International Journal in IT & Engineering (IJITE)
Volume 8 Issue 6, June 2020 ISSN: 2321-1776 Impact Factor: 6.493
Journal Homepage: http://ijmr.net.in, Email: irjmss@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal

# ADDRESSING CYBERSECURITY CHALLENGES OF HEALTH DATA IN THE COVID-19 PANDEMIC

**Kenneth Okereafor, PhD**
Deputy General Manager - Database Security,
Department of Information and Communications Technology,
National Health Insurance Scheme (NHIS) Abuja, NIGERIA.
nitelken@yahoo.com

**Alvin Marcelo**
University of the Philippines, Manila, PHILIPPINES.
admarcelo@up.edu.ph

## ABSTRACT

*Globally, the COVID-19 pandemic has given rise to the generation of huge health data directly from the source of crisis including medical diagnosis, hospitalization statistics, infection rate, comorbidity mapping, drug allergy, fatality rate, and other subtle metadata. While some of the data are common knowledge, others are sensitive and require proper management and security. Poorly secured health data could potentially harm reputation, lead to stigmatization, trigger misdiagnosis, or even undermine the current effort to contain the spread of COVID-19. Amid the threats and vulnerabilities that challenge health data, cybersecurity remains a sure way to address the effects and minimize impact of data compromise. This paper analyses the cybersecurity considerations over health data, and proposes mitigation actions to preserve their confidentiality, integrity, and availability.*

**Keywords:** *Cyberattack, cybersecurity, health data, information, record, sensitive, threat.*

## 1.0 INTRODUCTION

Whether it is the number of quarantined persons, the speed of contact tracing, email addresses of infected persons, or the financial cost of purchasing personal protective equipment for distribution among a list of isolation centres; the common term is that health-related data is being constantly generated, processed, transmitted across networks, stored and reused within and beyond the medical domain. Health data is sensitive because it deals with healthcare and well-being, and could be lifesaving.

Healthcare industry is one of the world's biggest and widest developing industries [1], and it is the sensitive nature of health data within the industry that makes it both valuable and attractive to cybercriminals who capitalize on porous networks and systems to steal confidential information, manipulate data, and disrupt operations within the health ecosystem. People want to access healthcare as quickly as they fall sick, they also want referral and billing to be as effective as possible using accurate health data. The efficacy of all these rely on reliable data.

It is a required cyber security practice to ensure access to reliable health data whenever required, guarantee the integrity of such health data and medical records, provide effective technology to manage the growing electronic health record (EHR), and be able to transfer health data from centers designated as isolation centers to nearby facilities. This paper discusses the cyber security dimension of health data as a highly-valuable tangible quantity that must be protected from illegal knowledge, unauthorized alteration, and delayed access.

The rest of the paper is structured as follows: **Section 2** gives a background on cyber security and health data and discusses their classifications. **Section 3** reviews the challenges associated with the confidentiality, integrity, and availability of health data. **Section 4** talks about threat impacts on health data. **Section 5** proposes a multi layered defense in-depth cyber security solution for managing health data. **Section 6** concludes the paper with summary remarks.

*Caveat*

*Throughout this paper, the terms "data", "information", and "record" are used interchangeably just for the sake of clarity and consistency.*

## 2.0 Cyber security of health data

As the COVID-19 pandemic rages on, the need for timely and accurate health data continues to rise. Frontline medics need health data for disease mapping, pharmaceutical companies rely on data to plan market segmentation, yet research institutes rely on health statistical records to strategize on clinical analysis of vaccine trials. In view of the importance of health data in the pandemic period and beyond, cybersecurity [2] is at the front burner to perform three major functions:

- Proactively ***detect*** planned or active cyberattacks against health data.
- Systematically ***prevent*** cyberattacks targeting vulnerable health information systems.
- Promptly ***respond*** to successfully executed cyberattacks and minimize their impact.

These detection, prevention and response roles define the cybersecurity dimension of managing health data in the COVID-19 era and beyond, and they rely heavily on the classification of health data [3] [1] based on value, sensitivity to privacy and criticality to life and living.

### 2.1 Classification

To classify data is to categorize or organize it based on its worth. A data classification approach to health data helps to determine the magnitude of protection that must be accorded to health data. Categorizing health data according to value and sensitivity recognizes that health data is related to life, well-being, health and living, which are critical components of life and survival.

Classification of health data also helps to estimate the strength and scope of cybersecurity systems that must be deployed to monitor its status based on value. Early and prompt data classification ensures the deployment of detective, preventive and deterrent controls, and that health data is placed under a high level of monitoring and surveillance to preserve its integrity and guarantee its availability at all times.

### 2.2 Sensitivity

Value-based classification associates health data with two important classes of information, namely: Personally Identifiable Information (PII) and Protected Health Information (PHI).

#### 2.2.1 Personally Identifiable Information (PII)

A PII is any piece of information that defines a person, distinguishes one person from another, or able to trace an individual's identity. Personally identifiable information is one of the most central concepts in information privacy regulation [4]. PII associates with an individual and can reveal a person's identity. E.g. first name, date of birth, national identification number, etc.

A patient's medical record is also classified as PII because it contains data that is personal such as diagnosis history, list of drugs administered and other sensitive private information. Due to their attachment to the individual's private information, PIIs are sensitive and require superior levels of security to manage their collation, storage, exchange, use and disposal.

#### 2.2.2 Protected Health Information (PHI)

A PHI is any set of information relating to healthcare, medical conditions, or clinical record collected of a person for the purpose of administering healthcare and monitoring health status. This includes ailment history, laboratory results, drug reaction, etc. PHIs provide an identity for potential patients to determine appropriate levels of healthcare. In some climes, patients have a legislative right of greater access to their electronic health record and insurance claims information [5] through safe management of their PHIs. As a result of their sensitive nature, PHIs are accorded a high degree of security.

Below is a partial listing of classified and sensitive health data requiring adequate protection through cybersecurity:

- Patient's health status
- Patient's registration profile
- Hospital visitation history
- Hospitalization record
- Disease treatment
- Information on medications
- Laboratory results
- Record of allergies
- Health insurance information

- Claims information
- Medical conditions
- Hospital preference, etc.

3.0 Challenges, threats, and vulnerabilities

As organizations attempt to find ways to improve patient safety and reduce the costs of care, storing health information in electronic form raises concerns about patient health, privacy, and safety [6]. Health data has been found to be vulnerable to a number of human and non-human threats, with potentials to cause data loss or have a certain degree of impact on confidentiality, integrity, and/or availability. Sources of threat [7] to health data vary with geography but would generally include vulnerable data, ignorant employees, infrastructure, medical devices, staff behaviour, automation, billing fraud, etc. Threats to health data typically manifest in the form of confidentiality, integrity, or availability issues.

### 3.1 Confidentiality issues

Loss of information confidentiality occurs when such an information or data is exposed to those who are not authorised to know or use it, or released prematurely (too early) ahead of its time of use or disclosure. Announcing a patient's health condition on social media is a typical example of such unprofessional and unethical release of data, and such act constitutes a major cyber security issues that bothers on confidentiality.

#### 3.1.1 Confidentiality-based threats

- Unauthorized access or knowledge of personal health record.
- Unauthorized disclosure of sensitive health information.
- Illegal view of patient's health profile.
- Illegal knowledge of patient's medical condition

### 3.2 Integrity issues

Health data is said to have lost its integrity when it is modified by unauthorised means or persons, or under illegal circumstances. Deletion of health data by persons who do not possess such authority compromises its integrity and places the original data owner at risk of misdiagnosis and medical emergencies. Table 1 shows variants of unauthorized modification of health data.

*Table 1: Various malicious manipulations to which health data is exposed*

| SN | Variant | Description |
|---|---|---|
| 1. | Data addition | Unauthorized inclusion of an item that was not originally part of the health data. This is also called data padding. |
| 2. | Data subtraction | Unauthorized removal of a part of data item that was meant to be part of the health data. |
| 3. | Data substitution | Unauthorized replacement of an item with another from an external source. |
| 4. | Data relocation | Movement of data to a new unintended or unapproved location. |
| 5. | Data swapping | Switching the positions of two or more data items, one in place of each other. |
| 6. | Data deletion | Removal of specific items from the health data of a living or dead person. |
| 7. | Record deletion | Removal of the entire health record of a specific living or dead person. |
| 8. | Data masking | Placing a deceptive or misleading label on data to give it a falsified meaning or manipulated interpretation. |
| 9. | Data pseudonymization | Masking data with a name or label that is unrelated to its content and that does not represent its content, e.g. nicknames or aliases. |
| 10. | Data anonymization | Deliberately making the origin, location, or identity of data unknown, or concealing its identity. |

### 3.2.1 Integrity-based threats

- Loss of patient's data or medical record to disaster
- Theft of health record of a patient
- Unauthorized modification of health data
- Accidental deletion of health information
- Malicious deletion of part/whole record of a patient's medical history
- Unplanned or unauthorized data duplication

### 3.3 Availability issues

Health information is said to have lost its availability when it proves difficult to access in a timely manner or when it is completely inaccessible due to system failure, virus, cyberattack, power supply, network fault, ransomware, sabotage, etc.

#### 3.3.1 Availability-based threats

- System shutdown
- Unfavourable environmental control system
- Network delays (high latency)
- Website highjack
- Internet slowness
- Database hacking
- Locked accounts
- Unstable access to network resources
- Power-related fluctuations

### 4.0 Impacts of threats

The effects of breach in confidentiality, integrity or availability of health information can vary depending on many factors including the attributes of the data, the systems that process them, the individual and groups, and the patients or direct beneficiaries of such data, etc. Almost all the time, an effect on confidentiality may end up affecting integrity and availability as well because health data relates to life, health and well-being and the attributes are all interwoven. Similarly, an attack can affect the organization's daily operations, reputation, and finances.

#### 4.0.1 Misdiagnosis

With an altered heath data, the tendency to misdiagnose is very high given the new wrong information, to the extent of the systems' ability to detect such unauthorized modification, otherwise the consequences may be grievous.

#### 4.0.2 Denial of service (DoS)

There is likelihood of disruptions in accessing electronic health records owing to altered processes and compromised procedures.

#### 4.0.3 System under-performance

Poor performance of medical devices may result from altered parameters, leading to inefficient operations which in turn could affect other operations within the healthcare ecosystem.

#### 4.0.4 Information safety

There is fear about safety of personal health information which in its current modified state is unusable.

#### 4.0.5 Reputational damage

On the part of the healthcare facility, hospital, or clinic whose data has been so compromised, the backlash in the media could become an image tarnishing critique capable of hunting the organization for a long time to come.

#### 4.0.6 Stolen health data used for future attacks

Cybercriminals gather valuable data in advance and often take time to aggregate them. Stolen health data becomes a potential weapon for future cyber-attacks.

#### 4.0.7 Litigations and conflicts

The outcome of privacy infringements occasioned by the leakage of confidential health information are veritable grounds for legal tussle which could spark a domino effect that could become scandalous and damaging.

#### 4.0.8 Fatality

Casualty may arise from delays occasioned by misdiagnosis or from wrong prescription due to reliance on altered drug administration information.

#### 4.0.9 Cost:

There are financial and non-financial costs associated with health data compromise. They include ransomware, fines, litigation, insurance, recovery, repairs, corrective actions, cybersecurity awareness, etc.

## 5.0 CYBERSECURITY SOLUTIONS TO HEALTH DATA CHALLENGES

One approach to managing security of health data is through risk assessment [7], and this is applied as a component of a larger more holistic approach – defense in-depth, which is the simultaneous application of multiple layers of security. Two of such layers are discussed here: governance framework, and cyber security countermeasures/controls.

### 5.1 Cyber security and IT Governance framework

According to Gartner, Information Technology (IT) governance is defined as the processes that ensure the effective and efficient use of IT in enabling an organization to achieve its goals [8]. Achieving a consistent set of secure and trustworthy health data is a deliberate IT goal of institutions and organizations that generate, use, process or interact with health data for their core functions such as hospitals, medical laboratories, pharmaceutical companies, and health insurers.

As evidenced in previous sections, cyber security is an important pillar for successful generation of trustworthy data from health information systems. An insecure system will not be acceptable to end users and will inevitably fail. Unfortunately, the health sector by itself is a very complex environment. While one patient may be under the care of several health workers who need to share healthcare data, they are also bound by privacy and confidentiality rules to keep this data within themselves, in compliance with professional ethics.

The very complex nature of health data sharing among multiple providers and stakeholders bound by strict regulatory laws makes cybersecurity mandatory. And when it comes to complex systems, structured governance and architecture frameworks are important to securing health information systems and the health data they produce.

Three IT governance frameworks applicable to healthcare are reviewed in order of complexity:

### 5.1.1 ISO/IEC 38500

ISO/IEC 38500 is a framework for the corporate governance of IT that started as a British standard and later became an International Organization for Standardization (ISO) standard. It provides a framework for effective governance of top-level IT in fulfilment of legal, regulatory, and ethical use of IT, including the management of health information systems. Specifically, by implementing ISO 27799:2016, healthcare organizations and other custodians of health information are able to ensure a minimum requisite level of security that is appropriate to their organization's circumstances and that will maintain the confidentiality, integrity and availability of health information [9]. The ISO/TC 215 is the technical committee in the field of health informatics that works on standardizing health IT systems.

### 5.1.2 COBIT5

COBIT5 is the Control Objectives for Information and related Technology (COBIT) business framework for the governance and management of enterprise IT managed by the Information Systems Audit and Control Association (ISACA). The COBIT5 Information Security Framework [10] [11] focuses on reducing cyberattacks on supply chain management system including enterprise health information systems.

### 5.1.3 ISO/TR 14639

The ISO/TR 14639 provides a guide to best practice business requirements and principles for countries and health authorities planning and implementing the use of IT to support the delivery and development of healthcare [12], specifically capacity-based eHealth architecture roadmap.

Regardless of the framework used, what is important in health information systems is that an organization formally adopts one as a guide in governing over its complex cybersecurity journey for the sake of trustworthy health data.

## 5.2 Cybersecurity countermeasures approach

The second layer is by the application of cybersecurity countermeasures. Countermeasures are intended to minimize the harm which a threat factor can potentially cause to the health sector. A countermeasure in this regard is any cybersecurity action, process, device, or system that can prevent or reduce the effects of threats, vulnerabilities, and attacks on health data due to its sensitive nature. They are the measures that a business deploys to manage threats targeting computer systems and networks.

The controls discussed below are constantly dynamic, and suitably adapting to an evolving cyber environment. A recommended approach is the application of multiple layers of these controls, earlier defined as defence in-depth cybersecurity posture.

### 5.2.1 Preventive controls

Preventive controls provide a platform to protect health data and related assets from unauthorized modification, theft, misuse, and illegal access. Some examples include:

- Well configured antivirus software
- Updated operating systems
- Well-motivated workforce
- Properly trained staff (on cybersecurity awareness, social engineering, phishing, etc)
- Fire suppressing systems
- Strategic data backup and disaster recovery plan, etc.

### 5.2.2 Deterrent controls

Deterrent controls are intended to warn a would-be attacker that they should not attack, and that their actions and activities are tracked should they go ahead. Anything that can delay or discourage a cyber attack on the infrastructure used for processing health data is regarded as a deterrent control. Some specific examples include:

- Alert systems on computer applications managing health data
- Warning posted on the wall of a computer room where health data is processed
- Prosecution notice displayed on a hospital's website
- Doors locks leading to data centre where patient/enrolee healthcare database is kept
- Barricade and mantrap at the entrance of a data centre
- Adequate lighting and CCTV systems

### 5.2.3 Detective controls

Detective controls identify and alert when cyberattacks are about to occur to facilitate effective technology management. They find errors after they have occurred. Some examples include:

- Events logging programs
- System monitoring utilities
- Alert system for suspicious data manipulation on the healthcare database

### 5.2.4 Corrective controls

Corrective controls are deployed to restore systems to normal state and minimize the effect after an unwanted or unauthorized activity has occurred. They repair any damage or restore resources and capabilities to their prior state following an unauthorized or unwanted activity [13] that could affect data. Examples of technical corrective controls include:

- Patching a system
- Scanning a network of virus
- Quarantining a virus
- Terminating a frozen process
- Archiving over-filled data storage buffers
- Refreshing log files
- Rebooting a system, etc.

## 6.0 CONCLUSION

All over the world, health data is classified as a critical asset because of its sensitivity and its relevance to life, living, well-being, health, and medical care. With the rising demands of accurate and trustworthy health data during the COVID-19 pandemic, the healthcare industry is facing an unprecedented number of cybersecurity issues, which have financial and reputational impacts on medical facilities, labs, hospitals, health insurers, health funders, governments, and other stakeholders. A proper understanding of the magnitude of these cybersecurity issues is needed to avoid the serious impacts associated with their poor protection.

This paper has proposed the adoption of relevant Information Technology governance frameworks to guide the corporate cybersecurity journey and build more trust into health information. In addition, the defence in-depth cybersecurity approach mitigates cyber threats against health data, which implies that all information systems that process health data (PIIs and PHIs) should be well protected.

It is beneficial to manage risks at the various stages of health data lifecycle including data collation, transmission, processing, and destruction. To effectively reduce the impacts of cyberattacks on health data, multiple layers of cybersecurity controls and safeguards are required, hence the proposed defence in-depth.

## ABOUT THE AUTHORS

**Kenneth Okereafor** is a United Nations trained Cybersecurity expert, and Deputy General Manager at the National Health Insurance Scheme (NHIS) Nigeria, where he oversees Database Security and Health Informatics. With a PhD in Cybersecurity & Biometrics from Azteca University Mexico, he has accumulated over two decades of professional ICT experience, and has acquired special skills in applying Cyber Threat Intelligence &

Mitigation Technologies to detect, prevent and respond to Cyberattacks in industry, government, and academia. Kenneth is a member of the International Organization for Standardization's Technical Committee on Health Informatics (ISO/TC-215), and he currently chairs ISO's Security and Privacy Working Group-4 in Nigeria, developing and adopting Cybersecurity standards for Nigeria's digital health ecosystem. He has research interests in, and publications on, Global Cybersecurity Operations, Incident Response, Multi-biometrics, Electronic Health Security, Computer Forensics, and Digital Identities; and may be reached at nitelken@yahoo.com.

**Alvin Marcelo** is a Professor of Surgery and Health Informatics at the University of the Philippines, Manila (on leave) and concurrently Chief Information Officer and Chief Medical Information Officer at St. Luke's Medical Centre.

## REFERENCES

[1] S. A and B. K. Rai, "Big Data in Healthcare Management: A Review of Literature," *American Journal of Theoretical and Applied Business,* vol. 4, no. 2, pp. 57-69, 2018.

[2] K. U. Okereafor and O. Adebola, "Tackling the Cybersecurity impacts of the coronavirus outbreak as a challenge to internet safety," *International Journal in IT and Engineering (IJITE),* vol. 8, no. 2, pp. 1-14, 2020.

[3] Z. Ansari, Q. H. Mateenuddin and A. Abdullah, "Performance Research on Medical Data Classification using Traditional and Soft Computing Techniques," *International Journal of Recent Technology and Engineering (IJRTE),* vol. 8, no. 2S3, pp. 990-995, 2019.

[4] "The PII Problem: Privacy and a New Concept of Personally Identifiable Information," *NEW YORK UNIVERSITY LAW REVIEW,* vol. 86, no. 1814, pp. 1814-1894, 2011.

[5] Michelle M. Mello, "HIPAA and Protecting Health Information in the 21st Century," *Journal of the American Medical Association (JAMA),* vol. doi:10.1001/jama.2018.5630, 2018.

[6] G. N. Samy, R. Ahmad and Z. Ismail, "Security threats categories in healthcare information systems," *Health Informatics Journal,* vol. 16, no. 201, p. DOI: 10.1177/1460458210377468, 2010.

[7] H. Pardue and P. Patidar, "THREATS TO HEALTHCARE DATA: A THREAT TREE FOR RISK ASSESSMENT," *Issues in Information Systems,* vol. 12, no. 1, pp. 106-113, 2011.

[8] Gartner, "IT governance (ITG)," Gartner, 2020. [Online]. Available: https://www.gartner.com/en/information-technology/glossary/it-governance. [Accessed 31 May 2020].

[9] ISO, "ISO 27799:2016. Health informatics — Information security management in health using ISO/IEC 27002," International Organization for Standardization, [Online]. Available: https://www.iso.org/standard/62777.html. [Accessed 31 May 2020].

[10] C. Dimitriadis and R. Stroud, "ISACA's Guide to COBIT5 for Information Security," ISACA Information Security Media Group, Princeton, New Jersey, 2012.

[11] COBIT, "Transforming Cybersecurity: Using COBIT5," ISACA, Meadows, Illinois, 2013.

[12] ISO, "ISO/TR 14639-2:2014 Health informatics — Capacity-based eHealth architecture roadmap — Part 2: Architectural components and maturity model," International Organization for Standardization, [Online]. Available: https://www.iso.org/standard/54903.html. [Accessed 31 May 2020].

[13] Debbie Walkowski, "What Are Security Controls?," F5 Labs, 22 August 2019. [Online]. Available: https://www.f5.com/labs/articles/education/what-are-security-controls. [Accessed 30 May 2020].