Healthcare and Medical Devices, Vol. 171, 2025, 50–58
https://doi.org/10.54941/ahfe1006190

Patients Over Process: Stratifying Risk in the Design, Development, and Deployment of Artificial Intelligence in Healthcare

Bryan Johnston[1] and Jawahar (Jay) Kalra[1,2]

[1]Department of Pathology and Laboratory Medicine, College of Medicine, University of Saskatchewan, Saskatoon, Canada
[2]Royal University Hospital, Saskatchewan Health Authority, 103 Hospital Drive, Saskatoon, Saskatchewan, S7N 0W8, Canada

ABSTRACT

The global focus on artificial intelligence (AI) in healthcare and medicine is on the rise. Despite remarkable progress in integrating AI into clinical workflows, gaps in regulation remain a prevalent issue within healthcare systems. Effective regulation of artificial intelligence in clinical practice is essential for managing medico-legal risk and ensuring patient safety. Numerous studies highlight the significant potential for medico-legal risk and the need for clear guidelines on the ethical and safe use of AI in clinical practice. Although there are various concerns that these guidelines must address, our work focused on researching best practices regarding patient-centred factors like patient autonomy, trust and transparency, privacy and security, equity and fairness, and ensuring human oversight. While challenges in AI workflow integration arise from many factors, including human interactions and system inadequacies, the focus on individuals rather than the system has fostered an unsuitable culture for enhancing patient-centred care. Key focus areas include risk stratification strategies and increasing transparency within this inherently complex system, as they play a crucial role in guiding clinical decisions in patient management. Proper integration of AI regulatory frameworks into clinical practice is essential for addressing gaps in the design, development, deployment, and long-term monitoring of AI solutions. Globally, the regulation of AI in clinical practice is continually evolving as governments and legal systems adapt to the rapid advances in AI as a medical device (AIaMD). In Canada, a strategic path forward prioritizes federal and provincial regulations; however, at this stage, they remain fragmented. We advocate for the establishment of uniform guidelines that address the risks, benefits, opportunities, and best practices as AI technologies are integrated into the clinical workflow. Achieving a national standard with clear guidance on the ethical and safe use of AI in clinical practice is recommended to move forward.

Keywords: Artificial intelligence, Medical ethics, Patient safety, Patient-centred care, Medico-legal risk, Regulation, Quality care

INTRODUCTION

Artificial intelligence (AI) is increasingly becoming an integral part of healthcare, transforming patient care and medical practice through its diverse applications (Spear, Ehrenfeld and Miller, 2023). The integration of AI into clinical workflows is advancing rapidly, but this progress is accompanied by significant challenges, particularly in terms of regulation (Dossabhoy *et al.*, 2023). Despite its promise, AI in healthcare is not without risk, especially when patient safety and medico-legal implications are considered (Cobianchi *et al.*, 2022). Gaps in regulation remain a prevalent issue within healthcare systems, raising concerns about the ethical and safe deployment of AI in clinical practice. Existing frameworks fail to adequately address patient-centred factors such as autonomy, trust, transparency, and equity. These gaps create a challenging environment for achieving patient-centred care, as they undermine the very principles of fairness and accountability that healthcare systems strive to uphold (Reddy *et al.*, 2021). In this context, risk stratification strategies and transparent regulatory frameworks are essential to guide the design, development, and deployment of AI solutions in clinical practice.

THE REGULATION OF AI IN HEALTHCARE

The rapid development of artificial intelligence in healthcare has outpaced the establishment of comprehensive regulatory frameworks, resulting in significant challenges in ensuring safe, effective, and ethical AI integration (Larson *et al.*, 2021). One of the primary issues is the fragmentation of regulations, resulting in regulatory disparities that complicate the adoption and oversight of AI systems. This fragmented approach creates challenges for healthcare providers and developers who must navigate a patchwork of rules, potentially delaying the adoption of AI technologies. Healthcare systems are inherently multifaceted, involving numerous stakeholders, from clinicians and administrators to patients and regulators. AI must operate within this intricate system while aligning with existing clinical practices and capabilities (Finkelstein *et al.*, 2024). Systemic barriers, such as resistance to change, lack of digital infrastructure, and limited technical literacy among healthcare professionals, hinder the seamless integration of AI technologies. Human interaction barriers further complicate AI adoption (Shevtsova *et al.*, 2024). Trust and transparency are essential for clinicians to embrace AI tools as reliable partners in patient care. However, the "black box" nature of many AI systems, where decision-making logic is not easily interpretable, undermines trust and raises ethical concerns (Wadden, 2022).

PATIENT-CENTRED FACTORS IN AI REGULATION

Patient-centred care is a cornerstone of effective healthcare, and integrating artificial intelligence into clinical workflows demands that its regulation prioritize patient autonomy, trust, privacy, equity, and human oversight (Teasdale, Mills and Costello, 2024). These factors are essential to ensure that AI technologies enhance healthcare delivery without compromising the foundational values of patient safety and dignity (Zhang *et al.*, 2021).

PATIENT AUTONOMY AND TRUST

To maintain patient autonomy, it is crucial to ensure informed consent and meaningful engagement with patients regarding the use of AI in their care. Patients should be fully informed about how AI technologies operate, their intended role in clinical decision-making, and their limitations (Zezza, 2025). Transparency is key to building trust. Clear communication about the capabilities and boundaries of AI tools enables patients to understand their contributions to diagnosis or treatment, thus empowering them to make educated decisions about their care (Awuah *et al.*, 2024).

PRIVACY AND SECURITY

Protecting patient data in AI systems presents significant challenges, especially as these technologies require access to large volumes of sensitive health information. Robust strategies are needed to safeguard patient privacy while enabling the effective use of data for AI training and deployment (Pesapane *et al.*, 2025). Cybersecurity risks, such as data breaches and unauthorized access, pose a major threat to patient trust and the ethical use of AI in healthcare (Kelly *et al.*, 2023).

EQUITY AND FAIRNESS

AI regulation must also tackle the issue of bias in AI algorithms to ensure equitable access to healthcare technologies (Chiruvella and Guddati, 2021; Kim *et al.*, 2024). Data biases, stemming from unrepresentative training datasets or historical disparities in healthcare delivery, can perpetuate or even exacerbate existing inequities (Chin *et al.*, 2023). Algorithms trained on predominantly affluent or homogenous populations may perform poorly when applied to diverse patient groups, leading to unequal outcomes (Hussain, Bresnahan and Zhuang, 2024). Mitigating bias requires deliberate and ongoing efforts, including the use of diverse training datasets, regular bias testing, and algorithmic debiasing. Equity-focused measures also extend to ensuring that AI technologies are accessible to all populations, regardless of socioeconomic or geographic barriers. Without such efforts, the benefits of AI risk being concentrated among certain groups, leaving underserved populations further marginalized.

HUMAN OVERSIGHT

Maintaining a human-in-the-loop approach is critical to the ethical use of AI in healthcare (Palaniappan *et al.*, 2024). While AI can augment clinical decision-making by providing data-driven insights, it should not replace human judgment. Physicians and healthcare providers must retain ultimate responsibility for patient care, validating AI-generated recommendations and ensuring they align with clinical realities. By mandating human oversight in the development, deployment, and monitoring of AI technologies, regulatory frameworks can ensure that these tools serve as extensions of human expertise rather than replacements for it (Reyes Gil, Pantanowitz and Rashidi, 2024).

DEFINING RISK LEVELS IN AI APPLICATIONS

The first step in risk stratification is categorizing AI risks in healthcare settings. Not all AI applications carry the same level of risk, and distinguishing between high-risk and low-risk interventions allows for tailored regulatory oversight (Ferrara *et al.*, 2024). High-risk applications include autonomous AI systems used in critical diagnostics, such as those used in detecting life-threatening conditions like sepsis or myocardial infarction (García-Gómez, Blanes-Selva and Doñate-Martínez, 2024; Patel *et al.*, 2024). These systems directly impact clinical decisions, and errors can have catastrophic consequences. Conversely, low-risk applications, such as AI tools for administrative purposes like appointment scheduling or resource allocation, have less direct patient impact and, therefore, require less stringent oversight (see Table 1). Regulation must also be tailored to each medical subspecialty; for instance, high-risk interventions include AI-powered diagnostic imaging systems that autonomously interpret mammograms or CT scans utilized in early cancer detection (Freeman *et al.*, 2021; Chustecki, 2024). While these tools hold tremendous potential for improving patient outcomes, any inaccuracies or biases in their algorithms could lead to misdiagnoses, delayed treatment, or unnecessary interventions.

Table 1: Risk stratification of AI in healthcare by human and patient-centred factors.

AI Tasks in Healthcare	Human Factors	Patient-Centred Care Factors
Low Risk		
- Appointment scheduling optimization	- User interface design for staff	- Improved access to care
- Administrative task automation	- Workflow integration	- Reduced wait times
- Patient engagement chatbots	- Patient-AI interaction	- Enhanced communication
Moderate Risk		
- Clinical decision support for non-critical conditions	- Situation awareness	- Personalized care recommendations
- Predictive analytics for hospital resource allocation	- Workload management	- Optimized resource availability
- Remote patient monitoring	- Human-AI team	- Continuous health tracking
High Risk		
- Diagnostic imaging analysis	- Automation bias prevention	- Earlier disease detection
- Autonomous treatment recommendations	- Explanation and trust	- Tailored treatment plans
- Risk stratification for critical care patients	- Ethical considerations	- Improved patient safety

RISK STRATIFICATION AND TRANSPARENCY

Transparent systems enable healthcare providers to critically evaluate AI recommendations, enhancing decision-making and minimizing the likelihood

of errors. Frameworks for transparent reporting of AI performance are critical in this context (Chaurasia *et al.*, 2024). These frameworks should include metrics such as sensitivity, specificity, and error rates, clearly communicated to end-users (Farrell, 2022). This approach ensures that clinicians and stakeholders are aware of the tool's capabilities and constraints, empowering them to make informed decisions about its use in patient care (Morley *et al.*, 2022).

COLLABORATIVE APPROACH

Collaboration between regulators, developers, and healthcare providers is essential for successful risk stratification (Van Buchem *et al.*, 2022). Regulators play a key role in establishing guidelines and standards that define acceptable levels of risk, while developers are responsible for designing AI systems that adhere to these standards (Massella, Dri and Gramaglia, 2022). Healthcare providers, as end-users, contribute valuable insights into the practical challenges of integrating AI into clinical workflows, ensuring that regulatory frameworks address real-world needs.

Patient feedback is another vital component of the risk stratification process (Mökander *et al.*, 2022). Patients are directly impacted by AI technologies and can provide critical perspectives on issues such as transparency, trust, and accessibility (Harvey and Gowda, 2020). Integrating patient voices into the design and deployment of AI systems helps ensure that these technologies align with patient-centred care principles (Labkoff *et al.*, 2024). For instance, incorporating patient advocacy groups in discussions about AI implementation can identify potential barriers to adoption and areas where patient education is needed.

REGULATORY FRAMEWORKS FOR AI

Globally, regulatory efforts seek to address safety, efficacy, and transparency in AI applications (Bogdanoski *et al.*, 2024; Zhou and Gattinger, 2024). However, the rapid pace of AI development often outstrips regulatory adaptation. The World Health Organization's framework outlines essential guidelines for evaluating artificial intelligence-based software as a medical device (AI-SaMD) in healthcare (*Global Strategy on Digital Health 2020–2025*. 1st ed, 2021). It emphasizes the importance of rigorous training, validation, and post-market surveillance to ensure AI tools' safety and effectiveness, particularly in low and middle-income countries (LMICs). The European Union (EU) has implemented the AI Act, which adopts a risk-based framework for AI regulation (Schmidt *et al.*, 2024). This framework categorizes AI systems by their potential risk level, imposing strict requirements for high-risk applications, such as those used in medical diagnostics or decision-making.

In the United States, the Food and Drug Administration (FDA) plays a central role in regulating AI as a medical device (AIaMD) (Perlis and Abbasi, 2024). The FDA has introduced guidelines for AI technologies, emphasizing pre-market evaluation, real-world performance monitoring,

and post-market surveillance. Notably, the FDA's adaptive approach to regulation accommodates the unique characteristics of continuously learning AI systems, ensuring they remain safe and effective over time (Harvey and Gowda, 2020). In the UK the Medicines and Healthcare Products Regulatory Agency (MHRA), which has implemented an "AI Airlock" pilot scheme aims to streamline their route to market while ensuring safety and efficacy (Medicines and Healthcare products Regulatory Agency, 2024). In Australia, there is no dedicated AI legislation, instead voluntary AI Ethics Principles guide AI development while policymakers consider regulatory reforms (Chau, 2024). These international examples highlight the importance of tailoring regulatory practices to regional healthcare needs while maintaining universal principles of safety, transparency, and accountability.

POSITIONING CANADA AS A GLOBAL LEADER

Canada has a unique opportunity to set benchmarks for ethical AI integration, leveraging its strong healthcare infrastructure and commitment to patient-centred care. By adopting a leadership role in AI regulation, Canada can establish itself as a model for other nations navigating the complexities of AI in healthcare. This requires the development of comprehensive national standards that prioritize ethical considerations, such as equity, transparency, and accountability. Collaboration with international regulatory bodies is another critical step in positioning Canada as a global leader (Silva *et al.*, 2022). By aligning its regulatory frameworks with those of other nations, Canada can facilitate the cross-border exchange of AI technologies and expertise. Joint initiatives, such as harmonized standards for AI performance evaluation or shared databases for training algorithms, can accelerate the development and deployment of safe and effective AI systems. These efforts must address regulatory fragmentation while prioritizing patient safety and system transparency (The College of Physicians and Surgeons of Manitoba, 2024).

CONCLUSION

The integration of artificial intelligence into healthcare presents a transformative opportunity, but its success depends on placing patients at the center of regulatory and implementation efforts. Throughout this discussion, the importance of prioritizing patients over processes has been emphasized, highlighting the critical need to address gaps in regulation while fostering transparency, accountability, and equity in AI applications. By focusing on patient-centred factors such as autonomy, trust, privacy, and fairness, AI regulation can ensure that technological advancements align with the fundamental values of healthcare.

Risk stratification strategies are vital in navigating the complexities of AI deployment, enabling tailored oversight based on the potential impact of different AI systems. Transparent frameworks further reinforce trust among stakeholders by providing clarity on AI performance and limitations. These elements are foundational to creating a regulatory environment

that promotes patient safety while fostering innovation. Establishing uniform national standards will provide a cohesive framework for AI adoption, ensuring that all stakeholders operate under consistent and enforceable guidelines. Policymakers must prioritize the development of ethical and patient-centred regulatory frameworks that balance innovation with accountability. The path forward for AI in healthcare requires a commitment to ethical, patient-centred regulation that not only addresses current challenges but also anticipates future needs. By prioritizing patients over processes, healthcare systems can harness the full potential of AI to deliver safer, more effective, and more equitable care.

REFERENCES

Awuah, W. A. et al. (2024) 'Empowering health care consumers & understanding patients' perspectives on AI integration in oncology and surgery: A perspective', Health Science Reports, 7(7), p. e2268.

Bogdanoski, G. et al. (2024) 'Translating the regulatory landscape of medical devices to create fit-for-purpose artificial intelligence (AI) cytometry solutions', Cytometry Part B: Clinical Cytometry, 106(4), pp. 294–307.

Chau, M. (2024) 'Ethical, legal, and regulatory landscape of artificial intelligence in Australian healthcare and ethical integration in radiography: A narrative review', Journal of Medical Imaging and Radiation Sciences, 55(4), p. 101733.

Chaurasia, A. et al. (2024) 'Use of artificial intelligence in obstetric and gynaecological diagnostics: A protocol for a systematic review and meta-analysis', BMJ Open, 14(5), p. e082287.

Chin, M. H. et al. (2023) 'Guiding Principles to Address the Impact of Algorithm Bias on Racial and Ethnic Disparities in Health and Health Care', JAMA Network Open, 6(12), p. e2345050.

Chiruvella, V. and Guddati, A. K. (2021) 'Ethical Issues in Patient Data Ownership', Interactive Journal of Medical Research, 10(2), p. e22269.

Chustecki, M. (2024) 'Benefits and Risks of AI in Health Care: Narrative Review', Interactive Journal of Medical Research, 13, p. e53616.

Cobianchi, L. et al. (2022) 'Artificial Intelligence and Surgery: Ethical Dilemmas and Open Issues', Journal of the American College of Surgeons, 235(2), pp. 268–275.

Dossabhoy, S. S. et al. (2023) 'Artificial intelligence in clinical workflow processes in vascular surgery and beyond', Seminars in Vascular Surgery, 36(3), pp. 401–412.

Farrell, C.-J. L. (2022) 'Decision support or autonomous artificial intelligence? The case of wrong blood in tube errors', Clinical Chemistry and Laboratory Medicine (CCLM), 60(12), pp. 1993–1997.

Ferrara, M. et al. (2024) 'Risk Management and Patient Safety in the Artificial Intelligence Era: A Systematic Review', Healthcare, 12(5), p. 549.

Finkelstein, J. et al. (2024) 'Identifying Facilitators and Barriers to Implementation of AI-Assisted Clinical Decision Support in an Electronic Health Record System', Journal of Medical Systems, 48(1), p. 89.

Freeman, K. et al. (2021) 'Use of artificial intelligence for image analysis in breast cancer screening programmes: Systematic review of test accuracy', BMJ, p. n1872.

García-Gómez, J. M., Blanes-Selva, V. and Doñate-Martínez, A. (2024) 'Proposing an AI Passport as a Mitigating Action of Risk Associated to Artificial Intelligence in Healthcare', in J. Mantas et al. (eds) Studies in Health Technology and Informatics. IOS Press.

Global Strategy on Digital Health 2020–2025. 1st ed (2021). Geneva: World Health Organization.

Harvey, H. B. and Gowda, V. (2020) 'How the FDA Regulates AI', Academic Radiology, 27(1), pp. 58–61.

Hussain, S. A., Bresnahan, M. and Zhuang, J. (2024) 'The bias algorithm: How AI in healthcare exacerbates ethnic and racial disparities – a scoping review', Ethnicity & Health, pp. 1–18.

Kelly, B. S. et al. (2023) 'Cybersecurity considerations for radiology departments involved with artificial intelligence', European Radiology, 33(12), pp. 8833–8841.

Kim, J. Y. et al. (2024) 'Development and preliminary testing of Health Equity Across the AI Lifecycle (HEAAL): A framework for healthcare delivery organizations to mitigate the risk of AI solutions worsening health inequities', PLOS Digital Health. Edited by G. Guillot, 3(5), p. e0000390.

Labkoff, S. et al. (2024) 'Toward a responsible future: Recommendations for AI-enabled clinical decision support', Journal of the American Medical Informatics Association, 31(11), pp. 2730–2739.

Larson, D. B. et al. (2021) 'Regulatory Frameworks for Development and Evaluation of Artificial Intelligence–Based Diagnostic Imaging Algorithms: Summary and Recommendations', Journal of the American College of Radiology, 18(3), pp. 413–424.

Massella, M., Dri, D. A. and Gramaglia, D. (2022) 'Regulatory Considerations on the use of Machine Learning based tools in Clinical Trials', Health and Technology, 12(6), pp. 1085–1096.

Medicines and Healthcare Products Regulatory Agency (2024) 'AI Airlock: The regulatory sandbox for AIaMD'.

Mökander, J. et al. (2022) 'Conformity Assessments and Post-market Monitoring: A Guide to the Role of Auditing in the Proposed European AI Regulation', Minds and Machines, 32(2), pp. 241–268.

Morley, J. et al. (2022) 'Governing Data and Artificial Intelligence for Health Care: Developing an International Understanding', JMIR Formative Research, 6(1), p. e31623.

Palaniappan, K. et al. (2024) 'Gaps in the Global Regulatory Frameworks for the Use of Artificial Intelligence (AI) in the Healthcare Services Sector and Key Recommendations', Healthcare, 12(17), p. 1730.

Patel, S. J. et al. (2024) 'Advancements in Artificial Intelligence for Precision Diagnosis and Treatment of Myocardial Infarction: A Comprehensive Review of Clinical Trials and Randomized Controlled Trials', Cureus [Preprint].

Perlis, R. and Abbasi, J. (2024) 'FDA Commissioner Robert Califf on Setting Guardrails for AI in Health Care', JAMA, 332(23), p. 1963.

Pesapane, F. et al. (2025) 'The translation of in-house imaging AI research into a medical device ensuring ethical and regulatory integrity', European Journal of Radiology, 182, p. 111852.

Reddy, S. et al. (2021) 'Evaluation framework to guide implementation of AI systems into healthcare settings', BMJ Health & Care Informatics, 28(1), p. e100444.

Reyes Gil, M., Pantanowitz, J. and Rashidi, H. H. (2024) 'Venous thromboembolism in the era of machine learning and artificial intelligence in medicine', Thrombosis Research, 242, p. 109121.

Schmidt, J. et al. (2024) 'Mapping the regulatory landscape for artificial intelligence in health within the European Union', npj Digital Medicine, 7(1), p. 229.

Shevtsova, D. et al. (2024) 'Trust in and Acceptance of Artificial Intelligence Applications in Medicine: Mixed Methods Study', JMIR Human Factors, 11, p. e47031.
Silva, M. et al. (2022) 'Regulating the Safety of Health-Related Artificial Intelligence', Healthcare Policy | Politiques de Santé, 17(4), pp. 63–77.
Spear, J., Ehrenfeld, J. M. and Miller, B. J. (2023) 'Applications of Artificial Intelligence in Health Care Delivery', Journal of Medical Systems, 47(1), p. 121.
Teasdale, A., Mills, L. and Costello, R. (2024) 'Artificial Intelligence-Powered Surgical Consent: Patient Insights', Cureus [Preprint].
The College of Physicians and Surgeons of Manitoba (2024) 'Advice to the Profession: Responsible use of Artificial Intelligence in the Practice of Medicine'.
Van Buchem, M. M. et al. (2022) 'Analyzing patient experiences using natural language processing: development and validation of the artificial intelligence patient reported experience measure (AI-PREM)', BMC Medical Informatics and Decision Making, 22(1), p. 183.
Wadden, J. J. (2022) 'Defining the undefinable: the black box problem in healthcare artificial intelligence', Journal of Medical Ethics, 48(10), pp. 764–768.
Zezza, M. (2025) 'The promise of AI in healthcare: transforming communication and decision-making for patients', Journal of Communication in Healthcare, pp. 1–4.
Zhang, Z. et al. (2021) 'Effect of AI Explanations on Human Perceptions of Patient-Facing AI-Powered Healthcare Systems', Journal of Medical Systems, 45(6), p. 64.
Zhou, K. and Gattinger, G. (2024) 'The Evolving Regulatory Paradigm of AI in MedTech: A Review of Perspectives and Where We Are Today', Therapeutic Innovation & Regulatory Science, 58(3), pp. 456–464.