Received: 21 November 2024 | Revised: 6 January 2025 | Accepted: 23 January 2025

DOI: 10.1002/jhrm.70000

ASHRM AMERICAN SOCIETY FOR HEALTH CARE RISK MANAGEMENT

RESEARCH ARTICLE

Integrating enterprise risk management to address AI-related risks in healthcare: Strategies for effective risk mitigation and implementation

Gianmarco Di Palma MD[1] | **Roberto Scendoni MD, PhD**[2,3] | **Vittoradolfo Tambone MD, PhD**[4] | **Rossana Alloni MD**[1,4] | **Francesco De Micco MD, PhD**[1,4]

[1]Department of Clinical Affair, Fondazione Policlinico Universitario Campus Bio-Medico, Rome, Italy

[2]Department of Law, Institute of Legal Medicine, University of Macerata, Macerata, Italy

[3]Italian Network for Safety in Healthcare (INSH), Coordination of Marche Region, Macerata, Italy

[4]Research Unit of Bioethics and Humanities, Department of Medicine and Surgery, Università Campus Bio-Medico di Roma, Rome, Italy

Correspondence:
Roberto Scendoni, Department of Law, Institute of Legal Medicine, University of Macerata, Macerata, Italy.
Email: r.scendoni@unimc.it

Abstract

The incorporation of artificial intelligence (AI) in health care offers revolutionary enhancements in patient diagnostics, clinical processes, and overall access to services. Nevertheless, this technological transition brings forth various new, intricate risks that pose challenges to current safety and ethical norms. This research explores the ability of enterprise risk management as an all-encompassing framework to tackle these arising risks, providing both a forward-looking and responsive strategy designed for the health care industry. At the core of this method are instruments that together seek to proactively uncover and address AI-related weaknesses like algorithmic bias, system failures, and data privacy issues. On the reactive side, it incorporates incident reporting systems and root cause analysis, tools that enable health care providers to quickly address unexpected events and consistently improve AI implementation procedures. However, some application difficulties still exist. The unclear, "black box" characteristics of numerous AI models hinder transparency and responsibility, prompting inquiries about the clarity of AI-generated choices and their adherence to ethical benchmarks in patient treatment. The research highlights that with the progress of AI technologies, the enterprise risk management framework also needs to evolve, addressing these new complexities while promoting a culture focused on safety in health care settings.

INTRODUCTION

The rise of artificial intelligence (AI) in health care

AI represents one of the essential tools in the professional experience of numerous practitioners. Similarly, the health care sector has embraced this wave of innovation, adopting AI systems across various domains, from administrative management to advanced diagnostics. In this context, the European Union has opted to contribute scientifically by highlighting key risks to ensure the protection of patients and their families from potential harm arising from the use of intelligent systems in clinical practice.[1] AI can bring numerous benefits to hospital settings, increasing diagnostic speed and accuracy. An algorithm can process vast amounts of data extremely quickly, facilitating the work of physicians and improving the quality of health care.[2–4]

Challenges in managing AI-related risks

Intelligent systems can support health care personnel in many ways, such as real-time monitoring of vital signs with alert capabilities in case of changes, enabling more rapid interventions.[5] Another application concerns the management of chronic diseases with algorithms that can suggest lifestyle modifications or propose adjustments to the therapeutic plan to better control the disease.[6,7] Furthermore, AI can facilitate access to health care by enabling health care professionals to reach patients in disadvantaged areas, indeed, services such as telemedicine allow for remote consultations and diagnoses.[8,9]

Challenges on the horizon

Despite the evident advantages, it is necessary to consider that the "indiscriminate" use of such systems, in the absence of controls, can create various issues, potentially shaping health care organizations negatively. For instance, intelligent systems can fall victim to so-called "hallucinations," that is, outputs containing totally or partially false information; thus, meticulous oversight of generated responses is required.[10,11] The introduction of such systems in health care demands an unconventional approach to risk management, and it is imperative for risk managers to develop new mitigation strategies to address the new risks associated with intelligent systems.[12] As long as these algorithms operate in isolation, performing simple tasks independently, risks remain contained; issues are more likely to arise when intelligent systems must communicate with each other, managing complex tasks and exchanging data with a risk of error propagation.[13]

Purpose and objectives of the study

In this scenario, enterprise risk management (ERM) presents itself as a promising solution to tackle the multiple and heterogeneous risks associated with the use of intelligent systems in health care settings. In 2004, the Committee of Sponsoring Organizations introduced the concept of ERM, presenting it as a viable alternative to optimize risk assessment and management using a holistic model as opposed to the traditional approach. The motto of this approach is "risk is capital," underscoring how risks can generate profit opportunities rather than merely represent an economic threat. This approach was then adapted from the traditional corporate sector to health care, where the progressive corporatization of hospitals has made this model a valid alternative to the traditional clinical risk management approach.[14] One of the fundamental concepts is that risk can be managed holistically and multidisciplinarily, avoiding management fragmentation that could lead to risk underestimation.[15–17] The purpose of this article is to evaluate the risks associated with the introduction of intelligent systems in everyday clinical-assistive practice, proposing an innovative management framework, such as ERM, to ensure appropriate mitigation of emerging risks.

EMERGING RISKS OF AI IN HEALTH CARE SETTINGS

Clinical risks

The risk of patient harm following the introduction of complex systems in health care represents a tangible threat; indeed, the integration of highly sophisticated technological tools within a high-complexity setting, such as health care, can increase the likelihood of error. In the event of a malfunction in a traditional system, the damage can be contained; however, if highly integrated AI systems used for functions such as scheduling appointments, interpreting lab results or radiological images, and managing medical records experience failures, the consequences can be catastrophic.[18] Despite technological advances in machine learning, clinical applications of AI systems still pose significant risks to patient safety, with multiple potential sources of error. First, predictive accuracy can be affected by "noise" in the input data; an error in the initial information is unlikely to be detected by the system, leading to error propagation. Even minor errors in the initial data can result in inaccurate clinical decisions. For instance, if the data distribution used as input in everyday clinical practice differs from that used to initially train the algorithm, erroneous outputs may be produced. This issue is particularly pronounced when AI systems are applied in settings different from those they were originally trained on.[19] In fact, a study on pneumonia diagnosis found a decrease in AI performance when tested in a hospital setting different from the experimental one, while this suggests potential bias related to individual hospitals, the harm that can arise when evaluating different ethnicities is markedly higher.[20] Moreover, unlike humans, AI cannot quickly adapt to unforeseen circumstances in clinical settings; for example, artifacts in a radiographic image can generate false positives if assessed by an intelligent system. This type of error highlights the need for such systems to undergo thorough training to be considered effective and safe, with diverse and inclusive datasets covering various target populations, as well as prolonged validation in real-world settings. In any case, they should be regarded solely as support tools rather than being entrusted with decision-making roles in place of physicians, preserving human oversight to minimize the possibility of error.[21] It is also clear that, for the system to maintain its validity over time, it must undergo dynamic training that accounts for new clinical scenarios. Therefore, a balance must be maintained between continuous algorithm updates and human supervision to ensure the timely identification of emerging issues.[21,22] This could be the example of an AI algorithm that was implemented in a hospital to help diagnose pneumonia. The system, initially validated in controlled environments, demonstrated high accuracy during testing. However, once implemented in a real clinical setting, the algorithm began to produce numerous false negatives. The problem was traced to noisy input data, including radiographic images with artifacts, which the system struggled to interpret correctly. As a result, several patients with pneumonia were misdiagnosed or their conditions were not detected in a timely manner. This led to delays in initiating appropriate treatment, causing serious medical complications and increasing pressure on the clinical team. The hospital's reliance on the algorithm without further validation highlighted the limitations of AI systems when moving from the experimental to the practical stage.

Technical risks

Another risk associated with the use of intelligent systems in health care lies in the limited transparency of such tools. For

example, the underlying logic of deep neural networks is often difficult, if not impossible, to interpret for both physicians and technical staff. This phenomenon is known as the "black box," where these technologies, lacking "transparency," complicate the verification of the accuracy of diagnostic or therapeutic processes, reducing health care personnel's trust in AI systems. This issue is typical of deep learning models, which, while able to generate statistically correct algorithmic decisions, may lack a clear and accessible explanation for physicians, who, due to their role, must be able to justify and implement therapeutic choices. Additionally, in cases where an error occurs, the lack of transparency may hinder understanding of the cause, risking a repetition of the same failure. Therefore, this lack of transparency not only endangers patient health but also raises ethical concerns regarding accountability in the event of errors.[23,24]

Another technical risk involves potential bias within the training data of the intelligent system. If the system is trained on limited or unrepresentative data, this may lead to errors in diagnosis and proposed treatments, resulting in discrimination based on gender, ethnicity, or age. Many studies have highlighted that AI can perpetuate inequalities if not adequately trained. For instance, if an algorithm is trained on a predominantly white population sample, it may be less accurate in diagnosing diseases in other ethnicities. Thus, the quality of the data used by AI is essential for ensuring accurate clinical decisions.[25]

At the same time, health care data may be incomplete, erroneous, or even manipulated, introducing vulnerabilities into a system. In this context, cyberattacks pose a serious threat, as the processing and continuous exchange of large volumes of data is central to the operation of intelligent systems. A cyberattack or security breach could compromise data integrity, with potentially disastrous effects on the patient care process.[26]

The increasing use of AI in hospitals can lead to cognitive overload for physicians, who are required to manage an ever-growing amount of information generated by intelligent systems. While automation in generating clinical responses may be helpful, it could reduce the physician's ability to critically exercise professional judgment. This risk is particularly evident in emergency settings, where health care personnel must make rapid and complex decisions. If the automated system fails, physicians may not be ready to intervene with alternative solutions, thereby increasing the risk of errors.

Another area of concern is the long-term reliability of AI systems. Many technologies require continuous updates and maintenance to remain accurate. The lack of system updates or hardware obsolescence can reduce effectiveness over time.[20] Thus, an algorithm that is not retrained with new data could become obsolete, resulting in incorrect diagnoses. Additionally, maintaining intelligent systems requires resources and technical expertise that may not always be available in hospital settings, adding a further layer of complexity.[2,21,27] An illustrative example is that of an AI-based telemedicine platform implemented to improve patient monitoring and remote consultations. At an early stage, the system functioned suitably, demonstrating efficiency in processing and managing large volumes of data regarding patients. Subsequently, it was the subject of a cybersecurity breach, during which unauthorized individuals gained access to sensitive patient data. The cyber-attack exploited vulnerabilities in the system's data encryption protocols, resulting in, among other things, the disruption of clinical operations, with medical staff forced to revert to manual processes, significantly slowing patient care and increasing operational strain. This resulted in reputational damage to the company as well as legal repercussions for any damages achieved by patients.

Socio-ethical risks

As previously discussed, AI systems in health care rely on large datasets that may reflect societal biases. Historically marginalized groups, such as ethnic and racial minorities, may be underrepresented or misrepresented in these datasets, resulting in suboptimal care. Unbalanced datasets may lead to inaccurate predictions for these groups, exacerbating existing health disparities. Gender biases are also evident: reports of pain from female patients are often downplayed, and these biases could be perpetuated or amplified by AI. Another crucial issue is that of "health care deserts," or areas where vulnerable populations lack access to health care services. This issue could be worsened by the digital divide, as populations without access to technology or digital literacy may be excluded from clinical advances such as telemedicine. Thus, marginalized communities may lack the infrastructure or resources needed to benefit from AI technologies, turning these advancements into barriers for certain populations.[28–30] Another key element for ensuring AI systems are fair is to build social trust. Without transparency, explainability, and ethical marketing, people may distrust these technologies. It is essential to design AI systems with user-centered values and to communicate their capabilities clearly to the public to encourage acceptance. Addressing AI biases requires interdisciplinary collaboration, continuous evaluation, and transparency. Diverse teams of AI developers, clinicians, ethicists, and even patient advocacy groups can help ensure that AI tools are designed inclusively and equitably. Proper representation in datasets, careful data labeling, and monitoring AI systems in real health care settings is essential to prevent the perpetuation of biases.[21,31–33] One notable example might involve an AI system designed to prioritize emergency patients based on the severity of their conditions. Although the system performed well in controlled test environments, its implementation in a real-world hospital setting revealed consistent problems; in fact, the algorithm underestimated the severity of conditions of patients of some specific ethnicities due to the lack of heterogeneity in the training dataset. These biases stemmed from the lack of representation of different populations, leading to delays in care for some patients and exacerbating existing health disparities. This failure not only put affected patients at risk but also raised ethical concerns about fairness and equity in AI-driven health care decisions. The incident highlighted the crucial need for diverse and representative datasets to ensure equitable care outcomes for all demographic groups.

Privacy and data protection

The increasing use of AI in health care, accelerated during the SARS-CoV-2 pandemic, has raised concerns about data privacy, confidentiality, and the protection of patients' rights. These risks include the exposure and potential misuse of sensitive data, which could lead to violations of individual rights as well as nonmedical uses of personal information. Informed consent remains a critical issue in this context, as patients must be adequately informed to make informed decisions about sharing their health data. Specifically, the introduction of nontransparent AI algorithms, along with complex consent forms, may limit patient autonomy, creating an obstacle to the shared decision-making process with physicians. Such complex algorithms not only make it difficult for patients to understand data-sharing practices but also create new vulnerabilities in terms of cybersecurity. The risk of cyberattacks, as well as unauthorized data usage, raises particular concerns, with issues including privacy breaches, identity theft, and the potential targeting of AI-controlled medical devices. Some studies highlight the need for ethical and transparent integration of AI technologies in compliance with GDPR to ensure that data are used securely and in accordance with regulations, thereby reducing the risk of system failures or algorithmic biases.[34] Another significant risk relates to privacy violations involving the use of big data. Many health care facilities have already encountered issues related to data protection during cyberattacks; some have shared or stored sensitive health data on publicly accessible servers, resulting in unauthorized exposure of information[35] Although data used for research purposes should be anonymized, elements such as treatment dates, medical notes, or personal information have sometimes not been removed. This type of exposure poses a significant risk, particularly with the use of AI, which requires large amounts of data to function effectively. Unauthorized access to health data can also facilitate insurance fraud, highlighting the vulnerability of these infrastructures.[36] Obtaining informed consent for the use of patient data is a topic of intense ethical debate. Generally, patients provide consent for data usage when admitted for medical treatment or billing purposes. However, when data are used for research or purposes unrelated to therapy, separate consent is required. If data have been anonymized, GDPR no longer applies, allowing health care facilities to use these data without additional legal restrictions.[37] However, even though anonymized data should theoretically be protected, there are concerns that combining these data with other sources could make it possible to identify individuals.[38]

Proposed solutions include enhancing the security of cloud storage systems and raising awareness among health care providers about the risks associated with these technologies to create a health care environment compatible with the safe and responsible use of AI. The growing complexity requires collaboration among researchers, policymakers, and health care professionals to develop more robust regulations and more effective security mechanisms to prevent privacy breaches and ensure the ethical and secure use of AI in health care.[39,40] Nonetheless, if data have been anonymized, GDPR no longer applies, allowing health care facilities to use such data without further legal restrictions.[37] Still, there are concerns that combining anonymized data with other sources might enable the re-identification of individuals.[38] One possible example involves the implementation of an AI system to monitor COVID-19 patients, optimizing care delivery. Although the system improved patient management, a critical problem emerged when sensitive health data was mistakenly shared with an unauthorized third party. The breach occurred due to poor data access controls and led to the exposure of the personal medical information of thousands of patients; this not only resulted in fines for noncompliance with data protection regulations but also undermined patients' confidence in the hospital's ability to safeguard their information. The situation has highlighted the urgent need for robust data governance, including stricter access protocols and increased staff training, to ensure compliance with privacy standards and protect sensitive health data.

DISCUSSION

Enterprise risk management: A holistic approach to AI challenges

To integrate enterprise risk management (ERM) into managing the risks associated with using intelligent systems in health care, the key principles of ERM can be utilized to address emerging challenges related to AI technologies.[41] ERM provides a holistic approach to identifying AI-related risks, including those associated with system malfunctions, privacy breaches, and algorithmic biases, which can compromise patient safety and data integrity (Figure 1). ERM's management approach involves shared tools like brainstorming, focus groups, and questionnaires to gather insights from various operational units involved. This approach can help identify potential risks across different areas, from clinical to technological and financial risks, allowing for a detailed risk analysis.

Preventing and correcting AI-related errors: An ERM perspective

Proactive methods such as FMECA (Failure Mode, Effects, and Criticality Analysis) are essential for anticipating and mitigating potential AI system failures, ensuring the safe and responsible use of technology. This tool operates on the premise that some errors are inevitable and may occur with certain probabilities; through in-depth analysis, the primary goal is to identify all potential ways an AI system failure, error, or malfunction might occur. This enables prioritizing critical situations and optimizing the system or process to improve safety and reduce failure risks.[42] Applying FMECA minimizes defects or malfunctions that may arise from inadequate design or risk oversight during the initial phase, thus strengthening the safety and reliability of the intelligent system by anticipating operational issues and enhancing overall performance. Simultaneously, reactive methods like

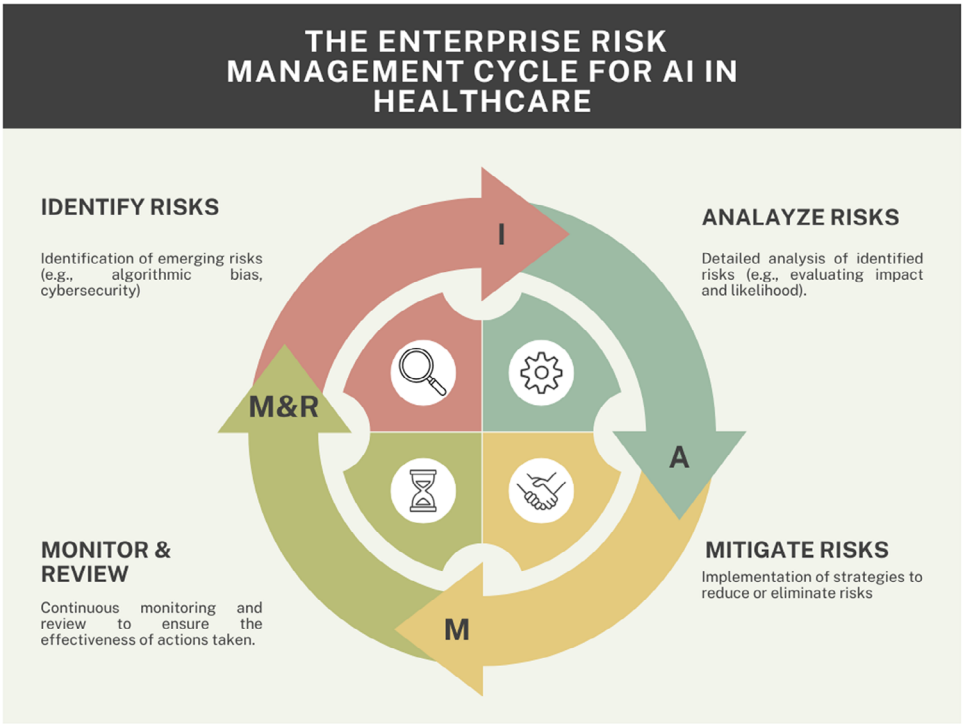

FIGURE 1 Illustrates the stages of ERM—identification, analysis, mitigation, and monitoring—applied to AI-related risks in health care.

incident reporting and peer review help identify and mitigate risks once they materialize, enabling a swift response and a process analysis to prevent recurrence. Other reactive methods include root cause analysis (RCA), which identifies the root causes of potential AI system malfunctions, and Hazard and Operability Study (HAZOP), which can be applied to analyze complex scenarios, ensuring that each potential risk is identified and mitigated before negatively impacting clinical safety.

Risk mapping in AI-enhanced health care systems

Such tools are crucial for managing the risks associated with AI implementation. Once AI-related risks are identified, ERM techniques should be applied for risk mapping, especially as less transparent technologies, like deep neural networks, pose serious clinical and decision-making liability risks.[43] (Figure 2). Mapping these risks and prioritizing them based on a scale of probability and impact is crucial for developing appropriate mitigation strategies. For example, cybersecurity risks may receive a high priority score due to potential legal and reputational implications. ERM offers various risk control techniques to address AI-related risks, such as risk control (limiting AI use in clinically sensitive areas), risk reduction (introducing backup systems and manual checks), and risk transfer (insurance against AI-related errors or damage from cyberattacks). This approach is essential to reduce the negative impact of potential AI malfunctions, preventing catastrophic errors, and protecting patients' sensitive data.[44]

Driving safety through monitoring and collaboration

Finally, the ERM process concludes with implementing corrective measures; however, it is necessary to constantly monitor key risk indicators (KRIs) to detect any changes in the probability and impact of AI-related risks promptly. KRIs are metrics used to monitor and measure an organization's risk level over time, providing early warning signals of potential risk events that could negatively impact operations or objectives.[45] KRIs are fundamental to risk management, allowing the risk manager to identify and address potential issues before they occur. Additionally, KRIs differ from Key Performance Indicators (KPIs) as KRIs focus on risk management, whereas KPIs monitor performance. In essence, KRIs measure factors that may negatively impact goal achievement and are often linked to specific risks. For instance, monitoring AI algorithm performance over time is essential, ensuring that systems are updated and operate on datasets representative of the target population to avoid biases or clinical errors.[46] Mitigating secondary risks associated with intelligent systems requires collaboration among various professionals, including Data Protection Officers (DPOs) for GDPR compliance and data breach risk management, AI experts and computer engineers for technical maintenance and model training, and ethicists to address the moral implications of using AI in health care. In this multidisciplinary context, the Risk Manager plays a crucial role as a facilitator and a link between different disciplines, coordinating risk management processes, fostering cooperation among working groups, and

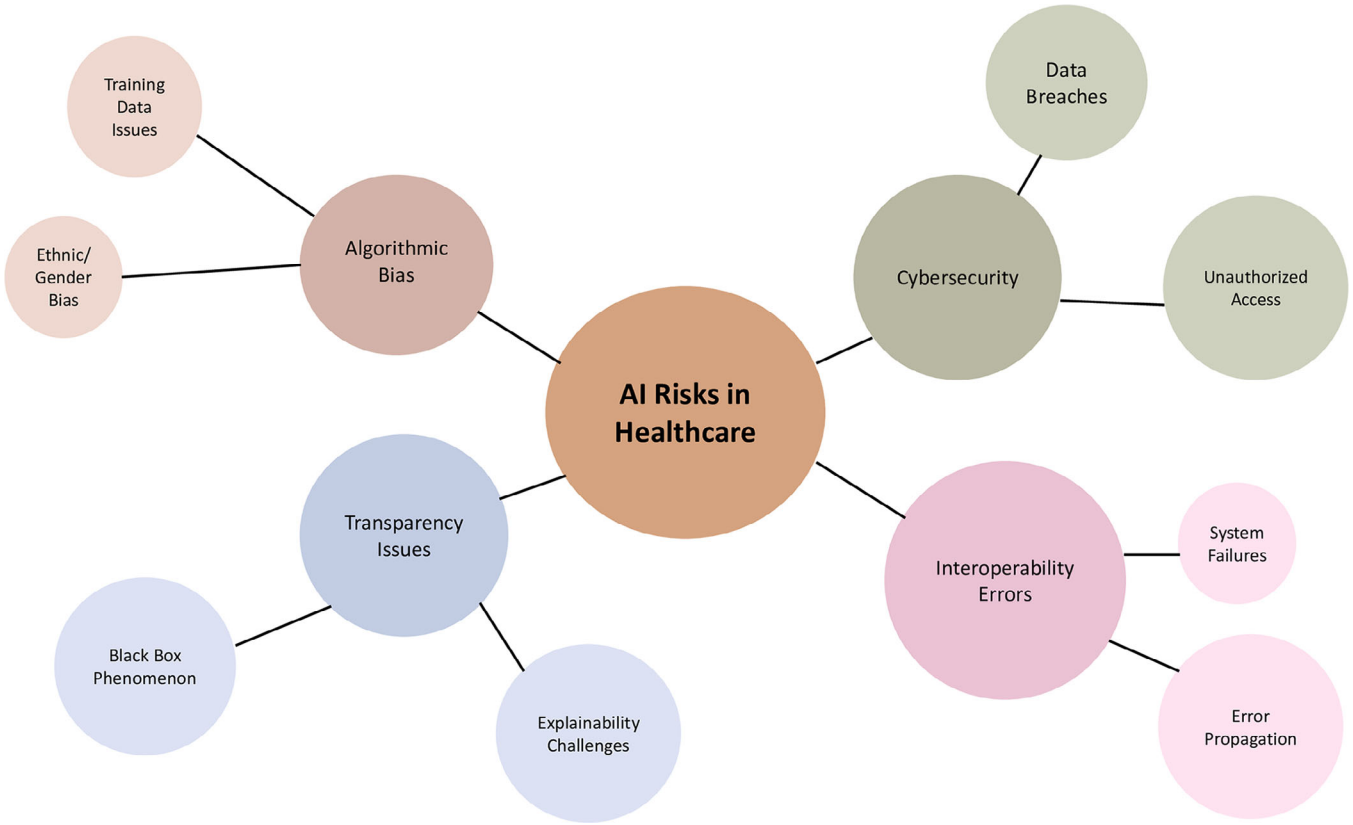

FIGURE 2 Depicts key AI-related risks in health care, including algorithmic bias, cybersecurity, and transparency issues, with their subcategories.

promoting a safety-oriented organizational culture that supports innovation.[47]

Future directions for ERM in AI risk management

However, despite ERM providing numerous tools and resources for a thorough management of risks associated with intelligent systems in health care, several open issues remain that this management approach will need to address in its initial application stages. Specifically, ERM should pay particular attention to the interpretability of AI models; while deep neural networks offer high performance, their inherent complexity makes it difficult to understand the rationale behind decisions. Therefore, it is crucial to develop techniques to make models more transparent and interpretable, facilitating risk assessment and increasing end-user trust.[48] ERM must also proactively address algorithmic bias to prevent discrimination and unfair decisions, implementing mechanisms to identify and mitigate biases in training datasets and within the models themselves, ensuring fairness and inclusivity. Additionally, the use of AI in health care raises important issues regarding legal liability. ERM must help define accountability in cases of errors or harm caused by AI systems, highlighting the need for clear legal and ethical frameworks to ensure transparency and accountability.[49] Indeed, such a risk management model should promote a safety culture within the organization, engaging all employees in training and awareness activities about AI-related risks. ERM is not a static process but a dynamic one, necessitating continuous evaluation of the effectiveness of risk control measures and updating the risk management plan based on new information and technological advancements. Finally, ERM itself can benefit from using advanced tools like big data analysis and machine learning to identify emerging risks and improve adverse event prediction.[50,51]

CONCLUSIONS

The use of ERM within health care organizations that have chosen to implement AI in their clinical assistance pathways could be an effective solution for better managing the risks associated with the use of intelligent systems in daily practice. ERM's holistic approach enables the identification, evaluation, and mitigation of all risks arising from the use of AI systems, including technical malfunctions, patient privacy breaches, and algorithmic biases. In this context, the use of both proactive risk assessment tools, such as FMECA, which anticipate potential problems, and reactive tools like RCA and HAZOP, which correct errors within internal processes, will be increasingly effective in optimizing clinical-assistance processes in response to critical events. Continuous monitoring of KRIs will be essential in this context, allowing timely detection of changes in the likelihood and impact of risk events. This approach provides an ongoing

assurance of safety, aligning with the highest clinical standards in real-time.[52–54]

However, despite the clear benefits, many significant challenges lie ahead. For instance, increasingly complex neural networks will make result interpretation more difficult, with less transparent clinical decision-making pathways due to the "black box" phenomenon. This system "opacity" could lead to increased medical-legal disputes, necessitating the development of algorithms that operate more "transparently." Another crucial issue to address is algorithmic bias, as this could paradoxically lead to increased health care disparities.[55–57] This suggests that in the future, the risk management framework will need to embrace technological advancements, incorporating intelligent systems capable of providing more accurate risk analyses.[58] Lastly, this impending change will require regulatory review to define liability limits in case of errors.

Thus, while ERM represents an essential and comprehensive approach to managing risks associated with AI implementation in health care, much remains to be done to ensure the safe integration of intelligent systems in hospital settings.

ACKNOWLEDGMENTS

Thanks to the National Ph.D. in Artificial Intelligence, XL cycle, a course on Health and Life Sciences, organized by University`Campus Bio-Medico of Rome, to which Gianmarco Di Palma is enrolled as a Ph.D. student.

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no conflicts of interest.

ORCID

Gianmarco Di Palma MD https://orcid.org/0009-0009-6638-4151
Roberto Scendoni MD, PhD https://orcid.org/0000-0003-1910-2405
Vittoradolfo Tambone MD, PhD https://orcid.org/0000-0002-9231-9523
Rossana Alloni MD https://orcid.org/0000-0001-8877-2141
Francesco De Micco MD, PhD https://orcid.org/0000-0001-8483-3787

REFERENCES

1. European Parliament. Artificial intelligence in healthcare: Applications, risks, and ethical and societal impacts. EPRS; 2022. Accessed September 16, 2024. https://www.europarl.europa.eu/thinktank/en/document/EPRS_STU(2022)729512
2. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med*. 2019;25(1):44-56. https://doi.org/10.1038/s41591-018-0300-7
3. Jiang F, Jiang Y, Zhi H, et al. Artificial intelligence in healthcare: past, present and future. *Stroke Vasc Neurol*. 2017;2(4):230-243. https://doi.org/10.1136/svn-2017-000101
4. Yu KH, Beam AL, Kohane IS. Artificial intelligence in healthcare. *Nat Biomed Eng*. 2018;2(10):719-731. https://doi.org/10.1038/s41551-018-0305-z
5. Areia C, Biggs C, Santos M, et al. The impact of wearable continuous vital sign monitoring on deterioration detection and clinical outcomes in hospitalised patients: a systematic review and meta-analysis. *Crit Care*. 2021;25(1):351. https://doi.org/10.1186/s13054-021-03766-4
6. Steinhubl SR, Muse ED, Topol EJ. Can mobile health technologies transform health care? *JAMA*. 2013;310(22):2395-2396. https://doi.org/10.1001/jama.2013.281078
7. Singareddy S, Sn VP, Jaramillo AP, et al. Artificial intelligence and its role in the management of chronic medical conditions: a systematic review. *Cureus*. 2023;15(9):e46066. https://doi.org/10.7759/cureus.46066
8. Verma P, Sood SK. Cloud-centric IoT based disease diagnosis healthcare framework. *J Parallel Distrib Comput*. 2018;116:27-38. https://doi.org/10.1016/j.jpdc.2017.11.018
9. Bhaskar S, Bradley S, Chattu VK, et al. Telemedicine across the globe-position paper from the COVID-19 pandemic health system resilience PROGRAM (REPROGRAM) International Consortium (Part 1). *Front Public Health*. 2020;8:556720. https://doi.org/10.3389/fpubh.2020.556720
10. Amann J, Blasimme A, Vayena E, Frey D, Madai VI, Precise4Q consortium. Explainability for artificial intelligence in healthcare: a multidisciplinary perspective. *BMC Med Inform Decis Mak*. 2020;20(1):310. https://doi.org/10.1186/s12911-020-01332-6
11. Loh E. Medicine and the rise of the robots: a qualitative review of recent advances of artificial intelligence in health. *Leader*. 2018;2(2):59-63. https://doi.org/10.1136/leader-2018-000071
12. Maleki Varnosfaderani S, Forouzanfar M. The role of AI in hospitals and clinics: transforming healthcare in the 21st century. *Bioengineering*. 2024;11(4):337. https://doi.org/10.3390/bioengineering11040337
13. Mennella C, Maniscalco U, De Pietro G, Esposito M. Ethical and regulatory challenges of AI technologies in healthcare: a narrative review. *Heliyon*. 2024;10(4):e26297. https://doi.org/10.1016/j.heliyon.2024.e26297
14. Altakhaineh Y, Abdulrahman DSA, Rameli DN, Ibrahim DRM, Afthanorhan WMA. The effect of enterprise risk management on firm performance in Jordan, the mediating role of supply chain management practices. *IOSR-JBM*. 2024;26:49–57
15. Grace MF, Leverty JT, Phillips RD, Shimpi P. The value of investing in enterprise risk management. *JRI*. 2015;82(2):289-316. https://doi.org/10.1111/jori.12022
16. Burnaby P, Hass S. Ten steps to enterprise-wide risk management. *Int J Corp Gov*. 2009;9(5):539-550. https://doi.org/10.1108/14720700910998111
17. Cagliano AC, Grimaldi S, Rafele C. Choosing project risk management techniques. A theoretical framework. *J Risk Res*. 2015;18(2):232-248. https://doi.org/10.1080/13669877.2014.896398
18. Sajithkumar A, Thomas J, Saji AM, et al. Artificial Intelligence in pathology: current applications, limitations, and future directions. *Ir J Med Sci*. 2024;193(2):1117-1121. https://doi.org/10.1007/s11845-023-03479-3
19. Ardila D, Kiraly AP, Bharadwaj S, et al. End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. *Nat Med*. 2019;25(6):954-961. https://doi.org/10.1038/s41591-019-0447-x
20. Challen R, Denny J, Pitt M, Gompels L, Edwards T, Tsaneva-Atanasova K. Artificial intelligence, bias and clinical safety. *BMJ Qual Saf*. 2019;28(3):231-237. https://doi.org/10.1136/bmjqs-2018-008370
21. Muley A, Muzumdar P, Kurian G, Basyal GP. Risk of AI in healthcare: a comprehensive literature review and study framework. *Am J Med & Health* 2023;21:276-291. doi:10.48550/arXiv.2309.14530
22. Gerke S, Minssen T, Cohen G. Ethical and legal challenges of artificial intelligence-driven healthcare. *Artificial Intelligence in Healthcare*. 2020; pp. 295–336. https://doi.org/10.1016/B978-0-12-818438-7.00012-5
23. Price II, NW. Risks and remedies for artificial intelligence in health care. Brookings; 2019. Accessed September 16, 2024. https://www.brookings.edu/articles/risks-and-remedies-for-artificial-intelligence-in-health-care/
24. Khan B, Fatima H, Qureshi A, et al. Drawbacks of artificial intelligence and their potential solutions in the healthcare sector. *Biomed Mater Devices*. 2023; 1(2):1–8. https://doi.org/10.1007/s44174-023-00063-2
25. Mittermaier M, Raza MM, Kvedar JC. Bias in AI-based models for medical applications: challenges and mitigation strategies. *NPJ Digit Med*. 2023;6(1):113. https://doi.org/10.1038/s41746-023-00858-z
26. Malatji M, Tolah A. Artificial intelligence (AI) cybersecurity dimensions: a comprehensive framework for understanding adversarial and offensive AI. *AI Ethics*. 2024; 1–28. https://doi.org/10.1007/s43681-024-00427-4

27. Knight W. The dark secret at the heart of AI. MIT Technology Review; 2024. Accessed September 16. https://www.technologyreview.com/2017/04/11/5113/the-dark-secret-at-the-heart-of-ai/
28. Khosravi M, Zare Z, Morteza MS, Izadi R. Ethical challenges of using artificial intelligence in healthcare delivery: a thematic analysis of a systematic review of reviews. *J Public Health (Berl.)*, 2024; 1–11. https://link.springer.com/article/10.1007/s10389-024-02219-w
29. Holčapek T, Šolc M, Šustek P. Telemedicine and the standard of care: a call for a new approach? *Front Public Health*. 2023;11:1184971. https://doi.org/10.3389/fpubh.2023.1184971
30. De Micco F, Tambone V, Frati P, Cingolani M, Scendoni R. Disability 4.0: bioethical considerations on the use of embodied artificial intelligence. *Front Med*. 2024;11:1437280. https://doi.org/10.3389/fmed.2024.1437280
31. Yang Y, Lin M, Zhao H, Peng Y, Huang F, Lu Z. A survey of recent methods for addressing AI fairness and bias in biomedicine. *J Biomed Inform*. 2024;154:104646. https://doi.org/10.1016/j.jbi.2024.104646
32. Ueda D, Kakinuma T, Fujita S, et al. Fairness of artificial intelligence in healthcare: review and recommendations. *Jpn J Radiol*. 2024;42(1):3-15. https://doi.org/10.1007/s11604-023-01474-3
33. Labkoff S, Oladimeji B, Kannry J, et al. Toward a responsible future: recommendations for AI-enabled clinical decision support. *J Am Med Inform Assoc*. 2024;31(11):2730-2739. https://doi.org/10.1093/jamia/ocae209
34. Mazur J. Artificial intelligence vs data protection: how the GDPR can help to develop a precautionary regulatory approach to AI? In: Kornilakis A, Nouskalis G, Pergantis V, Tzimas T, eds. *Artificial Intelligence and Normative Challenges: International and Comparative Legal Perspectives*. Springer International Publishing; 2023:215-233. https://doi.org/10.1007/978-3-031-41081-9_12
35. Yadav N, Pandey S, Gupta A, Dudani P, Gupta S, Rangarajan K. Data privacy in healthcare: in the era of artificial intelligence. *Indian Dermatol Online J*. 2023;14(6):788-792. https://doi.org/10.4103/idoj.idoj_543_23
36. Seh AH, Zarour M, Alenezi M, et al. Healthcare data breaches: insights and implications. *Healthcare*. 2020;8(2):133. https://doi.org/10.3390/healthcare8020133
37. Rumbold JMM, Pierscionek B. The effect of the general data protection regulation on medical research. *J Med Internet Res*. 2017;19(2):e47. https://doi.org/10.2196/jmir.7108
38. Langarizadeh M, Orooji A, Sheikhtaheri A. Effectiveness of anonymization methods in preserving patients' privacy: a systematic literature review. *Stud Health Technol Inform*. 2018;248:80-87.
39. Mohammad Amini M, Jesus M, Fanaei Sheikholeslami D, Alves P, Hassanzadeh Benam A, Hariri F. Artificial intelligence ethics and challenges in healthcare applications: a comprehensive review in the context of the European GDPR mandate. *Mach Learn Knowl Extr*. 2023;5(3):1023-1035. https://doi.org/10.3390/make5030053
40. Banja J. How might artificial intelligence applications impact risk management? *AMA J Ethics*. 2020;22(11):E945-E951. https://doi.org/10.1001/amajethics.2020.945
41. Frigo ML, Anderson RJ. Strategic risk management: A foundation for improving enterprise risk management and governance. *J Corp Account Finance*. 2011;22(3):81-88. https://doi.org/10.1002/jcaf.20677
42. Weber L, Schulze I, Jaehde U. Using failure mode and effects analysis to increase patient safety in cancer chemotherapy. *Res Social Adm Pharm*. 2022;18(8):3386-3393. https://doi.org/10.1016/j.sapharm.2021.11.009
43. Delgado J, de Manuel A, Parra I, et al. Bias in algorithms of AI systems developed for COVID-19: A scoping review. *J Bioeth Inq*. 2022;19(3):407-419. https://doi.org/10.1007/s11673-022-10200-z
44. van Kolfschooten H, van Oirschot J. The EU Artificial Intelligence Act (2024): implications for healthcare. *Health Policy*. 2024;149:105152. https://doi.org/10.1016/j.healthpol.2024.105152
45. Davenport T, Kalakota R. The potential for artificial intelligence in healthcare. *Future Healthc J*. 2019;6(2):94-98. https://doi.org/10.7861/futurehosp.6-2-94
46. Snoek J, Rippel O, Swersky K, et al. Scalable Bayesian optimization using deep neural networks. In: *Proceedings of the 32nd International Conference on Machine Learning*. PMLR; 2015:2171–2180. Accessed September 16, 2024. https://proceedings.mlr.press/v37/snoek15.html
47. Chen JH, Asch SM. Machine learning and prediction in medicine—beyond the peak of inflated expectations. *N Engl J Med*. 2017;376(26):2507-2509. https://doi.org/10.1056/NEJMp1702071
48. London AJ. Artificial intelligence and black-box medical decisions: accuracy versus explainability. *Hastings Cent Rep*. 2019;49(1):15-21. https://doi.org/10.1002/hast.973
49. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. *Science*. 2019;366(6464):447-453. https://doi.org/10.1126/science.aax2342
50. Morley J, Floridi L, Kinsey L, Elhalal A. From what to how: an initial review of publicly available AI ethics tools, methods and research to translate principles into practices. *Sci Eng Ethics*. 2020;26(4):2141-2168. https://doi.org/10.1007/s11948-019-00165-5
51. Rahman MA, Victoros E, Ernest J, Davis R, Shanjana Y, Islam MR. Impact of artificial intelligence (AI) technology in healthcare sector: a critical evaluation of both sides of the coin. *Clin Pathol*. 2024;17:2632010X241226887. https://doi.org/10.1177/2632010X241226887
52. Reddy S, Allan S, Coghlan S, Cooper P. A governance model for the application of AI in health care. *J Am Med Inform Assoc*. 2020;27(3):491-497. https://doi.org/10.1093/jamia/ocz192
53. Caruana R, Lou Y, Gehrke J, Koch P, Sturm M, Elhadad N. Intelligible models for healthcare: predicting pneumonia risk and hospital 30-day readmission. In: *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. KDD '15. Association for Computing Machinery; 2015:1721-1730. https://doi.org/10.1145/2783258.2788613
54. Balthazar P, Harri P, Prater A, Safdar NM. Protecting your patients' interests in the era of big data, artificial intelligence, and predictive analytics. *J Am Coll Radiol*. 2018;15(3 Pt B):580-586. https://doi.org/10.1016/j.jacr.2017.11.035
55. Kononenko I. Machine learning for medical diagnosis: history, state of the art and perspective. *Artif Intell Med*. 2001;23(1):89-109. https://doi.org/10.1016/s0933-3657(01)00077-x
56. Parikh RB, Teeple S, Navathe AS. Addressing bias in artificial intelligence in health care. *JAMA*. 2019;322(24):2377-2378. https://doi.org/10.1001/jama.2019.18058
57. Thomasian NM, Eickhoff C, Adashi EY. Advancing health equity with artificial intelligence. *J Public Health Pol*. 2021;42(4):602-611. https://doi.org/10.1057/s41271-021-00319-5
58. De Micco F, Grassi S, Tomassini L, Di Palma G, Ricchezze G, Scendoni R. Robotics and AI into healthcare from the perspective of European regulation: who is responsible for medical malpractice? *Front Med*. 2024;11:1428504. https://doi.org/10.3389/fmed.2024.1428504

How to cite this article: Di Palma G, Scendoni R, Tambone V, Alloni R, De Micco F. Integrating enterprise risk management to address AI-related risks in healthcare: Strategies for effective risk mitigation and implementation. *J Healthc Risk Manag*. 2025;1-9. https://doi.org/10.1002/jhrm.70000

AUTHOR BIOGRAPHIES

Dr. Gianmarco Di Palma, MD, is a final-year resident in legal medicine at the University of Pavia, currently training at the Clinical Directorate of Fondazione Policlinico Universitario Campus Bio-Medico. He is also a PhD student in the National PhD Program in artificial intelligence (XL cycle).

Prof. Roberto Scendoni, MD, PhD, has been a tenure track researcher in legal medicine at University of Macerata since 2022, he holds a PhD in juridical sciences. He is a consultant for forensic autopsies and medical malpractice and he conducts research activities in forensic medicine, forensic anthropology, forensic toxicology and risk management. He has been involved in many national and international training events as speaker or organizer.

Prof. Vittoradolfo Tambone is a full professor at Università Campus Bio-Medico di Roma, specializing in bioethics and legal medicine. He holds a PhD in Bioethics, a doctorate in systematic moral theology, and has coordinated national and European research projects in ethics and medicine.

Prof. Rossana Alloni is an associate professor in general surgery (MED/18) at Università Campus Bio-Medico di Roma and Clinical Director of the university hospital since 2013. She has been part of the institution since 1992, holding various roles, including faculty board member and e-learning coordinator.

Prof. Francesco De Micco, MD, PhD, has been a researcher in legal medicine at Università Campus Bio-Medico di Roma since 2022 and a clinical risk manager at Fondazione Policlinico Universitario Campus Bio-Medico since 2021. He collaborates as a forensic expert and serves on several institutional committees. He holds a PhD in translational and clinical medicine and specializes in legal medicine.