The Importance of Data Accuracy and Management in Healthcare

Terry L. Howard
University of Detroit Mercy

Gregory W. Ulferts
University of Detroit Mercy

Mary White
University of Detroit Mercy

Kate A. Kohn-Parrott
University of Detroit Mercy

Mithu Bhattacharya
University of Detroit Mercy

Vaishnavi Srivatsan
University of Detroit Mercy

ABSTRACT

Effective usage and storage of data are significant in the healthcare industry because this information can influence healthcare services, patients' satisfaction with a healthcare provider or organization, the efficiency of the working process of healthcare providers, and the safe storage of patients' data. Health data help doctors to maintain a comprehensive picture of patient's overall health, including previous treatments and diseases. Electronic Health Records (EHRs), used by 90% of healthcare organizations in the United States, are the most widely used system for storing health data. EHRs include but are not limited to data about administrative and clinical issues, demographics, treatments, symptoms, immunizations, prescription drugs, and radiology information. As such, EHRs can be powerful tools in achieving the Healthcare Quadruple Aim: improved health for patients and populations; enhanced patient experiences; lower healthcare costs; delivery of equitable care; and the well-being of medical professionals. Yet, for all the potential benefits offered by EHRs, healthcare providers may face challenges while using these tools. These problems relate to the confidentiality of data, ethical risks, records integration across multiple EHRs, data sets, management, and training costs required to use the supporting software.

Kewords: data storage, healthcare, electronic health records, ERH, data security, records integration

INTRODUCTION

The evidence supports the conclusion that data accuracy and management are paramount in the healthcare industry. The health of patients depends on these phenomena. Information accuracy is essential because new treatment methods can emerge from the experiences of other patients with similar conditions. Data accuracy in healthcare depends on technological advancements that have completely changed the discipline (Dash et al., 2019). Data accuracy helps provide a comprehensive approach to treating a particular patient, reducing barriers that previously hindered providers' ability to access a patient's health history and treatment protocols. Yet, hospitals, healthcare systems, and providers may struggle to implement technological advancements from operational and administrative perspectives (Batko and Siezak, 2022).

Problems may exist with the privacy of patients' data from an ethical perspective. So, every healthcare organization must be aware of processes in place to save data from cyber-attacks. Healthcare organizations apply different approaches to protect patients' data; however, cases of misuse continue. For example, out of 100 healthcare records, seven are mismatched or duplicated (Senthilkumar et al., 2018). While this number represents an average, this number can be more significant in some healthcare organizations. Problems with mismatched and duplicate records may arise from inconsistencies and shortcomings in data quality as well as the use of information technology (IT) systems that lack adequate integration. Further, there are cases where the electronic health record (EHR) is misused, thereby propagating misinformation. In addition, once data are entered, changing them in the next steps in EHRs can be challenging (Ismail et al., 2020), perpetuating data errors and quality concerns.

More than 90% of hospitals in the United States use EHRs (Kriegova et al., 2021). Such popularity of health data electronic systems creates unified information management systems. More than 50% of patients' records exist with the help of electronic systems (Kriegova et al., 2021). Therefore, data quality management is necessary for the healthcare industry, and all stakeholders must thoroughly consider this issue. High-quality data are vital for the country's social and public welfare and contribute to understanding the drivers of population health, allowing the development of policies and treatment protocols to improve health outcomes. The stakeholders in this situation are patients, medical and clinical staff, administrative staff, social care workers, government, and scientists. Conversely, low data quality can influence all healthcare organizations' departments (Ismail et al., 2020). The consequences can be patient mistreatment, poor decision-making, employee frustration, and increased bottlenecks.

Every healthcare organization is unique and has measures to ensure the importance of healthcare data quality. The structures and methods applied differ based on the organization's needs (Sieck et al., 2022). Choosing the most appropriate data systems is essential for a healthcare organization.

The evidence supports the conclusion that delivering the best possible care and improving health outcomes result because of data quality. Data management in healthcare is about using a systematic approach that assures the validity of data sources, supports data quality, and completes several redundant operations to decrease mistakes in healthcare data (Uslu and Stausberg, 2021). With the help of effective data management in healthcare, obtaining data that are accurate, valid, reliable, and easy to comprehend is possible. For example, healthcare data in EHRs include administrative records, patient records, health surveys, and results of clinical trial (Bossen et al., 2019). While these sources are full of different information with errors, employing a list of efficient processes can help deal with these mistakes and improve data quality, resulting in quality data management.

WHY DATA ACCURACY IS SIGNIFICANT IN HEALTHCARE

Because healthcare providers work with patients' medical records daily, this information needs regulations and precise protective measures to protect the information. Healthcare data include patients' medical history, previous treatment, and future needs (Khairat et al., 2021). Healthcare workers recognize the implications of collecting data and its influence on patients. Hospitals must be aware that getting precise medical data can influence patient treatment (Iyamu and Nunu, 2021). However, it is also significant for healthcare organizations to provide a high level of data security to maintain the integrity of the collected medication information. For healthcare organizations, it is essential not to forget to constantly train healthcare workers on updates in the technology sphere on health data storage.

A healthcare organization can be competitive if the data quality level is high. Maintaining high data quality has various advantages for healthcare organizations, including informed decision-making, better patient targeting, improved patient-physician relations, straightforward data implementation, and higher profitability (Schopf et al., 2019). The evidence supports the conjecture that if data are accurate, healthcare workers do not have to spend much time checking records and correcting mistakes. Therefore, high-quality data can help organize a hospital's record-keeping process more effectively. The more precise data an organization has, the better healthcare services it can provide. Accurate information allows marketers to focus on a specific group instead of targeting a broad range of people (Fanelli et al., 2022). Collecting quality patient data might make it possible to find potential customers using appropriate marketing campaigns and healthcare operations.

Good relations with customers are crucial to success in healthcare facilities. Collecting patient data and acting according to their preferences makes it possible to know patients better (Sieck et al., 2022). With the use of high-quality data, there appears to be a chance to increase the efficiency of a healthcare organization. However, time is needed to correct inconsistent or incomplete data so that it is usable. The challenge for organizations is using insights they obtain from data analysis. However, higher profitability can result from good patient relations, informed decisions, and improved marketing campaigns (Chen et al., 2020). When healthcare organizations' operations are cost-effective, their profits will likely increase.

One of the primary reasons health data accuracies are significant relates to patients receiving high-quality care. Doctors can pay more attention to patients when they are not searching for data about previous treatments, tests, allergies, and other factors. Moreover, this increases the likelihood that doctors will order appropriate treatments, including prescribing effective, non-contradicted drugs for each patient. This outcome is important because inappropriate treatments, including prescribing ineffective drugs, contribute to medical issues (Tapuria et al., 2021). Therefore, avoiding potentially harmful treatment of patients can be achieved owing to accurate health data.

A high level of health data results from paying attention to entering data at the initial stage correctly because the technology systems for storing this information are rapidly updated. When entering data incorrectly, there can be a threat that a healthcare provider from another organization can base treatment processes on incorrect data (Fanelli et al., 2022) that may not be in the patient's best interests. Hence, better practice operations are paramount and possible with high-quality health data in electronic systems.

In 2005, the domestic automakers (The Chrysler Group (now Stellantis), General Motors, and Ford Motor Company), in concert with the United Auto Workers, created the Southeast Michigan e-Prescribing Initiative (SEMI) recruiting health plans, health systems, and medical providers to join the initiative through a commitment to adopt e-prescribing as a means of ordering patient prescriptions. Within a few years, SEMI successfully catapulted

Michigan to 3rd in the nation regarding the number of medications dispensed electronically. Through this initiative, as with other e-prescribing initiatives, providers and pharmacists receive an immediate notification of patient drug allergies and potential adverse interactions with other medicines that patients are taking. The data tracked by e-prescribing help improve patient safety and assure better health outcomes. E-prescribing also helps reduce the costs of care. For example, after just a few years of use, the Chrysler Group reported an average of almost $5 in cash savings for every prescription issued electronically.

Healthcare Data Quality

Data quality is a scale to which the data completes any intended aim. In healthcare organizations, medical facilities apply data for different purposes: supporting EHRs, diagnosing and treating diseases, conducting research, analyzing patient histories, developing medical rules and procedures, and keeping patient records for community health surveillance (Westphal and Seitz, 2021). Data quality helps to maintain effective healthcare services. If data quality is low, patients may receive maltreatment or misdiagnosis, threatening the trust between patients and hospitals.

The available evidence supports the need to pay attention to the investigation of data accuracy in healthcare (Khairat et al., 2021). For example, EHRs contain information on patients' diseases, previous treatment, and physical and mental states. In addition, there is also data about patients' abnormal conditions. However, the main problem with EHRs is the duplication of records, meaning that two records exist for the same patient, which makes it difficult to get a comprehensive picture of the patient (Zhang et al., 2022). However, this problem goes away if a patient matching algorithm that identifies two records about the same patient is applied.

Another benefit of data quality comes from doctors reaching the correct diagnosis promptly using the data from nurses, patient history, vital patient records, and other sources. EHRs help in the treatment process of patients and the ability to conduct studies. By analyzing data in EHRs, finding more information regarding current trends and diseases becomes possible (Australian government, 2019). Data quality provides precise outcomes that can help to maintain clinical decision-making, find new treatments, or complete clinical trials for new medication.

The ICD-10 classification system in EHRs provides codes and symbols for diseases and symptoms. Hence, it is less challenging for healthcare providers to make mistakes because they know how to interpret information about patients (Wang et al., 2019). Such interoperability is an advantage of electronic systems that store patients' data. Furthermore, ICD-10 codes reduce healthcare costs and enable fair reimbursement policies (Wang et al., 2019). With the help of high-quality data, trustworthy relationships can form between healthcare providers and patients. Healthcare organizations that invest in healthcare data are more likely to provide valuable services to patients and increase patient loyalty.

Investigating data quality in healthcare requires attention to the Health Insurance Portability and Accountability Act (HIPAA), which collects healthcare data to revise data management approaches. Reaching these standards requires securing patients' data and accessing this information (Honavar, 2020). In addition, HIPAA also includes measures connected with healthcare providers' responsibility for following principles of transparency, accountability, confidentiality, and storage limitation. Hospitals and clinics must adhere to these principles if they want to avoid lawsuits and penalties. The available evidence provides a premise to assume that EHRs positively influence the time effectiveness of healthcare providers, minimizing time spent correcting mistakes manually.

Health Data Management

As stated above, health data management is integral to quality healthcare services. Health data management is gathering, housing, protecting, and transferring patient information, making it possible for healthcare providers to make appropriate decisions in patient treatment (Hoover, 2017). One main advantage of effective health data management is that any healthcare provider can get the necessary information about a particular patient regardless of their past treatment.

Health data management might be confused with Digital Healthcare data management or EHRs. However, the management of healthcare data is not limited to these phenomena and includes patients' health records, treatment efficacy, or medical tool logs that are crucial to supply chain management (Uslu and Stausberg, 2021). In addition, health data management is about operational and financial records covering healthcare providers and investors who deal with healthcare programs.

With the help of thoroughly organized health data management, it becomes possible to decrease healthcare expenses. For example, a cloud data management system that contains information can reduce costs in different areas, such as data administration and software support (Uslu and Stausberg, 2021). Moreover, healthcare providers do not have to deal with duplicates or spend time identifying whether a particular data source is reliable.

Health data management also helps researchers conduct studies in healthcare because it becomes possible to obtain information about diseases and symptoms in a particular area or age group of patients (Jabali et al., 2022). Therefore, there appear to be opportunities to consider preventive measures that can benefit the entire population, with achievement occurring if recommendations assist people or global actions result. These information analytics can positively influence the whole system of providing healthcare services to the public.

At the same time, benefits exist for patients when health data management is high quality. For instance, patients do not have to duplicate clinical forms or have duplicative medical tests (Upadhyay and Hu, 2022). Furthermore, patients also get access to their records, with such engagement improving records' accuracy.

Health data management boosts interoperability, and various healthcare organizations can transfer, interpret, and use information that decreases repeat testing, miscommunication, and involuntary treatment cases (Westphal and Seitz, 2021). This interoperability can also positively influence integration with patients' health-connected actions. In addition, different mobile applications provide patients with access to healthcare systems.

The evidence supports the conclusion that health data management is beneficial for integrating healthcare data with modern technologies. For example, artificial intelligence makes it possible to diagnose illness without visiting a doctor. Therefore, online chatbots can help self-service patients based on medical data management systems.

Quality Metrics for Healthcare Data

Poor data quality might lead to decreased effectiveness and mistreatment of patients. Therefore, it is imperative to investigate and use quality metrics that yield accuracy, precision, validity, reliability, cohesiveness, availability, completeness, identifiability, security, usability, and uniqueness. One of the main features of data quality is consistency. Data accuracy relates to the use of high-quality information. Misleading data can lead to severe complications (Schopf et al., 2019). The evidence supports the concept that investment in data quality management is necessary.

The validity of information is about the requirements needed for data. For example, most hospitals have a department that identifies a set of rules that allow an organization to determine the appropriateness of the information in their records. Moreover, the data collection process might be complete with the help of different sources (Rudin et al., 2020). At the same time, these sources should not contradict other systems that collect information and support the consistency of the collected data.

No differences exist in data that are incomplete or inaccurate. Missing information could lead to a contradiction of reality. Without a complete picture of a patient, inappropriate actions can be dangerous. Therefore, it is crucial to recognize the requirements that identify a cohesive information set to prevent such a situation (Chen et al., 2020). A healthcare organization must meet all needs for data collection. In collecting data, giving consideration to details is of supreme importance. If some pieces are missing, wrong decisions can occur that could negatively affect the whole healthcare organization (Sieck et al., 2022). However, a thoroughly organized information set can result in a different interpretation of identical information offered with inconsistencies at a lower level.

All healthcare data must be available, and healthcare providers must have free and rapid access. Metrics such as completeness are about the fact that prescriptions for a patient must contain the name of a doctor, drugs, and the time and date of their medication (Kriegova et al., 2021). All data must be unique and contain no duplicates. The treatment process and future data analytics can be informative to other cases. Metrics, including security or confidentiality, are important in data usage in healthcare organizations. Only authorized healthcare providers can access this data. In addition, data needs to be available in a format understandable to all stakeholders who might use the data (Tapuria et al., 2021). For example, EHR and manual records have codes, abbreviations, and symbols that doctors and nurses can read and interpret.

Quality Standards for Healthcare Data

Healthcare data standards for quality tightly connect with Health Level 7 (HL7) integration that offers a way to exchange and keep information. HL7 is about transferring information using standards for transferring administrative and clinical data (Alami et al., 2022). These regulations are also devoted to describing the structure of knowledge, its type, and language. HL7 also supports operations, management, and thoughts on data. The main aim of HL7 requirements is to provide seamless integration from one system to another, using structured project labeling, clinical information structures, messaging standards, and care documentation continuation (Alami et al., 2022).

When sharing patient health information, it is important to focus on protocols that protect patient data. This focus requires that software is updated regularly and observed by IT professionals. If patients' information becomes known, healthcare organizations can face lawsuits and penalties, leading to patient mistrust towards healthcare organizations (Bossen et al., 2019). Different programs that aim to effectively explain to healthcare providers how to deal with patient health information are available. The most widely used initiatives about the quality of health data are the International Classification of Diseases 10th Revision (ICD-10), Logical Observation Identifiers Names and Codes (LOINC), Consolidated Clinical Document Architecture (C-CDA), Minimum Data Set (MDS), Diagnostic and Statistical Manual, Fifth Edition (DSM-5), Systematized Nomenclature of Medicine -- Clinical Terms (SNOMED CT), Data elements for emergency department systems (DEEDS), and RxNorm. With the help of these standards, healthcare providers can prioritize and know precisely how to organize the process of collecting and storing patient health data (Wang et al., 2019). These

standards also maintain high interoperability between healthcare providers, positively influencing the level of providing healthcare services and exchanging important information.

The health data quality standards depend on symbols and codes used to hide information from nonauthorized users and for interpretation by healthcare providers. These symbols and codes' standards verify the accuracy of health information. Healthcare organizations should approve abbreviations to enter health data (Khairat et al., 2021). Available evidence supports the supposition that entering data into electronic health data systems is necessary to address international units.

Data Quality Tools

Hospitals and healthcare systems may employ a series of processes and plans, known as data quality tools, to bolster their efforts to obtain and support high levels of data quality and improve care delivery. Such quality-improving tools focus on the patient, system processes, and teamwork, with the recognition that all functions in healthcare organizations aim to meet patients' needs (Dash et al., 2019). Of course, patient satisfaction is an outflow of meeting patients' needs. Hence, using quality-improving tools may minimize or eliminate actions that result in poor levels of patient satisfaction. The evidence supports the inference that quality management results from systems that organize activities, healthcare workers, and sources (Johnson, 2021). In addition, teamwork is crucial to success because completed systems are connected better and deal with data quality in case there is communication between doctors and nurses.

Data profiling is about assessing current data and uncovering hidden information. A data profiling algorithm evaluates information and identifies possible data cleansing. Data profiling is discovering if missing or duplicated data, incorrect data patterns, data outside of acceptable values, or data written with the wrong measuring unit are in the files. In addition, adding missing data is necessary (Khairat et al., 2021). Contacting appropriate staff or patients for the needed data allows easy correction.

Data cleansing and standardization are important ways to improve the quality and reliability of healthcare data. Through these processes, there can be the removal of empty values, transforming of letter cases, combining similar columns, updating entries to conform with the precise patterns needed, and deletion of repetitive keywords, among other improvements to ensure data validity, accuracy, and shareability. Data cleansing and standardization thus decrease inappropriate and invalid data enabling the combination of patient data records from different sources (Bossen et al., 2019).

Duplication of patient data also creates a problem with EHRs, which requires mitigation through patient data matching, whereby there is a comparison of two or more similar patient records, and a determination made as to whether these records are about the same person. Patient record matching helps identify whether potential duplicate records apply to the same patient (Westphal and Seitz, 2021; Ismail et al., 2020). At the same time, if unique identifiers are absent, it becomes possible to use fuzzy matching algorithms.

Cases of Low Data Quality in Healthcare

Despite all the benefits associated with electronic health systems used for collecting and storing patients' data, there have been situations when their low quality has negatively influenced patients, hospitals, and the whole healthcare system. For example, the United States government identified patient misidentification as one of the leading causes of death. The number of duplicative records is rapidly growing at hospitals. As a result, approximately 10% of incoming patients cannot receive comprehensive treatment (Uslu and Stausberg,

2021). Moreover, exchanging data from various sources can lead to additional duplicate records in the future. Healthcare organizations may encounter difficulties analyzing the information they get, leading to lower data quality. For instance, the Ohio health department kept data in different silos, with no opportunity to organize information, resulting in doctors being unable to make appropriate patient decisions. As a result, following the Network for Excellence in Health Innovation, wrong prescriptions cost $21 billion annually and cause 7,000 deaths annually (Upadhyay and Hu, 2022). The evidence supports that healthcare organizations must improve the data quality they store and use while providing healthcare services.

Electronic Health Records (EHR)

EHRs are of paramount importance in the healthcare system. Through EHRs, opportunities exist to easily access information about a patient's health history. Then, it becomes less challenging to provide healthcare services to this patient. EHRs are the electronic version of patients' paper records (Ismail et al., 2020; Dash et al., 2019). EHRs are records that are accessible by authorized users. EHRs include information on diseases and treatment records of patients, creating a more inclusive and broader perspective of patients' care (Sieck et al., 2022). In addition, EHRs are an integral part of the healthcare IT sphere because healthcare workers have access to information they can use when making decisions about providing healthcare.

One more advantage of EHRs is that they help manage problem lists that describe the main challenges a particular healthcare organization must overcome. In addition, EHRs also describe guidelines, procedures, and patient-specific plans for care (Sieck et al., 2022). Therefore, providing healthcare services more effectively improves because there is no need to spend time searching data on a particular clinical case in other sources. EHRs also help organize clinical documents and notes. Due to specific signs and codes used in these systems, every healthcare provider can quickly interpret data entered by their colleagues.

EHRs are essential in healthcare because healthcare providers from various hospitals can get needed data if they are authorized users (Khairat et al., 2021; Chen et al., 2020). In addition, EHRs are available to healthcare providers from laboratories, hospitals, pharmacies, emergency facilities, and clinics.

In investigating data quality and healthcare management, paying attention to the benefits of EHRs, which are integral to providing healthcare services, is essential (Senthilkumar et al., 2018). These advantages include providing accurate data about patient treatment, diseases, and allergies; giving quick access to these data; providing confidentiality in exchanging these data with other healthcare organizations; helping healthcare providers offer care; improving communication among healthcare providers; ensuring more reliable treatment; promoting streamlined coding and billing; improving the productivity of healthcare providers; ensuring achieving of business goals of clinics; and decreasing expenses (Dash et al., 2019).

EHRs create an opportunity to conduct clinical studies investigating the most widely spread diseases or a group of people in a particular area. The available evidence provides a premise to use data from EHRs and help the population from a healthcare perspective (Wang et al., 2019). EHRs positively influence the whole healthcare system because they help improve different aspects of patients' treatment. As stated earlier, expenses on healthcare are likely to decrease. For the advantages and disadvantages of EHRs, the most relevant penalty relates to nonauthorized access to health data. These cases may happen because of cyber-attacks, with most issues occurring in Brazil and Mexico (Fanelli et al., 2022).

EHRs can be confused with Electronic Medical Records (EMRs). There are fine distinctions between the two. EMRs focus on offering healthcare providers the ability to diagnose and prescribe drugs. Furthermore, EMRs consist of data about patients not usually used outside the practice. Hence, EMRs are more challenging to access than EHRs.

Epic is an example of a cloud-based system that facilitates sharing of patient data across multiple health systems and users, thereby helping to reduce medical errors that may arise when patients obtain care from more than one hospital system. Epic functions as an EHR and an EMR and contains "more than 250 million electronic reports of patients" (Danao et al., 2022). In 2015, Henry Ford Health (HFH), a comprehensive, integrated health system headquartered in Detroit, Michigan, began a pilot to embed Choosing Wisely recommendations into its Epic EMR workflows. This pilot, conducted in collaboration with Stanson Health, consisted of turning on seven Choosing Wisely Best Practice Alerts (BPA) that affect outpatient and inpatient orders. These BPAs popped up when a provider attempted to order a test, medication, or procedure classified as "low value," according to Choosing Wisely, offering the ordering physician a reminder to reconsider the action. Upon successfully completing the pilot, HFH added more than 70 BPAs to its Epic workflow" (Kohn-Parrott and Panella, 2017). Accordingly, HFH successfully improved the delivery of high-value healthcare services and increased patient satisfaction. This outcome is just one example of how a health system can improve healthcare delivery by using an augmented EHR/EMR.

Challenges of Data Accuracy and Management in Healthcare

The healthcare data housing market reached $3.08 billion in 2021 and is anticipated to grow by 10.7% annually (Iyamu & Nunu, 2021). However, healthcare organizations face various challenges in this sphere, including integrating reliable sources, data verification, and data transformation to recipients, which are some of the main aims of exchanging healthcare data. At the same time, challenges exist in this sphere, such as establishing accurate analytics. In addition, getting accurate and consistent information is a challenge (Hoover, 2017). A high level of providing healthcare services is one of the advantages of constant information drift detection.

As stated above, the privacy of patients' data is critical in data usage. If privacy has been negatively affected, doctors may not have a complete picture of the patient, and diagnosing a disease or prescribing appropriate treatment becomes challenging (Rudin et al., 2020). Regular support, reliable antivirus hardware, and multi-factor authorization can help make the data storage system more confidential. A 6-month investigation conducted in the United States, the United Kingdom, France, and Germany revealed the theft of 45 million medical images from networked-attached storage systems (NAS) to 2,000 servers (Batko and Siezak, 2022). Data must meet storage requirements following the HIPAA principles, and the system must respond to government regulations on access (Chen et al., 2020). Data leaked because nonauthorized users entered the system without the required login information. Protecting patients' data from cyber-attacks is a high priority that requires comprehensive software support.

The available evidence supports the need for an extensive infrastructure that can cope with and maintain industry changes and efficient data management. Therefore, using comprehensive hardware to host the servers is a necessity. In addition, storing healthcare data also requires constant maintenance and upgrades, with healthcare providers observing the state of electronic systems and willing to call professionals when help is needed (Alami et al., 2022). At the same time, constant checks of software are also compulsory to ensure adequate data storage of healthcare records.

Another challenge in healthcare data storage is the functional integration for doctors, nurses, patients, payers, and the government. If data integration fails, the chances to combine different healthcare information decreases rapidly, and it becomes more complicated to get benefits for public health in that situation (Iyamu & Nunu, 2021; Schopf et al., 2019).

Organizing a proper catalog of datasets is also an issue in the health data industry. Different electronic systems store healthcare data, and it can be challenging for healthcare providers to deal with tagging and taxonomies of each design (Kriegova et al., 2021). Furthermore, because of these differences between electronic healthcare data systems, duplications and other errors can appear that lead to improper treatment and payment fraud.

Data Quality Improvement in Healthcare

Data analytics integration is vital to ensuring high-quality data. This process includes methods for quantifying and qualifying data. The quality of sources requires attention to improving data quality within healthcare organizations. Therefore, it is significant to investigate strategies that can help improve data analysis in a healthcare organization when examining the data quality problem (Schopf et al., 2019; Uslu and Stausberg, 2021). Data quality improvement is continuous and should never stop in healthcare organizations. A plan that includes the top data quality practices needs to be in place to ensure that data are of the highest quality. Integrated data analytics helps decrease errors, automate information workflow, and provide data management. The process consists of three phases: receiving, formatting, and transferring data. Another method is to apply tools for quantifying and qualifying data. This method assesses data sets that involve selecting appropriate tools and metrics. The correct data format is also critical in working with healthcare data (Rudin et al., 2020). Delivering timely data helps to organize the working process effectively.

The introduction of systematic data quality management at hospitals is a robust method that can help improve data quality (Alami et al., 2022). Implementing a data quality management structure is needed to ensure the collection of comprehensive data in the future. Its main aim is to check and fix information before being housed in the destination resource.

Other methods of improving data quality are available that ensure the data are secured, comprehensively organized, and stored in the correct place for use when needed. Furthermore, ensuring that healthcare data are of high quality in a healthcare organization is significant. For example, checking healthcare records when the working period is non-busy could be helpful. In addition, verifying and evaluating all information collected is a necessary step in the process (Khairat et al., 2021). For example, validation of demographic information can occur before and after the consultation. Another way of data verification is auditing to identify challenges in a healthcare organization. The audit preparation includes the aims and objectives required when the audit finishes. Audits can be quick and deal with an overview of data quality status. Audits can also be in-depth to identify strengths, opportunities, weaknesses, and threats (Wang et al., 2019). Recommendations and suggestions for improving data quality occur after the competition of audits.

The main reasons for poor data quality require identification with the audit results to avoid the same errors from happening in the future. With the help of a proactive approach, healthcare providers can decrease the time spent on fixing errors in data quality (Bossen et al., 2019). Analyzing the leading causes of data errors requires scrutinizing reports on the latest investigation of data quality and collaborating with other healthcare providers to understand how to prevent these mistakes in the future. In addition, keeping the information in the healthcare organization updated and maintaining archives of inactive or deceased patients are important (Jabali et al., 2022).

It is also important to involve healthcare leaders and managers to improve data quality. Making changes in any healthcare organization is impossible without an agreement among the leaders and higher-level management. Healthcare organizations hire data quality officers to deal with the launching of data quality practices that decrease data loss and increase data quality (Senthilkumar et al., 2018). Health data managers ensure that patients' data are accurate, confidential, and updated. These professionals supervise data analysts and information technologists to ensure accuracy and help with data tasks. Health data managers, who generally have primary responsibility for ensuring the high quality of healthcare data, may work at medical research facilities, hospitals, clinics, and other similar care facilities. Following the statistics, one out of five patients faces the problem of mistakes in their records (Fanelli et al., 2022). Incorrect test results and outdated treatment can result in low data quality. Therefore, health data managers must provide a high level of security for healthcare data and protect them from hacker attacks. Furthermore, health data managers must cope with data analysis as they gather and visualize information to make it easier for administrators and executives to make final decisions.

Training and educating healthcare teams is crucial because they must read, understand, and analyze the available information to support their treatment decisions. In addition, acclimating health care professionals to the use of electronic tools—including how to enter and manipulate data and what each data element means--and achieving the value of high-quality, accurate data through participation in courses with a focus on data literacy can lead to success (Batko and Siezak, 2022).

It is crucial to update technological software often to maintain a high data quality level. Investment in adopting a technical system can help overcome errors in information about patients. In 2020, the American medical sphere saved $122 billion annually due to automating administrative transactions (CAQH Explorations, February 2021). In addition to these savings, there is also a possibility to save $16 billion by moving to complete electronic transactions (CAQH Explorations, February 2021).

Results of Low Data Quality in Healthcare

Low-quality data quality may lead to consequences that negatively affect the healthcare organization. For example, it may not be easy to edit data once recorded in an EHR (Australian government, 2019). Hence, the chances of mistreatment and misdiagnosis rapidly increase, and a patient's level of satisfaction with a particular healthcare organization can decrease.

Cases exist where manual interventions happen in fulfilling data, increasing the chances of errors. Such manual interventions are ineffective and require time (Bossen et al., 2019). Systematic approaches can help this situation by decreasing the number of inconsistent operations. Furthermore, poor decision-making can result from low data quality (Jabali et al., 2022). Doctors and nurses can become reliant on EHRs with errors in data records, provoking mistreatment and misdiagnosis.

Understandably, healthcare providers are frustrated when they must manually correct data, resulting in decreasing trust in technology systems. Employees can complete reports manually, leading to mistakes and inconsistencies. In case of low health data quality, chances of conducting clinical studies decrease because investigating diseases, symptoms, and treatment using inaccurate information from healthcare organizations is difficult (Kriegova et al., 2021). This situation can negatively influence the whole health state system. Furthermore, there will be no chances to conduct preventive measures among the population.

Data Storage Options for Healthcare Organizations

Successful data management is complex without health data storage. However, cloud and on-premises decisions can help with this task. Storage area networks (SAN) are storage devices that different users can access with the help of fiber optics channel pipelines. In addition, SANs allow the transfer of data from storage to workstations without current hospital servers (Honavar, 2020). Accordingly, there may be more timely decision-making, increased data exchange, increased patient checking, and improved hospital workflow.

A network-attached storage system (NAS) is a separate storage device that gives information access to all devices on the same network. With the help of this system, hospitals can get information from centralized disk capacity. The benefits of NAS are high speed of information exchange, robust security, and an easy installment process. There are also external storage devices used by 62% of mid and low-level hospitals (Honavar, 2020). These devices include disks, tapes, and SSDs, but may be susceptible to hacking.

Storing information on-premises storage facilities can be costly. Transferring to servers governed by other hospitals off-site can decrease expenses. A recent study found that 93% of hospitals view outsourcing as a solution to data centers and IT infrastructure (Sieck et al., 2022). However, avoiding high expenses is not the main reason healthcare organizations use outsourced storage devices. In addition, facilitated by off-site data storage, there are HIPAA, Health Information Technology for Economic and Clinical Health (HITECH), and Payment Card Industry Data Security Standard (PCI DSS) that hospitals use to protect patients' information (Upadhyay & Hu, 2022). The government accepts these standards, development organizations, and other healthcare professionals to promote comprehensive healthcare information. Continuous improvement of these standards is essential. However, hospitals should use a single shared information management system to make providing healthcare services less challenging. All healthcare providers must be aware of a unique interpretation system of codes and symbols used in health data systems, making it less difficult to identify diseases and make prescriptions (Schopf et al., 2019).

Hospitals use cloud data storage to help decrease expenses on supporting on-premises physical data storage. Moreover, there are opportunities to buy more space for storage. As a result, healthcare cloud storage might grow to $40 billion by 2026 (Alami et al., 2022). Cloud storages have different benefits, including high speed, easy access, adjustment to state regulations on information security, quick recovery, automated operations, and accessible widening opportunities. However, transferring all data to cloud storage can be challenging and make healthcare organizations rethink infrastructure and data storage.

Another data storage system worthy of consideration is a vendor-neutral archive (VNA) that houses images from medical devices and scanners. It provides an opportunity to exchange these images between various stores, and healthcare providers can view them on multiple screens or devices (Jabali et al., 2022).

CONCLUSION

This paper investigated the state of health data and its management. One of healthcare's primary purposes is to provide high-quality services and establish trust between patients and the providers of healthcare services. To achieve this aim, healthcare providers must have a comprehensive picture of a patient. Healthcare data provide information about patients' previous diseases and treatments. Moreover, different healthcare organizations have collected information about patients that, when aggregated, allows for conducting clinical research that can benefit the public health sector.

Different systems provide storage of health information such as SAN, NAS, on-premises, and cloud data. One widely used system for storing health data is EHRs, which are famous for including a broad perspective on the patient's care. For example, EHRs have data about administrative and clinical issues, demographics, treatment, symptoms, immunization, and radiology information. Moreover, EHRs belong to the healthcare IT sphere, providing access to resources that healthcare providers can use in conducting healthcare services.

The evidence presented in this paper identified advantages of health data, including interoperability among healthcare providers, access to patient data, the possibility to conduct medical studies, and ensuring a proper decision-making process. In addition, health data quality improves because of such metrics as precision, validity, reliability, cohesiveness, availability, completeness, identifiability, security, usability, and uniqueness. Moreover, accurate health data help save money for healthcare organizations and organize working processes effectively. However, healthcare providers face challenges while using EHRs. These problems relate to the confidentiality of data, records integration, managing data sets, and supporting electronic records software.

Beyond these challenges, the delivery of healthcare services is ineffective when information about patients is not of high quality. If healthcare data are from different sources, chances exist that some of them can be of poor quality. Poor data quality can negatively affect trust between a patient and a particular healthcare organization and ultimately damage a hospital's reputation. In addition, poor-quality data can lead to errors in conducting studies on diseases, symptoms, or treatment.

To get benefits from healthcare data, they must be accurate, complete, and reliable. The achievement of these benefits results from the different strategies, such as supporting IT infrastructure used at a healthcare organization. Remembering that the market for electronic health records is rapidly developing, updating current technology to reflect these changes is important. The introduction of systematic data quality management at hospitals is one of the most crucial tools that can help improve data quality. To enhance data quality, checking the accuracy of healthcare records is important when the working period is non-busy. All collected health data must be verified, and its origin evaluated. One more way of verification can be data audits that aim to identify challenges in a healthcare organization. After completing audits, identifying the main reasons for the poor data quality is necessary, and establishing procedures to avoid these errors in the future is essential. Additional research is required to test the validity of information in the EHRs empirically.

REFERENCES

Alami, J., Hammonds, C., Hensien, E., Khraibani, J., Borowitz, S., Hellems, M., & Riggs, S.L. (2022). Usability challenges with electronic health records (EHRs) during prerounding on pediatric inpatients. *JAMIA Open, 5*(1), 1-19.

Australian Government. (2019). *Management of data and information in research.* Sydney: National Health and Medical Research Council.

Batko, K., & Siezak, A. (2022). The use of big data analytics in healthcare. *Journal of Big Data, 9*(1), 3-5.

Bossen, C., Pine, K.H., Cabitza, F., Ellingsen, G., & Piras, E.M. (2019). Data work in healthcare: An introduction. *Health Informatics Journal, 25*(3), 465-474.

CAQH Explorations (February 2021). Closing the gap: The industry continues to improve, but opportunities for automation remain. *2020 CAQH Index.* *https://www.caqh.org/sites/default/files/explorations/index/2020-caqh-index.pdf*

Chen, P.T., Lin, C.L., & Wu, W.N. (2020). Big data management in healthcare: Adoption challenges and implications. *International Journal of Information Management, 53*(1), 1-15.

Danao, M., Main, K. Epic EMT review (2023): features, pricing, & more. *Forbes Advisor.*

Dash, S., Shakyawar, S.K., Sharma, M., & Kaushik, S. (2019). Big data in healthcare: Management, analysis and future prospects. *Journal of Big Data, 6*(54), 1-19.

Fanelli, S., Pratici, L., Salvatore, F.P., Donelli, C.C., & Zangrandi, A. (2022). Big data analysis for decision-making processes: Challenges and opportunities for the management of healthcare organizations. *Management Research Review, 5*(1), 1-13.

Honavar, S.G. (2020). Electronic medical records - the good, the bad and the ugly. *Indian Journal of Ophthalmology, 68*(3), 417-418.

Hoover, R. (2017). Benefits of using an electronic health record. *Nursing Critical Care, 12*(1), 9-10.

Ismail, L., Materwala, H., Karduck, A.P., & Adem, A. (2020). Requirements of health data management systems for biomedical care and research: Scoping review. *Journal of Medical Internet Research, 22*(7), 1-15.

Iyamu, T., & Nunu, K. (2021). Healthcare data management conceptual framework for service delivery. *Education and Information Technologies, 26*(1), 3513-3527.

Jabali, A.K., Waris, A., Khan, D.I., Ahmed, S., & Hourani, R.J. (2022). Electronic health records: Three decades of bibliometric research productivity analysis and some insights. *Informatics in Medicine Unlocked, 29*(1), 4-16.

Johnson, R.J. (2021). A comprehensive research study literature review of EPIC in terms of enabling healthcare agility: A report card. *Journal of Medical Informatics and Decision Making, 1*(4), 26-55.

Khairat, S., Zalla, L., Gartland, A., & Seashore, C. (2021). Association between proficiency and efficiency in electronic health records among pediatricians at a major academic health system. *Frontiers in Digital Health, 6*(1), 15-44.

Kohn-Parrott, K., Panella, R. (2017). A learning guide to social impact assessment for community health. *Academy Health* (published as part of the Community Health Peer Learning Program.

Kriegova, E., Kudelka, M., Radvansky, M., & Gallo, J. (2021). A theoretical model of health management using data-driven decision-making: The future of precision medicine and health. *Journal of Translational Medicine, 19*(1), 68-75.

Rudin, R.S., Friedberg, M.W., Shekelle, P., & Shah, N. (2020). Getting value from electronic health records: Research needed to improve practice. *Annals of Internet Medicine, 13*(1), 1-12.

Schopf, T.R., Nedrebo, B., Hufthammer, K.O., Daphu, I.K., & Larum, H. (2019). How well is the electronic health record supporting the clinical tasks of hospital physicians? A survey of physicians at three Norwegian hospitals. *BMC Health Services Research, 19*(1), 25-45.

Senthilkumar, S.A., Rai, B., & Gunasekaran, A. (2018). Big data in healthcare management: A review of literature. *American Journal of Theoretical and Applied Business, 4*(2), 1-15.

Sieck, C.J., Henriksen, B., Scott, S., Kurien, N., & Rastetter, M. (2022). Training to improve patient-centered electronic health record (EHR) use. *Journal of Hospital Management and Health Policy, 6*(1), 1-18.

Tapuria, A., Porat, T., Kalra, D., Dsouza, G., Xiaohui, S., & Curcin, V. (2021). Impact of patient access to their electronic health record: Systematic review. *Informatics for Health and Social Care, 46*(2), 194-206.

Upadhyay, S., & Hu, H.F. (2022). A qualitative analysis of the impact of electronic health records (EHRs) on healthcare quality and safety: Clinicians' lived experiences. *Health Services Insights, 15*(1), 1-15.

Uslu, A., & Stausberg, J. (2021). Value of the electronic medical record for hospital care: Update from the literature. *Journal of Medical Internet Research, 23*(12), 26-32.

Wang, X., Williams, C., Liu, Z.H., & Croghan, J. (2019). Big data management challenges in health research - a literature review. *Briefings in Bioinformatics, 20*(1), 156-167.

Westphal, E., & Seitz, H. (2021). Digital and decentralized management of patient data in healthcare using blockchain implementations. *Frontiers in Blockchain, 5*(1), 35-56.

Zhang, J., Symons, J., Agapow, P., Teo, J.T., Paxton, C.A., Abdi, J., Mattie, H., Davie, C., Torres, A.Z., Folarin, A., Sood, H., Celi, L.A., Halamka, J., Eapen, S., & Budhdeo, S. (2022). Best practices in the real-world data life cycle. *PLOS Digital Health, 1*(1), 37-45.