Nnamdi Azikiwe University, Awka Journal of Private and Property Law Volume 2(1) April, 2025

NAVIGATING ISSUES AND CHALLENGES IN HEALTHCARE DATA MANAGEMENT SYSTEM

Ikechukwu Chime, Chidiebere Okoroafor & Akachi Nworgu-Ikojo[*]

Abstract

Healthcare data management systems have evolved significantly over the past century, with their value, velocity, and volume increasing and becoming more digitized. This evolution encompasses everything from Electronic Health Records (EHR) to Electronic Medical Records (EMR) obtained from doctor visits to handwritten medical notes scanned into digital repositories. Collecting data from diverse healthcare sources, along with storage, protection, and analysis to ensure accuracy, accessibility, consistency, and timeliness in healthcare, is essential for improving patient care and operational efficiency while maintaining security and data privacy. In this context, technological advancement and big data management awareness have led to breakthroughs in the medical field from paper to computer, the Internet of Things (IoT), data analysis, etc. This article aims to contribute to the efforts at understanding the management systems of healthcare data. It adopts the analytical approach of doctrinal research methodology and commences with the contextualization of the importance of big data in healthcare and its collection approach. It further considers the mechanisms for advancing the reliability of healthcare records while amplifying the integration of cutting-edge technologies as the future direction of healthcare data gathering. The challenges involved in healthcare data gathering were equally considered, with privacy and security, data discrimination, and job displacement being highlighted as potential risks. The article further analyzes the ethical concerns in healthcare data gathering. It concludes with a summary of findings and recommendations, including the establishment of precise policies and measures for secure and conscientious patients' data management, stakeholders' collaboration, continuous education on cyber security, etc.

Keywords: Big Data, Data Analysis, Data Protection, Electronic Health Records, Healthcare

1. Introduction

Comprehensive and accurate medical records play an important role in the advancement of healthcare, providing an extensive view of patients' health history and aiding improved care and outcomes. Over time, health information management has evolved from managing paper records to handling electronic data across various entities, promoting collaborative healthcare practices. From the earliest medical records documented by ancient Greeks[1] to the modern era of Electronic Health Records (EHRs) and data analytics, the healthcare sector has seen substantial technological advancements and standardization efforts. The future of health information management therefore emphasizes data integration, harmonization, and accessibility across expanding health systems, reinforcing the need for efficient data-sharing and informed decision-making to drive improved patient outcomes and operational processes.

As healthcare data is the lifeblood of modern day healthcare systems, "Big data" in healthcare refers to the vast volumes of information generated by the adoption of digital technologies for collecting patient records, improving patient care, and overseeing hospital operations.[2] According to

---

[*] **Ikechukwu Chime,** PhD, Faculty of Law, University of Nigeria, Enugu Campus. ike.chime@unn.edu.ng

[*] **Chidiebere Okoroafor Esq**, Senior Associate, Lex Telios LLP, chidiebere.okoroafo@yahoo.com

[*] **Akachi Nworgu-Ikojo,** Faculty of Law, University of Nigeria, Enugu Campus. akachi.nwogu-ikoo@unn.edu.ng

[1] Jacek Lorkowski, and Mieczysław Pokorski, 'Medical Records: A Historical Narrative' <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9599146/#:~:text=The%20book%2C%20a%20compilation%20of,including%20physicians'%20recommendations%2C%20descriptions%20> accessed 11 March 2025.

[2] VinBrain, 'Top 10 Big Data Applications in Healthcare – Part 1' (VinBtrain, 13 March 2024) <https://vinbrain.net/top-10-big-data-application-in-healthcare-part-1#:~:text=by%20business%20entrepreneurs.-,What%20Is%20Big%20Data%20In%20Healthcare?,treat%20diseases%2C%20and%20reduce%20costs.> accessed 10 March 2025.

Mckinsey, it entails datasets whose size is beyond the ability of typical database software tools to capture, store, manage, and analyze.[3] These datasets cannot feasibly be stored manually as they are often too immense and complex for traditional or manual forms of healthcare data management. Driven by advancement in healthcare to suit the pace at which the world is developing, the emergence of big data is propelling medical research, healthcare policies, shaping clinical decisions, and contributing to improving patient outcomes especially in the future.

2. Importance of Big Data in Healthcare

The use of paper-based medical records, despite its historical impact, is admittedly inefficient and insecure. With approximately 421 million hospitalizations in the world annually,[4] the disadvantages of relying on paper for documenting patient data are apparent. Accordingly, challenges in accessing and sharing information, inefficiencies in retrieval, and security concerns are propelling a shift towards EHRs which offer improved accessibility, efficiency, cost-effectiveness, and security in healthcare data management, effectively tackling the shortcomings in traditional paper-based methods. Despite the privacy concerns linked to the shift to EHRs, it signifies a significant step forward in modern healthcare, paving the way for efficient information management and improved patient care. Ensuring smooth data exchange among various medical systems, an advantage of EHRs, remains essential for ensuring effective interoperability. Although developing countries like Nigeria are still a long way from catching up, EHRs are the most common sources of big data in healthcare. The Big Data revolution in healthcare is well underway with its potential to completely transform healthcare in the future.[5] Defined as large volumes of high velocity, complex, and variable data that require advanced techniques and technologies to enable the capture, storage, distribution, management and analysis of the information,[6] Big Data has a tremendous impact on how healthcare is provided and received. Each carefully documented and analyzed interaction with patient produces a data point that can advance healthcare administration for the future; directing public policy, improving patient outcomes while reducing healthcare costs, furthering preventive medicine, in addition to countless other benefits for public health and healthcare administration.

3. Approaches to Health Care Data Gathering

Approaches to healthcare data gathering refers to the methodologies and techniques employed in collecting, storing, analyzing healthcare related data, which includes medical history, patient information, vital signs, insurance claims, consent forms, hospital length of stay, hospital deaths,[7] progress notes - for use in advancing medical research, healthcare policies and administration, etc. Effective collection or gathering of healthcare data is essential for the enhancement and optimization of healthcare delivery and this approach ranges from traditional to non-traditional data collection methods.

3.1 Traditional Approach to Healthcare Data Gathering

Effective data gathering is instrumental in understanding population health trends, evaluating treatment

---

[3] J Manyika and others, ‘Big Data: The Next Frontier for Innovation, Competition, and Productivity’ <https://www.mckinsey.com/~/media/mckinsey/business%20functions/mckinsey%20digital/our%20insights/big%20data%20the%20next%20frontier%20for%20innovation/mgi_big_data_exec_summary.pdf> accessed 10 March 2025.

[4] World Health Organisation, ‘Patient Safety’ (2019) <https://www.who.int/news-room/facts-in-pictures/detail/patient-safety> accessed 11 March 2025.

[5] K Liddell, D Simon, and A Lucassen, 'Patient Data Ownership: Who Owns Your Health?' [2021] 8 *Journal of Law and the Biosciences.*

[6] W Raghupathi and V Raghupathi, ‘IHTT: Transforming Health Care through Big Data Strategies for Leveraging Big Data in the Healthcare Industry’ <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4341817/#CR6> accessed 11 March 2025.

[7] M S and others, ‘Data Collection Methods in Health Services Research: Hospital Length of Stay and Discharge Destination’ <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4377563/> accessed 13 March 2025.

outcomes, and identifying areas for improvement within healthcare systems. At its core, innovation in healthcare involves finding more efficient ways to improve patient outcomes, and eliminate systemic inefficiencies. While traditional means remain an important source of patient information, the process is slow, arduous, and subject to responder bias. Traditional methods of healthcare data gathering, which involve paper-based systems, manual data collection, and human transcription, have been fundamental for most of the breakthroughs in the field of medicine till date.

Traditional methods include the use of questionnaires, surveys, and observational studies which are also used by researchers and healthcare organizations to collect specific health-related data from individuals and gain valuable insights into health risks, disease prevalence in demographics or communities, etc.[8] Healthcare surveys cover a wide range of topics; quality of care, access to care, healthcare provider performance ratings, healthcare costs, healthcare outcomes, etc. While these methods may have been the lifeblood of healthcare data, they have been characterized by so many shortcomings among which is the fact the process is time consuming, susceptible to human errors, and occasion inefficiencies in data management due to its manual collection and storage procedure.

Particularly, manual records pose a difficulty in accessibility and data sharing among healthcare providers, especially in emergencies. This poses a threat to the optimization of healthcare delivery. Paper-based records are prone to loss, damage, and unauthorized access and the poor medical data processing systems are the most significant causes of medical errors which also serve as the most authoritative argument for electronic data management methods, despite privacy and security concerns.

3.2 Non-Traditional Approaches to Healthcare Data Gathering

As technology advances, more healthcare institutions are adopting automated data collection systems to simplify procedures and enhance data accuracy. Thus, non-traditional means of healthcare data gathering refers to data and information obtained from sources outside the scope of the methods discussed above, usually with an infusion of technology. These sources span from electronic health records,[9] to activity patterns accumulated by smart devices and e-health apps, and aggregated metadata from social media and the internet. While reliance on these technological means of healthcare data gathering is still remotely behind in many countries and areas, the world is fast-adjusting to more accurate, and easy-to-obtain sources of healthcare data. The major non-traditional approaches to healthcare data gathering are Electronic Health Records and Electronic Medical Records (EMRs) which are digital and electronic records of patient information; medical diagnosis, history, treatment plans, medications, immunization dates, allergies, laboratory and test results, etc. They are typically maintained by healthcare providers, and are easily accessible to authorized personnel and stakeholders.

Another critical non-traditional approach to obtaining healthcare data is through Health Information Exchanges (HIE) which enable the transfer of patient information among healthcare systems and organizations. Through this method, healthcare providers can access data and information across different institutions. The extra advantage that the HIE offers is that the material/information is not limited to location or institution.

Furthermore, patient portals serve as very instrumental means of obtaining healthcare data. Patients are able to access and update their information, thereby contributing to data gathering with self-reported data. These portals are consequently a source of big healthcare data, contributing to the

---

[8] A Kastsiukovich, 'Why Data Collection in Healthcare is a Crucial Point Now' (Cleverdev Software, 14 December 2023)<https://www.cleverdevsoftware.com/blog/data-collection-in-healthcare#:~:text=Methods%20for%20healthcare%20data%20collection,%2C%20treatment%2C%20and%20administration.%22> accessed 15 March 2025.

[9] Mary Dellosa, 'Healthcare Data Collection: Transforming Patient Care and Efficiency' (Help Squad, 2 May 2024) <https://helpsquad.com/healthcare-data-collection/#:~:text=Electronic%20Health%20Records%20(EHRs)%20have,safety%20and%20outcome%20of%20health.> accessed 16 March 2025.

collection of comprehensive patient information for research, analysis, and ultimately the improvement of healthcare delivery. With the proliferation of wearable health devices (Fitness trackers, smart watches, etc.) and health-related mobile apps, individuals are able to continuously collect and monitor their health data easily, including heart rate, activity levels, sleep patterns, and other relevant health indicators. These devices provide real-time patient data, which facilitates preventive and personalized healthcare interventions.

Additionally, healthcare registries, population health data analysis, and insurance claims containing valuable healthcare information and costs associated with patient care also serve as healthcare data gathering sources. These data are often used for billing and research. Advances in genetics have also led to the collection of genomic data and genetic sequences which can be used to understand disease susceptibility, develop personalized treatment plans, and conduct genetic research. Meanwhile, data on air quality, water quality, and other environmental factors are equally very relevant as environmental factors can have a significant impact on health.

4. Data Integrity and Quality Control in Healthcare Data Gathering

Data integrity refers to the reliability of healthcare records and data. Upholding data integrity means that the information handled by the healthcare organization remains accurate, reliable, and consistent. As important as data integrity is for safeguarding patients and the organization, it is also important for ensuring precision in life sciences and bio-technological advancement and breakthroughs. Although certain aspects of data collection, such as physical examination results like height and weight or laboratory findings such as white blood cell counts, are relatively straightforward, the process of abstraction typically requires different degrees of assessment and interpretation.

Managing data integrity to ensure information accuracy and prevent its alteration remains an ongoing challenge within the contemporary healthcare industry.[10] This is because inaccurate data poses significant risks to patient well-being and places a heavy burden on healthcare practitioners, leading to issues such as fraud, malpractice, substandard care, and data breaches.[11] Such precarious situations significantly complicate the management of healthcare data. Extracting information from raw data, such as hospital charts, can be a challenging and time-consuming task, with lots of possibilities for mistakes, particularly when dealing with free text, and requiring human interpretation. The Interpretation would involve deciphering unclear handwriting, translating obscure acronyms and abbreviations, and understanding clinical context to accurately extract the necessary information. More so, the personnel saddled with the responsibility for such interpretation could be anyone; either a trained medical records analyst, or other medical professional, or even an untrained or inexperienced research assistant.

The above possibilities that may compromise data integrity and quality reemphasizes the need to establish clear guidelines, documentation processes, and standardized coding forms for data analysts and reviewers in order to maintain accuracy and consistency within registries. These guidelines typically outline specific data types to be included in the registry, providing instructions for coding using standardized codes or registry-specific codes. However, discrepancies may arise between different data abstracts due to the intrinsic inconsistencies and ‘noise’ in the raw hard-copy documents and the inherent imprecision of natural language, all of which are potential means of introducing errors into the registry.

To mitigate this risk, thorough training on the registry protocols, procedures, data sources, and definitions is essential, particularly for non-clinician data abstracts.[12] Training methods may include live meetings, online sessions, and interactive multimedia tutorials, often supplemented with test abstractions using sample charts. Standardized test charts can be used to assess and ensure that abstracts meet predefined performance standards. In evaluating a registry's efficacy for decision-making, understanding

---

[10] M I Zarour and others, ‘Ensuring Data Integrity of Healthcare Information in the Era of Digital Health’ [2021] 8(1) *Healthcare Technology Letters 66-77.*

[11] Ibid.

[12] Kirsten J Parker and others, ‘The Science of Clinical Quality Registries’ [2023] 22(2) *European Journal of Cardiovascular Nursing 220-225.*

the data acquisition procedures and the stored data quality is paramount. As registries serve as key evidence for decision-making processes, ensuring data quality is crucial for assessing their appropriateness.

Registry planners should thus, devise quality assurance plans to ensure data reliability and report on these activities. These quality assurance tasks typically fall within three primary classifications: (1) ensuring good quality of quality, (2) overseeing registry procedures, and (3) monitoring computerized systems. Notably, quality assurance methods vary based on the registry's intended purpose, with those serving for decision-making requiring more stringent measures compared to those documenting disease progression. While quality assurance can support healthcare data reliability, budgetary constraints can limit large registries' ability to implement comprehensive quality assurance protocols.[13] To strike a balance, a risk-based approach to quality assurance is recommended, focusing on the critical sources of error or procedural gaps, according to the registry's specific purpose. Standardizing quality assurance methods to align with specific goals may become more prevalent to enable result comparisons across various registries.

Importantly, the extent of data cleansing in registries is dependent on various factors, including the registry's objectives, the type and sources of data, and the registry's duration. These factors often influence the frequency and types of data queries, which can be generated both electronically and manually. Trained personnel can review these queries to identify error trends and provide additional site training where necessary. Audits, which may be conducted onsite or remotely, can also be used to verify the accuracy of the data entered into the registry, focusing on specific data fields in accordance with the registry's objectives. The audit approach should reflect the most significant sources of error specific to the registry's purpose and may involve various sampling techniques and periodic evaluations, depending on the registry's duration. Upon the conclusion of such audit, communication of findings and necessary corrective actions to the site personnel will also ensure compliance and data accuracy.

Therefore, structures, processes, policies, and procedures need to be put in place to ascertain the quality of the data in the registry and to ensure against several types of errors, including error in data entry, interpretation, or transfer so as to improve the reliability of healthcare records and data.

4.1 Future Directions in Healthcare Data Gathering

Spurred by technological advancements and the appreciation of the promises of data in patient care and public health, the healthcare data landscape is rapidly evolving, and promises to shape the future of healthcare.

One of the most interesting features of technology in healthcare data gathering is the feasibility of real-time data streams through wearable devices that enable continuous monitoring of patient health and timely intervention.[14] As this technology evolves, remote monitoring from smart watches, fitness trackers, sensor-embedded medical devices etc. will play a more prominent role in chronic disease management[15] and healthcare generally.

Again, Big Data has birthed interest into the field of "computational embryology" in revolutionizing fertility and reproduction. While In Vitro Fertilization (IVF) is highly dependent on expertise, and maybe chance, a machine that generates fresh insights into developmental milestones, identifying characteristics that are imperceptible to the naked eye, incorporating data from various sources, and yielding a result that surpasses the accuracy of any human embryologist is essential. This marks the era of personalized treatment through computational embryology, allowing machines to perform tasks beyond human capabilities.

---

[13] US National Library of Medicine, 'Data Collection and Quality Assurance' <https://www.ncbi.nlm.nih.gov/books/NBK208601/> accessed 14 March 2025.

[14] Dellosa (n 10).

[15] K Waddill, 'How Remote Patient Monitoring is Changing Disease Management' <https://healthcareexecintelligence.healthitanalytics.com/news/how-remote-patient-monitoring-is-changing-chronic-disease-management> accessed 15 March 2025.

Big data is also poised to play a transformative role in drug discovery.[16] Artificial Intelligence (AI) algorithms built with big data have the capacity to expedite the identification of potential drug targets by analyzing vast amounts of biological data, enabling researchers to pinpoint specific proteins or genes associated with diseases. These AI algorithms can also efficiently screen large databases of chemical compounds to identify those with the potential to act as effective drugs, reducing the time and expenses involved in traditional methods. According to the World Economic Forum,[17] from the advancement of drugs and vaccines to the enhancement of medical diagnostics and therapy, AI technologies are utilized at every step of the value chain, enhancing efficiencies throughout the entire healthcare system, with a large impact on future global health outcomes.

Overall, the future of healthcare data gathering is characterized by an interconnected ecosystem of technologies, methodologies, and data sources. As the healthcare industry embraces digital transformation, the integration of cutting-edge technologies and data-driven approaches will play a pivotal role in improving patient outcomes, optimizing healthcare delivery, and fostering evidence-based decision-making.

5. Data Risks and Challenges

Big data is a tool for the future, but it is not without its challenges. Three of the biggest risks that big data poses are in the areas of Data Privacy, Data Security, and Data Discrimination.[18] Perhaps an understated risk posed by big data is the possibility that doctors will be replaced to some extent, by technology.[19]

Data Privacy and Security

Big Healthcare data encompasses very sensitive bits of information; fingerprint records, DNA samples, insurance and financial information, personal contact information of patients and family with biographic details, etc,[20] and hackers, for instance, could find this beneficial for fraud and commerce. According to IBM,[21] The global average cost of a data breach in 2023 was USD 4.45 million, a 15% increase over 3 years. The growth in the value of data is continuous. As there has been an increase in reliance on cloud computing, as well as telehealth, especially since COVID-19,[22] Cyber-attacks on hospitals and the healthcare industry at large have recorded an all-time high. The incidence with Magellan health is a firm case study of where patient information was exposed and stolen in a phishing attack.[23] Through this attack, almost 300,000 patients were affected and the leaked information included sensitive data like social security numbers, emails and email attachments

---

[16] Kamila Dymalla, 'Revolutionizing Drug Development Through Artificial Intelligence, and Machine Learning' <https://www.pharmacytimes.com/view/revolutionizing-drug-development-through-artificial-intelligence-machine-learning> accessed 15 March 2025.

[17] S Yoon and A Amadiegwu, 'Emerging Tech, Like AI, is Poised to Make Healthcare More Accurate, Accessible and Sustainable' <https://www.weforum.org/agenda/2023/06/emerging-tech-like-ai-are-poised-to-make-healthcare-more-accurate-accessible-and-sustainable/#:~:text=Emerging%20technologies%20like%20AI%20and,precise%20diagnoses%20and%20timely%20interventions.> accessed 15 March 2025.

[18] G Alexandru, 'Big Data in Healthcare - Opportunities and Challenges' <https://www.researchgate.net/publication/326263025_Big_Data_in_Healthcare_Opportunities_and_Challenges#fullTextFileContent> accessed 16 March 2025.

[19] Sai Balasubramanian, 'Can Doctors Really be Replaced by Technology' <https://www.forbes.com/sites/saibala/2021/09/22/can-doctors-truly-be-replaced-by-technology/> Accessed 16 March 2025.

[20] S Abhisheck, 'Healthcare Data Security; A Brief Overview of Challenges and Safety Measures' <https://cynoteck.com/blog-post/healthcare-data-security/> accessed 16 March 2025.

[21] IBM Security, 'Cost of Data Breach Report' < https://www.ibm.com/reports/data-breach> accessed 16 March 2025.

[22] Rena Christina Tabatha, 'The Future Challenges of Big Data in Healthcare' <https://www.forbes.com/sites/forbestechcouncil/2021/06/18/the-future-challenges-of-big-data-in-healthcare/?sh=526e8bc346b2> accessed 16 March 2025.

[23] *Dearing v Magellan Health Inc. et al*. 28 U.S.C. § 1332.

with protected information. This reinforced the duty of hospitals and healthcare service providers alike to implement cyber security measures to protect patient data. So, while healthcare data and analysis provide all the advantages for public health enumerated and discussed above, it also calls for thorough protection from 'anti-bodies'.

Privacy and security are items of concern in healthcare data management due to the sensitive nature of the information involved.[24] Perhaps, the most significant advantages of healthcare data management is the advancement of preventive medicine; tracking of public health risks before they occur, optimization of healthcare services generally, policy making, and innovation in medicine. Healthcare service providers and professionals consequently have a legal[25] and ethical duty to ensure the protection of this information and adopt robust privacy guidelines, with legal consequences attached for breaches.[26] Protecting healthcare data includes adopting safe measures in the training of employees, and carefully selecting the software employed as well the method through which patients are informed about their medical data.[27] Since at the root of big data in healthcare is patients' willingness to release sensitive information into the system, there is an overwhelming need to ensure compliance with safety measures and thorough practices in handling these data in the future.

More so, staff education on cyber security can be considered as one of the strategies that healthcare service providers may adopt for protection of patient data.[28] This is so because continuing education is the backbone of most successful organizations. Staff education on cyber security can therefore support the establishment and implementation of controls on data usage, data encryption risk assessment, concerns over data security, while addressing language barriers and lack of skills in data collection and processing.

5.2 Data Discrimination

Data discrimination refers to biases or prejudices in the collection and management of data which may result in underrepresentation of certain demographics, and disparities in healthcare delivery outcomes. This sort of bias affects Artificial Intelligence, and they are so significant because they may not only reflect inequalities, but may exacerbate them.[29] Bias in data collection can occur at many points in the creation of an algorithm, and consequences have trickle-down effects on the desired use of the data collected in the first place.

The undesired impact of Algorithm bias is increasingly recognized in healthcare data management, because algorithms inherit bias from data from which they are trained. If the algorithms are not regularly evaluated for inclusivity and fairness, they can perpetuate discriminatory practices in areas such as patient diagnosis, treatment plans, and resource allocation. A study conducted in 2019 detected that a clinical algorithm used to determine the patients who are in most need of urgent care showed racial bias, as a result of defective data used in training the algorithm.[30] In view of the above limitation, companies dealing with healthcare data are taking proactive steps to promote diversity and inclusion in data collection systems. Incomplete data could also indirectly result in bias, specifically when specific groups or demographics are under-represented at the data collection stage. Globally, the legal framework for healthcare data management is unfortunately largely silent on data

---

[24] P Shojaei, Elena Vlahu-Gjorgievska and Yand-Wai Chow, 'Security and Privacy of Technologies in Health Information Systems: A Systematic Literature Review' [2024] 13(2) *Computers* 1-25.

[25] National Health Act (NHA) 2014, Ss 26-27; The HIV and AIDS (Anti-Discrimination) Act, 2014, S 14.

[26] Dellosa (n 10).

[27] Abigail Sims, 'What is Data Privacy in Healthcare' <https://www.tonic.ai/blog/what-is-data-privacy-in-healthcare-everything-you-need-to-know> accessed 17 March 2025.

[28] T Morgan, 'The Data Challenges Faced by the Healthcare Industry' <https://thejournalofmhealth.com/the-data-security-challenges-faced-by-the-healthcare-industry/> accessed 17 March 2025.

[29] K Igoe, 'Algorithimic Bias in Healthcare Exacerbated Social Inequities - How to Prevent It' https://www.hsph.harvard.edu/ecpe/how-to-prevent-algorithmic-bias-in-health-care/> accessed 17 March 2025.

[30] Zaid Oberyemer and others, 'Dissecting Racial Bias in an Algorithm Used to Manage the Health of Populations' <https://www.science.org/doi/10.1126/science.aax2342> accessed 18 March 2025.

bias. Perhaps, one of the only few regulatory instruments addressing this is the Food and Drug Administration (FDA)[31] which requires that healthcare device manufacturers test for any forms of bias, before marketing to the public. However, there is no system for the enforcement of this requirement.

Notably, in many welfare states, Algorithms determine who gets housing, financial aid, etc.[32] thus, if Algorithms are empowered to make healthcare decisions and significantly impact healthcare administration, AI in healthcare must promote equality, not exacerbate bias.

5.3 Replacement of Healthcare Professionals by Technology

At present, big data is not so advanced as to function independent of human control. Thus, technology in health and medicine is integrated with human-centered and humanly executed care. There is, however, a valid fear in the prediction that development in big data could replace medical personnel to the extent that it could pose a threat to the advantages of human execution of ethically-delivered healthcare solutions, raise concerns about patient data security, among some other apprehensions.

Workforce displacement is a primary cause of concern. For instance, it has been recorded that there is a much higher number of monthly visitors on health websites, than consultations of medical doctors.[33] Accordingly, as prices fall and technologies improve, devices and apps may completely replace physical visits to healthcare practitioners in the near future.[34]

As the management and administration of healthcare is increasingly replaced by the integration of AI and Automation, some resolution can be found in the fact that Big Data and technology calls for some new skills for healthcare professionals to adapt with the changes and maintain relevance in the future.

6. Privacy and Ethical Considerations in Healthcare Data Gathering

Healthcare data and informatics is increasingly relevant in the advancement of healthcare operations. However, privacy might be regarded as a heavy price to pay for these benefits, with the potential risk of patients losing confidence in the healthcare system and the goal of healthcare optimization suffering as a consequence. This is particularly so when the underlying effects of capitalism is weighed, as the profit-driven objective of the private-owned enterprises pose challenges to patient data and amplifies ethical concerns about the exploitation of data for commercial purposes.[35] Therefore, with big data, millions of patients are at risk, and the altruistic intentions behind the sourcing of these data will certainly be diverted.

Arguments about the Ethical considerations for healthcare data revolve around patient privacy being the cost of digitization. Specifically, a study reports that patients are willing to share their personal information to health workers, but they are increasingly concerned about their data being used for commercial purposes.[36] Healthcare data management must therefore, find a healthy balance between upholding patient privacy and autonomy, and advancing the advantages that pursuing big data promises.

---

[31] A Hammond, Bhav Jain, and Fatima Stanford, 'An Extension to the FDA Approval Process is Needed to Achieve AI Equity' <https://www.nature.com/articles/s42256-023-00614-8> accessed 18 March 2025.

[32] C Grant, 'Algorithms are Making Decisions about Healthcare, Which May Only Worsen Medical Racism' <https://www.aclu.org/news/privacy-technology/algorithms-in-health-care-may-worsen-medical-racism> accessed 18 March 2025.

[33] R Susskind and D Susskind, 'Technology Will Replace Many Doctors, Lawyers, and Other Professionals' <https://hbr.org/2016/10/robots-will-replace-doctors-lawyers-and-other-professionals> accessed 18 March 2025.

[34] M Musiol, 'Why Leveraging Non-Traditional Health Sources Will Improve Health Outcomes' <https://www.marsdd.com/news/leveraging-non-traditional-data-sources-will-improve-health-outcomes/> accessed 19 March 2025.

[35] V Chiruvella and A Guddati, 'Ethical Issues in Patient Data Ownership' <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8178732/> accessed 19 March 2025.

[36] E Bell, L Ohno-Machado, and M Grando, 'Sharing My Health Data: A Survey of Data Sharing Preferences of Healthy Individuals' [2014] *AMIA Annual Symposium Proceedings 1699–708.*

6.1 Ethical Dilemmas in Healthcare Data Gathering

One central dilemma in healthcare data management is the handling of the tension between the ethical need to respect patient data privacy, and the requirement of patient data for research and healthcare advancement. Related to this is the obtaining of consent for the use of the said data, and how consent forms should be designed to ensure that the patients understand how their data will be used, including the potential risks and benefits of sharing their data.

6.2 Data Ownership in Healthcare

The questions of whether patient data is a property, and who owns such data is a recurring ethical issue. With the growth in data economy, there is a protracted ethical issue regarding data ownership among stakeholders; patients, hospitals/healthcare providers, and data custodians. These concerns are driven by issues regarding how organizations handle the data, issues surrounding the data as a digital asset,[37] and lawsuit.[38] In a bid to promote trust and transparency in the healthcare system and empower patients to make informed decisions about their health, there is a shift towards patient-centered care. However, while patients may consequently have more control over their data, they lack the means to exercise independent control over it. At this juncture, a distinction must be made between Data ownership and Data custodianship. Patients are the subjects of healthcare data, and are therefore, the rightful data owners while other stakeholders such as healthcare providers, private health insurance providers, clinical software operators and developers, etc, are merely custodians of data. As custodians, they are regulated by data protection laws which set the limits to which acquired patient data may be used or transferred.[39]

While the above distinction might appear to be very straightforward, data ownership in healthcare is complicated by practical issues; it remains a complex and evolving issue in healthcare data management despite ongoing legal, ethical, and technological considerations. As healthcare systems continue to evolve digitally, it is important that these issues of patient's rights in relation to data, and custodian obligations evolve to serve an increasingly data-driven healthcare management landscape.

6.3 Patient Rights, Consent and Autonomy

Patient rights are a subset of human rights.[40] Also rooted in the principle of respect for individual's autonomy is patient consent. Incidental to this right is the right to make informed decisions about how their data is used and dispersed. Accordingly, the main purpose of the consent process in obtaining healthcare data is to ensure that patients have an understanding of the purpose and risks, which allows them to make a well-informed decision, even though the voluminous and 'endless' use of healthcare data, and the denseness of consent forms makes it difficult for patients to fully appreciate the usage of the data, and the risks involved. Despite this major shortcoming, the consent process is the best means for upholding the ethical principle of autonomy.[41]

---

[37] J Ritter and A Mayer, 'Regulating Data as Property: A New Construct for Moving Forward' [2017] 16(1) *Duke Law & Technology Review 220-277,* 221–223, 226–227.

[38] D Wakabayashi, 'Google and the University of Chicago Are Sued Over Data Sharing' *The New York Times* (USA, 26 June 2019) <https://www.nytimes.com/2019/06/26/technology/google-university-chicago-data-sharing-lawsuit.html> accessed 19 March 2025.

[39] Australian Medical Association, 'Data Governance and Patient Privacy in Healthcare' <https://www.ama.com.au/sites/default/files/2023-01/Data%20Governance%20Position%20Statement%20-%20FINAL.pdf> accessed 19 March 2025.

[40] Jacob Olejarczyk and Michael Young, 'Patient Rights and Ethics' <https://www.ncbi.nlm.nih.gov/books/NBK538279/> accessed 19 March 2025.

[41] FAbujarad and others, 'Building an Informed Consent Tool Starting With the Patient: The Patient-Centered Virtual Multimedia Interactive Informed Consent' (2018) <https://pubmed.ncbi.nlm.nih.gov/29854101/> Last accessed 19 March 2025.

In modern medicine, on the one hand, patients' rights include the right to informed consent,[42] confidentiality, timely access to detailed and accurate medical records, right to decline care, and the one most relevant to this part of this paper, the right to decline or consent to participate in medical research, experimental procedures, or clinical trials.[43] On the other hand, the ethical obligation to healthcare data custodians to ensure that patients are informed about how their data is collected, stored, shared, as well as the potential consequences of sharing their data might be impossible to fulfill, even for the researcher who uses the data directly, to determine what the findings could be. It then poses a problem in situations where the use of a patient's data is antithetical to his moral standing.[44] Illustratively, a study used recorded images of over 70,000 people to train an analytics model to predict sexual orientation based on facial features.[45] It is reasonable to presume that a significant number of individuals are unaware that their photos could provide this kind of information, and many would not consent to the use of their posted pictures in in forming such an analytical model.

Measures could therefore, be implemented to automatically alert individuals when their information is utilized in research, along with the study's name included. Although many individuals might disregard these notifications or choose to opt out entirely, at least it would be the participants' choice to make.

6.4 Data Monetization

Data is the currency of the digital realm. Data monetization refers to sale of patient data by healthcare institutions to third parties. While selling this data might help fund research, it raises concerns for the privacy and ownership of patient data because there must be a balance between economic value, and respecting privacy rights. Data monetization poses risks for the ethical use of healthcare data, especially because it leaves room for data to be used outside of the intent for which it was given and collected in the first place. As a matter of fact, monetization highlights every other ethical concern in the use of healthcare data and once this act is done, the question of who owns the data becomes even more complex, and raises concerns around consent, safeguarding user information, avoiding discrimination, etc.[46]Ultimately, consent, transparency, and safeguarding user information should be a priority in obtaining and using healthcare data and must be at the fore of data monetization.

6.5 Equitable Use of Healthcare Data

Healthcare disparity is among one of the biggest issues in healthcare administration worldwide. Where big data is mined from overrepresented demographics, predominantly obtained from a single group based on factors like gender, age, ethnicity, socio-economic class, or race, the conclusions reflect their characteristic, and it poses a challenge to the ethical principle of equality. In addition to inequality and discrimination, this poses a risk for feasibility and where these findings are applied to other groups, the results often do not translate.[47]

In medicine, inaccurate data, diagnosis, and conclusions can lead to public harm; people might be harmed where irrelevant findings are applied to underrepresented groups. Equally, a particular treatment plan tailored to a demographic could produce adverse effects in another. A human genetics

---

[42] *Canterbury v Spence* 464 F.2d. 77s, 782 D.C; Cobb v. Grant 8 Cal.3d 229; *Wilkinson v Vesey* 295 A.2d 676.

[43] Z Hashim, 'Patients' Bill of Rights: Making Health a Human Right in Nigeria' *Premium Times* (Nigeria, 14 April 2019) <https://www.premiumtimesng.com/news/headlines/325291-patients-bill-of-rights-making-health-a-human-right-in-nigeria.html?tztc=1> accessed 17 March 2025.

[44] M Froomkin, 'Big Data: Destroyer of Informed Consent' [2019] 18(3) *Yale Journal of Health Policy, Law, and Ethics 122.*

[45] Y Wang and M Kosinski, 'Deep Neural Networks are More Accurate Than Humans at Detecting Sexual Orientation From Facial Images' [2018] 114(2) *Journal of Personality and Social Psychology 246–257.*

[46]The Cable, 'The Ethics of Data Monetisation: Balancing Personal Privacy with Economic Value' <https://www.thecable.ng/the-ethics-of-data-monetisation-balancing-personal-privacy-with-economic-value> accessed 20 March 2025.

[47] F Wang, L Casalino, and D Khullar. 'Deep Learning in Medicine: Promise, Progress, and Challenges' [2019] 179(3) *JAMA Intern Med.* 293–294.

project[48] found that certain genetics studies employed date from mostly Euro-Americans from middle to upper socio-economic classes. The resultant findings from projects like these are homogenous and will hardly be applicable to other sects, widening the gap in medical findings and medical practice.

Even where Big Data is wholly representative, findings can be inequitable if input data is biased. An algorithm built on biased data produces biased conclusions.[49] Where such algorithm is trained with biased inputs, it perpetuates the bias with which it is fed.[50]Another area of concern for equity in big data is compensation as Organizations conducting research on specific populations may neglect to distribute the benefits of their research with the population involved in their study and may also inadequately compensate their participants.

6.6 Data Retention

Data retention pertains to the responsible and appropriate management and maintenance of data over time. It involves long term storage of patient information which includes sensitive health data. Ethical issues therefore arise when healthcare organizations or entities do not follow the recommended protocols for data retention, potentially resulting in patient privacy breaches, vulnerabilities in data security, and the improper or unauthorized use of confidential information. Undeniably, consent expires. The retention of patient information afterwards without such understanding from the patient is an infringement upon patient autonomy; the right to use and control their health information. Meanwhile, the prolonged retention of such data may pose a threat to patient privacy and expose sensitive information to unauthorized parties and may also result in data breaches. Dissatisfactory data security measures may also lead to unauthorized access, compromising patient confidentiality and privacy.

To address the aforementioned ethical issues linked with healthcare data retention, organizations ought to establish precise policies and guidelines for the secure and conscientious management of patient data. This includes implementing strong data security measures, adhering to data retention regulations, securing informed consent from patients for data retention purposes, and routinely assessing and removing obsolete or irrelevant patient information. By prioritizing patient privacy, data security, and adherence to regulations, healthcare organizations can uphold ethical norms in their data retention procedures and preserve patient trust and confidentiality.

7. Conclusion

The healthcare data management systems have evolved from paper charts to technology-based systems, and its landscape is significantly being transformed by the emergence of big data, which enhances innovation across various sectors, including EHRS and predictive analytics. This transformation reemphasizes the need for the management systems to satisfy the requirements of medical record data, real-time access, patient participation, data sharing, data security, and patient identity privacy. The future of healthcare data gathering is poised for significant advancements, driven by the integration of innovative technologies and a deeper understanding of the potential of data to enhance patient care. Accordingly, as real-time monitoring through wearable devices becomes more prevalent, and fields like computational embryology evolve, the landscape of personalized medicine will continue to expand. However, while big data is revolutionizing healthcare, its promise is accompanied by critical challenges, particularly concerning data privacy, security, and potential biases in data management. In this regard, the rapid acquisition and analysis of large healthcare data sets calls for more emphasis on patient data safety and security because data is sensitive and prone

---

[48] C Lewis and others, 'The Human Microbiome Project: Lessons from Human Genomics' (2011) <https://www.researchgate.net/publication/51824720_The_Human_Microbiome_Project_Lessons_from_human_genomics> accessed 20 March 2025.

[49] E Howe and F Elenberg, 'Ethical Challenges Posed by Big Data' (2020) <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7819582/> accessed 20 March 2025.

[50] J Angwin and others, 'Machine Bias' (2016) <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing> accessed 20 March 2025.

to malware, data theft, viruses, hacker strategies, etc. Addressing these challenges is essential to ensure that the benefits of big data are realized without compromising patients' rights and autonomy or exacerbating existing disparities in healthcare. As the healthcare industry navigates these complexities, a balanced approach that prioritizes ethical considerations, data protection, and equitable use of information is crucial. Accordingly, by entrenching a culture of transparency and accountability in healthcare data management, the healthcare sector can leverage big data to improve outcomes while safeguarding the rights and privacy of patients.

Above all, as data privacy has become paramount in today's world, there is also a need to implement measures and establish precise policies and guidelines for the secure and conscientious management of patient data. This may include the automatic notification to data subjects (i.e. the patients) when their information is utilized by data custodians e.g. for research purposes, with the study's name included. The integration of transparency and accountability culture in healthcare data management also requires collaboration among stakeholders, continuous education on cyber security, and a commitment to ethical standards, paving the way for a more effective and inclusive healthcare system.