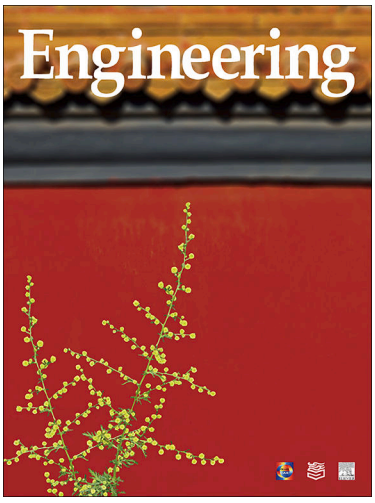
Journal Pre-proofs



Perspectives

Artificial Intelligence Empowers Regulatory Compliance Science for Medical Devices: From Premarket Approval to Post-market Surveillance, How Far Are We There?

Yu-Sheng Hsueh, Regis J. O'Keefe, Stuart Goodman, Péter Hegyi, R. Geoff Richards, Minghao Zheng, Wenxue Tong, Jiankun Xu, Christer Rolf, Patrick Shu-Hang Yung, Ling Qin, MRN working group,

PII: S2095-8099(26)00611-9
DOI: https://doi.org/10.1016/j.eng.2026.10.012
Reference: ENG 2495

To appear in: *Engineering*

Received Date: 22 April 2026
Revised Date: 21 August 2026
Accepted Date: 26 October 2026

Please cite this article as: Y-S. Hsueh, R.J. O'Keefe, S. Goodman, P. Hegyi, R. Geoff Richards, M. Zheng, W. Tong, J. Xu, C. Rolf, P.S-H. Yung, L. Qin, MRN working group, Artificial Intelligence Empowers Regulatory Compliance Science for Medical Devices: From Premarket Approval to Post-market Surveillance, How Far Are We There?, *Engineering* (2026), doi: https://doi.org/10.1016/j.eng.2026.10.012

This is a PDF of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability. This version will undergo additional copyediting, typesetting and review before it is published in its final form. As such, this version is no longer the Accepted Manuscript, but it is not yet the definitive Version of Record; we are providing this early version to give early visibility of the article. Please note that Elsevier's sharing policy for the Published Journal Article applies to this version, see: https://www.elsevier.com/about/policies-and-standards/sharing#4-published-journal-article. Please also note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2026 THE AUTHORS. Published by Elsevier LTD on behalf of Chinese Academy of Engineering and Higher Education Press Limited Company

Journal Pre-proofs

Research

AI, BCI, and Engineering—Perspective

Artificial Intelligence Empowers Regulatory Compliance Science for Medical Devices: From Premarket Approval to Post-market Surveillance, How Far Are We There?

Yu-Sheng Hsueh [a,b], Regis J O'Keefe [c], Stuart Goodman [d], Péter Hegyi [e], R. Geoff Richards [f], Minghao Zheng [g], Wenxue Tong [a,b], Jiankun Xu [a,b], Christer Rolf [h], Patrick Shu-Hang Yung [a,*], Ling Qin [a,b,*], MRN working group [i]

[a] Department of Orthopaedics and Traumatology, The Chinese University of Hong Kong, Hong Kong 999077, China

[b] AoE Centre for Musculoskeletal Degeneration and Regeneration, Department of Orthopaedics & Traumatology, Faculty of Medicine, The Chinese University of Hong Kong, Hong Kong 999077, China

[c] Department of Orthopaedic Surgery, School of Medicine, Washington University, St. Louis, MO 63110, USA

[d] Department of Orthopedic Surgery, Stanford Medicine, Stanford University, Redwood City, CA 94305, USA

[e] Centre for Translational Medicine, Semmelweis University, Budapest 1085, Hungary

[f] AO Research Institute Davos, Davos 7270, Switzerland

[g] Perron Institute for Neurological and Translational Science, University of Western Australia, Nedlands, WA 6009, Australia

[h] Department of Clinical Science, Intervention and Technology, Karolinska Institute, Huddinge SE-141 52, Sweden

[i] The Musculoskeletal Regeneration Research Network (MRN) Working Group

*Corresponding authors.

E-mail addresses: lingqin@cuhk.edu.hk (L. Qin), patrickyung@cuhk.edu.hk (P.S.-H. Yung).

ARTICLE INFO

Article history:

Received 22 Apr 2026

2095-8099/© 2026 THE AUTHORS. Published by Elsevier LTD on behalf of Chinese Academy of Engineering and Higher Education Press Limited Company. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Journal Pre-proofs

Revised 21 August 2026

Journal Pre-proofs

Accepted 26 October 2026

Available online

Keywords

Artificial intelligence governance

Review and approval

Medical device

Regulatory authority

Clinical compliance

Abstract

While artificial intelligence (AI) in medicine has been extensively studied as a medical device, its role as a governance tool for regulatory authorities remains largely unexplored. The rapid influx of innovative, AI-integrated medical devices has revealed critical bottlenecks in the current human-designed regulatory system (Hud-RS), including approval delays and resource constraints. In this work, we conceptualize a comprehensive dynamic framework for regulatory modernization, termed the “Artificial Intelligence Governance Regulatory System (Aig-RS).” The framework covers the full product lifecycle, from pre-clinical and clinical review to post-market surveillance. We examine current practices across regulatory agencies and propose strategies to optimize existing measures to streamline review processes. Although the US Food and Drug Administration (FDA) has made significant advances with its agency-wide tool (enterprise language support assistant (ELSA)), timely translation of innovative medical devices into clinical practice remains limited. Accordingly, Aig-RS is intended to accelerate regulatory decision-making safely by leveraging AI and to clarify how regulatory agencies can implement AI across the review continuum.

1. Background

The rapid influx of innovative, highly complex artificial intelligence (AI) medical devices has far outpaced the capacity of the current lagging, paper-based, and human-designed regulatory system (Hud-RS) responsible for overseeing them [1]. Many researchers encounter complex regulatory hurdles in the development of medical devices. These challenges may arise from classification, intended use, applicable standards, guidelines, and compliance requirements. A significant gap in regulatory transparency and available pathways persists. Although

Journal Pre-proofs

regulatory authorities (RAs) are transitioning to electronic documentation, many still rely on paper-based procedures to review devices. The focus of AI in medical devices is often placed on products rather than on the RAs that oversee them. In 1976, the US Food and Drug Administration (FDA) initiated a premarket review and approval process for medical devices. The European Union (EU) enacted the Medical Device Regulation (MDR) (Regulation 2017/745) and the *In Vitro* Diagnostic Medical Device Regulation (IVDR) (Regulation 2017/746) with the goal of modernizing the regulatory system from three legacy directives [2].

As of March 4, 2026, the FDA lists at least 1430 AI and machine learning (AI/ML)-enabled devices in its official announcement, “Artificial Intelligence-Enabled Medical Devices.” More than 500 000 medical devices circulate in the European market. Although these devices are categorized into different subcategories, their sheer volume poses challenges in establishing unified regulatory management standards [2].

While regulatory frameworks across jurisdictions continue to evolve, the complexity of Hud-RS continues to affect the final approval of these devices, with considerable room for improvement [3,4]. Furthermore, appraisers’ principles outside their areas of expertise may lead to judgments that carry assessment biases [5,6]. RAs have devised strategies to enhance regulatory ecosystems and explore the feasibility of AI [1,7,8].

AI has been deployed as medical devices (AIaMDs) [9,10]. The FDA list of 1430 AIaMDs primarily focuses on diagnosis and prognosis (e.g., AI imaging software in radiology and neurology). More than 80% of these devices have obtained clearance through the premarket notification 510(k) pathway [11]. In Europe, the European Medicines Agency (EMA) relies on certified notified bodies (NBs) for approval, and over 140 AI/ML devices receive Conformité Européenne (CE) markings or MDR approval. Currently, most regulatory agencies rely on paper-based workflows, owing to the rapid advancement of AI. As AI reshapes the healthcare industry and becomes a key driver of translational medicine, the concept of AI governance regulatory system (Aig-RS) offers an agile pathway for reforming existing regulatory processes, as depicted in Fig. 1. This concept also explores the potential of AI governance to facilitate market access, transforming from traditional review processes to total product life cycle (TPLC), and ultimately advancing regulatory modernization [7,12,13]. Unlike the existing literature, which largely focuses on AI as an AI/ML medical device, our novel approach integrates regulatory practice with scientific and AI methods. In doing so, it provides a framework for regulatory agencies to modernize the regulatory system itself as well as the specific devices they regulate.

Journal Pre-proofs

Journal Pre-proofs

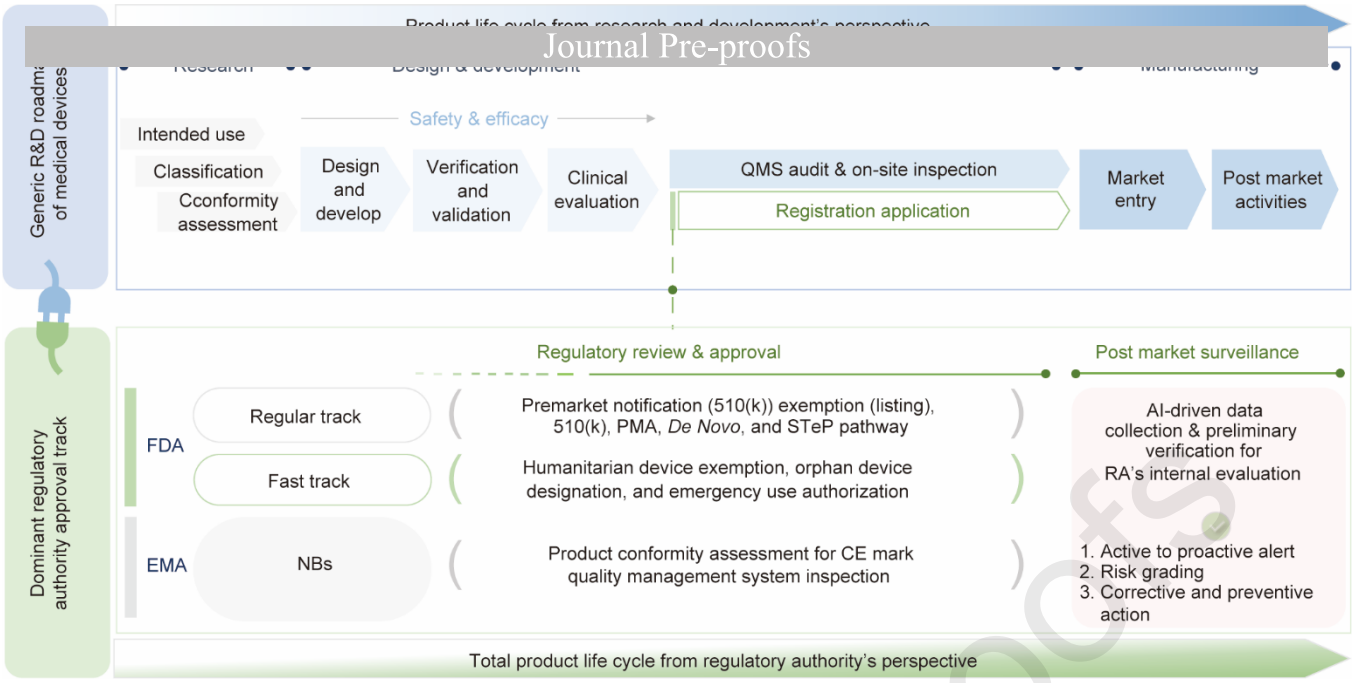


Fig. 1. Modern product life cycle from the perspectives of RA and research and development (R&D) workflows and stages. It provides an overview of the medical device life cycle, highlighting potential areas for AI deployment. PMA: premarket approval; STeP: safer technologies program; NMPA: the National Medical Products Administration of the People's Republic of China; QMS: quality management system.

2. Current Hud-RS bottlenecks to Aig-RS

A comparative regulatory roadmap illustrates the differences between the centralized (FDA and the National Medical Products Administration of the People's Republic of China (NMPA)) and decentralized (EMA/NB) approaches, as shown in Fig. 2. The most commonly used FDA pathway is the 510(k) application (e.g., titanium fixation screw), which requires substantial equivalence to an existing legally marketed device. Innovative devices, such as those classified under the *De Novo* process (e.g., biodegradable metal implants), often require additional evaluations to determine their classification, resulting in relatively longer review times. Class III devices (e.g., active implantable pacemakers) requiring premarket approval (PMA) typically take longer [14] and comply with the US §814.45 of the Code of Federal Regulations (CFR), with iterative medical device user fee amendments (MDUFA) communications posing further delay, particularly when submitted data are deemed unacceptable. Determining the authenticity of testing certificates during technical examinations is a key issue, particularly when suspicious documents are involved.

Journal Pre-proofs

Journal Pre-proofs

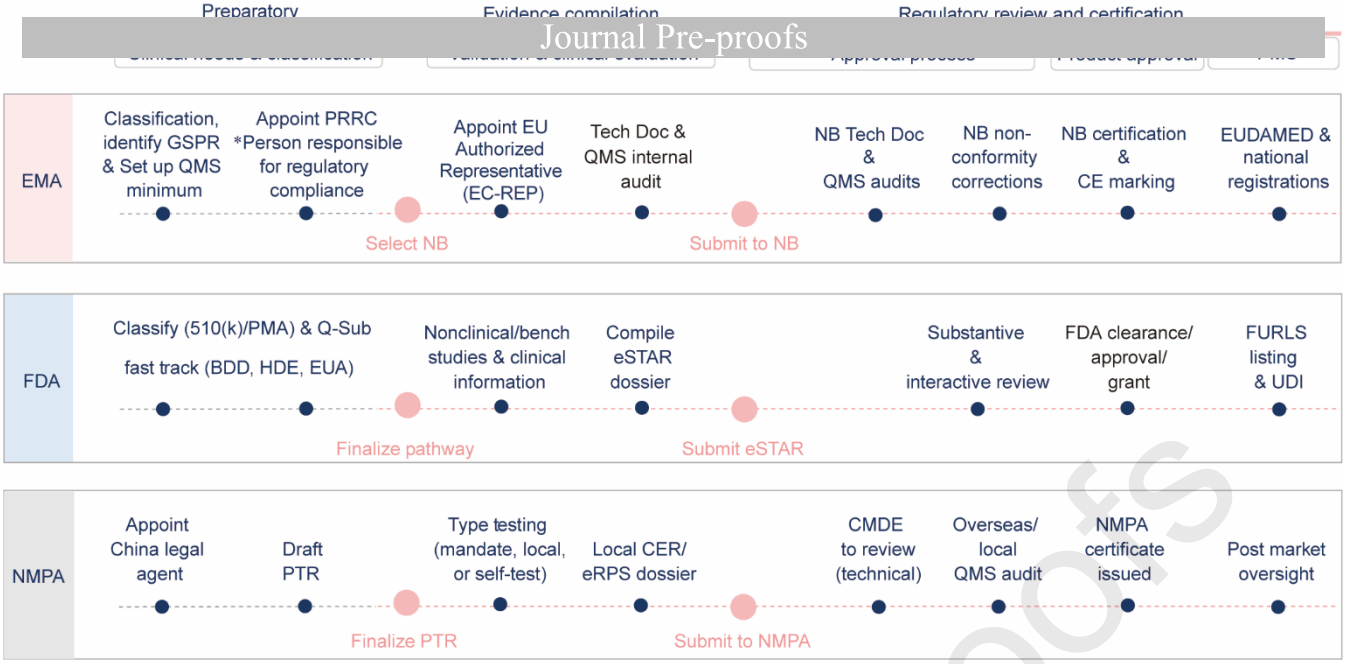


Fig. 2. Comparative regulatory roadmap for medical device approval across the EMA, FDA, and NMPA. The diagram outlines the complete device certification workflow, structured into three main stages: preparation, clinical and technical evidence compilation, and regulatory assessment and authorization. The black nodes represent standard routine procedures, while the pink nodes denote critical decision gateways throughout the process. GSPR: general safety and performance requirements ; PRRC: person responsible for regulatory compliance; EUDAMED: European Database for Medical Devices; Q-Sub: Q-submission; BDD: breakthrough device designation ; HDE: humanitarian device exemption; EUA: Emergency Use Authorization; eSTAR: electronic submission template and resource ; FURLS: FDA unified registration and listing system ; UDI: unique device identification; PTR: product technical requirements; CER: clinical evaluation report; eRPS: regulated product submission; CMDE: Center for Medical Device Evaluation; Tech Doc: technical documentation.

Under the Hud-RS system, accelerated regulatory pathways, including priority review, breakthrough device designation, humanitarian device exemption (HDE), and orphan device designation, are expected to decrease timelines [15]. However, the eligibility criteria limit the scope of use of many medical devices. According to the MDUFA review timeline (Fig. 3), PMA had the longest review time. Notably, the review time for *De Novo* devices is similar to that for PMA, which may be attributed to the classification process, despite *De Novo* devices being categorized as having a low to moderate risk [16] (Fig. 3(a)). Under the EU MDR/NB framework, the constrained review capacity among NBs has led to significant delays in the assessment and certification of medical devices [17], as illustrated in Fig. 3(c) and documented in recent EU NB surveys [18,19]. Moreover, the timelines shown in Fig. 3(d) are illustrative ranges synthesized from public MDR technical documentation and conformity assessment materials provided by three large NBs: British Standards Institution (BSI), Technischer Überwachungsverein Süd (TÜV SÜD), and Technischer Überwachungsverein Rheinland (TÜV Rheinland). To complement the aggregate EU certification timeline distribution shown in Fig. 3(c), Fig. 3(d) provides a risk-class-specific breakdown of NB certification duration using these representative timelines. Collectively, Figs. 3(c) and

Journal Pre-proofs

(d) identify the key regulatory challenges that an AI-governed framework should be designed to address, including approval delays, certificate authenticity risks, and systemic resource constraints.

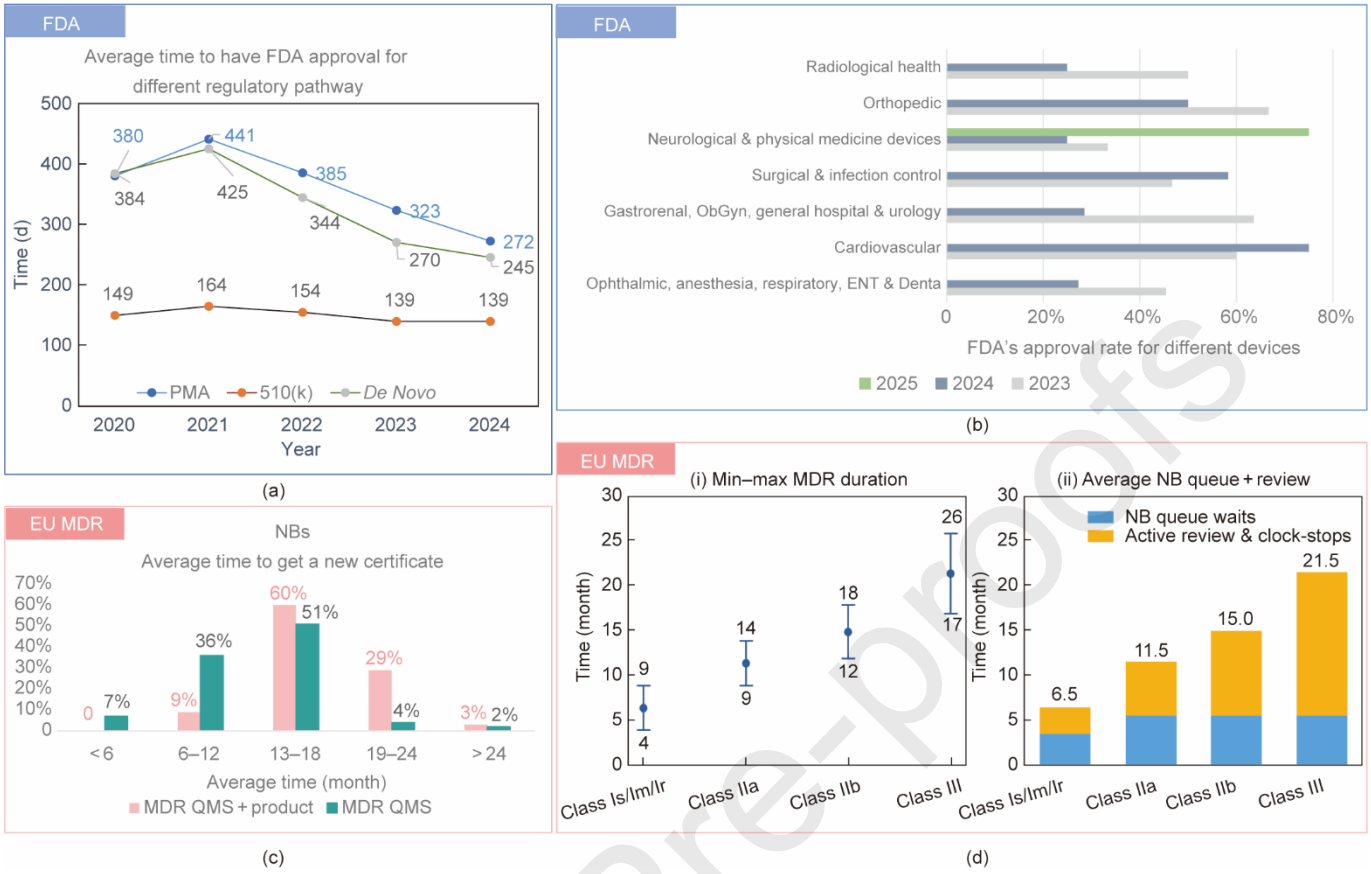


Fig. 3. Comparison of regulatory review timelines and approval rates for medical devices across two principal regulatory frameworks. (a) Mean total time (in days) for the FDA to reach a decision under the PMA, 510(k), and *De Novo* pathways between 2020 and 2024. (b) FDA approval rates by medical device specialty between 2023 and 2025. (c) Distribution of the average time required to obtain a new certificate from an EU designated NB under the EU MDR, stratified by QMS assessment alone versus combined QMS plus product assessment. *Y*-axis: percentage of total number of NB that have issued certificates. (d) NB certification timelines under EU MDR by device risk class. (i) The minimum and maximum total certification durations for Class Is (sterile)/Im (measuring)/Ir (reusable) devices, Class IIa, Class IIb (non-implantable/standard), and Class III (implantable) devices. (ii) Total certification time into NB queue waiting periods and active review time, including clock-stops for each risk class, with total duration indicated above each bar. All values are derived from representative publicly available timelines and conformity assessment documentation from BSI, TÜV SÜD, and TÜV Rheinland. QbGyn: obstetrics and gynecology; ENT: ear, nose, and throat; Denta: denta/dentistry.

Therefore, we propose an Aig-RS, as shown in Fig. 4. This framework outlines an approach for regulatory acceleration that encompasses three key phases: classification, preclinical compliance with new approach methodologies (NAMs) matching, and clinical compliance. In the preclinical compliance phase, AI technologies are integrated to validate study protocols and testing criteria such as biocompatibility, imaging assessments, and electromagnetic compatibility (EMC). The clinical phase leverages AI to support the literature review, clinical trial

Journal Pre-proofs

design and real-world data analysis to meet regulatory requirements. It presents two models: Model 1 applies AI at specific stages, whereas Model 2 implements a comprehensive end-to-end AI technology review. Model 1 represents a human-in-the-loop (HITL) staged approach that uses modular AI tools. Model 2 provides a fully automated pipeline for low-risk devices. Neither model relied on a single deep learning network. Instead, they integrate multiple specialized algorithms such as decision trees, natural language processing (NLP), anomaly detection, and computer vision. This included a supervised NLP model with a fine-tuned bidirectional encoder representations from transformers (BERT)/decision tree for classification. The performance of each algorithm can be tested by comparing it with human reviewer benchmarks using metrics such as the F1-score, which helps ensure the robustness and reliability of the evaluation results. See Case Study Box 1 for AI classification profiles.

Case study box 1. Technical specification for an AI classification module.

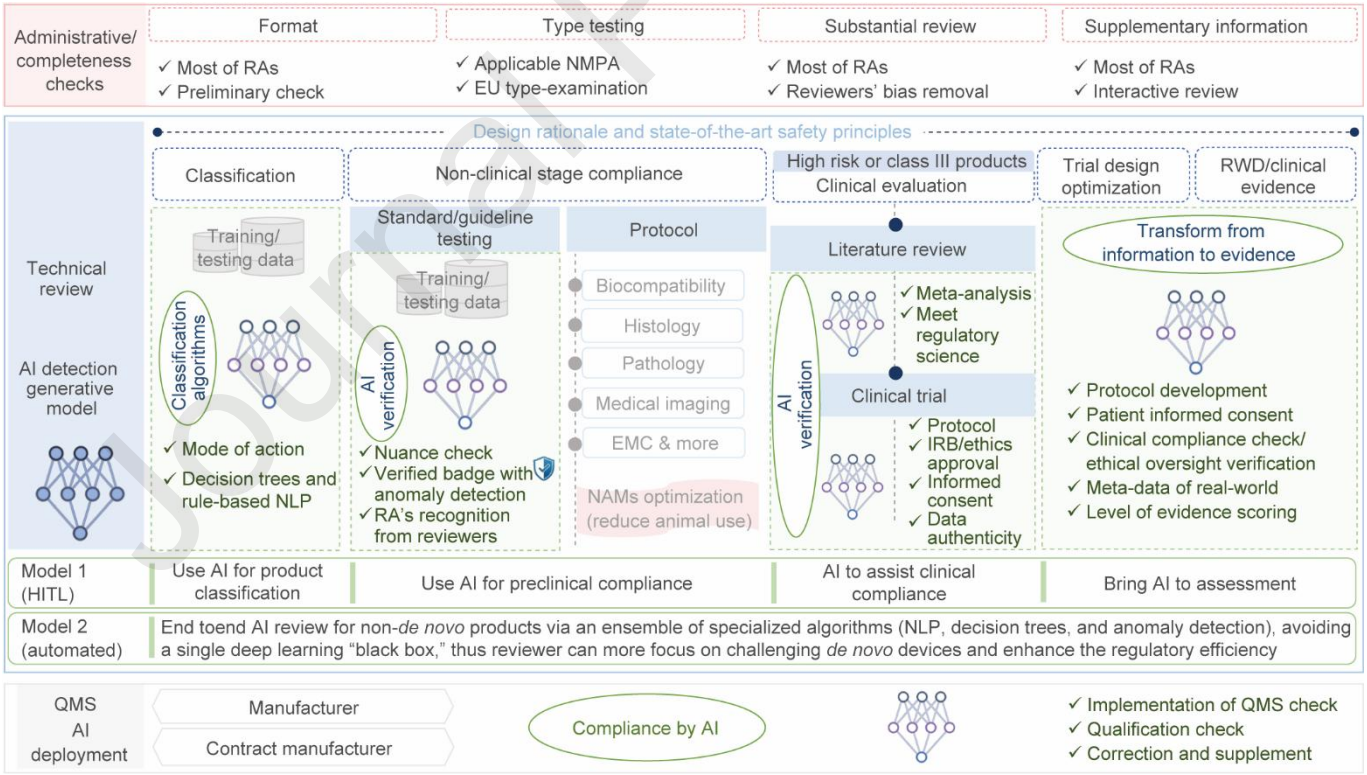
1. Input data types: structured metadata and unstructured narrative text from regulatory submissions (e.g., indications, device description)
2. Preprocessing pipeline: optical character recognition → text normalization → tokenization (readable pieces) → embedding
3. Model architecture: fine-tuned BERT-base encoder with a downstream decision tree classifier for assigning regulatory category
4. Proposed training dataset: public data (e.g., FDA 510(k)/PMA) decision summaries and related documentation used as a surrogate labelled corpus
5. Evaluation metrics: F1-score, precision, and recall compared against a human reviewer considered as gold standard, as recommended for medical AI evaluation

Journal Pre-proofs

robustness

Remark: conceptual technical specification intended to guide future collaboration with public sector/regulatory bodies; not yet implemented or validated in a live review setting

Regarding the role of ethics committees, in Model 1, NLP automatically extracts and verifies ethics committee approval letters, flagging any discrepancies for human reviewers. In Model 2, the automated system must include a dedicated ethics verification module when processing clinical evidence. This module verifies the digital signatures of the Institutional Review Board (IRB) or Independent Ethics Committee (IEC) and ensures that the clinical trial registration number (e.g., ClinicalTrials.gov) matches the corresponding ethics approval documents. The automated review will proceed only after these checks are successfully confirmed. Furthermore, the quality management system (QMS) incorporates AI for compliance management and production oversight.

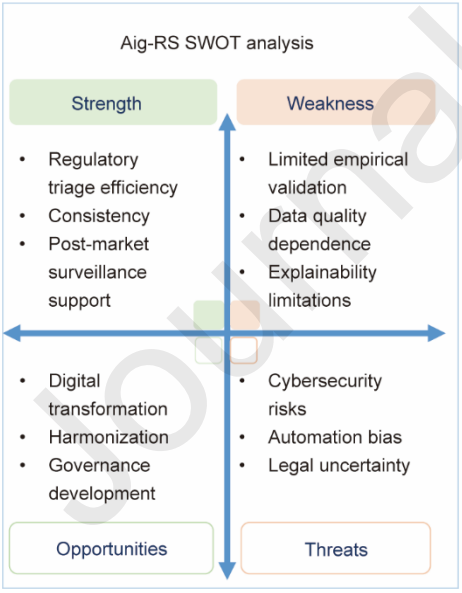


Journal Pre-proofs

Fig. 4. Comparison of the Hud-RS and the Aig-RS in terms of regulatory workflow automation, data training, deep learning, and AI decision-making. Model 1 includes administrative and completeness checks, literature review & meta-analysis, NAMs, and human-in-the-loop (HITL) process. Because the final decision is made by a human reviewer, this model avoids the major legal and ethical challenges associated with fully automated decision-making. This proposed architecture also avoids relying on a single deep learning "black box." Instead, it uses an ensemble of specialized algorithms, such as NLP, decision trees, anomaly detection, and computer vision. Model 2, by contrast, envisions a highly automated pipeline designed for low-risk devices. NLP: natural language processing; EMC: electromagnetic compatibility; NAMs: new approach methodologies; IRB: institutional review board; RWD: real-world data.

While empirical data on Aig-RS implementation is still emerging, the FDA's enterprise language support assistant (ELSA) pilot indicates that automation may deliver meaningful benefits [13,20]. As summarized in Fig. 5, transitioning from a Hud-RS to an Aig-RS model is estimated to reduce administrative time and lessen overall review-related logistics burdens. Consequently, this approach can result in substantial cost savings for manufacturers and improve patient accessibility.

Consider a Class II AI/ML software-as-a-medical-device (SaMD) that is submitted via the FDA's 510(k) pathway. Under the Hud-RS framework, review timelines often exceed 90 days because reviewers must conduct side-by-side predicate comparisons. In contrast, in Aig-RS, the system uses NLP technology to cross-reference the device under review using the FDA database of predicate devices. This reduces both the administrative and technical review times.



Strategic situation	Integrated SWOT category	Strategy for Aig-RS
Maxi–Maxi (S–O)	Strengths + opportunities	Leverage AI-enabled triage, NLP-based document screening, explainable decision support, and real-world data analytics to accelerate regulatory review and align with ongoing digital transformation in regulatory agencies
Mini–Maxi (W–O)	Weaknesses + opportunities	Use emerging public regulatory databases, pilot initiatives, and harmonization efforts to address the current lack of empirical validation, limited benchmarking, and the need for standardized datasets and governance training
Maxi–Mini (S–T)	Strengths + threats	Deploy modular AI architecture, HITL oversight, and explainable AI-based traceability to mitigate automation bias, black-box risks, and cybersecurity threats while keeping legal accountability
Mini–Mini (W–T)	Weaknesses + threats	Restrict early deployment to low-risk administrative tasks and phased regulatory sandbox testing to minimize weaknesses in validation, transparency, and infrastructure before confronting broader cybersecurity and legal uncertainties

Fig. 5. Strengths, weaknesses, opportunities, and threats (SWOT) analysis of the proposed Aig-RS for medical device approval.

NB certification lengthens the overall certification timeline, with over 92% of applications taking 13–18 months for approval. Notably, the processing time of the manufacturers accounted for 58% of the total duration. The main reasons for refusals include “applications not complete” (39%) and “outside the scope of NB designation” (26%) [18]. NBs prioritize devices with high public health importance and significant innovation [21]. In China, the NMPA issued a revised “Special Review Procedure for Innovative Medical Devices” in 2018 to accelerate translation and innovation. This revision aims to enhance the review and approval processes without merely increasing the human power involved, as indicated in the NMPA’s response to proposal No. 5113 of the National People’s Congress of the People’s Republic of China [22]. AI can supplement professional manpower and accelerate the process, particularly in the face of budget cuts.

2.1. AI-driven classification: Resolving jurisdictional complexity

Classification is a prerequisite for determining regulatory pathways, and its complexity should not be underestimated. Although risk classification has been largely harmonized globally via the International Medical Device Regulators Forum (IMDRF), important differences still exist across jurisdictions at the level of specific regulatory requirements. Many device developers fail to prioritize classification early in the development process, which can create significant regulatory hurdles during product translation. For example, compared to the FDA guidelines, Annex VIII and Rule 11 of the EU MDR can systematically reclassify many software applications, such as SaMD, into higher-risk categories. Furthermore, the FDA’s centralized agency model contrasts sharply with the EU’s decentralized NB framework.

The EMA established decision tree guidelines under the MDR framework, with NBs leading to CE marking compliance. An increasing number of RAs have published online classification rules such as the FDA’s AI Software Navigator and the Hong Kong’s Medical Device Classification Online Tool [23,24].

2.2. Preclinical compliance: From animal models to AI-assisted NAMs

The preliminary safety and efficacy of novel devices are critical and are typically evaluated through biological assessment *in vitro* and *in vivo* as a proof-of-concept prior to human use [25]. Animal models from rodents for early safety screening of larger animals (e.g., pigs, dogs, and sheep) for cardiovascular, orthopedic, or implantable devices provide indispensable biological validation before human trials [26–28]. However, variability in animal physiology, disease progression, and age introduces translational uncertainty [29]. Furthermore, animal models often fail to accurately replicate human conditions owing to differences in organ metabolism, immune function, and tissue morphology.

To address these limitations, NAMs have been developed, including *in silico* models, multiple organoids, and organ-on-a-chip systems. These models simulate human physiology and disease, offering animal welfare, ethical, and cost-effective options. However, challenges persist in scaling and standardizing organoids and organs-on-a-

Journal Pre-proofs

chip, including issues related to reproducibility, vascularization, nutrient uptake, and biological maturity. Moreover, the lack of consensus on protocols and validation methods limits their widespread adoption.

We suggest that RAs consider animal testing appropriate for evaluating the efficacy and consistency of safety. By leveraging NAMs, the FDA is advancing its goal of promoting the animal reduce, refine, replace (3R) rule under FDA Modernization 2.0 [30]. However, this primarily entails the removal of the requirement for animal evidence as part of the process of obtaining a medicinal approval license for biological products. International standards for organoids and organ-on-a-chip remain under development (ISO/AWI 25430-1, ISO/WD 25693, and ISO/WD 25448 as of July 2025) [31]. Similar to the benefit-risk analysis in AI/ML medical devices, AI training datasets should comply with internationally harmonized standards/guidelines (e.g., International Organization for Standardization (ISO)) concerning efficacy, safety, and risk [32]. AI has the potential to integrate vast datasets to predict human responses to substances, thereby improving the reliability of toxicity and efficacy tests. AI is used to supplement and improve the use of animals rather than to function as a replacement [33]. Moreover, it can integrate data from NAM technologies, animal models, and clinical studies to improve predictive accuracy. This holds potential for evaluating model construction, research design, and ethical compliance based on internal training data [34]. We can consider the evaluation of osseointegration of novel orthopedic implants, for example. Currently, RAs rely heavily on sheep models; however, the physiological differences between ovine and human bone remodeling often lead to limited clinical translation. An AI-assisted cross-species prediction model could help address this gap. By inputting multi-omics and transcriptomic data from ovine models with detailed information on the material properties of the implant, a trained neural network can output the expected human inflammatory and osteogenic responses. Validation of such a model requires retrospective benchmarking against historical clinical trial data to demonstrate its ability to reliably bridge interspecies physiological differences. Furthermore, “organs-on-a-chip” systems offer the capability to simulate human-specific immune responses, providing a robust non-animal alternative for ISO-10993 biocompatibility testing. For example, three-dimensional (3D) biomimetic intestinal mucus models can be used to evaluate bacterial endotoxins during pyrogenicity assessments [35]. Incorporating NAMs into review and approval processes can further expand their application in preclinical use.

2.3. Clinical evaluation: AI optimization and its regulatory limits

Clinical evaluation is essential to ensure the safety and effectiveness of medical devices, particularly those classified as high risk (Table 1 [36–42]). Despite their importance, clinical evaluation processes remain insufficiently optimized across jurisdictions [36,43]. Compliance with established standards, including ISO 13485 and ISO 14155:2020, is fundamental for maintaining quality, scientific rigor, and data reliability throughout clinical investigations. AI technologies have significant potential to optimize clinical study designs. For instance, NLP can analyze the scientific literature and clinical trial databases to identify research gaps, generate hypotheses, and optimize study endpoints. During protocol development, AI-driven predictive analytics can simulate study scenarios to optimize sample sizes and improve endpoint selection. AI can also enhance the diversity of the inclusion criteria by analyzing population data encompassing socioeconomic, genetic, and geographic variables.

Journal Pre-proofs

Overall, it accelerates subject recruitment, reduces costs, increases population diversity, supports real-time monitoring, and improves research design accuracy, making it versatile for clinical evaluation.

Although AI has demonstrated potential for process optimization, its implementation still poses some challenges. RAs remain cautious regarding AI-assisted evidence intended to support human trials. For example, if a device developer uses predictive analytics to reduce clinical trial sample sizes (e.g., 30%), regulators may raise concerns about whether such designs maintain sufficient statistical power and the risk of bias introduced by differences in the diversity of AI-generated study populations. We believe that adopting a Bayesian adaptive design is a viable solution [44,45].

Table 1

Clinical compliance: classification in global view and trends in clinical evaluation reports (CERs) with clinical trials.

Classification	Regions & regulatory bodies	Clinical evaluation								
		Clinical assessment reporting					Clinical trial			
		Country / region	CER required?	Regulatory body	Key points	Regulation source	Comparison	(Multi) center	Decentralized	Refs.
Class I, II, III	• USA (FDA) • China (NMPA) • Japan (PMDA) • Republic of Korea (MFDS)	EU	☑ Yes	European Commission/NBs	• Clinical evaluation plan • CER • Post-market clinical follow-up if applicable	MEDDEV 2.7/1 Rev. 4 and EU MDR 2017/745, Article 61	Pros	Traditionally recognized	• Better retention • Patient accessibility • Better for devices	[36–39]
Class I, IIa, IIb, III	• EU (EMA)	USA	× Not always formally required	FDA	• Emphasis on risk assessment and mitigation	• Recommendations for the use of clinical data in premarket	Cons	Higher drop-out rate	Agile for retention	[36–38]

Journal Pre-proofs

	(TGA) • UK (MHRA)				increasingly important • Clinical evaluations are crucial	submissions • Study of sex differences in the clinical evaluation of medical products				
Class A, B, C, D	• IMDRF (global model) • Singapore (HSA) • India (CDSCO) • Many ASEAN nations	China	☑ Yes (class II/III)	NMPA	• Clinical trials are mandatory for innovative or high-risk devices •.Higher class III devices classified lead to more clinical trial	Regulation on the supervision and administration of medical devices (decree of the state council No. 650);	Suitable specialties	• First-in-human trials • Surgical or interventional studies • Trials requiring frequent imaging (e.g., CT/MRI) • Complex oncology trials with intravenous infusions	• Chronic diseases neurology and psychiatry rare diseases • COVID-19 and infectious disease • Pain management and musculoskeletal disorders	[36–41]
Class I, II, III, IV	• Canada (Health Canada) • Hong Kong, China (MDD) • Brazil (ANVISA)	Canada	☑ Yes (class III/IV)	Health Canada	• Demonstration of safety and effectiveness • Critical evaluation of clinical data	Guidance on clinical evidence requirements for medical devices: submitting clinical evidence				
		Japan	☑ Yes (for premarket approval)	PMDA	• Demonstration of clinical safety and performance • Use of valid scientific evidence and justification	Pharmaceuticals and Medical Devices Act	Trends	Current standard	Regulatory preferred	[40–42]

PMDA: Pharmaceuticals and Medical Devices Agency; MFDS: Ministry of Food and Drug Safety; TGA: Therapeutic Goods Administration; MHRA: Medicines and Healthcare products Regulatory Agency; HSA: Health Sciences Authority; CDSCO: Central Drugs Standard Control Organisation; ASEAN: Association of Southeast Asian Nations; MDD: Medical Device Division; ANVISA:

Journal Pre-proofs

2.4. Key factors influencing review and approval timelines (FDA and EMA)

Approval timelines diverge significantly across jurisdictions, with documentation gaps and limited EMA/NBs capacity being the primary drivers of delay [46]. Globally, over 90% of medical devices are approved by major regulators such as the FDA, EMA/NBs, and NMPA. The NMPA has established a “Special Review Procedure for Innovative Medical Devices” as a green channel to expedite novel technologies. Modern regulatory frameworks have established structured pathways based on the principles of safety and efficacy. However, regulatory systems struggle to keep pace with rapid technological innovation, particularly in AI-driven devices. AI holds potential for automating and standardizing protocol development, thereby enhancing the consistency and efficiency of regulatory submissions.

A notable example of the outer limits of regulatory velocity is the coronavirus disease 2019 (COVID-19) pandemic. Under the Emergency Use Authorization (EUA) provisions (21 CFR Part 312), the US FDA rapidly authorized COVID-19 diagnostic tests and messenger RNA (mRNA) vaccines within months, whereas standard regulatory pathways typically require several years. EUAs are issued when no adequately approved alternatives exist and when the known and potential benefits outweigh the known and potential risks. This experience underscores that, with structured pre-submission data standards and automated completeness assessments, AI-assisted regulatory review could potentially compress conventional review timelines toward EUA-equivalent speeds, even for non-emergency submissions, particularly for well-characterized device categories [47–49].

2.5. Post-market surveillance (PMS): Real-time AI signal detection

A PMS is a critical component of the TPLC, which is designed to evaluate a product’s performance and safety in real-world settings after its launch into the marketplace, thereby addressing gaps that may exist in premarket assessments. Although medical devices have been developed with stringent design controls, the actual usage variability can lead to previously unidentified adverse events. An effective PMS relies on robust reporting mechanisms and timely regulatory actions. However, challenges, such as underreporting, data inconsistencies, and delayed signal detection, persist. RAs leverage PMS data to monitor safety trends, assess risks, and support decisions, such as recalls, particularly for high-risk products, such as implants and surgical robots. The integration of AI into PMS enables automated signal detection, risk stratification, and early warning systems. For instance, the implementation of a unique device identification (UDI) system for implantable devices, which is an obligatory tracking mechanism, significantly improves proactive monitoring when combined with the patient’s electronic health record (EHR). NLP can be further applied to reporting systems, such as the FDA’s Manufacturer and User Facility Device Experience (MAUDE) database or the European Database for Medical Devices (EUDAMED), by continuously parsing unstructured clinical narratives. This enables the proactive detection of adverse event signals that support corrective and preventive actions (CAPA). By automating signal extraction, NLP can identify emerging trends in device malfunctions earlier than human-led reviews. Furthermore, regulatory agencies are increasingly

Journal Pre-proofs

adopting real-world evidence (RWE) to provide a broader context for device performance, thereby enhancing PMS efficacy and product innovation. The convergence of AI and RWE has significant potential for advancing post-market risk management and patient safety [50].

3. The current adoption and challenge of AI/ML in regulatory system

There is a growing trend among global RAs toward digital transformation. As shown in Table 2, agencies such as Health Canada, Therapeutic Goods Administration (TGA, Australia), Medicines and Healthcare products Regulatory Agency (MHRA, UK), Pharmaceuticals and Medical Devices Agency (PMDA, Japan), and Health Sciences Authority (HAS, Singapore) have initiated digital modernization strategies. The proposed Aig-RS model, grounded in a regulatory science framework, is presented as a conceptual reference model to inform and support similar initiatives undertaken by the relevant authorities. The adoption of AI rests on evolving legal frameworks and a robust regulatory infrastructure that ensures safety, transparency, and efficacy. Reports indicate that the current AI development faces critical issues that pose challenges to regulatory adoption [51,52].

Table 2

Global progress in AI implementation across different jurisdiction. Three stages of AI governance: digitalization, automation, and AI deployment.

Regulatory bodies that primarily use E-systems	Regulatory bodies that may not fully use E-systems		
	Case study	Regulatory body	Details
FDA (USA) √ Uses CDRH eSubmitter, ESG EMA (EU) √ Uses EUDAMED for medical devices (gradually mandatory from 2026) Health Canada √ Uses the Common Electronic Submissions Gateway TGA (Australia) √ Uses the TGA Business Services Portal MHRA (UK)	Indonesia	Badan Pengawas Obat dan Makanan	Some electronic systems exist, but paper-based submission is still common
	Philippines	FDA Philippines	Uses online systems for some processes, but still reliant on manual steps
	Vietnam	Drug Administration of Vietnam	Transitioning to digital, but many processes are still manual
	Pakistan	Drug Regulatory Authority of Pakistan	Partial digitization; still uses physical submissions for device registrations

Journal Pre-proofs

√ Uses the Device Online Registration System PMDA (Japan) √ Uses the Gateway system and PMDA electronic application		Drug Administration and Control	and email submissions are still in use
HSA (Singapore) √ Uses the Singapore Health Product Access and Regulatory E-System, 2025; previously the Medical Device Information and Communication System platform	Kenya	Pharmacy and Poisons Board	Manual submissions still common; online portal under development
NMPA (China) √ Uses eRPS and online portals for submissions, 2019	Bangladesh	Directorate General of Drug Administration	Mostly manual processes for medical device registration
Medical Device Division (Hong Kong, China) √ Uses the Medical Device Information System, 2024	Sri Lanka	National Medicine Regulatory Authority	Still reliant on paper-based systems and face-to-face interactions

CDRH: Center for Devices and Radiological Health; ESG: Electronic Submission Gateway.

3.1. Current adoption of AI by global jurisdictions

Contrary to the Aig-RS framework proposed above, it is instructive to assess the extent to which the current RA adoption of AI has progressed. In 2017, the FDA launched the Precertification (Pre-Cert) Pilot Program in collaboration with developers of Digital Health Software [53]. AI's capacity for data analytics and process automation offers clear benefits for regulatory workflows, particularly in PMS signal detection and premarket reviews [54,55]. The current progress in AI adoption is summarized in Fig. 6., which highlights the overall trend of regulators increasingly embracing AI.

In the United States, the 21st Century Cures Act (2016) permits the use of RWE in regulatory decisions, thereby supporting the integration of AI into healthcare. The FDA's Digital Health Innovation Action Plan (2017) and AI/ML-based Software as a Medical Devices Action Plan (2021) introduced the TPLC model and good machine learning practice (GMLP). The regulatory framework for AI/ML-based Software as a SaMD modification moves beyond traditional review, allowing safe evolution through predetermined change control plans (PCCPs) and algorithm change protocols (ACPs). Although the FDA has not yet fully adopted AI for regulatory decisions, it has deployed an ELSA built on Anthropic's Claude technology, restricted to agency-wide use only, and launched in April 2024. The system cannot process industrial data, and despite incorporating mechanisms to prevent hallucinations, concerns remain regarding its accuracy and ability to engage critically with applicants. At present, detailed performance metrics and validation results from the FDA's ELSA pilot remain proprietary and have not been publicly disclosed, thereby precluding rigorous external benchmarking against proposed AI-governed regulatory systems, such as Aig-RS. To enable meaningful comparative evaluation and independent methodological research, it would be beneficial to summarize and disseminate the outcomes of the ELSA pilot through existing

Journal Pre-proofs

MDUFA transparency and reporting mechanisms. Such disclosures could include aggregate data on triage accuracy, impact on review timelines, and effects on reviewer workload. The EMA also highlights AI applications in its "NDSG Work Plan 2025–2028: Data and Artificial Intelligence in Medicines Regulation." The NMPA issued a draft guidance on AI-based medical devices, emphasizing the importance of training data quality, algorithm transparency, clinical validation, and post-market surveillance. The IMDRF has issued ten guiding principles for GMLP for medical device development, with some RAs (e.g., PMDA and Health Canada) aligned with the GMLP and international harmonization.

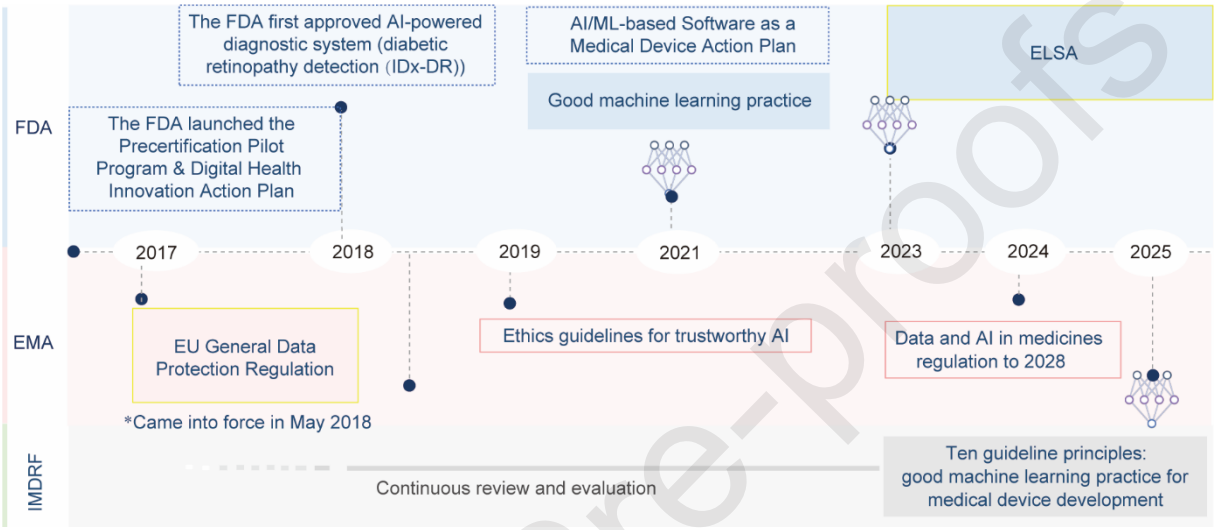


Fig. 6. Current adoption and status of regulation of AI/ML development in dominant RAs and organization.

3.2. The challenge of AI adoption in regulatory system for product approval

The principal barrier to the adoption of Aig-RS is not merely technical but also institutional, specifically, the issue of legal liability arising from erroneous AI-assisted regulatory decisions. Establishing robust governance frameworks is essential to achieve ethical, transparent, and secure AI deployment. AI algorithms, which commonly include supervised, unsupervised, and reinforcement learning, must align with the types of data and regulatory objectives. For example, supervised learning offers practical value in compliance monitoring using decision trees. However, regulatory data are commercially sensitive and cannot be processed by public AI models, necessitating localized and secure solutions.

Drawing on the provisions of the EU AI Act, we propose a strict risk-tiered decision-making framework for regulatory bodies. For low-risk administrative tasks (e.g., completeness checks), full AI automation may be permissible. For medium-risk tasks (e.g., predicate device matching), AI outputs require explicit human reviewer confirmation. For high-risk decisions (e.g., evaluating clinical trial endpoints), AI must be strictly restricted to an assistive role. To avoid the "black-box certifying a black-box" paradox, all Aig-RS components must operate within an explainable AI (XAI) framework that provides legally sufficient explanatory detail for administrative accountability while ensuring that robust error-correction processes are maintained. Critically, high-risk decisions

must operate further under strict HITL requirements so that the final legal decision-making authority remains firmly in the hands of human reviewers. Finally, each jurisdiction should ensure standardized risk assessments and consistent regulatory reporting, particularly for high-risk products that require clinical evaluation.

Journal Pre-proofs

Another systemic challenge is cross-jurisdictional data-sharing friction. The General Data Protection Regulation (GDPR) constraints limit the transfer of EU patient-level PMS data to train Aig-RS models across borders. At the same time, the FDA's PCCP/ACP adaptive update framework is structurally incompatible with the EU MDR requirements for conformity reassessment, resulting in divergent AI update pathways across jurisdictions. The IMDRF regulatory sandbox concept offers a promising mechanism for piloting international Aig-RS interoperability under controlled conditions, and its development should be prioritized.

The transition to Aig-RS requires deliberate and well-organized organizational change. Regulatory reviewers would not be displaced, but rather re-professionalized. Experienced staff can dedicate greater attention to novel technology, particularly AI/ML-enabled devices and other complex submissions, whereas AI systems manage routine administrative triage and preliminary assessments. Accordingly, new core competencies are required, including AI output auditing, the interpretation of explainability frameworks, and training in structured override protocols. Therefore, regulatory agencies should develop formal structured AI governance training programs aligned with the WHO's Global Benchmarking Tool (GBT) capacity-building benchmarks and broader governance guidance [56,57].

AI can recognize complex patterns and achieve a high level of accuracy when appropriately trained [58]. However, in the context of regulatory approval, its role should be conceptualized as augmenting rather than replacing human expertise. Evidence from prior research on human and AI collaboration indicates that combined decision-making can outperform either humans or AI operating independently [59]. Accordingly, AI-generated regulatory recommendations must remain fully traceable and auditable, whereas the final legal authorization should rest with qualified human reviewers. Human oversight should be structured according to the risk tier. AI assistance may be appropriate for low-risk administrative tasks, medium-risk determinations require human confirmation, and high-risk decisions must remain under direct human leadership. This tiered approach aligns with the WHO guidelines and promotes cross-jurisdictional regulatory consistency [57].

3.2.1. Cybersecurity and operational risks of regulatory AI systems

Regulatory AI platforms must operate within secure, agency-controlled environments, with cybersecurity embedded as a core element of AI governance, rather than being treated as a supplementary safeguard. Besides the National Institute of Standards and Technology (NIST), the Artificial Intelligence Risk Management Framework (AI RMF) 1.0, ISO/IEC 27001, FDA cybersecurity guidance, and IEC 81001-5-1 are particularly relevant, as they more directly address the medical device and health software context. Accordingly, the development of AI in regulatory systems should adhere to a cybersecurity-by-design approach. These include robust access controls, comprehensive audit logging, secure data management practices, and clearly defined mechanisms for human

oversight. Such measures are essential for mitigating risks such as adversarial manipulation, data poisoning, and the unauthorized disclosure of confidential regulatory submissions.

Journal Pre-proofs

4. Limitations and future validation

The operational value of Aig-RS must be established through prospective, risk-stratified validation rather than inferred from conceptual feasibility alone. As Aig-RS has not yet been evaluated using real regulatory submissions, this perspective makes no quantitative claim regarding its classification accuracy relative to experienced regulatory reviewers or its ability to shorten regulatory review timelines. Therefore, AI-assisted pathway classification should be benchmarked against experienced regulatory reviewers using prespecified measures of accuracy, precision, recall, F1-score, calibration, reviewer disagreement, false-negative rate, and review-time impact. Second, model reliability and generalizability depend on the quality, completeness, representativeness, provenance, and currency of the training data. Therefore, external and temporal validation, bias assessment, dataset-shift monitoring, independent audit, and post-deployment surveillance are essential and consistent with the IMDRF GMLP principles [60]. Third, human oversight should be proportional to device risk and decision criticality. AI may initially support low-risk administrative tasks such as triage, completeness checks, information extraction, and preliminary routing. However, clinical evidence assessment, benefit–risk evaluation, approval or refusal decisions, high-risk or implantable devices, and post-market safety actions require final expert review and accountability. Review time reductions should be accepted only if the prospective evaluation shows no unacceptable increase in false-negative decisions, missed safety signals, or patient safety risks. Transparency principles emphasize effective human–AI collaboration, while the FDA's ELSA 4.0 illustrates AI-enabled support for internal regulatory workflows rather than autonomous decision-making [20,61].

5. Conclusion

Regulatory approval is an appraisal process based on regulatory science and staged compliance. The rise in AI presents significant challenges for policymakers worldwide, including the increasing number of AI-integrated medical devices and their adoption [62,63]. Although numerous AI/ML devices have been approved, regulatory system adoption by RAs remains limited, primarily because of concerns about explainability, transparency, and potential biases associated with hallucinations in AI outputs. XAI offers opportunities for creating transparent, trustworthy regulations, and comprehensive oversight mechanisms [64]. AI can further support human reviewers by reducing expertise-related assessment biases and improving cross-disciplinary consistency in high-stakes regulatory decisions. The shift from Hud-RS to Aig-RS is not a matter of whether, but of when and how. RAs should pilot Aig-RS Model 1 (HITL) using sandboxed internal data starting with low-liability administrative triage and classification tasks. Engineers and AI developers should prioritize XAI-compliant federated learning architectures that can process commercially sensitive regulatory data without centralization. Policymakers should collaborate through the IMDRF to establish harmonized interoperability standards, enabling the Aig-RS frameworks validated in one jurisdiction to be adopted across others.

Journal Pre-proofs

Thus, the proposed AiqRS should be understood as a conceptual governance framework rather than a validated operational system. Although sensitivity analyses, simulation outputs, and scenario-based modeling are essential for empirical validation, such investigations are beyond the scope of the present perspective. Future research should systematically evaluate how the system performance varies under different assumptions, including review workload, device risk class, data quality, and thresholds for human reviewer intervention. These assessments should be conducted using formal simulation methodologies and structured sensitivity analyses to rigorously test the robustness and generalizability of the proposed framework.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the Areas of Excellence (AoE) from the Research Grants Council of Hong Kong (AoE/M-402/20) and the Direct Grants from Research Committee of the Chinese University of Hong Kong (2025.101; 4054939).

References

[1] Darrow JJ, Avorn J, Kesselheim AS. FDA regulation and approval of medical devices: 1976–2020. JAMA 2021;326(5):420–32.

[2] The European Medicines Agency. Overview of medical devices and *in vitro* diagnostic medical devices (IVDs) [Internet]. Amsterdam: The European Medicines Agency; [cited 2025 Jul 21]. Available from: https://health.ec.europa.eu/medical-devices-sector/overview_en.

[3] Freyer O, Jayabalan S, Kather JN, Gilbert S. Overcoming regulatory barriers to the implementation of AI agents in healthcare. Nat Med 2025;31(10):3239–43.

[4] Alam M, Shi VJ, Maisel-Campbell A, Cressey BD, Nadir U, Koza E, et al. US FDA advisory panel members' assessment of premarket approval process and suggestions for improvement. JAMA Netw Open 2024;7(10):e2436066.

[5] Bauchner H, Fontanarosa P. Bias among peer reviewers. JAMA 2017;318(8):755–6.

[6] Link AM. US and non-US submissions: an analysis of reviewer bias. JAMA 1998;280(3):246–7.

[7] The European Medicines Agency. Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence [Internet]. Amsterdam: The European Medicines Agency; 2024 [cited 2025 Aug 18]. Available from: https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai.

[8] Rajpurkar P, Chen E, Banerjee O, Topol EJ. AI in health and medicine. Nat Med 2022;28(1):31–8.

[9] Wankhede PR, Bhuyar D, Zanwar S, Pawar R, Jadhav MR, Gandhewar N, et al. Artificial intelligence for noninvasive health diagnostics. ACS Sens 2025;10(11):8217–55.

[10] Cuocolo R, Bernardini D, Pinto Dos Santos D, Klontzas ME, Akinci D'Antonoli T, Semedo LC, et al.; the European Society of Radiology. AI medical device post-market surveillance regulations: consensus recommendations by the European Society of Radiology. Insights Imaging 2025;16(1):275.

[11] Clark P, Kim J, Aphinyanaphongs Y. Marketing and US Food and Drug Administration clearance of artificial intelligence and machine learning enabled software in and as medical devices: a systematic review. JAMA Netw Open 2023;6(7):e2321792.

[12] The European Medicines Agency. Artificial intelligence [Internet]. Amsterdam: The European Medicines Agency; [cited 2025 Jul 30]. Available from: https://www.ema.europa.eu/en/about-us/how-we-work/data-regulation-big-data-other-sources/artificial-intelligence.

Journal Pre-proofs

[13] US Food and Drug Administration. FDA launches agency-wide AI tool to optimize performance for the American people [Internet]. Silver Spring: US F ency-wide-ai-tool-optimize-performance-american-people.

[14] US Food and Drug Administration. PMA review process [Internet]. Silver Spring: US Food and Drug Administration; [cited 2025 Aug 22]. Available from: https://www.fda.gov/medical-devices/premarket-approval-pma/pma-review-process.

[15] Shea M, Ostermann L, Hohman R, Roberts S, Kozak M, Dull R, et al. Regulatory watch: impact of breakthrough therapy designation on cancer drug development. Nat Rev Drug Discov 2016;15(3):152.

[16] US Food and Drug Administration. August 27, 2025 MDUFA V performance report (FY 2023 through FY 2027) [Internet]. Silver Spring: US Food and Drug Administration; [cited 2025 Aug 28]. Available from: https://www.fda.gov/about-fda/user-fee-performance-reports/mdufa-performance-reports.

[17] The European Medicines Agency. Notified bodies for medical devices—European Commission [Internet]. Amsterdam: The European Medicines Agency; [cited 2025 Jul 31]. Available from: https://health.ec.europa.eu/medical-devices-topics-interest/notified-bodies-medical-devices_en.

[18] European Commission. Study supporting the monitoring of availability of medical devices on the EU market—survey results of the 14th NB survey (MDR/IVDR) with data status 28 February 2025 [Internet]. Brussels: European Commission; [cited 2025 Aug 22]. Available from: https://health.ec.europa.eu/study-supporting-monitoring-availability-medical-devices-eu-market_en.

[19] Stange AF, Julian E. Conformity assessment of medical devices: an overview from a notified body. J Mark Access Health Policy 2026;14(1):4.

[20] US Food and Drug Administration. FDA expands AI capabilities and completes data platform consolidation [Internet]. Silver Spring: US Food and Drug Administration; [cited 2026 May 19]. Available from: https://www.fda.gov/news-events/press-announcements/fda-expands-ai-capabilities-and-completes-data-platform-consolidation.

[21] Windecker S, Fraser AG, Szymanski P, Gilard M, Lüscher TF, Abid L, et al. Priorities for medical device regulatory approval: a report from the European Society of Cardiology Cardiovascular Round Table. Eur Heart J 2025;46(16):1469–79.

[22] National Medical Products Administration. Response to proposal No. 5113 of the Second Meeting of the 14th National People's Congress [Internet]. Beijing: National Medical Products Administration; 2024 Oct 15 [cited 2025 Aug 23]. Available from: https://www.nmpa.gov.cn/zwgk/jyta/rdjy/20241015182848162.html?type=pc&m=. Chinese.

[23] US Food and Drug Administration. Digital health policy navigator [Internet]. Silver Spring: US Food and Drug Administration; [cited 2026 Jan 20]. Available from: https://www.fda.gov/medical-devices/digital-health-center-excellence/digital-health-policy-navigator.

[24] Medical Device Division, Department of Health, HKSAR. General medical device classification program [Internet]. Hongkong: Medical Device Division; [cited 2026 Jan 20]. Available from: https://www.mdd.gov.hk/en/mdacs/online-tools/general-medical-device-classification-program/index.html.

[25] Suarez-Sharp S, Li M, Duan J, Shah H, Seo P. Regulatory experience with *in vivo in vitro* correlations (IVIVC) in new drug applications. AAPS J 2016;18(6):1379–90.

[26] US Food and Drug Administration. General considerations for animal studies intended to evaluate medical devices [Internet]. Silver Spring: US Food and Drug Administration; [cited 2025 Aug 23]. Available from: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/general-considerations-animal-studies-intended-evaluate-medical-devices.

[27] Huang S, Wen J, Zhang Y, Bai X, Cui ZK. Choosing the right animal model for osteomyelitis research: considerations and challenges. J Orthop Translat 2023;43:47–65.

[28] Swearengen JR. Choosing the right animal model for infectious disease research. Animal Model Exp Med 2018;1(2):100–8.

[29] Cheng M, Liu W, Zhang J, Zhang S, Guo Z, Liu L, et al. Regulatory considerations for animal studies of biomaterial products. Bioact Mater 2021;11:52–6.

[30] FDA pushes to replace animal testing. Nat Biotechnol 2025;43(5):655.

[31] ISO. ISO/AWI 25430-1, biotechnology—organoids [Internet]. Geneva: ISO; [cited 2025 Aug 3]. https://www.iso.org/standard/90322.html.

[32] Lin JC, Jain B, Iyer JM, Rola I, Srinivasan AR, Kang C, et al. Benefit-risk reporting for FDA-cleared artificial intelligence-enabled medical devices. JAMA Health Forum 2025;6(9):e253351.

[33] European Animal Research Association Team. When will animals be replaced in research? EARA feature [Internet]. London: European Animal Research Association Team; [cited 2025 Aug 23]. Available from: https://www.eara.eu/news/when-will-animals-be-replaced-in-research-eara-feature.

[34] Hossein-Zadeh NG. Artificial intelligence in veterinary and animal science: applications, challenges, and future prospects. Comput Electron Agric 2025;235:235.

[35] Sardelli L, Campanile M, Boeri L, Donnaloja F, Fanizza F, Perottoni S, et al. A novel on-a-chip system with a 3D-bioinspired gut mucus suitable to inv
Journal Pre-proofs

[36] The European Medicines Agency. Regulatory science research needs, EMA/510633/2024 [Internet]. Amsterdam: The European Medicines Agency; [cited 2025 Jul 12]. Available from: https://www.ema.europa.eu/en/about-us/what-we-do/regulatory-science-research/regulatory-science-research-needs.

[37] Jean-Louis G, Seixas AA. The value of decentralized clinical trials: inclusion, accessibility, and innovation. Science 2024;385(6711):eadq4994.

[38] Jiang Y, Leng Y, Wu Q, Hou Y, Dong X, Zeng C, et al. Understanding the gap between expectations and reality in decentralized clinical trials. NPJ Digit Med 2025;8(1):408.

[39] De Brouwer W, Patel CJ, Manrai AK, Rodriguez-Chavez IR, Shah NR. Empowering clinical research in a decentralized world. NPJ Digit Med 2021;4(1):102.

[40] Bauchner H, Fontanarosa PB. Randomized clinical trials and COVID-19: managing expectations. JAMA 2020;323(22):2262–3.

[41] Underhill C, Sabesan S, Wilson E. Decentralised clinical trials: a game changer for improved access to clinical trials. Lancet Oncol 2024;25(12):e625.

[42] Kunkoski E, Saha A, Kam MM. FDA perspective on the importance of digital health technologies in clinical trials. JAMA Health Forum 2025;6(4):e250165.

[43] Hegyi P, Garami A, Agusti A, Agyemang C, Anadon A, Balla J, et al.; the Academia Europaea Translational Medicine Working Group. Academia Europaea's guidelines for the visualization of clinical outcomes. Nat Med 2025;31(12):3976–80.

[44] Lee JJ, Harrell FE Jr, LaVange LM, Spiegelhalter DJ. Embracing Bayesian methods in clinical trials: FDA's long-awaited draft guidance. JAMA 2026;335(19):1664.

[45] Campbell G, Irony T, Pennello G, Thompson L. Bayesian statistics for medical devices: progress since 2010. Ther Innov Regul Sci 2023;57(3):453–63.

[46] Lu Q, Hsueh YS, Tong W, Zhang Y, Xu J, Qin L. Importance to understand medical device regulations for accelerating clinical translation. J Orthop Translat 2025;51:290–7.

[47] Bishnoi M, Sonker A. Emergency use authorization of medicines: history and ethical dilemma. Perspect Clin Res 2023;14(2):49–55.

[48] Simpson S, Chakrabarti A, Robinson D, Chirgwin K, Lumpkin M. Navigating facilitated regulatory pathways during a disease X pandemic. NPJ Vaccines 2020;5(1):101.

[49] Ghadanian M, Schafheutle E. Comparison between European Medicines Agency and US Food and Drug Administration in granting accelerated marketing authorizations for COVID-19 medicines and their utilized regulations. Ther Innov Regul Sci 2024;58(1):79–113.

[50] Simpson C, Gillis AM. The pacemaker and implantable cardioverter defibrillator recall issue—Canadian perspective. Can J Cardiol 2006;22(6):467–71.

[51] Colasacco CJ, Born HL. A case of artificial intelligence chatbot hallucination. JAMA Otolaryngol Head Neck Surg 2024;150(6):457–8.

[52] Khera R, Simon MA, Ross JS. Automation bias and assistive AI: risk of harm from AI-driven clinical decision support. JAMA 2023;330(23):2255–7.

[53] US Food and Drug Administration. Digital Health Software Precertification (Pre-Cert) Pilot Program [Internet]. Silver Spring: US Food and Drug Administration; [cited 2025 Aug 19]. Available from: https://www.fda.gov/medical-devices/digital-health-center-excellence/digital-health-software-precertification-pre-cert-pilot-program.

[54] Brewer T, Colditz GA. Postmarketing surveillance and adverse drug reactions: current perspectives and future needs. JAMA 1999;281(9):824–9.

[55] Matheny ME, Yang J, Smith JC, Walsh CG, Al-Garadi MA, Davis SE, et al. Enhancing postmarketing surveillance of medical products with large language models. JAMA Netw Open 2024;7(8):e2428276.

[56] World Health Organization. Ethics and governance of artificial intelligence for health [Internet]. Geneva: World Health Organization [cited 2026 Jul 1]. Available from: https://www.who.int/publications/i/item/9789240029200.

[57] Broojerdi AK, Salvati AL, Abdelfattah MR, Dehaghi ROA, Sillo HB, Gaspar R. WHO-listed authorities (WLA) framework: transparent evidence-based approach for promoting regulatory reliance towards increased access to quality-assured medical products. Front Med 2024;11:1467229.

[58] Korteling JEH, van de Boer-Visschedijk GC, Blankendaal RAM, Boonekamp RC, Eikelboom AR. Human versus artificial intelligence. Front Artif Intell
Journal Pre-proofs

[59] Vaccaro M, Almaatouq A, Malone T. When combinations of humans and AI are useful: a systematic review and meta-analysis. Nat Hum Behav 2024;8(12):2293–303.

[60] The International Medical Device Regulators Forum (IMDRF). Good machine learning practice for medical device development: guiding principles (IMDRF/AIML WG/N88 FINAL:2025) [Internet]. The International Medical Device Regulators Forum; [cited 2026 Aug 17]. Available from: https://www.imdrf.org/documents/good-machine-learning-practice-medical-device-development-guiding-principles.

[61] US Food and Drug Administration. Predetermined change control plans for machine learning-enabled medical devices: guiding principles [Internet]. Silver Spring: US Food and Drug Administration; [cited 2026 Jul 15]. Available from: https://www.fda.gov/medical-devices/software-medical-device-samd/predetermined-change-control-plans-machine-learning-enabled-medical-devices-guiding-principles.

[62] Makary MA, Prasad V. Priorities for a new FDA. JAMA 2025;334(7):565–6.

[63] Mathias R, Schmitt A, Campos M, Vasey B, Lorenz S, McCulloch P, et al. Evaluation of human factors-related risks in AI-enabled medical devices—a practical guide. NEJM AI 2026;3(4):3.

[64] Longo L, Brcic M, Cabitza F, Choi J, Confalonieri R, Ser JD, et al. Explainable artificial intelligence (XAI) 2.0: a manifesto of open challenges and interdisciplinary research directions. Inf Fusion 2024;106:106.

Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐ The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Journal Pre-proofs