ABHI

HealthTech for Life

ABHI Response to the MHRA Call for Evidence: Regulation of AI in Healthcare

February 2026

Executive Summary

Artificial Intelligence is already reshaping diagnostics, clinical workflow and service productivity across the NHS. The UK's regulator-led, principles-based stance is a sound foundation; however, the regulatory system is not yet optimised for adaptive, data-driven AI at NHS scale. Industry faces ambiguity about borderlines, scope and risk classification, clinical evidence requirements, variability in application of rules, and low-visibility enforcement in the areas where clarity matters most. The result is delay, additional service-led processes and uneven competition. Companies who fail to invest in regulatory compliance and clinical evaluation not only increase patient and public risk, they create a commercial disadvantage for compliant and responsible companies.

The lack of certainty threatens to undermine the realisation of the UK government 's desire to harness the power of AI to support the health system and the ambitions of the MHRA to be a world leading regulator of AI medical devices. Having the right regulatory framework is one part of the requirement for UK to lead in this area. MHRA also needs to build the leadership, workforce capacity and capability to support delivery of the ambitious goals laid out in government and NHS plans. HealthTech generally and AI particularly needs to build higher visibility and focus with MHRA.

Regulation does not finish with the fixation of an assurance mark, rather it extends across the lifecycle of the product through post market surveillance, corrective actions, ongoing clinical evaluation to evidence that the device remains clinically safe and effective, plus change control and re-certification. This emphasis on lifecycle gains greater importance with software and AI based technologies with the need to address post-deployment performance.

We do not recommend immediate wholesale legislative upheaval nor AI-exceptionalism. Rather, ABHI propose decisive guidance and proportionate practice that removes ambiguity, aligns responsibilities along the whole AI value chain, and carries the learning from sandboxes into a predictable pathway. We believe that a successful regulatory framework must be as dynamic as the technology it oversees.

AI is not a homogeneous sector and regulation needs to be risk proportionate and clearly differentiate between risks that apply only to a subset of AI medical devices versus those that may apply across all AI and all medical devices. Consequently, expectation on control measures need to be aligned to risk, based on the function of the AI component specifically.

An explicit embrace of international reliance routes and designated standards reduce duplication and signal that the UK is a first-launch market for high-quality AI medical technologies. If the MHRA adopts this approach, patients will benefit sooner, clinicians will have confidence in deployment, and the UK will be better placed to attract investment, jobs and export-ready growth in HealthTech.

Our recommendations converge on the need for clarity of scope, consistency of application, proportionality of requirements and an integrated approach to responsibilities across manufacturers, integrators, deployers and professionals.

Whether the UK framework is sufficient

The UK's existing medical device legislation remains sufficient for AI at this point. We will continue to monitor how the landscape develops. Its practical application can be enhanced through non-legislative measures in four key areas as it transitions toward a lifecycle-based approach that prioritises international alignment and technical agility. The current framework faces a gap in its ability to manage the iterative nature of AI, particularly concerning continuously learning algorithms. In particular:

1. There is unresolved ambiguity at the borderline between medical and non-medical scope, and between simple rules-based AI and adaptive ML. Ambient voice assistants and similar "clinical scribe" tools illustrate the problem. They influence clinical records, decisions and medico-legal risk, yet the criteria for medical device status and class remain contested. Innovators who choose to treat such systems as medical devices and pursue CE/UKCA under a higher risk class do so to take a conservative and ethical approach rather than because of clear domestic guidance, and can then find themselves competing in tenders against products making similar claims without the same level of device oversight. That is corrosive to trust and investment.
2. Risk stratification and change control for adaptive AI lack specificity. The difference between locked algorithms and continuously learning systems is not operationalised. In UK guidance the criteria for significant changes (similar to the MDCG 2020-3), pre-determined change control plans, thresholds and evidence expectation need to be clarified. Currently, manufacturers face unpredictability on what must be documented, monitored, notified, and/or reassessed
3. Enforcement and assurance could be strengthened, particularly where software is concerned. Traditional device enforcement is geared to tangible failures; the harm pathway for information systems—misleading outputs, over-trust, bias or drift—is indirect and requires performance monitoring and real-world analytics rather than "smoking gun" events. Adverse event reporting frameworks should support effective reporting for AI and AI-enabled devices. In class I/self-declared territories, enforcement has been especially low-visibility and presents challenges for the regulatory system. Without credible, risk-proportionate enforcement, regulation loses its levelling function.
4. Whilst many AI as a Medical Device (AIaMD) products placed on the GB market fall under MHRA premarket oversight and UK MDR post market rules, CE marked devices entering the GB market via recognition operate differently. These devices undergo EU premarket scrutiny, but there is no effective or formal mechanism for systematic sharing of post market data between the EU and UK. Additionally, there is a discrepancy in the templates (e.g., MIR, trend reporting), mechanism of reporting (e.g., EUDAMED, MORE), and differences in regulation with further changes to EU MDR and IVDR in planning. This creates practical differences in vigilance expectations, signal detection and market surveillance activity, which the framework should recognise explicitly.
5. It is important that post-market surveillance continues to be tailored to the specific device, intended use, functionality and associated risks and that a one-size-fits-all

approach is not adopted for all AI. For example, locked algorithms do not change autonomously over time and therefore carry risks comparable to traditional software devices. The risks for such devices may be adequately controlled with existing PMS controls. Conversely, adaptive algorithms that update autonomously may require additional monitoring to proactively evaluate drift or performance degradation. Some higher-risk AI systems may require performance tracking through meaningful, real-time metrics rather than relying on incident reports alone, dashboards tracking patterns of use, accuracy, drift and subgroup performance. Because AI tools are woven into clinical workflows, a change in performance may reflect problems with the product, data, workflow or training. AI system transparency for the deployer requires that the manufacturer provide risk appropriate transparency tools to enable the deployer to assess, for example, concept and data drift. This implies (1) that manufacturers must provide meaningful, use-case-specific metrics and tools, and (2) deployers must abide by harmonized data governance practices to avoid AI systems being connected to feedback loops with unreliable reference values. Otherwise, it can create a noisy, reputationally punitive environment that drives risk to less capable deployers.

6. Northern Ireland continues to operate EU MDR and IVDR, but (as far as current understanding goes) the EU AI Act will not apply in the same way. The consequences of this divergence require further analysis, especially for manufacturers operating across GB, NI and Ireland. Issues may include differing definitions, classification logic, conformity routes and PMS expectations, all of which could undermine clarity for the system and industry.

In summary the regulatory legislation is currently sufficient but will require monitoring as the technology capabilities increase. The application of the legislation needs to be addressed through other elements of the wider regulatory framework, with decisive guidance and clear standards, in areas such as transparency, change control processes and risk stratification, with modern enforcement focused on performance and outcomes.

How the framework should be improved to ensure fast access to safe, effective AI

Improvement should be practical and should begin with clear answers to the questions the market is asking.

Given rapid technological change in AI, there is a substantial risk that any AI-specific regulatory provisions could date quickly. The UK framework therefore needs to emphasise adaptability, with guidance and designated standards doing more of the work than rigid legislation. This inevitably makes enforcement more complex, but it is likely to be the only sustainable way to keep pace with frontier AI models, adaptive systems and emerging clinical use cases.

1. The AI Commission should recommend a structured “Borderlines Manual” (or build on existing publications) that clarifies, with examples, when AI-enabled software is in scope as a medical device, how classes apply, and what features tip borderline cases into device territory. This manual should be updated on a cadence but anchored in principles that survive iteration: the intended purpose, proximity to diagnosis or treatment

decisions, degree of autonomy, and foreseeable impact on patient safety or medico-legal accountability. Companies cannot plan when scope is opaque. We recommend that MHRA actively engage with industry and other stakeholders during the development and ongoing updates of this manual. This will ensure guidance is both grounded in regulatory principles but also reflects the practical challenges and nuances faced by those designing, developing, and deploying AI-enabled medical devices.

2. MHRA should operationalise adaptive AI change control with explicit mechanisms for use of pre-determined change control plans (PCCP), when appropriate. Manufacturers should be able to agree, at authorisation, the categories of change that may be implemented within declared controls without re-submission, those that require notification, and those that require re-assessment. Modifications included in a PCCP would not require resubmission or notification. For manufacturers, much of the value of PCCPs consists in the ability to execute modifications, record as a part of QMS and risk management processes, but not otherwise seek authorisation. PCCPs are a voluntary tool that can offer broad benefits for all medical devices, including AI-enabled devices, by enabling more efficient and timely release of important device updates to patient and healthcare providers. They can also be valuable for supporting adaptive AI by allowing manufacturers to define a range of automatic changes that can be made to the algorithm in accordance with the PCCP. Evidence expectations should be aligned to Good Machine Learning Practice. By adopting a PCCP framework, MHRA can help streamline regulatory processes, shorten innovation cycles, reduce "stop-start" adoption, and still protect patients. To maximize its effectiveness and global utility, we encourage MHRA to align its PCCP framework with other jurisdictions and the IMDRF PCCP guiding principles.
3. The AI Airlock must become a pipeline, not a one-off pilot. Where an Airlock cohort resolves a regulatory challenge, MHRA should publish a technical note and a reusable evidence template, then provide a graduated pathway into authorisation.
4. IDAP should scale up to provide a more comprehensive route for priority categories to come to market. It should build on the learnings from the AI Airlock, enabling pre-market learning to support predictable-authorisation routes.
5. The UK should push expansion of international reliance for AI, complementing the broader medical devices reform and building on the existing recognition of the EU CE mark. For software and AI, duplication of technical assessment adds cost and delay without adding safety. Clear criteria for reliance on trusted approvals, coupled with UK-specific post-market obligations and labelling, will position the UK as a fast, attractive market while safeguarding patient interests. Recognition and reliance requires UK leadership at international forums such as IMDRF. Aligning with international standards ensures the UK remains a priority market for global developers.
6. Duplication in NHS assurance should be eliminated through a 'passport' approach. Developers should assemble a single evidence set—regulatory dossier, clinical and economic evidence (aligned to NICE's framework), cybersecurity posture and information governance artefacts—that is accepted across national and local procurement and

deployment programmes. The AI and Digital Regulations Service (AIDRS) already maps the pathway; the passport, alongside effective post market surveillance, is the operational mechanism that turns “map” into “one-and-done” assurance.

7. Organisational AI governance should be recognised explicitly. For high-impact deployments, NHS organisations should be encouraged to adopt certification against relevant AI Management System and data governance standards, with procurement incentives to signal the system’s expectations on transparency, bias control, documentation and monitoring. Deployer certification would enable reliable and effective transparency and ensure that training and testing data are state-of-the-art, so they can also be leveraged for secondary use to train and test AI systems in compliance with data governance that allows manufacturers to meet their obligations.
8. Finally, capacity and advice must track ambition. Early scientific advice services for AI—integrated with NICE and NHS evaluation teams—will reduce avoidable failure, especially for SMEs. The technology is fast-moving; predictable advice lowers the cost of doing the right thing.

Approaches to checking safety once AI devices are in use

The guiding principle is risk proportionality: AI that has a higher risk of performance degradation or drift (e.g., certain adaptive AI systems) *and* a higher clinical risk should carry heavier monitoring obligations whereas other AI tools do not warrant the same level of scrutiny, but they should still be visible in monitoring tools so that emerging risks can be caught early. MHRA should build off existing medical device requirements and produce guidance on how they translate in the context of AI, safety monitoring should be built from the devices intended purpose and associated risks and integrated with clinical governance. We are advocating for a regulatory framework that transforms post-market surveillance (PMS) from a passive reporting exercise into a dynamic, total product lifecycle monitoring system.

Monitoring methods for general risks that arise with AI-based medical devices could include:

Performance monitoring and external quality assurance. We recommend that the MHRA produce guidance for risk based continuous performance monitoring that specifically track metrics like algorithm drift and bias detection across diverse patient subgroups to ensure that a tool’s safety and effectiveness do not degrade as clinical practices or patient demographics evolve over time. AI harms are indirect and aggregate: biased outputs across sub-populations, slow drift as data distributions change, over-trust by clinicians whose task patterns alter, or unchecked “personalisation” of templates. For certain types of AI, where standard post-market monitoring and event reporting are not sufficient to mitigate risks, proactive monitoring of relevant metrics that are interpretable and actionable may be warranted. To make post-market assurance actionable, a shared library of possible metrics that represent quality should be defined in collaboration between industry, NHS and MHRA. Developers and deployers may then leverage dashboards based on these metrics that are interpretable and actionable. It is important to acknowledge that the nature and frequency of the monitoring and the appropriate

metrics should be tailored to the specific device, its intended use, function and the type of monitoring required to mitigate device-specific risks effectively.

Privacy-preserving real-world evaluation. As Secure Data Environments mature, the NHS can stand up environments that enable industry to run federated performance monitoring and drift detection without exposing raw patient data. This enables pooled learning and faster safety signal detection in sensitive indications where local datasets are too small to be reliable. This infrastructure will support information sharing between healthcare providers, MHRA and manufacturers.

Integrated clinical governance. Product-level monitoring must dovetail with organisational and professional responsibilities. Many AI products are deployed as support tools for health care professionals, integrated into existing clinical pathways and IT infrastructure and with data feeds from several sources. A change in a performance metric may be a product issue, a workflow/data issue, or a training/use issue. Governance must allocate investigation and action to the right party, and it must provide a route for shared learning, so the same signal prompts the same response across sites using the same technology.

Clear roles and responsibilities. Post-market responsibilities in the UK MDR already distinguish clearly between manufacturers (vigilance and PMS), users (duty to report and operate devices as intended) and MHRA (market surveillance). Any post-market reforms or data-driven monitoring solutions should align with these roles. Where new approaches are required to accommodate AI-specific methods such as continuous performance monitoring, drift detection and federated analytics these should be implemented via guidance and build upon (or build into) international harmonisation principles.

Managing responsibility and liability between parties

ABHI considers the UK's current product liability framework to remain broadly fit for purpose. The Consumer Protection Act 1987 continues to give patients strong protection while providing industry with legal certainty to support innovation. UK courts already recognise that "products" include software within medical devices, and the existing strict-liability regime ensures clear routes for redress across the supply chain. Given the ongoing Law Commission review of UK product liability law, ABHI does not advocate further legislative change at this stage. Any reconsideration should follow completion of that review and the publication of resulting recommendations.

Clarity on responsibility is as vital as clarity on scope. In practical terms, responsibility should be apportioned across the following roles: the developer/manufacturer, the integrator, the deploying healthcare organisation, and the professional user. We advocate for a model that distributes responsibilities across the entire supply chain. This framework would mirror the collaborative nature of modern technology deployments. At the foundation, manufacturers are responsible for pre-market assurance and technical robustness, which includes validation using datasets that are representative of the intended use population in order to mitigate bias. The developer/manufacturer must ensure essential requirements are met, risks are controlled in design and the intended purpose and limitations are documented and shared with the deploying

organisation. This framework requires clear distinction between technological failure & human judgment. Developer/manufacturer must take responsibility for post market surveillance; provide performance-monitoring hooks, documentation for bias control and drift detection, and transparent information suitable for clinical use.

The integrator—whether internal or third-party—must ensure that system configuration, interfaces and localisations do not invalidate conformity, and that any integration choices that alter behaviour are assessed, documented and governed.

The deploying organisation must conduct local risk assessment in accordance with appropriate guidance, maintain information governance, train users proportionately, monitor performance in operation, and escalate to the manufacturer or regulator where thresholds are crossed. For high-impact deployments, the organisation should adopt an AI management system and publish its governance stance for staff and patients alike.

Manufacturers have a responsibility to clearly define the intended use of their devices and to provide training as appropriate to ensure safe and effective use. Professional users, in turn, play a vital role by exercising their clinical judgment and remaining accountable for their decisions, ensuring that these technologies are used appropriately to support high-quality patient care.

A governance framework involving the key actors mentioned above could be considered (and perhaps form part of the aforementioned guidance) which sets out clear roles and responsibilities, which could be tailored by local hospitals to align with the products they use, to help provide the clarity that is needed.

Economic growth from enabling AI to reach the NHS market faster

A faster, safer route from evaluation to routine adoption is not merely a regulatory win; it is an economic growth strategy. Global companies already view the NHS as a uniquely valuable testbed for real-world evidence. If the UK offers a coherent, internationally aligned path to authorisation and a predictable move into commissioning and scaled deployment, investment will follow. SMEs will be able to raise capital against a clearer path to revenue. Multinationals will base engineering and clinical validation teams here. The export case strengthens when products have demonstrably performed under the UK's governance and monitoring discipline. In the medium term, productivity-enhancing categories—ambient voice, diagnostic ML, operational AI—will contribute directly to NHS throughput and cost containment, reinforcing the Government's missions on growth and public service reform.

Recommendations

ABHI invites the Commission to consider the following practical measures, each of which maps directly to the issues and opportunities identified.

Publish a Borderlines Manual that answers, with examples and based on transparent principles, when AI is a medical device, how classes apply, and how borderline categories are treated. This needs to be a 'living document' updated when practice or evidence requires.

Operationalise adaptive change control using pre-determined change control plans agreed at authorisation, with thresholds for notification and re-assessment, anchored in designated standards for Good Machine Learning Practice.

Institutionalise a sandbox-to-market pipeline so that AI Airlock learning becomes reusable technical notes and evidence templates, and qualifying products follow a graduated path into authorisation and commissioning rather than starting afresh.

Implement an Assurance Passport across NHS procurement and deployment, turning AIDRS guidance into a single, re-used evidence set for regulatory, HTA and information-governance assurance. This should be incorporated within wider innovations passporting initiatives.

Recognise organisational AI and data governance for high-impact deployments, coupled with procurement incentives that reward responsible AI management and data governance.

Expand early scientific advice services for AI, integrated with NICE and NHS evaluation teams, with a focus on SMEs and frontier categories where early course-correction saves time and capital.

Modernise enforcement for software by prioritising risk-based performance monitoring, outcome measures and risk-proportionate interventions over narrow event-based triggers, particularly in class I/self-declared areas.

Develop a governance framework which sets out clear roles and responsibilities and can be tailored by local hospitals

Align international reliance for AI so trusted approvals from partner regulators reduce duplication and time-to-market, with UK-specific post-market obligations preserved.

Anchor post-market surveillance in real-world data using privacy-preserving analytics in Secure Data Environments, with national learning loops that turn signals into actions across sites.

Conclusion

Drawing on ABHI's and members sector expertise we have assessed the current regulatory framework, identified key areas for improvement, and proposed practical, proportionate solutions to enable safe, rapid adoption of AI medical technologies. The recommendations focus on clarity, risk-based regulation, international harmonisation, robust post-market surveillance, and clear allocation of responsibilities—always with an eye to supporting patient outcomes and economic growth and ensuring the UK remains at the forefront of health technology innovation.

The UK can lead safely and at speed if it chooses clarity over ambiguity, practice aligned to policy and pragmatism over perfection.

In closing, it is vital to recognise the depth and breadth of expertise that exists within the UK health technology industry. ABHI's membership and wider sector partners bring significant hands-on experience in developing, deploying, and regulating advanced medical technologies—including AI—across diverse clinical environments. This expertise is not only a resource for shaping robust, practical Commission recommendations, but also an essential asset for the ongoing work required to keep pace with rapid technological change. By actively involving

industry specialists in the design, piloting, and continual refinement of regulatory frameworks, the Commission can ensure its outputs remain grounded in real-world practice, anticipate emerging risks and opportunities, and adapt swiftly to innovation, ensuring that UK regulation and practice remain world leading. Sustained collaboration between regulators, industry, clinicians, and the NHS will be the key to maintaining the UK's leadership in safe, effective, and economically transformative AI for healthcare.

Additional Notes

A: Borderlines & Principles Manual—scope and examples

A manual should codify principles that determine scope and class, then illustrate them with examples. The principles are the intended medical purpose, the proximity of outputs to diagnosis or treatment, the autonomy of the system, and the foreseeable impact on patient safety and medico-legal accountability. Such a manual would greatly assist manufacturers in understanding the regulatory status and risk classification of various types of AI. Ambient voice assistants merit explicit treatment: if the technology generates or materially alters clinical records relied upon for diagnosis or treatment, and if foreseeable harms arise from systematic errors or over-trust, device classification is warranted. The manual should also delineate when "office-automation" use sits outside scope and document the evidence that warrants movement from outside scope to device status when functionality changes.

B: A practical library of performance-monitoring metrics for dashboards

To assist developers in post-market assurance, a shared library of proposed metrics should be defined. For example, in clinical scribe contexts, relevant metrics may include edit rates, escalation counts, time-to-final-note, agreement with reference standards and error-type distributions matter; each metric needs an interpretive discipline that treats unexpected improvements as potentially unsafe until investigated. In diagnostic ML, sensitivity/specificity by sub-group, calibration, decision curves and false-discovery patterns are crucial; drift detection should combine statistical methods with domain-expert review. In operational AI, throughput, queue dynamics, rework rates and exception handling metrics may be relevant.

Other examples include intervention rates (with an interpretive discipline that assumes a pessimistic hypothesis until disproved), escalation frequencies, turnaround intervals, agreement statistics with reference standards, error-type distributions, and signals of unsafe localisation or configuration.

C: NHS Assurance Passport—fields and artefacts

The passport should contain a single set of artefacts re-used across assurance steps: device conformity assessment; ESF-aligned clinical and economic evidence; cybersecurity posture; DPIA and lawful-basis statements; information-governance controls; bias-control and drift-detection plans; performance-monitoring hooks and thresholds; and training and competence plans. Acceptance of the passport in national procurements removes repetitive local gatekeeping and shortens time to patient benefit.