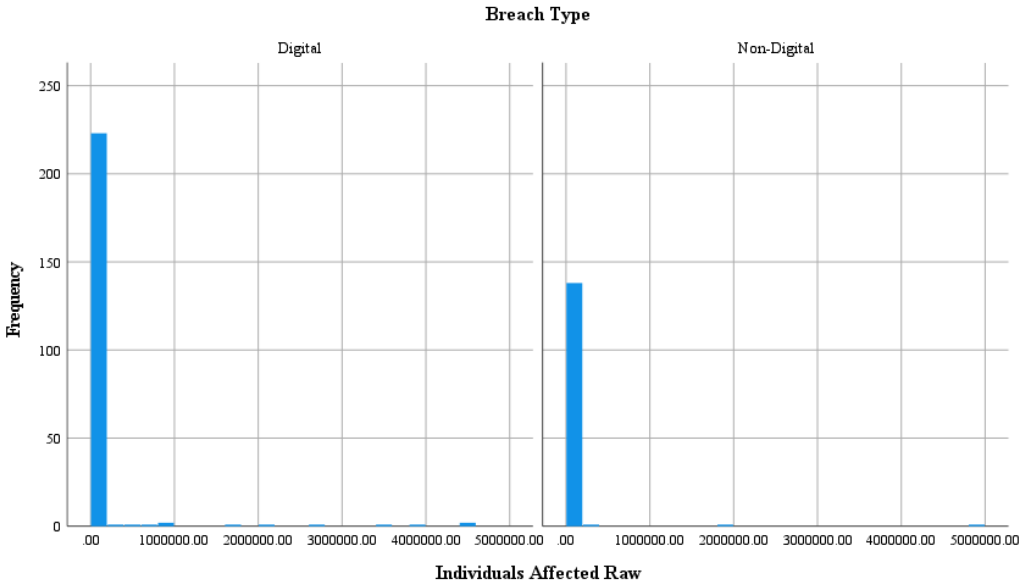
Independent Samples t*-Test Result for the Number of Individuals Affected Based on the Type of Breach for Raw Data*

	t-test for Equality of Means						
						90% Confidence Interval of the Difference	
	t	*df*	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Individuals affected	1.154	355.29	0.249	61,593.86	53,378.36	-26,435	149,623

The histogram of the data helped to identify the skewness of the data. As shown in Figure 7, it is evident that the data are very highly skewed.

Figure 7

Histogram of Raw Data of Health Care Clearing Houses



Because the raw data are highly skewed with a very long tail, I decided to trim the upper tail by removing the top 10% of the individual values. Table 29 shows the results

of the Levene's test for the equality of the variances of the raw data. The significance of Levene's test is 0.000, based on which the hypothesis of equal variances was rejected.

Table 19 shows the descriptive statistics of the original number of individuals affected in the healthcare provider entity. The results showed that the number of individuals affected is higher for digital types of the breach (M = 6,986.52, SD = 9,023.08) as compared to nondigital types of the breach (M = 4,592.81, SD = 5,935.37).

Table 19

Descriptive Statistics of the 10% Trimmed Raw Number of Individuals Affected Based on the Type of Breach

Type of Breach	*N*	*M*	*SD*	Lower 90% CI	Upper 90% CI
Digital	208	6,986.52	9,023.08	5,953	8,020
Nondigital	130	4,592.81	5,935.37	3,730	5,455

Based on this, as shown in Table 20, the independent samples t-test was conducted by considering that equal variances are not assumed. The result showed that there is a significant difference between the average number of individuals affected in digital and nondigital types of breaches (t = 2.941, p = .003). This also shows that there is sufficient evidence to reject the null hypothesis, which stated that there is no significant difference between the average number of individual patient records affected for digital breach and nondigital breach for healthcare clearinghouses. In Table 29, the 90% confidence interval for digital and nondigital are shown.

Table 20

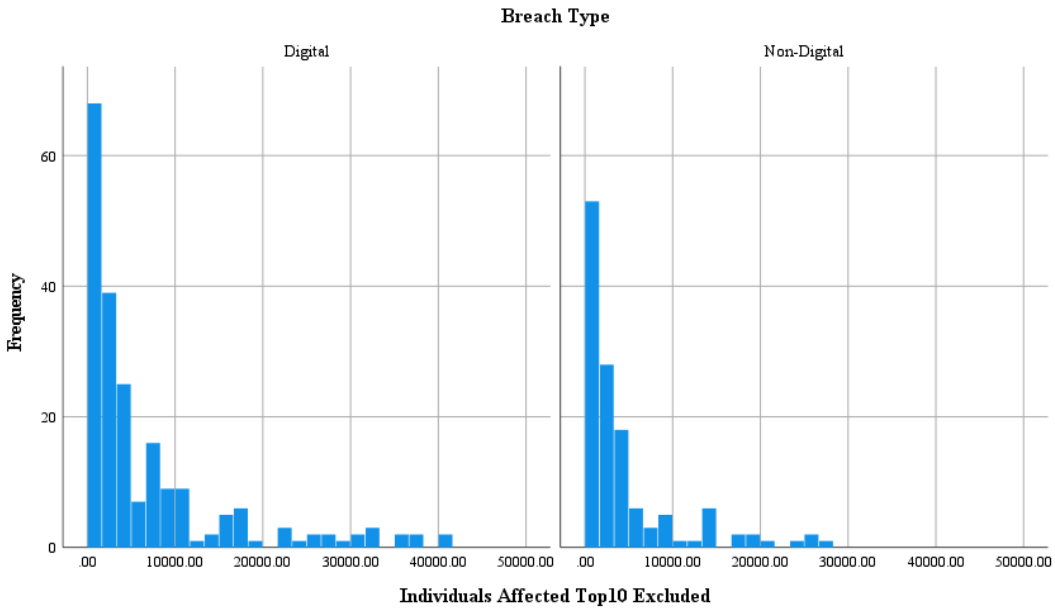
Independent Samples t-*Test Result for the Number of Individuals Affected Based on the Type of Breach for 10% Trimmed Raw Data With 90% Confidence Interval*

	t-test for Equality of Means						
						90% Confidence Interval of the Difference	
	t	*df*	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Individuals affected	2.941	335.103	0.003	1,658.83	193.60	1,279	2,038.66

While comparing this value with the raw data, the value of the 10% trimmed value is less skewed. The histogram of the data was analyzed to see the skewness of the data. As shown in Figure 8, the skewness is much better with the top 10% excluded in comparison to the original raw data.

Figure 8

Histogram of Top 10% Excluded Data



In order to improve this further, the loge of raw data excluding the top 10% of the values is considered. Since the trimmed data are still highly skewed, I decided to consider loge transformation.

As shown in Table 21, the loge of the raw data excluding top 10% was analyzed for the equaity of the variances. Levene's test for equality of variances had a significance of 0.007, based on which the null hypothesis of equal variances was rejected.

Table 21

Independent Samples t*-Test Result for the Number of Individuals Affected Based on the Type of Breach for Loge of 10% Trimmed Raw Data With 90% Confidence Interval*

	Levene's Test for Equality of Variances		t-test for Equality of Means						
								90% Confidence Interval of the Difference	
	F	Sig.	*t*	*df*	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Individuals affected	45.310	.181	2.726	289	0.007	0.35242	0.12930	0.13906	0.56578

Based on this, as shown in Table 22, the independent samples *t*-test was conducted by considering the equal variances are not assumed. The result showed that there is a significant difference between the average number of individuals affected in digital and nondigital types of breaches ($t = 2.726$, $p = .007$). This shows that there is sufficient evidence to reject the null hypothesis, which stated that there is no significant difference between the average number of individual patient records affected for digital breach and nondigital breach for healthcare providers.

Table 22

Descriptive Statistics of the Loge of 10% Trimmed Raw Number of Individuals Affected Based on the Type of Breach

Type of Breach	*N*	*M*	*SD*	Lower 90% CI	Upper 90% CI
Digital	208	8.1367	1.21	7.9980	8.2754
Nondigital	130	7.7843	1.12	7.6214	7.9472

Table 22 shows the descriptive statistics of the original number of individuals affected in the healthcare provider entity. The results showed that for the loge scale the number of individuals affected is higher for digital types of the breach ($M = 8.1367$, $SD = 1.21$) as compared to nondigital types of the breach ($M = 7.7843$, $SD = 1.12$). To bring

this back to the original scale, the exponential of the mean values of digital and nondigital breaches in the healthcare entity was obtained and the results are displayed in Table 35. The results of the original highly skewed raw data for digital types of the breach (90% lower CI = 58972, 90% upper CI = 185438) as compared to nondigital types of the breach (90% lower CI = -968, 90% upper CI = 122190). Whereas after the exponential of the loge of the 10% trimmed data for digital types of the breach (90% lower CI = 2981, 90% upper CI = 3944) as compared to nondigital types of the breach (90% lower CI = 2039, 90% upper CI = 2836).

Table 23

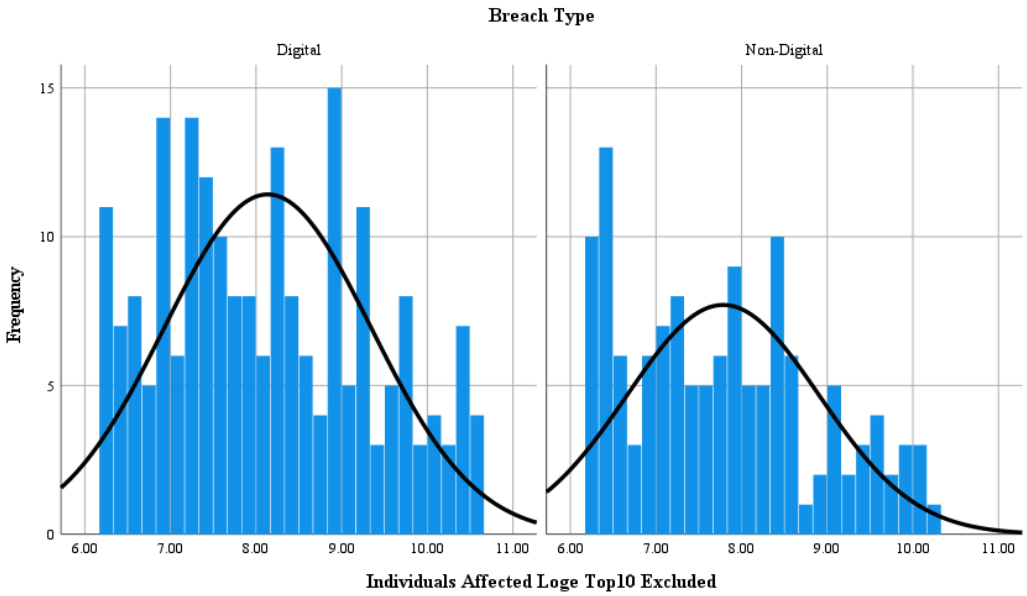
Descriptive Statistics of the Exponential of Loge of 10% Trimmed Raw Number of Individuals Affected Based on the Type of Breach

Type of Breach	*N*	*M*	Lower 90% CI	Upper 90% CI
Digital	208	3417.62	2981	3944
Nondigital	130	2402.58	2039	2836

In Table 23, the 90% confidence interval for digital and nondigital are shown.. As shown in Figure 9, there is skewness in the histogram with the loge of the top 10% excluded data. However, the skewness is lower on loge scale than the original scale.

Figure 9

Histogram of Loge of Top 10% Excluded Data



The last set of confidence intervals shown in Table 23 are the shortest confidence intervals, as they are based on the loge transformation, which yields the best histograms shown in Figure 9. This graph still shows some skewness in the two histograms, but it is much smaller than those based on the raw data. In addition, the normal probability curve shows the typical pattern for both breaches. For the digital breaches, the average frequency is around 12, whereas the average frequency is at 8 for the nondigital breaches as shown in Figure 9.

In Table 24, the summary is shown with the difference based on the raw data, top 10% trimmed data, and the transformed data. These values come from the Tables 17, 19, and 22. It is evident that the trimming of the top 10% data and the transformation helped to make the data less skewed, as shown in Figure 3.

Table 24

Summary Table for Healthcare Clearinghouses

Type of Breach		Digital	Nondigital
Raw Data	Mean	122,205	60,611
	90% Lower Limit	58,972	-968
	90% Upper Limit	185,438	122,190
	Width of 90% CI	126,466	123,158
10% Trimmed Data	Mean	6,983	4,593
	90% Lower Limit	5,953	3,730
	90% Upper Limit	8,020	5,455
	Width of 90% CI	2,067	1,725
Transformed Data	Mean	3,418	2,403
	90% Lower Limit	2,981	2,039
	90% Upper Limit	3,944	2,836
	Width of 90% CI	963	797

As shown in Table 36, the width of the 90% confidence interval has narrowed considerably by transforming the data. As the significance of the tests shown in Table 22 is 0.007, I reject the null hypothesis that there is no significant difference between the average number of individual patient records affected for digital breaches and nondigital breaches for healthcare clearinghouses.

Summary

The purpose of this quantitative study was to determine if there was a significant statistical difference between digital and nondigital breaches of individual patient records for each of the three types of healthcare entities in the United States. The examination of digital and nondigital breaches amongst the three healthcare entities is essential to both reducing the number of data breaches and ensuring proper allocation of resources to achieve that end. This study included reports beginning in 2010 to the most recent report available in the archival section of the breach portal, which was a total of 2,601 reports.

Most of the cases (n = 1876, 72.13%) were covered with healthcare providers, followed by healthcare clearinghouses (n = 376, 14.45%), and 349 cases (13.42%) were covered by health plan providers. When it comes to the healthcare providers, there is a significant difference between the average number of individuals affected in digital and nondigital types of breaches (t = 8.204, p = .000). For the health plan providers, there is a significant difference between the average number of individuals affected in digital and nondigital types of breaches (t = 2.979, p = .003). For the healthcare clearinghouses, there is a significant difference between the average number of individuals affected in digital and nondigital types of breaches (t = 2.726, p = .007).

Chapter 5: Discussion, Conclusions, and Recommendations

The purpose of this quantitative study was to determine if there is a significant difference between digital and nondigital breaches of individual patient records for each of the three types of U.S. health care entities. The results of this study are significant because understanding the differences between digital and nondigital breaches among the three types of health care entities can lead to information that can be used to reduce the overall number of breaches and better protect patient data. Furthermore, understanding the types of breaches that occur is useful in determining an appropriate allocation of security funding.

Chapter 4 contained the study findings. The present study uncovered several key findings related to health care data security breaches. First, 69.59% of all data breaches were digital (n = 1,810), while 30.41% of the cases were nondigital (n = 791). This suggests that the majority of data breaches were digital in origin. The number of breaches that occurred in health care providers is higher when comparing with health plan providers or health care clearinghouses.

RQ1: Is there a significant difference between the average number of individual patient records affected by digital breaches and nondigital breaches for health care providers? The results of the study were sufficient to reject the null hypothesis for RQ1. The null hypothesis for RQ1 stated that there was no significant difference between the average number of individual patient records affected by digital breaches and nondigital breaches for health care providers.

RQ2: Is there a significant difference between the average number of individual patient records affected by digital breaches and nondigital breaches for health plan providers? The results of the study were sufficient to reject the null hypothesis for RQ2. The null hypothesis for RQ2 stated there was no significant difference between the average number of individual patient records affected by digital breaches and nondigital breaches for health plan providers.

RQ3: Is there a significant difference between the average number of individual patient records affected for digital breaches and nondigital breaches for health care clearinghouses? The results of the study were sufficient to reject the null hypothesis for RQ3. The null hypothesis for RQ3 stated there was no significant difference between the average number of individual patient records affected by digital breaches and nondigital breaches for health care clearinghouses.

Interpretation of the Findings

The study findings indicated that 72% of all identified data breaches occurred through health care providers. Data from health plan providers accounted for 14% of data breaches, while 14% of data breaches occurred through health care clearinghouses. However, in terms of impact, more patient records are breached in the breaches at health plan providers followed by the health care providers and health care clearinghouses. Though numerous recent studies discussed security breaches in the context of health care, there is a gap in literature regarding the main source of health care data breaches (Martin et al., 2017). As previously mentioned, a number of health care-related entities have access to patient data, and the repetitious instances of individual private information

provides additional opportunities for breaches in security (Martin et al., 2017). However, previous studies did not comparatively determine the frequency of breaches across various health care agencies. Therefore, the present study extends the literature by providing evidence that health care providers may be the most susceptible or the most targeted for data breach attacks.

RQ1

The major finding of this study was that there is sufficient evidence to reject the null hypothesis that there was no significant difference between the average number of individual patient records affected for digital breaches and nondigital breaches for health care providers. This implies that may be a significant difference between the average number of individual patient records affected for digital breaches and nondigital breaches for healthcare providers. In addition to this major finding, the data related to RQ1 revealed a number of other findings supported by recent literature.

The study findings revealed that 75% of all data breaches occurred digitally, while 25% included physical records. The implication that virtual data may be more susceptible or more targeted for theft is supported by literature (Kimani et al., 2019). Kimani et al. (2019) determined that digital information is vulnerable to attack and more security is necessary every year. Kimani et al. found that design flaws within the data or the security systems can result in easy access for unauthorized persons, thus making digital data susceptible. Additionally, unlike physical data where a person must be physically present to access the data, digital data are theoretically accessible by a much larger group of individuals looking to obtain it (Kimani et al., 2019). The study findings

are, therefore, consistent with recent literature when implying that digital breaches are more common than physical breaches.

The first two findings are related to the type of breach and the type of health care organization breached. These findings provide useful background information to the major question of RQ1. nondigitalInstead of using all the data, I used the top 10% excluded data along with the transformation of log scale to reduce the skewness. The type of data breach was the independent variable, while the number of individuals impacted was the dependent variable. The results showed that there is a significant difference between the average number of individuals affected in digital and nondigital types of breaches ($t = 8.204$, $p = .000$). This shows that there is sufficient evidence to reject the null hypothesis, which stated that there is no significant difference between the average number of individual patient records affected for digital breach and nondigital breach for health care providers.

RQ2

The major finding of this study related to RQ2 was that there is sufficient evidence to reject the null hypothesis that there was no significant difference between the average number of individual patient records affected by digital breaches and nondigital breaches for health plan providers. This implies there may be a significant difference between the average number of individual patient records affected by digital breaches and nondigital breaches for health plan providers. In addition to this major finding, the data related to RQ2 revealed a number of other findings supported by recent literature.

The study findings revealed that 51% of all data breaches for health plan providers occurred digitally, while 49% included physical records. This finding suggests that the occurances of breaches in the health plan providers are equally distributed among the digital breaches and physical breaches. Previous research established that there are a number of different public and private agencies with access to sensitive patient information (Papanicolas et al., 2018). Additionally, previous research established that health care information is susceptible to attack due to its sensitive nature and the frequency of information sharing which occurs between health care agencies to ensure continuity of patient care (Bhavnani et al., 2016). Combined, the relevant studies revealed that there are different types of health care agencies and that the data from the different types of agencies is sensitive to theft (Bhavnani et al., 2016). The study findings are, therefore, consistent with recent literature when implying the digital and physical breaches have major effects on the entity.

The first two findings are related to the type of breach and the health care organization breached. These findings provide useful background information to the addressing RQ2nondigital To test the hypothesis, I performed an individual sample *t* test. Instead of using the data, the top 10% excluded data were used along with the transformation of log scale to reduce the skewness. The type of data breach was the independent variable, while the number of individuals impacted was the dependent variable. The results showed that there is a significant difference between the average number of individuals affected in digital and nondigital types of breaches ($t = 2.979$, $p = .003$). Therefore, there is sufficient evidence to reject the null hypothesis, which stated

that there is no significant difference between the average number of individual patient records affected by digital breach and nondigital breach for health plan providers.

RQ3

The major finding of this study related to RQ3 was that there is sufficient evidence to reject the null hypothesis that there was no significant difference between the average number of individual patient records affected by digital breaches and nondigital breaches for health care clearinghouses. This implies that may be a significant difference between the average number of individual patient records affected by digital breaches and nondigital breaches for health care clearinghouses. In addition to this major finding, the data related to RQ3 revealed a number of other findings supported by recent literature.

The study findings revealed that 63% of all data breaches through health care clearinghouse occurred digitally, while 37% included physical records.The implication that virtual data may be more susceptible or more targeted for theft is supported by literature (Kimani et al., 2019). Kimani et al. (2019) determined that digital information is vulnerable to attack and more security is required every year. Kimani et al. found that design flaws within the data or the security systems can result in easy access for unauthorized persons, thus making digital data susceptible. Additionally, unlike physical data where a person must be physically present to access the data, digital data are theoretically accessible to a much larger group of individuals looking to obtain it (Kimani et al., 2019). The study findings are, therefore, consistent with recent literature when implying that digital breaches are more common than physical breaches.

The first two findings are related to the type of breach and the health care organization breached. These findings provide useful background information to the major question of RQ3. nondigitalInstead of using all the data, I used the top 10% excluded data along with the transformation of log scale to reduce the skewness. The type of data breach was the independent variable, while the number of individuals impacted was the dependent variable. The results showed that there is a significant difference between the average number of individuals affected in digital and nondigital types of breaches ($t = 2.726$, $p = .007$). Therefore, there is sufficient evidence to reject the null hypothesis, which stated that there is no significant difference between the average number of individual patient records affected for digital breach and nondigital breach for health care clearinghouses.

This finding is significant because it extends and partially refutes recent literature on health care breaches (Woolhandler & Himmelstein, 2017). Woolhandler and Himmelstein (2017) found that health care databases contained mass amounts of sensitive personal data, including medical records, payment records, and personal identification data. Woolhandler and Himmelstein reported that the sheer quantity of data available digitally meant that there was a large opportunity for individuals looking to access a vast number of medical records. Though the study results indicate that more breaches occurred more often digitally than physically, the data results do not imply that the magnitude of the digital security breaches was greater than the physical security breaches. When combined, the study findings related to the three research questions extends literature by implying that, while there may be a greater opportunity for theft

through digital channels, the magnitude of the digital theft may not be higher than the magnitude of physical theft.

Limitations of the Study

There are several limitations associated with the study. Firstly, I gathered and utilized the data from databases of government agencies. While these databases provide a robust and accessible dataset, using one specific source for data breach records could have affected the findings. Though I sought to establish protocols that would ensure the included sample was unbiased, the results may not apply to all organizations given the various funding streams and record-keeping methodologies utilized.

There were two other limitations to the study. First, the variables were from databases of government agencies. This sample might have limited the insights gathered from the analysis as the data might not have reflected the general population. Second, the use of a nonprobability sampling procedure, such as purposive sampling, reduces the possibility of generalizing the results to a larger population. Though these limitations prevent the study findings from applying universally to all types of breaches, agencies, and patient groups, the results do provide useful information for practitioners regarding data security and data security prioritization.

Recommendations

Woolhandler and Himmelstein (2017) found that healthcare databases contained large amounts of sensitive personal data, including medical records, payment records, and personal identification data. This study results indicated that data breaches occurred both digitally and nondigitally. Recent literature on the topic of data breaches within a

healthcare context focused on digital data breaches (Woolhandler & Himmelstein, 2017). the study results indicated that, while digital data breaches were more common than physical data breaches, given the observed presence of nondigital data breaches and the impact of such breaches, future research could focus on determining how healthcare agencies can secure physical files in a decade when the focus is on digital security (Woolhandler & Himmelstein, 2017).

The study results indicated that a greater number of individuals were impacted due to health plan data breaches as opposed to health clearing house breaches and healthcare provider breaches. Though this result was established through the study, the study results did not indicate why health plan provider undergo huge record breaches compared to the other two entities. Future research could establish why more physical breaches are occurring with health plan providers as opposed to the other healthcare entities.

Implications

Based on the study results, I recommend that health plan providers review practices related to physical security as much as they allocate their budget for the digital security. The study results revealed that more individuals were impacted by digital data breaches from healthcare providers than from health plan providers or healthcare clearinghouses. It is recommended that individuals involved with data security at healthcare providers consider protocols for reducing security breach events and the magnitude of events should they occur.

Data security officials at healthcare organizations have more information on how to prioritize security spending. The study results indicate that more breaches occurred digitally, so security officials should consider whether it is appropriate to invest a greater proportion of funds towards preventing digital breaches as opposed to nondigital breaches However, the study results also indicated that nondigital breaches did still occur, and there was a significant difference in the number of individuals impacted. This information implies that the information security team in the healthcare entities should consider these facts when allocating their budget for digital and nondigital security.

The healthcare cost savings would result in positive social change. If healthcare agencies could implement protocols for improved data security and data management, it is possible that fewer data breaches would occur. Fewer data breaches would positively impact society by reducing losses associated with identity theft and theft of financial information. Furthermore, a reduction in data security breaches would benefit healthcare providers through an increase in reputation and a decrease in legal claims associated with data breaches.

Conclusion

The study results indicated that more breaches occurred digitally than nondigitally, but that the impact of the breaches was significantly associated with the type of breach. Additionally, the study found that health plan provider breaches resulted in a greater number of individuals impacted per incident. The study results both supported and extended recent literature on healthcare security breaches. Based on the study results and recent literature, future literature could focus on understanding physical breaches in

the modern era. Additionally, future research could consider why health plan providers appear to have larger breaches than other types of healthcare institutions. The study has positive implications for social change such as the possibility of providing health security officers with valuable information related to security prioritizations. Reducing security breaches would benefit individuals by reducing the harm associated with financial and identity theft.

References

Abouelmehdi, K., Beni-Hessane, A., & Khaloufi, H. (2018). Big healthcare data: Preserving security and privacy. *Journal of Big Data*, *5*(1), 1. https://link.springer.com/article/10.1186/s40537-017-0110-7

Adat, V., Dahiya, A., & Gupta, B. B. (2018). Economic incentive-based solution against distributed denial of service attacks for IoT customers. In *2018 IEEE International Conference on Consumer Electronics (ICCE)* (pp. 1-5). https://link.springer.com/article/10.1007/s11235-019-00599-z?shared-article-renderer

Allwood, C. M. (2012). The distinction between qualitative and quantitative research methods is problematic. *Quality & Quantity*, *46*(5), 1417-1429. https://doi.org/10.1007/s11135-011-9455-8

Banegas, M. P., Guy, G. P., de Moor, J. S., Ekwueme, D. U., Virgo, K. S., Kent, E. E., & Yabroff, K. R. (2016). For working-age cancer survivors, medical debt and bankruptcy create financial hardships. *Health Affairs*, *35*(1), 54-61. https://doi.org/10.1377/hlthaff.2015.0830

Barr, D. A. (2016). *Introduction to US health policy: The organization, financing, and delivery of health care in America*. JHU Press.

Başaran, S., & Hama, G. O. (2018). Exploring faculty members' views on adoption of cloud computing in education. *Society, Integration and Integration Confernce*, *5*, 227-237. https://doi.org/10.17770/sie2018vol1.3290

Berwick, D. M., & Gaines, M. E. (2018). How HIPAA harms care, and how to stop it.

Journal of the American Medical Association, 320(3), 215-245. https://doi.org/10.1001/jama.2018.8829

Bhavnani, S. P., Narula, J., & Sengupta, P. P. (2016). Mobile technology and the digitization of healthcare. *European Heart Journal*, *37*(18), 1428-1438. https://academic.oup.com/eurheartj

Biener, C., Eling, M., & Wirfs, J. H. (2015). Insurability of cyber risk: An empirical analysis. *The Geneva Papers on Risk and Insurance-Issues and Practice*, *40*(1), 131-158. https://www.palgrave.com/us/journal/41288

Bronfenbrenner, U. (1992). *Ecological systems theory*. Jessica Kingsley Publishers.

Chen, H. S., & Jai, T. M. C. (2019). Cyber alarm: Determining the impacts of hotel's data breach messages. *International Journal of Hospitality Management*, *82*, 326-334. https://doi.org/10.1016/j.ijhm.2018.10.002

Dietz, W. H., Douglas, C. E., & Brownson, R. C. (2016). Chronic disease prevention: Tobacco avoidance, physical activity, and nutrition for a healthy start. *JAMA*, *316*(16), 1645-1646. https://doi.org/10.1001/jama.2016.14370

Digital Guardian. (2019, January 03). *What is data loss prevention (DLP)? A definition of data loss prevention*. https://digitalguardian.com/blog/what-data-loss-prevention-dlp-definition-data-loss-prevention

Duffany, J. L. (2018). *Computer and network security essentials*. Springer.

Esser, M. (2018). Identify, protect, detect, respond, and recover: The NIST cybersecurity framework. https://www.nist.gov/blogs/taking-measure/identify-protect-detect-respond-and-recover-nist-cybersecurity-framework

Frost, J. (2019). Classical assumptions of ordinary least squares (OLS) linear regression. https://statisticsbyjim.com/regression/ols-linear-regression-assumptions/

Gall, M. D., Gall, J. P., & Borg, W. R. (2010). *Applying educational research: How to read, do, and use research to solve problems of practice* (6th ed.). Pearson.

Gomillion, D. L. (2017). Teaching case: ComprehensiveCare and the failed implementation of an electronic health records system. *Journal of Information Systems Education, 28*(2), 71-81. http://jise.org/

Health Care Cost Institute. (2019**)**. 2017 health care cost and utilization report. https://www.healthcostinstitute.org/images/pdfs/HCCI_2017_%20Health_%20Care_Cost_and_Utilization_Report

Hertler, S. C., Figueredo, A. J., Peñaherrera-Aguirre, M., & Fernandes, H. B. (2018). Urie Bronfenbrenner: Toward an evolutionary ecological systems theory. In *Life history evolution* (pp. 323-339). Palgrave Macmillan.

Hiller, J., & Russell, R. (2017). Privacy in crises: The NIST privacy framework. *Journal of Contingencies and Crisis Management, 25*(1), 31-38. https://doi.org/10.1111/1468-5973.12143

Irons, A., Savage, N., Maple, C., & Davies, A. (2016). Embedding cybersecurity in the computer science curriculum. *IT Now*, *1*, 56-57. https://academic.oup.com/itnow

Jarrett, H. M., Bailie, M. W., Hagen, E., & Judish, N. (2009). *Searching and seizing computers and obtaining electronic evidence in criminal investigations.* U.S. Department of Justice, Computer Crime and Intellectual Property Section Criminal Division. https://scholarworks.uno.edu/honors_theses/20/

Keller, J. P. (2017). Patient safety implications with the rapid adoption of IT-based health technologies. *Digital Medicine*, *3*(3), 115. https://doi.org/10.4103/digm.digm_20_17

Kim, D. D., & Basu, A. (2016). Estimating the medical care costs of obesity in the United States: Systematic review, meta-analysis, and empirical analysis. *Value in Health*, *19*(5), 602-613. https://doi.org/10.1016/j.jval.2016.02.008

Kimani, K., Oduol, V., & Langat, K. (2019). Cyber security challenges for IoT-based smart grid networks. *International Journal of Critical Infrastructure Protection*, *25*, 36-49. https://doi.org/10.1016/j.ijcip.2019.01.001

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. http://doi.org/10.3233/thc-161263

Langer, S. G. (2017). Cyber-security issues in healthcare information technology. *Journal of Digital Imaging*, *30*(1), 117–125. https://doi.org/10.1007/s10278-016-9913-x

Li, L., He, W., Xu, L., Ash, I., Anwar, M., & Yuan, X. (2019). Investigating the impact of cybersecurity policy awareness on employees' cybersecurity behavior. *International Journal of Information Management*, *45*, 13-24. https://doi.org/10.1016/j.ijinfomgt.2018.10.017

Lineberry, S. (2007). The human element: the weakest link in information security. *Journal of Accountancy*, *204*(5), 44–49. https://www.journalofaccountancy.com/

Mallela, S. S., & Jonnalagadda, S. K. (2018). *Microelectronics, electromagnetics, and telecommunications.* Springer.

Makridis, C., & Dean, B. (2018). Measuring the economic effects of data breaches on firm outcomes: Challenges and opportunities. *Journal of Economic and Social Measurement*, *43*(1-2), 59-83. https://doi.org/10.3233/JEM-180450

Margulis, S. (2003). On the status and contribution of Westin's and Altman's theories of privacy. *Journal of Social Issues, 59*(2). https://doi.org/10.1111/1540-4560.00071

Martin, G., Martin, P., Hankin, C., Darzi, A., & Kinross, J. (2017). Cybersecurity and healthcare: How safe are we?. *British Medical Journal*, *358*, j3179. https://doi.org/10.1136/bmj.j3179

Miron, W., & Muita, K. (2014). Cybersecurity capability maturity models for providers of critical infrastructure. *Technology Innovation Management Review*, *4*(10). https://timreview.ca/

Mishra, N., Sharma, T. K., Sharma, V., & Vimal, V. (2018). Secure framework for data security in cloud computing. In M. Pant, T. K. Sharma, O. P. Verma, R. Singla, & A. Sikander (Eds.). *Soft computing: Theories and applications* (pp. 61-71). Springer.

Mustafa, R. F. (2011). The P.O.E.M.s of educational research: A beginners' concise guide. *International Education Studies*, *4*(3), 23-30. http://www.ccsenet.org/journal/index.php/ies

Nimkar, S. (2016). Promoting individual health using information technology: Trends in the US health system. *Health Education Journal, 75*(6).

https://doi.org/10.1177/0017896916632790

Papanicolas, I., Woskie, L. R., & Jha, A. K. (2018). Health care spending in the United States and other high-income countries. *Jama*, *319*(10), 1024-1039. https://jamanetwork.com/journals/jama/article-abstract/2674671

Pekala, S. (2017). Privacy and user experience in 21st century library discovery. *Information Technology and Libraries*, *36*(2), 48. https://doi.org/10.6017/ital.v36i2.9817

Rice, S., Tamilselvan, G., Winter, S. R., Milner, M. N., Anania, E. C., Sperlak, L., & Marte, D. A. (2018). Public perception of UAS privacy concerns: A gender comparison. *Journal of Unmanned Vehicle Systems*, *6*(2), 83-99. https://www.nrcresearchpress.com/journal/juvs?mobileUi=0&

SearchSecurity. (2019). *What is a data breach?* https://searchsecurity.techtarget.com/definition/data-breach

Scofield, M. (2016). Benefiting from the NIST cybersecurity framework. *Information Management*, *50*(2), 25. https://www.information-management.com/

Scott, J. W., Raykar, N. P., Rose, J. A., Tsai, T. C., Zogg, C. K., Haider, A. H., Salim, A, Meara, J., & Shrime, M. G. (2018). Cured into destitution: Catastrophic health expenditure risk among uninsured trauma patients in the United States. *Annals of Surgery*, *267*(6), 1093-1099. https://doi.org/10.1097/SLA.0000000000002254

Shackelford, S. J., Proia, A. A., Martell, B., & Craig, A. N. (2015). Toward a global cybersecurity standard of care: Exploring the implications of the 2014 NIST

cybersecurity framework on shaping reasonable national and international cybersecurity practices. *Texas International Law Journal*. https://tilj.org/

Shahri, A. B., Ismail, Z., & Rahim, N. Z. A. B. (2012). Security effectiveness in health information system: through improving the human factors by education and training. *Australian Journal of Basic and Applied Sciences*, *6*(12), 226-233. http://ajbasweb.com/

Shen, L. (2014). The NIST cybersecurity framework: Overview and potential impacts. *Scitech Lawyer*, *10*(4), 16. https://www.americanbar.org/groups/science_technology/publications/scitech_lawyer/

Singh, D. A. (2015). *Essentials of the US health care system*. Jones & Bartlett.

Toé, C. A. (2013). *An examination of the explicit costs of sensitive information security breaches* [Doctoral dissertation, Capella University]. ProQuest Dissertations and Theses Global. https://search.proquest.com/docview/1427864219

Tuptuk, N., & Hailes, S. (2018). Security of smart manufacturing systems. *Journal of Manufacturing Systems*, *47*, 93-106. https://doi.org/10.1016/j.jmsy.2018.04.007

U.S. Department of Health and Human Services. (n.d.) *Cases currently under investigation*. https://ocrportal.hhs.gov/ocr/breach/breach_report.jsf

VanderWielen, L. M., & Ozcan, Y. A. (2015). An assessment of the health care safety net: Performance evaluation of free clinics. *Nonprofit and Voluntary Sector Quarterly*, *44*(3), 474-486. https://doi.org/10.1177/0899764013520235

Walser, R. (2018). Evaluation of cybersecurity management controls and metrics of critical infrastructures: A literature review considering the NIST cybersecurity framework. *Secure IT Systems: 23rd Nordic Conference Proceedings, 11252*, 369. Springer.

Woolhandler, S., & Himmelstein, D. U. (2017). The relationship of health insurance and mortality: Is lack of insurance deadly?. *Annals of Internal Medicine*, *167*(6), 424-431. https://doi.org/10.7326/M17-1403

World Health Organization. (2016). *World health statistics 2016: Monitoring health for the SDGs sustainable development goals*. https://apps.who.int/iris/bitstream/handle/10665/324835/9789241565707-eng.pdf

Wright, A., Aaron, S., & Bates, D. W. (2016). The big phish: Cyberattacks against U.S. healthcare systems. *Journal of General Internal Medicine, 31*(10), 1115-1118. http://doi.org/10.1007/s11606-016-3741-z

Yeganeh, H. (2019). An analysis of emerging trends and transformations in global healthcare. *International Journal of Health Governance*, *24*(2), 169-180. https://doi.org/10.1108/IJHG-02-2019-0012