


**THE IMPACT OF ARTIFICIAL INTELLIGENCE ON HEALTHCARE: PRIVACY
CHALLENGES AND CYBERSECURITY RISKS** <https://doi.org/10.63330/aurumpub.009-002>**Mirian Esquárcio Jabur¹****ABSTRACT**

This article examines the advancements and challenges of digital transformation in healthcare, focusing on the adoption of artificial intelligence (AI) and its demonstrated benefits—such as reductions in diagnostic time and medical errors. It also addresses critical issues surrounding information security, protection of sensitive data, and ethics in the use of emerging technologies. Real-world cases of cyber incidents are presented, the relevance of regulations such as Brazil's General Data Protection Law (LGPD) is discussed, and guidelines are proposed to ensure responsible innovation that balances technological progress with the safeguarding of patients' fundamental rights.

Keywords: Digital health; Artificial Intelligence; Information security; Data privacy; AI ethics; LGPD.

¹ Highest Academic Degree: Master's in Technological Education
Academic Institution: CEFET/MG
E-mail: mirianjabur@gmail.com



INTRODUCTION

Digital transformation has driven profound and irreversible changes in healthcare, propelled by emerging technologies such as artificial intelligence (AI), the Internet of Medical Things (IoMT), and big data analytics. These innovations offer significant potential to enhance care quality and efficiency, evidenced by up to a 44% reduction in diagnostic time in radiology (Esteva et al., 2017) and up to an 80% decrease in medical errors, particularly in critical processes like medication administration and misdiagnosis (Rajpurkar et al., 2018).

However, alongside these advances arise complex challenges related to information security and the privacy of patients' sensitive data. Incidents like the leakage of 1.5 million records from the SingHealth system in 2018 highlight the vulnerability of healthcare systems to cyber threats and underscore the urgent need for rigorous data protection practices (Singapore Ministry of Health, 2018). This alarming scenario reinforces the urgent need for robust governance to guarantee security and accountability in the use of these technologies.

In light of these challenges, it is essential to develop clear and effective guidelines that balance technological progress with the protection of fundamental rights, ensuring that innovation in healthcare remains ethical, secure, and patient-centered.

To contribute to this debate, this article is structured into six thematic axes: technological advances in the sector; the vulnerability of health data; the regulatory and legal framework; the main types of cyber threats; protection strategies and best practices; and, finally, ethical dilemmas associated with the use of AI in medicine. The conclusion reinforces the importance of ensuring that innovation is always guided by the principles of security, ethics, and respect for human dignity.

Technology evolves rapidly, bringing new challenges and opportunities. To understand these transformations, this study adopts a qualitative methodology.

This research is characterized by a qualitative, exploratory, and retrospective approach, grounded in a narrative literature review and document analysis. Secondary data were gathered from scientific articles, institutional reports, current legislation, and case studies.

METHODOLOGY

This study adopts a qualitative, exploratory, and retrospective approach, grounded in a narrative literature review and document analysis. Secondary data were obtained from scientific articles, institutional reports, current legislation, and case studies.

Sources were selected based on their recency (publications from 2018 onward), thematic relevance, and diversity of origin—public, private, and international. Practical case studies were chosen



for their representativeness in diagnostics, hospital management, prevention, and patient care, prioritizing nationally recognized institutions with publicly available results.

A regional diversity and varying levels of digital maturity were also sought to provide a comprehensive overview. Case information was validated through triangulation with reliable sources such as official documents, news reports, and scientific studies, ensuring consistency and credibility.

With this methodology, we explore the application of AI in healthcare, highlighting both advances and challenges.

COMMON MYTHS ABOUT AI: DEMYSTIFYING THE TECHNOLOGY

Many healthcare professionals may find the idea of using AI intimidating. Let us clarify some common myths and show how this technology is here to assist rather than replace.

1) Will AI replace doctors?

No. AI is a support tool. It can assist in analyzing tests or suggesting diagnoses, but the final decision always rests with the healthcare professional. This distinction is crucial: technology is not meant to substitute physicians but to expand their diagnostic and decision-making capacities.

2) Is AI too difficult to integrate into daily practice?

Although AI technology is advanced, user interfaces are designed to be intuitive. Many AI systems can integrate directly into existing workflows, such as electronic health records, facilitating adoption without requiring professionals to be technology experts.

3) Can I trust AI-generated diagnoses?

AI reliability depends on the data on which it was trained. AI systems are designed to complement human work, offering data-driven support under professional supervision. Thus, human judgment remains indispensable for ensuring safety and ethics in patient care.

Despite its potential, unregulated AI use can lead to overdiagnosis and unnecessary treatments, causing both costs and anxiety. Therefore, technical advances are impressive, but success hinges on responsible implementation that respects fundamental ethical principles of healthcare.

AI ADVANCES AND APPLICATIONS IN HEALTHCARE

AI is transforming healthcare with measurable results across various domains:

- 1) **Imaging Diagnosis:** Google's DeepMind Health system achieved 94% accuracy in detecting eye diseases, outperforming human specialists (De Fauw et al., 2018). In Brazil, Hospital Israelita Albert Einstein uses AI for CT scan analysis, reducing diagnostic time from 30 minutes to 5 minutes.



- 2) **Personalized Medicine:** The Brazilian startup Mendelics developed a genomic analysis system that reduced rare disease diagnosis time from years to weeks at 60% lower cost.
- 3) **Hospital Management:** Hospital das Clínicas of São Paulo implemented predictive algorithms that cut readmission rates by 40% (HC-FMUSP, 2021).
- 4) **Drug Research:** BenevolentAI identified potential COVID-19 treatments in 48 hours—processes that typically take months (Richardson et al., 2021).

These examples underscore AI's transformative potential but also demand robust technological infrastructure, data protection measures, and ethical governance.

Despite technological progress, safeguarding sensitive data remains fundamental to ensure patient privacy and security.

However, for these advances to be sustainable, it is essential to ensure the rigorous protection of sensitive data.

After exploring the promising applications of AI, it is necessary to understand the challenges and risks that accompany this technology.

Despite the encouraging progress and concrete applications that are transforming healthcare, it is crucial to understand the limitations and risks that come with this technology. Only by recognizing these limitations can we ensure that innovation unfolds in a safe, ethical, and effective manner.

CHALLENGES AND RISKS OF AI IN HEALTHCARE

While AI's benefits are clear, protecting sensitive data is vital. Patient health information is personal and requires extra care when collected, stored, or used.

PRIVACY AND REGULATION

In Brazil, the LGPD mandates that health data receive maximum protection and may only be used in specific, legally justified situations.

Both the LGPD (Law No. 13.709/2018) and the European GDPR require explicit patient consent as the primary legal basis for processing health data.

Thus, the data subject (the owner of the data) must clearly and specifically agree to the use of their data, being informed of the purpose for which it will be used. This consent must be freely given, presented in a clear and prominent manner, using accessible and objective language.

HOW SHOULD THIS DATA BE PROTECTED?

In addition to consent, institutions (such as hospitals, clinics, or researchers) must adopt additional safeguards to ensure that health data is used securely and responsibly. Some of these include:



- 1) **Pseudonymization:** In scientific research, it is mandatory to use techniques that “depersonalize” the data. This means processing data in a way that does not directly identify the individual. This is outlined in Article 89 of the GDPR..
- 2) **Retention Period:** Data may only be stored for as long as necessary to fulfill its intended purpose. After that, it must be securely deleted. The LGPD addresses this in Article 15.
- 3) **Restricted Access:** Access to information must be limited to authorized individuals, such as physicians or healthcare professionals directly involved in patient care. The Federal Council of Medicine (Resolution CFM 2.217/2018) mandates that all access must be logged.

Health data is not like any other type of data. It pertains to the most intimate aspects of a person’s life. Therefore, both the LGPD and other international regulations impose strict rules to protect this information. Whenever possible, it is essential that the patient remains in control through clear and informed consent.

With the methodology established, we now analyze the incidence and impact of cyberattacks in the healthcare sector.

Given the legal and ethical requirements for data protection, it is essential to understand the cyber threats that expose the healthcare sector to vulnerabilities.

RISKS OF USING AI IN ADMINISTRATIVE AND MEDICAL ACTIVITIES

In addition to general risks related to information security and the privacy of sensitive data, the use of artificial intelligence (AI) in healthcare presents specific challenges in both administrative and clinical practice. These risks require heightened attention to avoid operational, ethical, and legal consequences.

1) **Inadequate Process Automation**

Automating administrative tasks through bots and intelligent systems may fail to account for all variables and nuances of human activities. This can result in errors or incorrect decisions. For example, scheduling processes, document handling, or automated communication may produce distortions that compromise efficiency and generate rework, affecting hospital workflow and the patient experience.

2) **Lack of Transparency**

Many AI systems operate as “black boxes,” delivering results without clearly explaining the reasoning behind their decisions. This opacity hinders traceability and the justification of automated decisions, leading to a loss of trust among staff and patients, and complicating compliance with regulations that require auditability and transparency.

3) **Excessive Dependence on Automated System**



High dependence on automated systems can make organizations vulnerable to technical failures or cyberattacks. In cases of instability, the inability to perform tasks manually—such as financial control, human resources management, or patient support—can lead to the interruption of essential operations, jeopardizing the continuity of healthcare services.

4) Exposure to Cyber Threats

AI bots and automated systems are potential targets for cyberattacks, including phishing, malicious code injection, and other vulnerabilities exploited by hackers. Compromising these systems can result in unauthorized access, leakage of sensitive data, and disruption of administrative services, further exacerbating information security risks in hospital environments.

5) Errors in Data Interpretation

AI relies on the accurate interpretation of input data. Errors in reading or analyzing this information can lead to flawed administrative decisions, affecting organizational performance and strategic planning. In the medical field, incorrect interpretations may result in misdiagnoses, compromising patient treatment.

6) Data Bias

AI models learn from historical data, which, if tainted by social, economic, or cultural biases, can replicate and amplify these distortions. This may lead to unfair or discriminatory decisions, such as improper prioritization in administrative processes or inaccurate diagnoses for minority groups, undermining equity and ethics in healthcare.

7) Data Quality

Data quality is crucial for AI performance. Outdated, incomplete, or incorrect data impairs the accuracy of analyses, leading to distorted results. This can negatively impact hospital operations and the quality of medical care.

8) Incorrect Pattern Interpretation

AI may identify erroneous correlations or patterns in datasets, resulting in flawed interpretations. This can lead to the implementation of ineffective strategies, resource waste, and risks to patient safety.

9) Risk of Overfitting

Machine learning models that are overtrained on specific datasets may exhibit poor generalization, failing when applied to new data or scenarios. This reduces the effectiveness and reliability of AI systems in healthcare.

AI applied to imaging diagnostics often uses convolutional neural networks (CNNs), which are capable of identifying complex visual patterns.



In light of these complex challenges, it becomes imperative to establish a structured governance framework grounded in ethics and responsibility to mitigate risks and ensure the protection of patients and professionals involved. Among these challenges, ethical risks deserve special attention, as lack of transparency and bias can compromise fairness and equity in care.

AI GOVERNANCE IN HEALTHCARE: STRUCTURE, BENEFITS, AND CHALLENGES

Implementing a governance framework for artificial intelligence (AI) in healthcare is imperative in light of the ethical, technical, and operational challenges posed by this technology. More than a legal requirement, governance is a strategic necessity to ensure that AI systems are used safely, fairly, transparently, and in alignment with the interests of patients, healthcare professionals, and institutions.

GOVERNANCE STRUCTURE: CORE PILLARS AND FUNCTIONS

For governance to be effective, it must be based on interdependent pillars that work synergistically to ensure responsible innovation and risk mitigation:

1) Multidisciplinary Committees: Shared and Accountable Decision-Making

The creation of multidisciplinary committees is the first step toward ensuring ethical and technical governance. These committees should include professionals from the medical, technological, legal, and bioethical fields, as well as patient representatives, ensuring a broad and balanced perspective.

These groups act as guardians of ethics and safety in AI use, conducting prior evaluations of algorithms to identify potential biases, technical flaws, and negative impacts on clinical practice. They also continuously monitor system implementation, correcting course and adjusting processes as needed to preserve human-centered care and service quality. This oversight prevents erroneous or unjust automated decisions that could compromise patient health.

2) Periodic Audits: Continuous and Preventive Oversight

The mere existence of committees is not sufficient to guarantee safety and accountability. Regular and systematic audits—preferably quarterly—are essential to review the performance of AI systems, especially in critical environments such as medical report generation or clinical decision support.

These audits aim to detect ethical deviations, technical failures, biases, and potential threats to information security. They also assess system compliance with current legislation, including the LGPD and applicable international standards. Furthermore, they ensure that automated



decisions are explainable and uphold patients' rights to transparency—fundamental to autonomy and trust.

3) Transparency and Human Oversight: Ensuring Control and Trust

Transparency is a non-negotiable condition for the acceptance and success of AI in healthcare. Physicians, patients, and other stakeholders must receive clear and accessible information about how systems operate, the criteria used, and the nature of automated decisions. Comprehensible reports and effective communication foster an environment of trust and collaboration

Moreover, human oversight must be an integral part of the process. Healthcare professionals must be able to review, challenge, and, when necessary, override automated decisions that may endanger patient safety. Patients, in turn, should have channels to question AI-generated diagnoses or treatments and request second opinions when desired. This balance ensures that AI serves as a support tool—not an exclusive arbiter.

4) Benefits of Well-Structured Governance

When these pillars operate in an integrated and effective manner, the benefits are evident. Research indicates that the implementation of solid governance can reduce litigation involving automated diagnoses by up to 72%, while patient trust increases by approximately 40% when they are provided with clear and transparent information about the use of AI. Moreover, institutions that adopt rigorous governance practices protect themselves against fines and sanctions established by legislation such as the LGPD, thereby reducing legal risks and ensuring greater legal security. Governance thus becomes an instrument for institutional value, care quality, and economic sustainability.

5) Challenges and Risks of the Absence of Governance

Conversely, neglecting governance exposes healthcare organizations to serious consequences. AI systems without proper oversight can generate errors, biased and unjust decisions, and failures that compromise patient safety and data integrity. These failures result in financial losses, reputational damage, and a loss of trust from both professionals and patients. Professional distrust is a critical factor. A 2023 survey by CREMESP revealed that 68% of Brazilian physicians do not trust AI systems that are not subject to audits or ethical oversight. This skepticism can hinder the implementation and effective use of AI, limiting its transformative potential.

6) Clearly Defined Roles and Responsibilities: The Foundation of Governance

For governance to move beyond an abstract concept and become an effective practice, it is essential that responsibilities are clearly defined. Developers, technical teams, healthcare



professionals, managers, and regulatory bodies must work in coordination, each with specific roles in supervision, auditing, training, and communication.

This network of responsibilities creates a safe and trustworthy environment in which AI becomes an ally of human care—always accompanied by clinical and ethical judgment from professionals.

The article then presents a Responsibility Matrix, which clearly outlines the essential roles in AI governance in healthcare, ensuring that each stakeholder understands their duties for the ethical, safe, and effective use of the technology.

Responsibility Matrix and Challenges of Cyber Threats in Digital Health						
Activities / Roles	Developers	Multidisciplinary Committee	Healthcare Professionals	Hospital Managers	Regulatory Bodies	Patients
AI system development	R	C	I	I	I	I
Ethical and legal evaluation of the system	I	R	C	C	C	I
Clinical validation and testing	C	R	R	I	I	I
Continuous system monitoring	C	R	C	I	I	I
Periodic audits	I	R	C	C	A	I
Training of healthcare professionals	I	C	R	A	I	I
Communication and transparency with patients	I	C	C	R	I	A
Review and contestation of AI decisions	I	C	R	I	I	A
Compliance with LGPD and legal standards	C	R	I	A	A	I
Definition of internal AI policies	I	R	C	A	C	I

Legend:

- R (Responsible): Executes the activity A (Approver): Final authority that approves the activity
- C (Consulted): Provides input and is consulted I (Informed): Is kept informed about progress or results

Defining these responsibilities is fundamental to ensuring that AI governance in healthcare functions effectively, encompassing all technical, ethical, and legal dimensions involved.



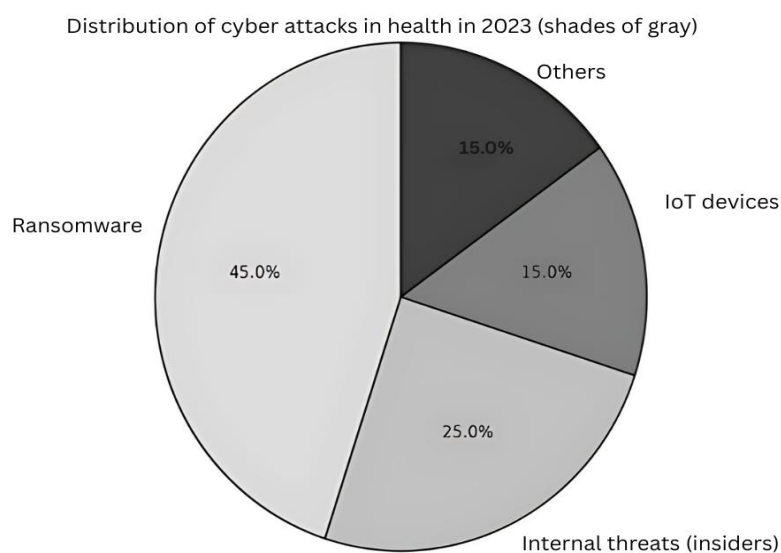
However, governance does not operate in isolation. It is a key component in addressing cyber threats, which represent one of the greatest challenges facing the healthcare sector. These attacks can compromise sensitive data, disrupt clinical service operations, and endanger patient safety.

Therefore, it is essential that technical teams, clinical staff, managers, and regulatory bodies act in a coordinated manner, with clearly defined responsibilities, to prevent, detect, and respond swiftly to digital threats

In this context, the responsibility matrix becomes a crucial tool for strengthening cybersecurity in healthcare, fostering a resilient, ethical, and secure digital environment in which artificial intelligence can truly serve as an ally to human-centered care.

CYBER THREATS

In parallel with the need for ethical governance, the healthcare sector faces a growing threat in the realm of digital security. In 2023, approximately 34% of cyberattacks targeted hospitals, clinics, and healthcare providers, highlighting their high vulnerability (CHECK POINT, 2023). Among the most common threats, ransomware leads with 45% of incidents, followed by insider threats and attacks on connected medical devices. These data reveal not only the frequency of threats but also their significant financial impact, which will be detailed below.

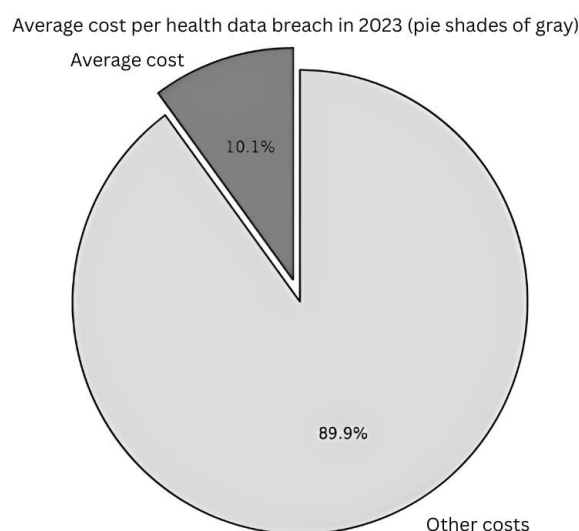


Source: CHECK POINT, 2023. Adapted by the author.

These figures not only demonstrate the frequency and variety of threats but also underscore their substantial financial consequences. In 2023, the average cost per data breach in the healthcare sector was estimated at USD 10.1 million, showing that such attacks compromise not only data security but also



cause considerable economic damage. This scenario reinforces the need for investments in governance and effective protection measures.



Source: CHECK POINT, 2023. Adapted by the author.

SPECIFIC RISKS FOR AI IN ADMINISTRATIVE AND MEDICAL ACTIVITIES

In addition to the general risks of security and privacy already discussed, the adoption of artificial intelligence introduces specific challenges that demand heightened attention in both administrative and clinical contexts.

1) Inadequate Automation Can Lead to Operational Errors

Many administrative processes are designed for human execution, where individuals can perceive nuances and exceptions. Bots and automated systems may fail to handle such variations, resulting in errors such as duplicate appointments, document processing failures, or incorrect automated responses. These mistakes can directly impact operational efficiency, generate rework, and even compromise the patient experience.

2) Lack of Transparency Hinders Auditing and Trust

IAI often functions as a “black box”: the system delivers a result without clearly explaining how it arrived at that conclusion. This opacity creates barriers for auditors, managers, and even healthcare professionals to understand and validate AI-driven decisions. Without transparency, distrust grows, and the risk of misuse or undetected errors increases.

3) Excessive Dependence May Compromise Operational Continuity

Total reliance on automated systems creates a single point of failure. If the AI system becomes unstable, is targeted by a cyberattack, or suffers a technical breakdown, the entire administrative or clinical routine may come to a halt, with staff unprepared to take over



manually. This vulnerability threatens service continuity, which in healthcare is critical, as delays can cause irreparable harm.

4) **Data Bias Can Reinforce Inequities**

Algorithms learn from historical data. If these data contain social, cultural, or economic biases, AI will replicate and potentially amplify these distortions. This can lead to unfair decisions, such as prioritizing certain groups in administrative processes or misdiagnosing minority patients, undermining equity in care and the ethical integrity of the system.

5) **Diagnostic Errors Due to Misinterpretation**

In the evaluation of reports and exams, AI may misinterpret images, signals, or clinical data—especially when the data are incomplete or of poor quality. These errors can result in incorrect diagnoses, leading to inappropriate treatments, delays, or harm to the patient's health..

6) **Lack of Human Oversight Can Endanger Lives**

Complete automation without proper medical supervision eliminates the critical layer of verification that ensures patient safety. Automated decisions, when not reviewed by qualified professionals, may fail to detect subtle signs or important alerts, increasing the risk of adverse events and endangering lives..

7) **Poorly Defined Legal and Ethical Issues**

Responsibility for AI errors and failures remains a gray area. It is unclear who is legally accountable when an AI system makes a mistake—the physician, the institution, or the software developer. This lack of clarity creates legal uncertainty for healthcare professionals and opens the door to ethical conflicts, which may reduce patient trust and hinder the safe adoption of the technology.

CAUSES OF VULNERABILITY

To understand why the healthcare sector is so susceptible to these attacks, it is essential to analyze the main factors that increase its exposure.

This vulnerability is not accidental. It stems from three primary factors:

- 1) **High Value of Medical Data:** Health information is extremely valuable. A complete medical record can be worth up to USD 1,000 on the black market (FBI, 2022). For cybercriminals, such data holds high value for illicit activities such as fraud, extortion, and illegal data trading.
- 2) **Outdated Technology:** Many hospitals still operate with legacy systems that no longer receive security updates. It is estimated that 68% of these institutions use obsolete software (HIPAA Journal, 2023), creating easy entry points for attackers.



- 3) **Pressure to Maintain Service Continuity:** Unlike other sectors, hospitals cannot afford downtime. This urgency makes healthcare institutions up to three times more likely to pay ransoms quickly in cases of data hijacking (ransomware) (Sophos, 2023).

MOST COMMON TYPES OF ATTACKS:

On average, organizations worldwide experienced 1,248 attacks per week in 2023 (Check Point).

In the healthcare sector, the most common types include:

- 1) **Ransomware:** In 2021, such an attack on Brazil's public health system (SUS) caused the shutdown of over 600 healthcare units for several weeks.
- 2) **Insider Threats:** In 2022, a notable case involved a nurse caught selling confidential data of celebrities.
- 3) **Connected Medical Devices (IoT):** Equipment such as pacemakers and insulin pumps can be hacked, raising serious safety concerns (FDA, 2023).

Given these factors, it is urgent to implement effective strategies that mitigate risks and strengthen security.

Given the risks outlined, it becomes essential to implement robust protection and governance strategies.

In addition to technical measures, digital transformation introduces significant ethical dilemmas that warrant in-depth analysis.

THE IMPORTANCE OF PREVENTION AND THREE EFFECTIVE STRATEGIES

Rather than reacting after an attack, the most strategic and secure approach is to act before the problem occurs. This is where the adoption of governance for security and privacy becomes essential—not as a cost, but as a necessary component for the safe continuity of services.

Adopting governance means establishing clear processes, defined responsibilities, and a culture of data security and protection that involves all sectors of the institution. It enables the anticipation and mitigation of risks, reduction of vulnerabilities, and rapid response to any signs of threat.

In addition, some practical and technological solutions have proven particularly effective:

- 1) **Encryption:** A kind of “digital lock” that protects data even if intercepted. This ensures that patient information remains secure.
- 2) **Microsegmentation:** This involves digitally dividing the hospital into isolated zones. If one system (e.g., the reception desk) is attacked, other critical areas—such as digital ICUs—remain protected.



- 3) **Artificial Intelligence Applied to Cybersecurity:** Intelligent systems can detect suspicious behavior in real time, such as unusual access or abnormal activity, and act before damage occurs (MITRE, 2023).
- 4) **Training and Awareness:** Despite all available technology, the human factor remains one of the main entry points for attacks. Therefore, it is essential to invest continuously in training all professionals involved—from administrative staff to medical and management teams.

In addition to technical and financial aspects, the adoption of artificial intelligence in healthcare brings with it complex ethical dilemmas.

ETHICAL DILEMMAS IN EMERGING TECHNOLOGIES

Ensuring ethical standards in digital health means guaranteeing that the use of technology is safe, fair, transparent, and, above all, human-centered and balanced. With the rapid advancement of technologies in healthcare—such as artificial intelligence (AI), predictive algorithms, and automated systems—new ethical dilemmas have emerged that cannot be ignored. Addressing these issues responsibly is not optional; it is a requirement. When it comes to people's lives and health, acting ethically is what ensures that technology becomes an ally rather than a risk.

Ignoring this debate could jeopardize trust, safety, and even the credibility of institutions. Below are some of the most pressing ethical dilemmas:

1) **Loss of Patient Trust**

Trust is the foundation of any relationship in healthcare. However, 62% of patients still distrust the use of artificial intelligence in diagnostics and treatments (PEW, 2023). If technology is not introduced with transparency and empathy—ensuring complete safety and integrity—it may alienate rather than assist patients.

2) **Algorithmic Bias**

AI systems learn from data. If these data are incomplete or poorly distributed—for example, lacking representation of certain ethnic or social groups—the results may be unfair or inaccurate. One study showed that models with such bias made errors in 34% of diagnoses (Obermeyer, 2023). This means that technology can reinforce inequalities rather than correct them.

3) **Responsibility for the Use of Technology**

If an automated diagnosis fails, who is held accountable? The physician? The hospital? The software developer? This question remains unresolved, and the lack of clarity can create legal and ethical uncertainty for all parties involved, including patients.



These threats demonstrate that governance must not be seen as a bureaucratic obligation, but as a vital strategy to protect patients, professionals, and institutions. Only through integrated security policies, continuous audits, and multidisciplinary collaboration will it be possible to navigate the complex digital risk landscape in healthcare.

CASE STUDY

Theory and guidelines are fundamental, but the true effectiveness of artificial intelligence in healthcare is measured in practice. The technical, ethical, and security challenges only gain real meaning when confronted with the actual experiences of institutions that have adopted these technologies.

This section presents concrete examples of how hospitals and clinics—both in Brazil and internationally—have integrated AI into their routines, addressing the complexities of the digital environment, ensuring the protection of sensitive data, and applying robust governance to mitigate risks.

These cases illustrate not only the tangible benefits of innovation, such as reduced diagnostic times and improved hospital management, but also demonstrate how the conscious adoption of technology can be a strategic differentiator for safety, efficiency, and quality of care.

By analyzing these experiences, it becomes clear that implementing AI in healthcare is not a linear path nor free of challenges. Therefore, understanding the successes and lessons learned from these institutions is essential for those seeking to navigate this rapidly evolving landscape with safety and responsibility.

CASE 01: DRGBRASIL

- **Challenge:** Difficulty in achieving efficient hospital management, preventing complications, and ensuring rational use of resources.
- **Solution:** Implementation of machine learning algorithms to analyze large volumes of clinical and operational data.
- **Result:** Early disease identification, optimization of bed usage and care processes, improved patient safety, and reduced waste.

Read more: <https://www.drgbrasil.com.br/valoremsaude/inteligencia-artificial-na-saude/>

CASE 02: SÍRIO-LIBANÊS HOSPITAL

- **Challenge:** Making medical care faster, safer, and free of bureaucracy.
- **Solution:** Development of the Sofya platform, using AI and voice recognition to automate anamnesis and fill out medical records.



- **Result:** Greater clinical efficiency, reduced errors and time spent on administrative tasks. Strengthened digital governance and ethics through the use of anonymized data.

Read more: <https://www.projetodraft.com/transformacao-digital-e-uma-questao-de-saude-o-sirio-libanes-vem-evoluindo-para-tornar-o-cuidado-medico-mais-agil-e-seguro>

CASE 03: CEU CLINIC

- **Challenge:** Improving accuracy and speed in imaging diagnostics, especially for early-stage diseases.
- **Solution:** Use of AI to analyze X-rays, CT scans, and correlate with other clinical data.
- **Result:** Earlier and more reliable diagnoses, enhanced safety in care, and support for continuous monitoring of chronic patients.

Read more: <https://www.clinicaceu.com.br/blog/ia-qual-seu-papel-no-rastreio-de-doencas/>

CASE 04: UNIMED-BH

- **Challenge:** Speeding up the authorization process for exams and procedures, reducing patient wait times.
- **Solution:** Use of AI with machine learning algorithms to automatically classify and approve requests based on historical patterns.
- **Result:** Instant approval of 30% of requests, with 99.8% accuracy and a reduction in average response time to 10 minutes. Increased user satisfaction.

Read more: <https://www.plano-de-saude-saopaulo.com.br/noticias-planos-de-saude/unimed-bh-agiliza-consultas-e-exames/>

CASE 05: HOSPITAL ALEMÃO OSWALDO CRUZ

- **Challenge:** Continuously monitoring patients with chronic diseases (e.g., diabetes, hypertension, heart failure) to reduce avoidable hospitalizations and improve treatment adherence.
- **Solution:** Development of a predictive AI system to identify clinical decompensation risks based on vital signs, lab tests, and electronic health records. The system alerts healthcare teams for timely remote or in-person interventions.
- **Result:** The integration of clinical engineering, artificial intelligence, and wearables contributes to care efficiency and safety, although specific quantitative results for the predictive system are not deta



Read more: <https://www.hospitaloswaldocruz.org.br/imprensa/releases/engenharia-clinica-inteligencia-artificial-e-wearables-trazem-maior-eficiencia-e-seguranca-no-atendimento-aos-pacientes/>

The adoption of technology does not bring only benefits. In recent years, the healthcare sector in Brazil has faced significant challenges related to information security, with several incidents involving the exploitation of vulnerabilities resulting in the compromise of sensitive data. Below, three recent cases are presented to illustrate these issues:

CASE 01: UNIVERSITY HOSPITAL OF USP (HU-USP)

- **Incident Description:** On March 23, 2024, the University Hospital of the University of São Paulo suffered a cyberattack that paralyzed essential services and affected public healthcare delivery.
- **Impact:** The attack disrupted critical hospital operations, compromised the provision of medical services, and exposed vulnerabilities in the institution's security systems.

Read more: <https://tecnoblog.net/noticias/hospital-da-usp-sofre-ataque-hacker-e-suspende-consultas-e-exames-de-rotina/>

CASE 02: MINISTRY OF HEALTH (BRAZIL)

- **Incident Description:** In November 2022, the illegal sale of administrative databases from government systems was discovered, including CADSUS (the Unified Health System User Registration System) and immunization data from e-SUS Notifica.
- **Impact:** Although it was not possible to confirm whether the illegally sold databases were current or outdated, the incident revealed serious failures in the protection of citizens' sensitive information.

Read more: <https://www.gov.br/saude/pt-br/aceso-a-informacao/lgpd/registro-de-incidentes-com-dados-pessoais>

CASE 03: BARRETOS CANCER HOSPITAL

- **Incident Description:** The Barretos Cancer Hospital was targeted by a cyberattack in which criminals used ransomware to encrypt data and demanded a ransom for its release.
- **Impact:** The attackers demanded USD 300 per affected machine, totaling a potential cost of USD 360,000 (approximately BRL 1.08 million) for the hospital.

Read more: <https://www.ufsm.br/app/uploads/sites/563/2019/09/5.22.pdf>

These cases underscore the importance of continuous investment in cybersecurity within the healthcare sector, aiming to protect information and ensure the continuity of services to the population.



However, for these actions to be effective and systematic, it is necessary to establish clear and comprehensive guidelines to govern the safe and ethical use of artificial intelligence in healthcare.

GUIDING PRINCIPLES AND SECURITY IN AI FOR HEALTHCARE

Artificial intelligence holds the promise of revolutionizing healthcare, but the adoption of this technology faces challenges that go far beyond the development of sophisticated algorithms. The central question is not how many intelligent systems we can create, but whether we can ensure that these systems do no harm—respecting the fundamental Hippocratic principle: *primum non nocere* (first, do no harm).

To navigate this complex landscape, three essential guidelines are proposed: Mandatory adoption of rigorous cybersecurity standards; Creation of ethical and technical guidelines with multidisciplinary participation for continuous auditing; Real transparency with patients regarding the use of their data and automated decisions that affect their treatment.

However, this is where the major practical dilemma lies: are we truly prepared to implement these measures effectively? The painful experience of the University of Vermont Medical Center, which suffered a ransomware attack that led to the cancellation of surgeries for weeks, serves as a harsh warning. Such incidents reveal that digital fragility in healthcare can cost lives—this is no longer a hypothetical scenario, but a brutal reality.

Moreover, the threat extends beyond operational disruption. Imagine if an attack not only halted exams but silently altered dosages in electronic medical records. Or if genetic data were leaked and used to discriminate against families in health and life insurance. These are not science fiction scenarios, but concrete risks that challenge our regulatory, technical, and ethical capacities.

The critical discussion must therefore address how prepared institutions and governments are to face this reality. The cost of ensuring security and ethics is high and involves investment, training, cultural change, and effective legislation. Without these, we risk turning the promise of AI into a source of new harm—widening inequalities and further eroding patient trust.

Ultimately, the true measure of progress in digital health will be our ability to prevent tragedies—not merely the speed at which we deploy technologies. The choice is clear: we can either be pioneers in building an ethical, safe, and responsible future, or passive spectators—and victims—of foreseeable crises.

GLOSSARY

- **AI Ethics:** A set of principles that guide the responsible development and use of artificial intelligence, ensuring respect for human dignity, fairness, and transparency.



- **Algorithmic Transparency:** The ability of AI systems to explain their decisions in a way that is understandable to humans, facilitating auditing and trust.
- **Artificial Intelligence (AI):** A field of computer science dedicated to developing systems capable of performing tasks that typically require human intelligence, such as pattern recognition, decision-making, and learning.
- **Audit (technical, ethical):** A systematic process for evaluating AI systems to verify technical, legal, and ethical compliance, identifying failures, biases, and risks.
- **Automated Decision:** A result generated by AI systems that influences clinical actions without direct human intervention.
- **Automated Diagnosis:** A process in which AI systems assist in analyzing and interpreting clinical data to suggest medical diagnoses.
- **Bias:** A systematic tendency or distortion in data or algorithms that can lead to unfair or inaccurate decisions.
- **Big Data:** A massive set of structured and unstructured data that, when analyzed with specific tools, reveals patterns and trends applicable to health research.
- **Blockchain:** A distributed ledger technology that stores information in encrypted and immutable blocks, used to ensure traceability and security in medical databases.
- **Clinical Validation:** The scientific and clinical confirmation of the effectiveness and safety of AI systems before and during their application in healthcare settings.
- **Convolutional Neural Networks (CNNs):** A type of neural network architecture particularly effective in processing visual data, widely used for medical image analysis and assisted diagnosis.
- **Continuous Audit:** An ongoing process of monitoring and reviewing AI systems to ensure their compliance and performance over time.
- **Cryptography:** An advanced method of data encryption that allows processing while keeping data encrypted, used for secure analysis without exposing sensitive information.
- **Cyber Vulnerability:** A weakness or flaw in information systems that can be exploited by malicious actors to compromise data or services.
- **Data Anonymization:** The process of removing or altering information that allows the direct or indirect identification of an individual, as defined in Article 12 of the LGPD.
- **Data Protection:** A set of policies and practices to ensure the privacy and security of personal information.



- **Federated Learning:** A distributed machine learning technique in which data remains on the original devices, allowing collaborative model training without centralizing sensitive information.
- **GDPR (General Data Protection Regulation):** The European regulation (EU 2016/679) that sets guidelines for the protection of personal data, serving as a basis for laws such as the LGPD.
- **Generative AI:** A subfield of artificial intelligence capable of generating new content (texts, medical images, etc.) from training data, such as GPT models.
- **Governance:** A set of processes, policies, and organizational structures that ensure the responsible, safe, and ethical use of artificial intelligence in healthcare.
- **Governance Framework:** An organizational and regulatory structure that defines policies, processes, and responsibilities for the use of AI.
- **Human Oversight:** The active involvement of professionals in monitoring and reviewing decisions made by automated systems.
- **Information Security:** A set of measures and practices to protect data and systems from unauthorized access, damage, or disruption.
- **Informed Consent:** The right of patients to receive clear and complete information about the use of their data and automated systems that influence their treatment, with the ability to accept or refuse.
- **IoMT (Internet of Medical Things):** A network of connected medical devices (such as wearables and hospital equipment) that collect and transmit data in real time.
- **Legal Security:** Assurance that the use of AI complies with current laws and regulations, minimizing legal risks for institutions and patients.
- **LGPD (General Data Protection Law):** Brazilian law (No. 13.709/2018) that regulates the processing of personal data, including requirements such as consent and breach notification.
- **Litigation:** Legal proceedings arising from damages or errors associated with the use of AI involving legal liability.
- **Machine Learning:** A subfield of artificial intelligence that enables systems to learn and improve automatically from data, without being explicitly programmed for specific tasks.
- **Multidisciplinary Committee:** A group composed of professionals from various fields (medicine, law, technology, bioethics, and patients) that oversees the development and application of AI to ensure balanced and ethical decisions.
- **Precision Medicine:** A medical approach that customizes treatments based on the patient's genetic, environmental, and lifestyle data.



- **Pseudonymization:** A technique that replaces identifiable personal data with pseudonyms, reducing the risk of direct identification during analysis and processing.
- **Ransomware:** A type of malware that encrypts data and demands payment for its release (e.g., attacks on hospitals).
- **Security Incident:** An event that compromises the confidentiality, integrity, or availability of data and systems.
- **Sensitive Data:** Personal information that requires greater protection due to its intimate nature, such as health, genetic, biometric data, and others.
- **Shared Responsibility:** The clear distribution of duties and obligations among all actors involved in the AI lifecycle to ensure safety and ethics.
- **Technical Failure:** An error or defect in AI systems that may affect their operation or produce incorrect results.
- **Transparency:** A principle requiring clarity and accessibility of information about how AI systems operate and make decisions, facilitating understanding by professionals and patients.
- **XAI (Explainable AI):** Artificial intelligence that provides transparency about its decisions through human-understandable interpretations.
- **Zero Trust Architecture:** A security model that assumes no user or device within the network is inherently trustworthy, requiring continuous verification.

CONCLUSION

Digital health is at the epicenter of a revolution that combines unprecedented technological advances with equally complex ethical, operational, and security challenges. Artificial intelligence, due to its ability to transform diagnostics, treatments, and hospital management, represents a powerful tool to raise the standard of care and efficiency in the sector. However, its positive impact depends directly on how this technology is adopted, governed, and controlled.

More than innovating, it is essential that professionals and institutions develop a solid culture of responsibility and transparency. The protection of sensitive data, respect for patients' rights, and constant human oversight are non-negotiable pillars to ensure that AI is a reliable ally and not a source of risks and harm.

The adoption of strict governance guidelines, continuous investment in cybersecurity, and the promotion of independent audits should not be seen as additional costs, but as essential elements for the sustainability of digital health. The sector must move forward with determination, aligning the speed of innovation with the prudence and ethics necessary to preserve public trust.



FINAL TIPS FOR THOSE WORKING IN DIGITAL HEALTH:

- **Consider the implementation of AI always with a focus on the principle *primum non nocere* (first, do no harm).** Before any innovation, ensure that the system will not cause unexpected or invisible harm.
- **Invest in multidisciplinary governance.** Technology, ethics, law, and medicine must go hand in hand to ensure safe and fair decisions.
- **Do not underestimate the importance of transparency.** Patients and professionals must understand how AI works, and be able to question and review automated decisions.
- **Be prepared to respond quickly to security incidents.** Contingency plans and ongoing training are essential.
- **Keep the human at the center of care.** AI is a tool, not a substitute for human intelligence, empathy, and clinical judgment.

Only in this way can artificial intelligence in healthcare fulfill its promise of beneficial transformation, balancing innovation with responsibility, and ensuring that digital evolution is an ally of life — not another risk to be feared.



REFERENCES

1. Brasil. (2018). Lei nº 13.709, de 14 de agosto de 2018. Lei Geral de Proteção de Dados Pessoais (LGPD) [Law No. 13,709 of August 14, 2018. General Data Protection Law (LGPD)]. Diário Oficial da União, Brasília.
2. Check Point. (2023). Cyber attack trends: 2023 mid-year report. Retrieved from <https://blog.checkpoint.com> (Accessed: March 22, 2025).
3. Conselho Federal de Medicina. (2022). Resolução CFM nº 2.324, de 8 de dezembro de 2022. Uso de IA na prática médica [CFM Resolution No. 2,324 of December 8, 2022. Use of AI in Medical Practice]. Brasília.
4. Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542(7639), 115–118. <https://doi.org/10.1038/nature21056>
5. Esteva, A., Robicquet, A., Ramsundar, B., Kuleshov, V., DePristo, M., Chou, K., ... & Dean, J. (2019). A guide to deep learning in healthcare. *Nature Medicine*.
6. FBI. (2022). Data breach investigations report. Retrieved from <https://www.fbi.gov> (Accessed: March 22, 2025).
7. Hospital Sírio-Libanês. (2023). Relatório de Governança em IA 2022 [AI Governance Report 2022]. São Paulo: HSL.
8. Holzinger, A., Biemann, C., Pattichis, C. S., & Kell, D. B. (2019). Causability and explainability of artificial intelligence in medicine. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*.
9. Organização Mundial da Saúde. (2021). Ethics and governance of artificial intelligence for health [Ética e governança da inteligência artificial para a saúde]. Geneva: WHO.
10. Pew Research Center. (2023). Public trust in artificial intelligence in health. Retrieved from <https://www.pewresearch.org> (Accessed: March 22, 2025).
11. Rajpurkar, P., Irvin, J., Zhu, K., Yang, B., Mehta, H., Duan, T., ... & Ng, A. Y. (2018). Deep learning for detecting pneumonia from chest X-rays. *arXiv preprint*, arXiv:1711.05225. Retrieved from <https://arxiv.org/abs/1711.05225> (Accessed: June 21, 2025).
12. Singapore Ministry of Health. (2018). SingHealth cyber attack – official statement. Retrieved from <https://www.healthcareitnews.com/news/singhealth-cyberattack-raises-concerns-over-health-data-security> (Accessed: June 21, 2025).
13. Smith, M. L., Chen, H., & Patel, V. (2023). Algorithmic accountability in healthcare. *Journal of Medical Ethics*, 49(3), 172–178.
14. Sophos. (2023). The state of ransomware in healthcare 2023. Retrieved from <https://www.sophos.com> (Accessed: March 22, 2025).



15. Topol, E. J. (2019). Deep medicine: How artificial intelligence can make healthcare human again. New York: Basic Books.
16. Wiens, J., & Shenoy, E. S. (2018). Machine learning for healthcare: On the verge of a major shift in healthcare epidemiology. Clinical Infectious Diseases.



APPENDIX A – DETAILED CHECKLIST FOR GOVERNANCE AND ETHICAL USE OF AI IN HEALTHCARE

GOVERNANCE AND ORGANIZATIONAL STRUCTURE

- 1) Form a multidisciplinary committee with representatives from technical, medical, legal, bioethics, and patient areas.
Detail: The committee should include experts offering diverse perspectives to assess technical risks, legal implications, ethical impacts, and ensure respect for patient rights.
- 2) Clearly define roles and responsibilities for all involved in AI governance.
Detail: Document who is responsible for development, approval, auditing, clinical oversight, communication, and incident response to avoid gaps and conflicts.
- 3) Establish formal processes for ethical and legal evaluation of AI systems before deployment.
Detail: Create protocols for prior algorithm analysis, including bias assessment, social impact, and compliance with current regulations.

LEGAL COMPLIANCE AND SECURITY

- 1) Ensure strict compliance with the LGPD and other applicable national and international regulations.
Detail: Implement privacy policies, data collection and usage practices that respect informed consent, secure storage, and data subject rights.
 - 2) Implement policies for the protection and privacy of patients' sensitive data.
Detail: Use techniques such as pseudonymization, encryption, and access control, and limit data retention time according to regulations.
 - 3) Create protocols for rapid response to security incidents and data breaches.
Detail: Develop clear contingency plans with designated personnel, communication flows, and immediate mitigation actions.
- 1) Auditing and Monitoring**
- 1) Conduct periodic audits (preferably quarterly) to review the operation and impact of AI systems.
Detail: Evaluate performance, accuracy, transparency, clinical impacts, and ongoing compliance, with detailed records of results.
 - 2) Continuously monitor systems to detect biases, technical failures, and ethical deviations.
Detail: Use monitoring and analysis tools to identify anomalies, recurring errors, or decisions that may cause discrimination.



- 3) Document audit results and implement corrective actions when necessary.
Detail: Ensure reports are formalized and improvements are applied within defined deadlines, with follow-up on corrections.
- 2) Transparency and Human Participation**
 - 1) Provide clear and accessible reports on automated decisions to healthcare professionals and patients.
 - 2) Detail: Ensure reports use understandable language, highlighting criteria used by AI and justifications for decisions.
 - 3) Establish mechanisms for human review of AI decisions, with the possibility of contestation.
 - 4) Detail: Allow physicians to review results and patients to request reassessment or a second opinion.
 - 5) Offer communication channels for patients to question and request second opinions on AI-based diagnoses or treatments.
 - 6) Detail: Provide direct channels, such as a help desk or specialized support, to handle questions, complaints, and appeals.
- 3) Training and Organizational Culture**
 - 1) Promote regular training for technical and clinical teams on ethical, safe, and effective AI use.
 - 2) Detail: Develop educational programs covering technical, legal, and ethical aspects, updating professionals on best practices.
 - 3) Encourage an organizational culture that values responsibility, ethics, and safety in technology adoption.
 - 4) Detail: Implement internal policies and campaigns that reinforce ethical values and commitment to patient safety.
 - 5) Foster continuous dialogue between technical, clinical, and administrative areas to align expectations and practices.
 - 6) Detail: Hold regular meetings and working groups to discuss challenges, updates, and improvements in AI governance.