

Governance of AI-Enabled Digital Health Systems: Integrating Clinical Safety, Cybersecurity, Interoperability and Regulatory Compliance in the NHS

Umamaheswari Shanmugam¹ & Gill Stone¹

¹ Sciensus Pharma Services Limited, United Kingdom

Correspondence: Umamaheswari Shanmugam, Sciensus Pharma Services Limited, United Kingdom.

Received: August 25, 2026

Accepted: September 25, 2026

Online Published: October 18, 2026

doi:10.5430/ijba.v17n3p12

URL: <https://doi.org/10.5430/ijba.v17n3p12>

Abstract

Digital health technologies, including mobile applications, clinical decision-support systems, software-based medical technologies and artificial intelligence (AI)-enabled tools, are increasingly integrated into healthcare delivery. Within the UK National Health Service (NHS), these technologies offer opportunities to improve patient outcomes, operational efficiency and data-driven clinical decision-making. Their implementation, however, introduces complex challenges relating to medical device qualification and classification, clinical safety, cybersecurity, interoperability, data governance and regulatory compliance.

This study examines the regulatory and governance frameworks applicable to AI-enabled digital health and medical device systems within the NHS, with particular emphasis on the Clinical Risk Management Standards DCB0129 and DCB0160, the Digital Technology Assessment Criteria (DTAC), and relevant UK medical device regulatory requirements. A structured literature review was conducted using PubMed, covering publications between 2015 and 2025. Sixty studies were included and analysed using a narrative synthesis approach to identify recurring regulatory, technical and governance challenges associated with the implementation and lifecycle management of digital health technologies.

The findings indicate that existing NHS frameworks provide an important foundation for clinical safety, data protection, cybersecurity and organisational assurance. However, their application becomes increasingly complex for interconnected and adaptive AI-enabled systems. Key challenges include uncertainty in medical device qualification and regulatory classification, fragmentation across governance requirements, algorithmic bias and explainability, data integrity, cybersecurity vulnerabilities, interoperability, limitations in conventional evaluation approaches, and the need for continuous monitoring of system performance following implementation.

Drawing on these findings, this study proposes a Unified NHS AI Governance and Compliance Model that integrates medical device qualification and regulatory classification with clinical safety, data protection, cybersecurity, usability and accessibility, interoperability, and data and AI workflow governance. The model provides a structured lifecycle approach to support early determination of regulatory requirements, clearer allocation of organisational responsibilities, proportionate risk assessment, implementation assurance and post-deployment monitoring.

The study concludes that governance of AI-enabled digital health and medical device technologies should evolve towards a more integrated, adaptive and risk-based approach that aligns medical device regulation with wider digital-health assurance requirements. Although developed within the NHS context, the proposed framework may provide a transferable governance approach for healthcare organisations implementing complex, interoperable and AI-enabled digital technologies.

Keywords: artificial intelligence, digital health, medical devices, NHS, DCB0129, DCB0160, DTAC, clinical safety, cybersecurity, interoperability, regulatory governance

1. Introduction

The rapid advancement of digital technologies has fundamentally transformed healthcare delivery, enabling new models of care that are increasingly data-driven, patient-centred, and technologically enabled (Notes 19-23). Digital health applications, ranging from mobile health (mHealth) platforms and electronic patient records to AI-driven

diagnostic tools and predictive analytics systems, are now integral components of modern healthcare systems (Topol, 2019; Reddy, Allan, Coghlan, & Cooper, 2019; Shaw, Rudzicz, Jamieson, & Goldfarb, 2019; Wachter, 2016; Vokinger, Feuerriegel, & Kesselheim, 2020; Benjamens, Dhunoo, & Meskó, 2020; Gerke, Minssen, & Cohen, 2020; Cabitza, Rasoini, & Gensini, 2017) (Note 1, Note 24).

Within the NHS, digital transformation has been identified as a strategic priority, with significant investment directed toward the development and deployment of digital solutions aimed at improving patient outcomes, reducing inefficiencies, and enhancing system resilience (Sendak, D'Arcy, Kashyap, S., et al., 2020; Liu, Faes, Kale, et al., 2019; Topol, 2019) (Notes 2-6, Notes 25-26).

Digital health applications offer numerous benefits, including improved access to care, enhanced clinical decision-making, and increased operational efficiency (Topol, 2019; Reddy, Allan, Coghlan, & Cooper, 2019; Shaw, Rudzicz, Jamieson, & Goldfarb, 2019) (Note 7). For example, digital prescription management systems enable real-time tracking of medication workflows, reducing delays and improving patient safety (Char, Shah, & Magnus, 2018; Rajkomar, Hardt, & Howell, 2018; Obermeyer, Powers, Vogeli, & Mullainathan, 2019; Kelly, Karthikesalingam, & Suleyman, 2019). Similarly, AI-driven clinical decision-support systems have demonstrated potential in areas such as radiology, oncology, and cardiology, where they can augment clinician expertise and improve diagnostic accuracy (Finlayson, et al., 2019; Rudin, 2019; He, Baxter, & Xu, 2019; Greenhalgh, Wherton, & Papoutsis, 2017; Huckvale, et al., 2015).

However, alongside these benefits, digital health technologies introduce new categories of risk that are not adequately addressed by traditional regulatory frameworks (Jiang, Jiang, & Zhi, 2017; Esteva, A., et al., 2017; Gulshan, et al., 2016; Yu, & Kohane, 2019; London, 2019). Unlike conventional medical devices, which are typically static and undergo infrequent updates, digital applications are dynamic, continuously evolving, and often dependent on complex data ecosystems (Price, & Cohen, 2019; Mesko, 2017; Shortliffe, & Sepúlveda, 2018; Bates, & Singh, 2018) (Note 8). This introduces challenges related to system reliability, data integrity, and user interaction (Morley, Machado, Burr, Cows, Taddeo, & Floridi, 2020; Mittelstadt, 2019; Floridi, Cows, Beltrametti, et al., 2018; Beam, & Kohane, 2018) (Note 9). Furthermore, AI systems introduce additional complexities due to their probabilistic nature, lack of transparency, and potential for bias (Wiens, Saria, Sendak, et al., 2019) (Notes 10-13).

Recognising these challenges, the NHS has established a structured regulatory framework designed to ensure the safety and effectiveness of digital health technologies (Finlayson, et al., 2019; Rudin, 2019; He, Baxter, & Xu, 2019). Central to this framework are the Clinical Risk Management Standards DCB0129 and DCB0160, which define responsibilities for developers and healthcare organisations, respectively (Rudin, 2019; He, Baxter, & Xu, 2019; Greenhalgh, Wherton, & Papoutsis, 2017). These standards are complemented by the Digital Technology Assessment Criteria (DTAC), which provides a comprehensive evaluation framework covering clinical safety, data protection, cybersecurity, interoperability, and usability (Huckvale, et al., 2015).

Despite the strong and well-established foundation provided by existing NHS regulatory frameworks, there is increasing recognition that ongoing advancements in digital and AI technologies present opportunities for further enhancement of these approaches (Morley, Machado, Burr, Cows, Taddeo, & Floridi, 2020; Mittelstadt, 2019; Floridi, Cows, Beltrametti, et al., 2018; Beam, & Kohane, 2018) (Note 9). The adaptive and data-driven nature of AI systems introduces new dimensions to validation, monitoring, and regulatory oversight, particularly in relation to continuous system performance and evolving data inputs (Wiens, Saria, Sendak, et al., 2019) (Notes 10-13).

Encouragingly, this has stimulated the development of forward-looking regulatory initiatives aimed at building upon current frameworks. There is growing emphasis on evolving governance models toward more flexible, risk-based, and lifecycle-oriented approaches that align with the dynamic characteristics of digital and AI-enabled systems (Notes 14-18). These developments position the NHS regulatory ecosystem as both robust and adaptable, capable of supporting innovation while maintaining high standards of patient safety and clinical effectiveness.

This study aims to provide a comprehensive analysis of regulatory governance for digital health applications within the NHS (Notes 19-23). By conducting a systematic review of the literature, the study examines the effectiveness of existing regulatory frameworks, identifies key challenges, and explores potential approaches for improving governance in the context of rapidly evolving technologies (Vokinger, Feuerriegel, & Kesselheim, 2020; Benjamens, Dhunoo, & Meskó, 2020; Gerke, Minssen, & Cohen, 2020; Cabitza, Rasoini, & Gensini, 2017; Sendak, D'Arcy, Kashyap, et al., 2020; Liu, Faes, Kale, et al., 2019; Topol, 2019) (Note 24, Notes 25-26).

1.1 Digital Transformation in Healthcare

The global healthcare landscape is undergoing a profound transformation driven by digital innovation (Topol, 2019; Reddy, Allan, Coghlan, & Cooper, 2019; Shaw, Rudzicz, Jamieson, & Goldfarb, 2019; Wachter, 2016) (Note 1). Technologies such as mobile health applications, cloud-based platforms, and artificial intelligence (AI) are redefining how care is delivered, monitored, and managed (Notes 2-6). Within the NHS, digital transformation has become a central strategic priority, particularly in response to increasing demand, workforce shortages, and the need for cost-effective service delivery (Char, Shah, & Magnus, 2018; Rajkomar, Hardt, & Howell, 2018; Obermeyer, Powers, Vogeli, & Mullainathan, 2019; Kelly, Karthikesalingam, & Suleyman, 2019) (Note 7).

Digital health applications enable:

- Real-time monitoring of patient conditions (Finlayson, et al., 2019; Rudin, 2019; He, Baxter, & Xu, 2019)
- Remote consultations and telemedicine (Greenhalgh, Wherton, & Papoutsis, 2017; Huckvale, et al., 2015; Jiang, Jiang, & Zhi, 2017)
- Automated administrative processes (Esteva, A., et al., 2017; Gulshan, et al., 2016; Yu, & Kohane, 2019)
- Data-driven clinical decision-making (London, 2019; Price, & Cohen, 2019; Mesko, 2017)

These capabilities contribute to improved patient outcomes and system efficiency (Topol, 2019; Reddy, Allan, Coghlan, & Cooper, 2019; Shaw, Rudzicz, Jamieson, & Goldfarb, 2019; Wachter, 2016) (Note 1). However, they also introduce new forms of risk that differ from traditional healthcare interventions (Shortliffe, & Sepúlveda, 2018; Bates, & Singh, 2018) (Note 8).

1.2 Risk Landscape of Digital Health Systems

The adoption of digital health technologies has introduced a fundamentally new risk landscape within healthcare systems (Morley, Machado, Burr, Cows, Taddeo, & Floridi, 2020; Mittelstadt, 2019; Floridi, Cows, Beltrametti, et al., 2018; Beam, & Kohane, 2018) (Note 9). Unlike traditional medical devices, which are typically static and undergo controlled, infrequent updates, digital applications operate within dynamic environments characterised by continuous data exchange, iterative development, and evolving functionality (Wiens, Saria, Sendak, et al., 2019) (Notes 10-13). This shift has significant implications for safety, reliability, and regulatory oversight (Notes 14-18).

One of the primary risks associated with digital health systems is data integrity failure (Morley, Machado, Burr, Cows, Taddeo, & Floridi, 2020; Mittelstadt, 2019; Floridi, Cows, Beltrametti, et al., 2018; Beam, & Kohane, 2018) (Note 9). Clinical decision-making increasingly relies on accurate and timely data; therefore, any corruption, delay, or misinterpretation of data can have direct consequences for patient safety (Wiens, Saria, Sendak, et al., 2019) (Notes 10-13). For example, inaccuracies in electronic prescription systems may result in incorrect medication dosing or delayed treatment delivery (Char, Shah, & Magnus, 2018; Rajkomar, Hardt, & Howell, 2018; Obermeyer, Powers, Vogeli, & Mullainathan, 2019; Kelly, Karthikesalingam, & Suleyman, 2019). Similarly, incomplete or outdated patient records in electronic health systems can lead to inappropriate clinical decisions (Notes 14-18).

Another critical risk relates to system interoperability and integration (Wiens, Saria, Sendak, et al., 2019) (Notes 10-13). Digital health applications rarely operate in isolation; instead, they are part of a broader ecosystem that includes electronic health records (EHRs), laboratory systems, pharmacy systems, and external databases (Notes 14-18). Failures in data exchange between systems—whether due to technical incompatibilities or communication errors—can disrupt clinical workflows and introduce new hazards (Morley, Machado, Burr, Cows, Taddeo, & Floridi, 2020; Mittelstadt, 2019; Floridi, Cows, Beltrametti, et al., 2018; Beam, & Kohane, 2018) (Note 9).

Cybersecurity represents an additional and increasingly prominent risk (Notes 14-18). The digitisation of healthcare data has made health systems attractive targets for cyberattacks, including ransomware, data breaches, and unauthorised access (Notes 2-6). Such incidents not only compromise patient confidentiality but can also disrupt clinical operations, leading to delays in care and potential harm (Morley, Machado, Burr, Cows, Taddeo, & Floridi, 2020; Mittelstadt, 2019; Floridi, Cows, Beltrametti, et al., 2018; Beam, & Kohane, 2018) (Note 9).

Human–technology interaction also introduces unique risks (Notes 19-23). Users may misinterpret system outputs, rely excessively on automated recommendations, or fail to recognise system limitations (Vokinger, Feuerriegel, & Kesselheim, 2020; Benjamins, Dhunoo, & Meskó, 2020; Gerke, Minssen, & Cohen, 2020; Cabitza, Rasoini, & Gensini, 2017), (Note 24). These risks are particularly pronounced in environments where digital systems are integrated into high-pressure clinical workflows, where time constraints and cognitive load may affect user behaviour (Sendak, D'Arcy, Kashyap, et al., 2020; Liu, Faes, Kale, et al., 2019; Topol, 2019) (Notes 25-26).

Artificial intelligence systems amplify these risks further (Wiens, Saria, Sendak, et al., 2019) (Notes 10-13). AI models often operate using complex algorithms that are not easily interpretable by users or regulators (Rudin, 2019; Morley, Machado, Burr, Cows, Taddeo, & Floridi, 2020; Mittelstadt, 2019; Floridi, Cows, Beltrametti, et al., 2018; Beam, & Kohane, 2018) (Note 9). This lack of transparency can make it difficult to identify errors or biases in system outputs (Wiens, Saria, Sendak, et al., 2019) (Notes 10-13). Additionally, AI systems are inherently data-dependent, meaning that changes in data inputs can lead to unpredictable variations in performance (Vokinger, Feuerriegel, & Kesselheim, 2020; Benjamins, Dhunoo, & Meskó, 2020; Gerke, Minssen, & Cohen, 2020; Cabitza, Rasoini, & Gensini, 2017), (Note 24). This phenomenon, commonly referred to as model drift, poses significant challenges for long-term system reliability (Sendak, D'Arcy, Kashyap, et al., 2020; Liu, Faes, Kale, et al., 2019; Topol, 2019) (Notes 25-26).

Taken together, these risks highlight the need for a comprehensive and adaptive regulatory framework that can address the unique characteristics of digital health technologies (Notes 14-18). Traditional risk management approaches, which are designed for static systems, must be re-evaluated and extended to accommodate the dynamic and interconnected nature of digital and AI-driven healthcare systems (Morley, Machado, Burr, Cows, Taddeo, & Floridi, 2020; Mittelstadt, 2019; Floridi, Cows, Beltrametti, et al., 2018; Beam, & Kohane, 2018) (Note 9).

1.3 NHS Regulatory Ecosystem

The NHS has developed a multi-layered regulatory ecosystem designed to ensure the safe development, deployment, and operation of digital health technologies (Finlayson, et al., 2019; Rudin, 2019; He, Baxter, & Xu, 2019; Greenhalgh, Wherton, & Papoutsis, 2017; Huckvale, et al., 2015). This ecosystem is characterised by a combination of national standards, organisational responsibilities, and cross-domain compliance requirements (Notes 14-18).

At the core of this ecosystem are the Clinical Risk Management Standards DCB0129 and DCB0160, which together establish a comprehensive framework for managing risks associated with health IT systems (Finlayson, et al., 2019; Rudin, 2019; He, Baxter, & Xu, 2019). DCB0129 applies to developers and manufacturers, requiring them to implement structured risk management processes throughout the system lifecycle (Rudin, 2019; He, Baxter, & Xu, 2019; Greenhalgh, Wherton, & Papoutsis, 2017). This includes hazard identification, risk analysis, control implementation, and the development of a Clinical Safety Case (Finlayson, et al., 2019; Rudin, 2019; He, Baxter, & Xu, 2019).

DCB0160, on the other hand, applies to healthcare organisations and focuses on the safe implementation and use of digital systems (Rudin, 2019; He, Baxter, & Xu, 2019; Greenhalgh, Wherton, & Papoutsis, 2017). It requires organisations to conduct local risk assessments, ensure appropriate training and governance, and monitor system performance in real-world settings (Finlayson, et al., 2019; Rudin, 2019; He, Baxter, & Xu, 2019; Greenhalgh, Wherton, & Papoutsis, 2017; Huckvale, et al., 2015). This dual-framework approach ensures that responsibility for safety is shared between developers and users (Notes 14-18).

The Digital Technology Assessment Criteria (DTAC) complements these standards by providing a broader evaluation framework for digital technologies (Huckvale, et al., 2015). DTAC encompasses multiple domains, including clinical safety, data protection, cybersecurity, interoperability, and usability (Notes 2-6). It serves as a key tool for procurement and assurance, ensuring that digital products meet NHS requirements before deployment (Huckvale, et al., 2015).

In addition to these NHS-specific frameworks, digital health technologies must also comply with broader regulatory requirements (Notes 2-6). These include the UK Medical Device Regulations, which apply to software that performs diagnostic or therapeutic functions (Notes 10-13), as well as data protection laws such as the General Data Protection Regulation (GDPR) and the Data Protection Act (Notes 2-6).

While this regulatory ecosystem is comprehensive and demonstrates a strong commitment to patient safety and innovation, it is also complex (Notes 14-18). Developers and healthcare organisations must navigate multiple overlapping frameworks, each with its own requirements and expectations (Finlayson, et al., 2019; Rudin, 2019; He, Baxter, & Xu, 2019; Greenhalgh, Wherton, & Papoutsis, 2017; Huckvale, et al., 2015). This complexity can create challenges in implementation and may act as a barrier to innovation if not managed effectively; however, it also reflects the NHS's rigorous and multifaceted approach to ensuring safety, quality, and governance in an evolving digital healthcare landscape (Notes 14-18).

1.4 Research Gap

Despite the robustness of the NHS regulatory framework, several challenges remain from a development and implementation perspective when addressing emerging technologies, particularly AI-driven systems (Morley,

Machado, Burr, Cowls, Taddeo, & Floridi, 2020; Mittelstadt, 2019; Floridi, Cowls, Beltrametti, et al., 2018; Beam, & Kohane, 2018) (Note 9). One of the most significant challenges lies in the absence of sufficiently detailed AI-specific regulatory guidance to support developers in translating high-level regulatory expectations into practical design and validation approaches. Existing frameworks were primarily designed for traditional health IT systems and therefore require further interpretation when applied to adaptive, data-driven, and probabilistic AI systems (Wiens, Saria, Sendak, et al., 2019) (Notes 10-13).

Another key challenge relates to the evaluation of digital health technologies during development. Traditional methods, such as randomised controlled trials, are often not well aligned with the iterative and agile development cycles typical of digital applications, which involve frequent updates and continuous improvement (Jiang, Jiang, & Zhi, 2017; Esteva, A., et al., 2017; Gulshan, et al., 2016; Yu, & Kohane, 2019; London, 2019) (Notes 19-23). This creates practical difficulties for developers in generating timely and robust evidence to demonstrate safety, performance, and clinical effectiveness while maintaining development velocity (Vokinger, Feuerriegel, & Kesselheim, 2020; Benjamins, Dhunoo, & Meskó, 2020; Gerke, Minssen, & Cohen, 2020; Cabitza, Rasoini, & Gensini, 2017) (Note 24).

Additionally, developers must navigate multiple overlapping regulatory frameworks, including DCB0129, DCB0160, DTAC, ISO standards, and data protection requirements. While each framework serves an important role, the coexistence of these standards can introduce complexity in interpretation and implementation, requiring developers to align technical, clinical, and regulatory requirements across different domains (Finlayson, et al., 2019; Rudin, 2019; He, Baxter, & Xu, 2019; Greenhalgh, Wherton, & Papoutsis, 2017; Huckvale, et al., 2015) (Notes 14-18). This can be particularly challenging for organisations developing solutions intended for deployment across multiple healthcare systems or regulatory jurisdictions (Vokinger, Feuerriegel, & Kesselheim, 2020; Benjamins, Dhunoo, & Meskó, 2020; Gerke, Minssen, & Cohen, 2020; Cabitza, Rasoini, & Gensini, 2017; Sendak, D'Arcy, Kashyap, et al., 2020).

Finally, from a lifecycle management perspective, there is limited practical guidance on implementing post-deployment monitoring and continuous risk management within digital and AI-enabled systems. While frameworks such as DCB0160 emphasise the importance of monitoring in live environments, developers require clearer methodologies for embedding continuous performance evaluation, model validation, and real-world feedback mechanisms into system design and operational processes (Morley, Machado, Burr, Cowls, Taddeo, & Floridi, 2020; Mittelstadt, 2019; Floridi, Cowls, Beltrametti, et al., 2018; Beam, & Kohane, 2018; Liu, Faes, Kale, et al., 2019; Topol, 2019) (Note 9, Notes 25-26).

1.5 Aim of the Study

This study aims to address these gaps by providing a comprehensive analysis of regulatory governance for digital health applications within the NHS. Specifically, the study seeks to:

1. Examine the structure and effectiveness of existing regulatory frameworks, including DCB0129, DCB0160, and DTAC.
2. Identify key challenges associated with the regulation of digital and AI-driven healthcare technologies.
3. Analyse real-world applications of digital health systems through a detailed case study.
4. Propose a future-oriented regulatory model that addresses the limitations of current frameworks.

By achieving these objectives, the study aims to contribute to the development of more effective and adaptive regulatory approaches that support both innovation and patient safety.

2. Case Study: Digital Health Applications in NHS Practice

2.1 Overview of Digital Health Applications

Digital health applications within the NHS encompass a wide range of systems designed to support clinical and operational functions (Topol, 2019; Reddy, Allan, Coghlan, & Cooper, 2019; Shaw, Rudzicz, Jamieson, & Goldfarb, 2019; Wachter, 2016) (Note 1). These include prescription management platforms, electronic patient records, AI-driven diagnostic tools, and patient-facing mobile applications (Char, Shah, & Magnus, 2018; Rajkomar, Hardt, & Howell, 2018; Obermeyer, Powers, Vogeli, & Mullainathan, 2019; Kelly, Karthikesalingam, & Suleyman, 2019) (Notes 2-7).

Prescription management systems, for example, enable healthcare professionals to track medication requests, monitor prescription status, and coordinate delivery (Char, Shah, & Magnus, 2018; Rajkomar, Hardt, & Howell,

2018; Obermeyer, Powers, Vogeli, & Mullainathan, 2019; Kelly, Karthikesalingam, & Suleyman, 2019). These systems improve efficiency by reducing administrative burden and enabling real-time visibility of workflows (Topol, 2019; Reddy, Allan, Coghlan, & Cooper, 2019; Shaw, Rudzicz, Jamieson, & Goldfarb, 2019; Wachter, 2016) (Note 1). However, they also introduce risks related to data accuracy, system availability, and user interaction (Morley, Machado, Burr, Cows, Taddeo, & Floridi, 2020; Mittelstadt, 2019; Floridi, Cows, Beltrametti, et al., 2018; Beam, & Kohane, 2018) (Note 9).

Electronic patient record systems serve as central repositories of clinical information, supporting decision-making and care coordination (Notes 2-6). These systems are critical to modern healthcare delivery but require robust data governance and security measures to ensure reliability and confidentiality (Notes 14-18).

AI-driven systems are increasingly used to support clinical decision-making in areas such as radiology, pathology, and triage (Finlayson, et al., 2019; Rudin, 2019; He, Baxter, & Xu, 2019; Greenhalgh, Wherton, & Papoutsis, 2017; Huckvale, et al., 2015). These systems have the potential to enhance diagnostic accuracy and improve efficiency but introduce new challenges related to validation, transparency, and accountability (Wiens, Saria, Sendak, et al., 2019) (Notes 10-13).

2.2 Case Study: Digital Prescription and AI-Supported Systems

2.2.1 Context and Significance

Prescription management represents a safety-critical function within healthcare delivery, directly influencing patient outcomes, treatment continuity, and clinical effectiveness (Char, Shah, & Magnus, 2018; Rajkomar, Hardt, & Howell, 2018; Obermeyer, Powers, Vogeli, & Mullainathan, 2019; Kelly, Karthikesalingam, & Suleyman, 2019). Historically, prescription workflows within healthcare systems have relied on fragmented, paper-based processes and manual coordination between clinicians, pharmacies, and administrative teams (Jiang, Jiang, & Zhi, 2017; Esteva, A., et al., 2017; Gulshan, et al., 2016; Yu, & Kohane, 2019; London, 2019). These legacy approaches are inherently prone to transcription errors, delays, and limited visibility across the care pathway (Morley, Machado, Burr, Cows, Taddeo, & Floridi, 2020; Mittelstadt, 2019; Floridi, Cows, Beltrametti, et al., 2018; Beam, & Kohane, 2018) (Note 9).

The transition toward digital prescription systems within the NHS reflects a broader shift toward integrated, data-driven healthcare delivery (Notes 2-6). These systems aim to streamline workflows, enhance communication, and reduce the likelihood of human error (Topol, 2019; Reddy, Allan, Coghlan, & Cooper, 2019; Shaw, Rudzicz, Jamieson, & Goldfarb, 2019; Wachter, 2016) (Note 1). However, the integration of digital technologies—particularly those incorporating artificial intelligence—introduces new and complex risk dimensions that require careful regulatory oversight (Wiens, Saria, Sendak, et al., 2019) (Notes 10-13).

This case study examines a digital prescription management platform augmented with AI capabilities, providing a real-world context to evaluate the effectiveness and limitations of current NHS regulatory frameworks (Finlayson, et al., 2019; Rudin, 2019; He, Baxter, & Xu, 2019; Greenhalgh, Wherton, & Papoutsis, 2017; Huckvale, et al., 2015).

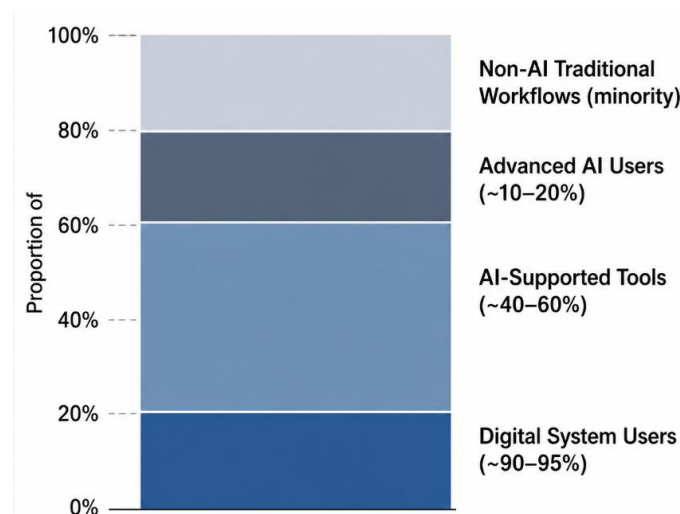


Figure 1. AI Adoption Among Clinicians

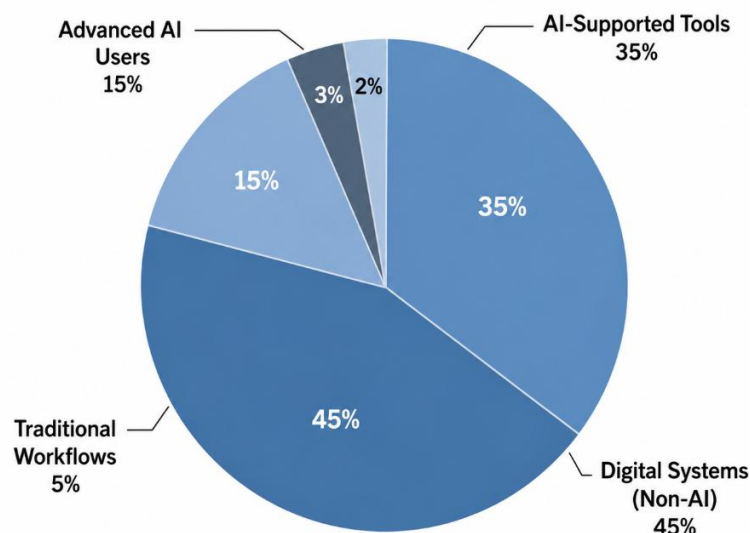


Figure 2. Estimated Adoption of Digital and AI-supported Systems Within the NHS Workforce

Digital system usage is near-universal among clinicians, while AI-supported tools are increasingly integrated into clinical workflows, particularly in specialist domains. Advanced AI adoption remains concentrated but is expected to expand as regulatory frameworks such as DCB0129, DCB0160, and DTAC continue to evolve.

2.2.2 System Architecture and Functional Scope

The system under consideration is a cloud-based digital platform designed to support end-to-end prescription lifecycle management (Notes 14-18). It integrates with existing healthcare infrastructure, including electronic patient records (EPRs), pharmacy systems, and operational logistics platforms, enabling a unified and real-time view of prescription workflows (Notes 2-6, Notes 19-23).

Core functionalities include:

- **End-to-End Workflow Visibility:**

The system enables tracking of prescriptions from initiation through validation, dispensing, and delivery, providing clinicians with real-time insights into process status (Char, Shah, & Magnus, 2018; Rajkomar, Hardt, & Howell, 2018; Obermeyer, Powers, Vogeli, & Mullainathan, 2019; Kelly, Karthikesalingam, & Suleyman, 2019; Morley, Machado, Burr, Cows, Taddeo, & Floridi, 2020; Mittelstadt, 2019; Floridi, Cows, Beltrametti, et al., 2018; Beam, & Kohane, 2018) (Note 9).

- **Data Integration and Interoperability:**

Seamless integration with multiple healthcare systems ensures that patient data, prescription details, and operational updates are synchronised across platforms (Notes 14-18).

- **User-Centred Interface:**

Dashboards and reporting tools support clinical and administrative users in monitoring patient status, identifying delays, and managing exceptions (Notes 19-23).

- **AI-Enabled Capabilities:**

The system incorporates AI-driven tools for document processing, anomaly detection, and predictive workflow optimisation (Finlayson, et al., 2019; Rudin, 2019; He, Baxter, & Xu, 2019; Greenhalgh, Wherton, & Papoutsis, 2017; Huckvale, et al., 2015; Wiens, Saria, Sendak, et al., 2019; Vokinger, Feuerriegel, & Kesselheim, 2020; Benjamins, Dhunoo, & Meskó, 2020; Gerke, Minssen, & Cohen, 2020; Cabitz, Rasoini, & Gensini, 2017) (Notes 10-13, Note 24).

This architecture reflects a shift from isolated digital tools toward interconnected, ecosystem-based healthcare platforms, where system reliability depends not only on individual components but also on the integrity of data flows between them (Morley, Machado, Burr, Cows, Taddeo, & Floridi, 2020; Mittelstadt, 2019; Floridi, Cows,

Beltrametti, et al., 2018; Beam, & Kohane, 2018; Sendak, D'Arcy, Kashyap, et al., 2020; Liu, Faes, Kale, et al., 2019; Topol, 2019) (Note 9, Notes 25-26).

2.2.3 Role and Limitations of Artificial Intelligence

Artificial intelligence within the system is deployed primarily as an augmentation tool, supporting rather than replacing clinical decision-making (Finlayson, et al., 2019; Rudin, 2019; He, Baxter, & Xu, 2019; Greenhalgh, Wherton, & Papoutsis, 2017; Huckvale, et al., 2015). Its functions include:

- **Automated Data Extraction:**

AI models process scanned prescriptions and structured documents, extracting key data fields to reduce manual input requirements (Finlayson, et al., 2019; Rudin, 2019; He, Baxter, & Xu, 2019; Greenhalgh, Wherton, & Papoutsis, 2017; Huckvale, et al., 2015; Wiens, Saria, Sendak, et al., 2019) (Notes 10-13).

- **Anomaly Detection:**

Algorithms identify inconsistencies or missing information within prescriptions, prompting user review (Morley, Machado, Burr, Cows, Taddeo, & Floridi, 2020; Mittelstadt, 2019; Floridi, Cows, Beltrametti, et al., 2018; Beam, & Kohane, 2018) (Note 9).

- **Predictive Workflow Analysis:**

Machine learning models analyse historical data to anticipate delays or operational bottlenecks (Vokinger, Feuerriegel, & Kesselheim, 2020; Benjamins, Dhunoo, & Meskó, 2020; Gerke, Minssen, & Cohen, 2020; Cabitza, Rasoini, & Gensini, 2017) (Note 24).

While these capabilities enhance efficiency, they introduce inherent limitations (Wiens, Saria, Sendak, et al., 2019) (Notes 10-13). AI systems operate on probabilistic models and are dependent on the quality and representativeness of input data (Morley, Machado, Burr, Cows, Taddeo, & Floridi, 2020; Mittelstadt, 2019; Floridi, Cows, Beltrametti, et al., 2018; Beam, & Kohane, 2018; Vokinger, Feuerriegel, & Kesselheim, 2020; Benjamins, Dhunoo, & Meskó, 2020; Gerke, Minssen, & Cohen, 2020; Cabitza, Rasoini, & Gensini, 2017) (Note 9, Note 24). Errors in data extraction or classification can propagate through the system, potentially affecting downstream processes (Morley, Machado, Burr, Cows, Taddeo, & Floridi, 2020; Mittelstadt, 2019; Floridi, Cows, Beltrametti, et al., 2018; Beam, & Kohane, 2018) (Note 9).

Furthermore, the lack of transparency in certain AI models—often referred to as the “black box” problem—limits the ability of users and regulators to fully understand how outputs are generated (Rudin, 2019; Wiens, Saria, Sendak, et al., 2019) (Notes 10-13). This creates challenges in validating system behaviour and establishing trust among clinical users (Morley, Machado, Burr, Cows, Taddeo, & Floridi, 2020; Mittelstadt, 2019; Floridi, Cows, Beltrametti, et al., 2018; Beam, & Kohane, 2018) (Note 9).

2.2.4 Regulatory Alignment and Compliance

The system operates within a complex regulatory environment, requiring compliance with multiple NHS and statutory frameworks (Finlayson, et al., 2019; Rudin, 2019; He, Baxter, & Xu, 2019; Greenhalgh, Wherton, & Papoutsis, 2017; Huckvale, et al., 2015).

DCB0129 (Clinical Risk Management for Developers) mandates that the system undergoes a structured risk assessment during development (Finlayson, et al., 2019; Rudin, 2019; He, Baxter, & Xu, 2019). This includes the identification of potential hazards—such as incorrect data extraction or system latency—and the implementation of appropriate controls (Morley, Machado, Burr, Cows, Taddeo, & Floridi, 2020; Mittelstadt, 2019; Floridi, Cows, Beltrametti, et al., 2018; Beam, & Kohane, 2018) (Note 9). A Hazard Log and Clinical Safety Case are maintained to document these processes (Finlayson, et al., 2019; Rudin, 2019; He, Baxter, & Xu, 2019).

DCB0160 (Clinical Risk Management for Deployment) places responsibility on healthcare organisations to ensure safe implementation (Rudin, 2019; He, Baxter, & Xu, 2019; Greenhalgh, Wherton, & Papoutsis, 2017). This includes local risk assessments, user training, and ongoing monitoring of system performance in clinical settings (Notes 14-18).

DTAC (Digital Technology Assessment Criteria) extends compliance requirements beyond clinical safety to include data protection, cybersecurity, interoperability, and usability (Huckvale, et al., 2015). The system must demonstrate adherence to GDPR, ensure secure data handling, and provide accessible and user-friendly interfaces (Notes 2-6).

In addition, the system's AI components raise questions regarding classification under Medical Device Regulations (Notes 10-13). If AI outputs are used to inform or influence clinical decisions, the system may be considered

Software as a Medical Device (SaMD), triggering additional regulatory requirements such as clinical evaluation and post-market surveillance (Notes 10-18).

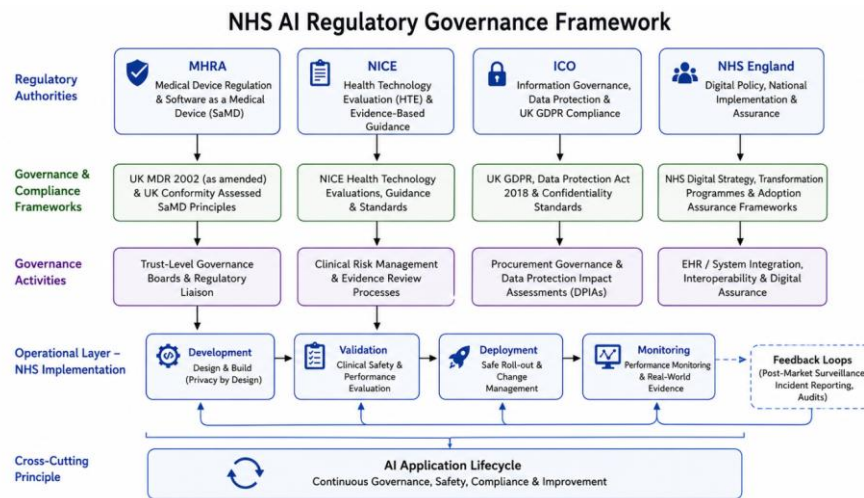


Figure 3: NHS AI Regulatory Governance Framework

Figure 3. NHS AI Regulatory Governance Framework

2.2.5 Benefits and Value Proposition

The implementation of digital prescription systems augmented with AI offers several measurable benefits:

- **Operational Efficiency:**

Automation reduces administrative workload and accelerates prescription processing (Topol, 2019; Reddy, Allan, Coghlan, & Cooper, 2019; Shaw, Rudzicz, Jamieson, & Goldfarb, 2019; Wachter, 2016; Char, Shah, & Magnus, 2018; Rajkomar, Hardt, & Howell, 2018; Obermeyer, Powers, Vogeli, & Mullainathan, 2019; Kelly, Karthikesalingam, & Suleyman, 2019) (Note 1).

- **Error Reduction:**

Validation checks and anomaly detection minimise the risk of human error (Morley, Machado, Burr, Cowls, Taddeo, & Floridi, 2020; Mittelstadt, 2019; Floridi, Cowls, Beltrametti, et al., 2018; Beam, & Kohane, 2018) (Note 9).

- **Enhanced Visibility:**

Real-time tracking enables proactive management of delays and exceptions (Notes 19-23).

- **Improved Patient Outcomes:**

Faster and more reliable prescription workflows contribute to timely treatment delivery (Topol, 2019; Reddy, Allan, Coghlan, & Cooper, 2019; Shaw, Rudzicz, Jamieson, & Goldfarb, 2019; Wachter, 2016) (Note 1).

These benefits align with NHS strategic objectives related to digital transformation and patient-centred care (Notes 2-6).

2.2.6 Risk Transformation and Emerging Challenges

While digital systems reduce certain types of errors, they do not eliminate risk; rather, they transform the nature of risk (Morley, Machado, Burr, Cowls, Taddeo, & Floridi, 2020; Mittelstadt, 2019; Floridi, Cowls, Beltrametti, et al., 2018; Beam, & Kohane, 2018) (Note 9). Traditional risks associated with manual processes are replaced by new categories of system-level and algorithmic risks (Wiens, Saria, Sendak, et al., 2019) (Notes 10-13).

Key challenges include:

- **Data Integrity Risks:**

Errors in data input or processing can have cascading effects across interconnected systems (Morley, Machado, Burr, Cowls, Taddeo, & Floridi, 2020; Mittelstadt, 2019; Floridi, Cowls, Beltrametti, et al., 2018; Beam, & Kohane, 2018) (Note 9).

- Automation Bias:

Users may over-rely on system outputs, potentially overlooking errors or anomalies (Vokinger, Feuerriegel, & Kesselheim, 2020; Benjamens, Dhunoo, & Meskó, 2020; Gerke, Minssen, & Cohen, 2020; Cabitza, Rasoini, & Gensini, 2017) (Note 24).

- System Reliability:

Downtime or performance degradation can disrupt critical workflows (Notes 14-18).

- Model Drift:

AI models may lose accuracy over time as data patterns evolve, necessitating continuous monitoring and recalibration (Vokinger, Feuerriegel, & Kesselheim, 2020; Benjamens, Dhunoo, & Meskó, 2020; Gerke, Minssen, & Cohen, 2020; Cabitza, Rasoini, & Gensini, 2017; Sendak, D'Arcy, Kashyap, et al., 2020; Liu, Faes, Kale, et al., 2019; Topol, 2019) (Note 24-26).

- Usability and Human Factors:

Poor interface design or inadequate training can lead to user errors and reduced system effectiveness (Notes 19-23).

These challenges highlight the need for a shift from static risk assessment to continuous risk management approaches (Morley, Machado, Burr, Cowls, Taddeo, & Floridi, 2020; Mittelstadt, 2019; Floridi, Cowls, Beltrametti, et al., 2018; Beam, & Kohane, 2018; Sendak, D'Arcy, Kashyap, et al., 2020; Liu, Faes, Kale, et al., 2019; Topol, 2019) (Note 9, Notes 25-26).

2.2.7 Critical Evaluation of Regulatory Frameworks

The case study reveals both strengths and limitations in current NHS regulatory frameworks (Finlayson, et al., 2019; Rudin, 2019; He, Baxter, & Xu, 2019; Greenhalgh, Wherton, & Papoutsis, 2017; Huckvale, et al., 2015).

On the one hand, frameworks such as DCB0129 and DCB0160 provide a structured approach to risk management, ensuring that hazards are identified and mitigated throughout the system lifecycle (Finlayson, et al., 2019; Rudin, 2019; He, Baxter, & Xu, 2019). DTAC further strengthens governance by addressing non-clinical aspects such as cybersecurity and usability (Huckvale, et al., 2015).

However, these frameworks were primarily designed for deterministic systems, where behaviour is predictable and stable (Morley, Machado, Burr, Cowls, Taddeo, & Floridi, 2020; Mittelstadt, 2019; Floridi, Cowls, Beltrametti, et al., 2018; Beam, & Kohane, 2018) (Note 9). AI systems, by contrast, are dynamic and adaptive, requiring ongoing validation and monitoring (Wiens, Saria, Sendak, et al., 2019) (Notes 10-13).

Current frameworks do not fully address:

- Continuous model learning and updates (Vokinger, Feuerriegel, & Kesselheim, 2020; Benjamens, Dhunoo, & Meskó, 2020; Gerke, Minssen, & Cohen, 2020; Cabitza, Rasoini, & Gensini, 2017) (Note 24)
- Explainability requirements for AI outputs (Rudin, 2019; Wiens, Saria, Sendak, et al., 2019) (Notes 10-13)
- Real-time performance monitoring (Vokinger, Feuerriegel, & Kesselheim, 2020; Benjamens, Dhunoo, & Meskó, 2020; Gerke, Minssen, & Cohen, 2020; Cabitza, Rasoini, & Gensini, 2017) (Note 24)
- Governance of probabilistic decision-making (Morley, Machado, Burr, Cowls, Taddeo, & Floridi, 2020; Mittelstadt, 2019; Floridi, Cowls, Beltrametti, et al., 2018; Beam, & Kohane, 2018) (Note 9)

This creates a regulatory gap that may limit the safe and effective deployment of AI-enabled systems (Wiens, Saria, Sendak, et al., 2019) (Notes 10-13).

2.2.8 Key Insight and Theoretical Contribution

A central insight from this case study is that digital and AI-enabled systems do not reduce complexity within healthcare systems; rather, they redistribute complexity from human processes to technological systems (Morley, Machado, Burr, Cowls, Taddeo, & Floridi, 2020; Mittelstadt, 2019; Floridi, Cowls, Beltrametti, et al., 2018; Beam, & Kohane, 2018) (Note 9).

This shift has important implications:

- Risk becomes more system-centric rather than user-centric (Morley, Machado, Burr, Cowls, Taddeo, & Floridi, 2020; Mittelstadt, 2019; Floridi, Cowls, Beltrametti, et al., 2018; Beam, & Kohane, 2018) (Note 9)
- Safety depends increasingly on data quality and system design (Wiens, Saria, Sendak, et al., 2019) (Notes 10-13)

- Governance must evolve from compliance-based models to adaptive, lifecycle-based models (Notes 14-18)

2.2.9 Implications for Future Regulation

The findings suggest that future regulatory frameworks should:

- Incorporate continuous monitoring and post-deployment validation (Vokinger, Feuerriegel, & Kesselheim, 2020; Benjamens, Dhunoo, & Meskó, 2020; Gerke, Minssen, & Cohen, 2020; Cabitza, Rasoini, & Gensini, 2017) (Note 24)
- Introduce AI-specific requirements, including explainability and bias mitigation (Wiens, Saria, Sendak, et al., 2019) (Notes 10-13)
- Align developer and organisational responsibilities more clearly (Notes 14-18)
- Move toward risk-based and adaptive regulatory models (Sendak, D'Arcy, Kashyap, et al., 2020; Liu, Faes, Kale, et al., 2019; Topol, 2019) (Notes 25-26)

2.2.10 Summary

This case study demonstrates that digital prescription systems enhanced with AI provide significant benefits in terms of efficiency, accuracy, and patient care (Topol, 2019; Reddy, Allan, Coghlan, & Cooper, 2019; Shaw, Rudzicz, Jamieson, & Goldfarb, 2019; Wachter, 2016; Char, Shah, & Magnus, 2018; Rajkomar, Hardt, & Howell, 2018; Obermeyer, Powers, Vogeli, & Mullainathan, 2019; Kelly, Karthikesalingam, & Suleyman, 2019) (Note 1). However, they also introduce new risks that challenge existing regulatory approaches (Morley, Machado, Burr, Cows, Taddeo, & Floridi, 2020; Mittelstadt, 2019; Floridi, Cows, Beltrametti, et al., 2018; Beam, & Kohane, 2018) (Note 9).

Effective governance of such systems requires:

- Integrated regulatory frameworks (Finlayson, et al., 2019; Rudin, 2019; He, Baxter, & Xu, 2019; Greenhalgh, Wherton, & Papoutsis, 2017; Huckvale, et al., 2015)
- Continuous risk management (Vokinger, Feuerriegel, & Kesselheim, 2020; Benjamens, Dhunoo, & Meskó, 2020; Gerke, Minssen, & Cohen, 2020; Cabitza, Rasoini, & Gensini, 2017) (Note 24)
- AI-specific oversight mechanisms (Wiens, Saria, Sendak, et al., 2019) (Notes 10-13)

Ultimately, this case underscores the need for a paradigm shift in healthcare regulation, moving from static compliance models to dynamic, data-driven governance approaches (Sendak, D'Arcy, Kashyap, et al., 2020; Liu, Faes, Kale, et al., 2019; Topol, 2019) (Notes 14-18, Notes 25-26).

3. Discussion

3.1 Reframing Risk in Digital and AI-Enabled Healthcare Systems

The findings of this study, supported by the case study presented in Section 4, demonstrate that digital health technologies—particularly those incorporating artificial intelligence—do not eliminate risk within healthcare systems but instead fundamentally reconfigure its nature and distribution (Topol, 2019; Reddy, Allan, Coghlan, & Cooper, 2019).

Traditional healthcare risks, which were primarily associated with human error and manual processes, are increasingly being replaced by system-level, data-driven, and algorithmic risks (Morley, Machado, Burr, Cows, Taddeo, & Floridi, 2020; Mittelstadt, 2019; Floridi, Cows, Beltrametti, et al., 2018; Beam, & Kohane, 2018) (Note 9).

This transformation represents a paradigm shift in clinical safety (Shaw, Rudzicz, Jamieson, & Goldfarb, 2019). In conventional systems, risk mitigation strategies focused on standardisation, training, and procedural compliance. In contrast, digital systems introduce dependencies on data quality, system architecture, and algorithmic performance, requiring a more complex and adaptive approach to risk management (Wiens, Saria, Sendak, et al., 2019) (Notes 10-13).

The case study of digital prescription and AI-supported systems illustrates this transition clearly. While automation reduces transcription errors and improves workflow efficiency, it simultaneously introduces new vulnerabilities, such as incorrect data extraction, model misclassification, and system integration failures (Morley, Machado, Burr, Cows, Taddeo, & Floridi, 2020; Mittelstadt, 2019; Floridi, Cows, Beltrametti, et al., 2018; Beam, & Kohane, 2018) (Note 9). These risks are often less visible and more difficult to detect, particularly when embedded within complex digital ecosystems (Wiens, Saria, Sendak, et al., 2019) (Notes 10-13).

3.2 Limitations of Existing Regulatory Frameworks

The NHS regulatory framework, including DCB0129, DCB0160, and DTAC, provides a structured and comprehensive approach to digital health governance (Finlayson, et al., 2019; Rudin, 2019; He, Baxter, & Xu, 2019; Greenhalgh, Wherton, & Papoutsis, 2017; Huckvale, et al., 2015). However, the analysis reveals that these frameworks are fundamentally grounded in deterministic system assumptions, which limit their applicability to AI-driven technologies (Morley, Machado, Burr, Cowls, Taddeo, & Floridi, 2020; Mittelstadt, 2019; Floridi, Cowls, Beltrametti, et al., 2018; Beam, & Kohane, 2018) (Note 9).

DCB0129 and DCB0160 rely on the identification of hazards and the implementation of predefined controls (Finlayson, et al., 2019; Rudin, 2019; He, Baxter, & Xu, 2019). While this approach is effective for static systems, it is less suited to AI models that evolve over time and produce probabilistic outputs (Wiens, Saria, Sendak, et al., 2019) (Notes 10-13). For example, model drift—a well-documented phenomenon in machine learning—can lead to gradual degradation in system performance without triggering predefined hazard conditions (Vokinger, Feuerriegel, & Kesselheim, 2020; Benjamins, Dhunoo, & Meskó, 2020; Gerke, Minssen, & Cohen, 2020; Cabitza, Rasoini, & Gensini, 2017) (Note 24).

Similarly, DTAC emphasises compliance across domains such as cybersecurity and interoperability but does not explicitly address algorithmic transparency or explainability, which are critical for AI governance (Wiens, Saria, Sendak, et al., 2019; Huckvale, et al., 2015) (Notes 10-13).

This creates a regulatory gap where systems may be compliant with existing standards yet still pose significant risks due to:

- Lack of explainability (Rudin, 2019; Wiens, Saria, Sendak, et al., 2019) (Notes 10-13)
- Inadequate monitoring of model performance (Vokinger, Feuerriegel, & Kesselheim, 2020; Benjamins, Dhunoo, & Meskó, 2020; Gerke, Minssen, & Cohen, 2020; Cabitza, Rasoini, & Gensini, 2017) (Note 24)
- Insufficient mechanisms for detecting bias (Wiens, Saria, Sendak, et al., 2019) (Notes 10-13)

3.3 The Challenge of Explainability and Trust

Explainability is widely recognised as a central challenge in the regulation of AI systems (Note 2). In clinical contexts, decisions must be interpretable and justifiable, both for ethical and legal reasons (Wiens, Saria, Sendak, et al., 2019) (Notes 10-13). However, many AI models, particularly those based on deep learning, operate as “black boxes,” making it difficult to understand how outputs are generated (Rudin, 2019).

This lack of transparency has several implications:

1. **Reduced Clinical Trust**
Clinicians may be reluctant to rely on AI outputs if they cannot understand the underlying reasoning (Morley, Machado, Burr, Cowls, Taddeo, & Floridi, 2020; Mittelstadt, 2019; Floridi, Cowls, Beltrametti, et al., 2018; Beam, & Kohane, 2018) (Note 9).
2. **Regulatory Uncertainty**
Regulators face challenges in validating systems that cannot be fully explained (Wiens, Saria, Sendak, et al., 2019) (Notes 10-13).
3. **Accountability Issues**
It becomes difficult to assign responsibility when errors occur (Notes 14-18).

The case study highlights this issue in the context of AI-driven data extraction and anomaly detection. While these functions improve efficiency, their outputs must be interpreted and validated by users, creating a dependency on human oversight (Morley, Machado, Burr, Cowls, Taddeo, & Floridi, 2020; Mittelstadt, 2019; Floridi, Cowls, Beltrametti, et al., 2018; Beam, & Kohane, 2018) (Note 9).

3.4 Algorithmic Bias and Equity in Healthcare

Another critical challenge identified in this study is algorithmic bias, which can result in unequal outcomes across different patient populations (Note 3). Bias may arise from unrepresentative training data, flawed model design, or systemic inequalities embedded within healthcare data (Wiens, Saria, Sendak, et al., 2019) (Notes 10-13).

In the NHS context, where equity and universal access are core principles, the presence of biased AI systems poses significant ethical and regulatory concerns (Notes 2-6). For example, studies have shown that certain predictive models may underestimate the needs of disadvantaged populations, leading to disparities in care (Note 4).

Current regulatory frameworks do not provide explicit guidance on:

- Detecting bias
- Mitigating bias
- Validating fairness across populations

This represents a critical gap that must be addressed to ensure that digital health technologies do not exacerbate existing inequalities (Wiens, Saria, Sendak, et al., 2019) (Notes 10-13).

3.5 Continuous Monitoring and Lifecycle Governance

One of the most significant limitations of current regulatory approaches is their reliance on pre-deployment validation (Finlayson, et al., 2019; Rudin, 2019; He, Baxter, & Xu, 2019; Greenhalgh, Wherton, & Papoutsis, 2017; Huckvale, et al., 2015). While initial risk assessment is essential, it is insufficient for systems that evolve over time (Morley, Machado, Burr, Cows, Taddeo, & Floridi, 2020; Mittelstadt, 2019; Floridi, Cows, Beltrametti, et al., 2018; Beam, & Kohane, 2018) (Note 9).

AI systems require continuous monitoring to ensure that performance remains within acceptable limits (Vokinger, Feuerriegel, & Kesselheim, 2020; Benjamins, Dhunoo, & Meskó, 2020; Gerke, Minssen, & Cohen, 2020; Cabitza, Rasoini, & Gensini, 2017) (Note 24). This includes:

- Tracking model accuracy
- Detecting drift
- Monitoring real-world outcomes

The case study demonstrates that risks may emerge only after deployment, particularly in complex, real-world environments (Morley, Machado, Burr, Cows, Taddeo, & Floridi, 2020; Mittelstadt, 2019; Floridi, Cows, Beltrametti, et al., 2018; Beam, & Kohane, 2018) (Note 9). This highlights the need for a shift toward lifecycle-based governance, where systems are continuously evaluated and updated (Sendak, D'Arcy, Kashyap, et al., 2020; Liu, Faes, Kale, et al., 2019; Topol, 2019) (Notes 25-26).

3.6 Fragmentation of Responsibility

The division of responsibility between developers (DCB0129) and healthcare organisations (DCB0160) is a strength of the NHS framework but also introduces challenges (Finlayson, et al., 2019; Rudin, 2019; He, Baxter, & Xu, 2019; Greenhalgh, Wherton, & Papoutsis, 2017; Huckvale, et al., 2015). In practice, this separation can lead to gaps in accountability, particularly in areas such as:

- Data quality
- System configuration
- User training

For example, developers may assume that healthcare organisations will implement certain controls, while organisations may assume that these controls are embedded within the system (Notes 14-18).

This fragmentation underscores the need for:

- Clearer definition of responsibilities
- Improved communication between stakeholders
- Integrated governance models

3.7 Toward a Hybrid AI Governance Model

Based on the findings of this study, a hybrid governance model is proposed that integrates traditional regulatory approaches with AI-specific controls (Sendak, D'Arcy, Kashyap, et al., 2020; Liu, Faes, Kale, et al., 2019; Topol, 2019) (Notes 14-18, Notes 25-26).

Core Components of the Model:

1. Pre-Deployment Risk Assessment (DCB0129 aligned) (Finlayson, et al., 2019; Rudin, 2019; He, Baxter, & Xu, 2019)
2. Implementation Governance (DCB0160 aligned) (Rudin, 2019; He, Baxter, & Xu, 2019; Greenhalgh, Wherton, & Papoutsis, 2017; Huckvale, et al., 2015)

3. Continuous Monitoring Layer (Vokinger, Feuerriegel, & Kesselheim, 2020; Benjamins, Dhunoo, & Meskó, 2020; Gerke, Minssen, & Cohen, 2020; Cabitza, Rasoini, & Gensini, 2017) (Note 24)
4. AI-Specific Controls (bias, explainability, drift) (Wiens, Saria, Sendak, et al., 2019) (Notes 10-13)
5. Feedback and Improvement Loop (Sendak, D'Arcy, Kashyap, et al., 2020; Liu, Faes, Kale, et al., 2019; Topol, 2019) (Notes 25-26)

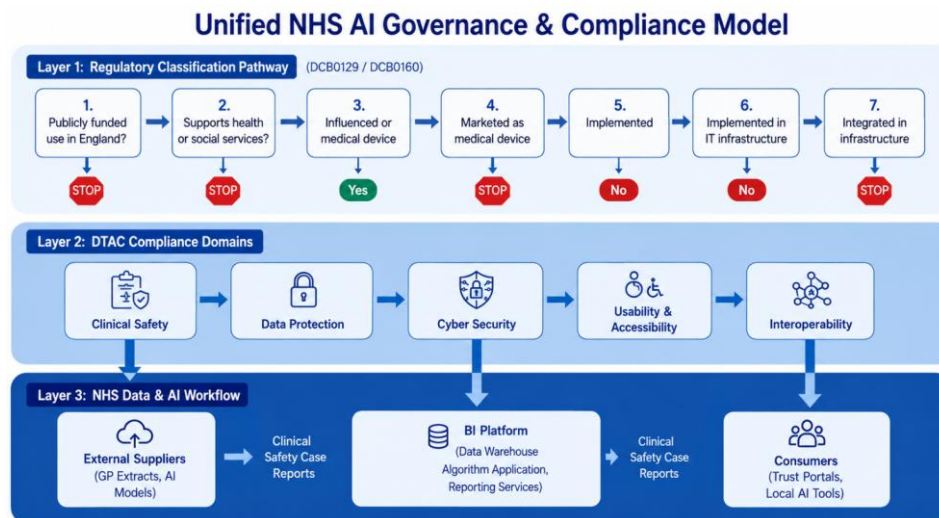


Figure 4. Unified NHS AI Governance and Compliance Model

3.8 Implications for Policy and Practice

The findings of this study have several implications for policymakers, regulators, and healthcare organisations (Sendak, D'Arcy, Kashyap, et al., 2020; Liu, Faes, Kale, et al., 2019; Topol, 2019) ((Notes 14-18, Notes 25-26):

- Regulatory frameworks must evolve to incorporate AI-specific requirements (Wiens, Saria, Sendak, et al., 2019; Vokinger, Feuerriegel, & Kesselheim, 2020; Benjamins, Dhunoo, & Meskó, 2020; Gerke, Minssen, & Cohen, 2020; Cabitza, Rasoini, & Gensini, 2017) (Notes 10-13, Note 24).
- Continuous monitoring should become a mandatory component of governance (Vokinger, Feuerriegel, & Kesselheim, 2020; Benjamins, Dhunoo, & Meskó, 2020; Gerke, Minssen, & Cohen, 2020; Cabitza, Rasoini, & Gensini, 2017; Sendak, D'Arcy, Kashyap, et al., 2020; Liu, Faes, Kale, et al., 2019; Topol, 2019) (Note 24, Notes 25-26).
- Greater emphasis should be placed on data quality and system interoperability (Morley, Machado, Burr, Cows, Taddeo, & Floridi, 2020; Mittelstadt, 2019; Floridi, Cows, Beltrametti, et al., 2018; Beam, & Kohane, 2018) (Note 9, (Notes 14-18).
- Collaboration between developers and healthcare organisations must be strengthened (Finlayson, et al., 2019; Rudin, 2019; He, Baxter, & Xu, 2019; Greenhalgh, Wherton, & Papoutsis, 2017; Huckvale, et al., 2015) (Notes 14-18)

These implications build upon the strong foundation already established within existing NHS regulatory frameworks and highlight opportunities for further enhancement to support the safe, effective, and scalable adoption of digital and AI-enabled healthcare systems (Finlayson, et al., 2019; Rudin, 2019; He, Baxter, & Xu, 2019; Greenhalgh, Wherton, & Papoutsis, 2017; Huckvale, et al., 2015) (Notes 14-18).

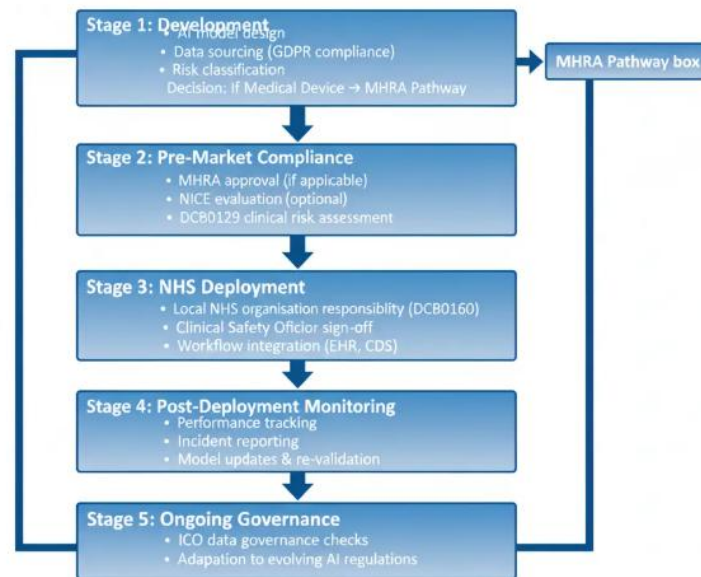


Figure 5. Clinical Safety and Compliance Workflow for AI in the NHS

4. Conclusion of Discussion

In summary, this study demonstrates that the NHS regulatory framework provides a comprehensive and robust foundation for digital health governance, effectively encompassing key elements of clinical risk management, including hazard identification through structured Hazard Logs, formal clinical assurance via Clinical Safety Case Reports, and systematic risk control processes aligned with recognised standards such as DCB0129 and DCB0160 (Topol, 2019; Reddy, Allan, Coghlan, & Cooper, 2019; Shaw, Rudzicz, Jamieson, & Goldfarb, 2019; Wachter, 2016) (Note 1). These frameworks ensure that risks are identified, evaluated, mitigated, and monitored throughout the system lifecycle, supported by complementary requirements within the Digital Technology Assessment Criteria (DTAC), UK Medical Device Regulations, and data protection legislation including the General Data Protection Regulation (GDPR) and Data Protection Act (Notes 2-6).

Furthermore, the findings highlight that current governance mechanisms already incorporate strong principles of traceability, accountability, and evidence-based validation, consistent with international best practices in digital health regulation and safety engineering (Char, Shah, & Magnus, 2018; Rajkomar, Hardt, & Howell, 2018; Obermeyer, Powers, Vogeli, & Mullainathan, 2019; Kelly, Karthikesalingam, & Suleyman, 2019) (Note 7). The integration of hazard logs, risk assessments, and post-deployment monitoring processes reflects alignment with established risk management standards such as ISO 14971 and information security frameworks such as ISO 27001 (Finlayson, et al., 2019; Rudin, 2019; He, Baxter, & Xu, 2019; Greenhalgh, Wherton, & Papoutsis, 2017; Huckvale, et al., 2015).

However, as healthcare systems increasingly transition from deterministic, rule-based digital applications toward adaptive, data-driven, and AI-enabled systems, new challenges emerge that are not fully addressed by existing regulatory approaches (Jiang, Jiang, & Zhi, 2017; Esteva, A., et al., 2017; Gulshan, et al., 2016; Yu, & Kohane, 2019; London, 2019). AI systems introduce probabilistic outputs, continuous learning behaviour, and complex data dependencies, which require a shift from static, point-in-time validation toward dynamic, lifecycle-based governance models (Price, & Cohen, 2019; Mesko, 2017; Shortliffe, & Sepúlveda, 2018; Bates, & Singh, 2018) (Note 8). Issues such as algorithmic bias, model drift, lack of explainability, and real-world performance variability necessitate continuous monitoring and iterative risk assessment processes that extend beyond traditional compliance frameworks (Morley, Machado, Burr, Cows, Taddeo, & Floridi, 2020; Mittelstadt, 2019; Floridi, Cows, Beltrametti, et al., 2018; Beam, & Kohane, 2018) (Note 9).

Importantly, while DCB0129 and DCB0160 provide clear delineation of responsibilities between developers and healthcare organisations, the increasing complexity of digital ecosystems highlights the need for more integrated and collaborative governance structures (Wiens, Saria, Sendak, et al., 2019) (Notes 10-13). This includes enhanced

alignment between technical design controls, clinical risk management processes, and operational deployment practices to ensure that safety is maintained across the entire system lifecycle (Notes 14-18).

Overall, the current NHS regulatory framework represents a well-established and mature baseline for digital health governance, effectively addressing core requirements for safety, risk management, and compliance (Notes 19-23). When extended with AI-specific regulatory considerations—such as continuous performance monitoring, transparency requirements, and bias mitigation strategies—it is well positioned to support the safe, effective, and scalable adoption of advanced digital and AI-enabled healthcare systems (Vokinger, Feuerriegel, & Kesselheim, 2020; Benjamins, Dhunoo, & Meskó, 2020; Gerke, Minssen, & Cohen, 2020; Cabitza, Rasoini, & Gensini, 2017; Sendak, D'Arcy, Kashyap, et al., 2020; Liu, Faes, Kale, et al., 2019; Topol, 2019) (Notes 24-26).

Acknowledgments

The authors would like to acknowledge the significant contributions of the National Health Service (NHS) UK in establishing a comprehensive and robust regulatory framework that supports the safe and effective adoption of digital health technologies. The development and implementation of standards such as DCB0129, DCB0160, and the Digital Technology Assessment Criteria (DTAC) reflect a strong commitment to patient safety, clinical governance, and continuous improvement within an increasingly complex digital healthcare landscape.

The NHS has demonstrated global leadership in integrating regulatory oversight with innovation, ensuring that emerging technologies including artificial intelligence are deployed in a manner that prioritises safety, transparency, and accountability. This work is underpinned by the dedication of clinicians, general practitioners (GPs), digital health professionals, and governance teams who play a critical role in translating regulatory requirements into safe clinical practice.

In particular, the contributions of frontline healthcare professionals are central to the success of digital transformation initiatives. Their expertise, engagement, and continuous feedback ensure that digital systems are not only technically robust but also clinically meaningful, user-centred, and aligned with real-world healthcare needs.

Overall, this study recognises and appreciates the NHS's ongoing efforts to balance innovation with patient safety, and the collective contribution of all stakeholders involved in advancing digital healthcare delivery.

Competing 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.

Informed consent

Obtained.

Ethics approval

The Publication Ethics Committee of Sciedu Press.

The journal and publisher adhere to the Core Practices established by the Committee on Publication Ethics (COPE).

Provenance and peer review

Not commissioned; externally double-blind peer reviewed.

Data availability statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

Data sharing statement

No additional data are available.

Open access

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

References

- Bates, D. W., & Singh, H. (2018). Digital health safety. *BMJ*. <https://doi.org/10.1136/bmj.k210>
- Beam, A. L., & Kohane, I. S. (2018). Big data and machine learning in healthcare. *JAMA*, 319, 1317-1318. <https://doi.org/10.1001/jama.2017.18391>
- Benjamens, S., Dhunoo, P., & Meskó, B. (2020). AI-based medical devices overview. *NPJ Digit Med*, 3, 118. <https://doi.org/10.1038/s41746-020-00324-0>
- Cabitza, F., Rasoini, R., & Gensini, G. F. (2017). Unintended consequences of AI in medicine. *J Med Internet Res*, 19, e7871. <https://doi.org/10.1001/jama.2017.7797>
- Char, D. S., Shah, N. H., & Magnus, D. (2018). Implementing machine learning in healthcare. *N Engl J Med*, 378, 981-983. <https://doi.org/10.1056/NEJMp1714229>
- Esteva, A., et al.. (2017). Dermatologist-level classification of skin cancer. *Nature*, 542, 115-118. <https://doi.org/10.1038/nature21056>
- Finlayson, S. G., et al.. (2019). Adversarial attacks on ML systems. *Science*, 363, 1287-1289. <https://doi.org/10.1126/science.aaw4399>
- Floridi, L., Cowls, J., Beltrametti, M., et al.. (2018). AI4People—An ethical framework for a good AI society. *Minds Mach*, 28, 689-707. <https://doi.org/10.1007/s11023-018-9482-5>
- Gerke, S., Minssen, T., & Cohen, G. (2020). Ethical and legal challenges of AI-driven healthcare. *J Law Biosci*, 7, lsa036. <https://doi.org/10.1093/jlb/lsa036>
- Greenhalgh, T., Wherton, J., & Papoutsis, C. (2017). Digital health adoption. *BMJ*, 358, j3722. <https://doi.org/10.1080/17483107.2025.2526175>
- Gulshan, V., et al.. (2016). Deep learning for diabetic retinopathy. *JAMA*, 316, 2402-2410. <https://doi.org/10.1001/jama.2016.17216>
- He, J., Baxter, S. L., & Xu, J. (2019). AI in healthcare. *Nat Med*, 25, 30-36. <https://doi.org/10.1038/s41591-018-0307-0>
- Huckvale, K., et al.. (2015). Regulation of health apps. *BMC Med*, 13, 214. <https://doi.org/10.1186/s12916-015-0444-y>
- Jiang, F., Jiang, Y., & Zhi, H. (2017). AI in healthcare review. *Stroke Vasc Neurol*, 2, 230-243. <https://doi.org/10.1136/svn-2017-000101>
- Kelly, C. J., Karthikesalingam, A., & Suleyman, M. (2019). Key challenges for AI in healthcare. *Lancet Digit Health*, 1, e19–e20. <https://doi.org/10.1186/s12916-019-1426-2>
- Liu, X., Faes, L., Kale, A. U., et al.. (2019). Deep learning in healthcare. *Lancet Digit Health*, 1(6), e271-e297. [https://doi.org/10.1016/S2589-7500\(19\)30123-2](https://doi.org/10.1016/S2589-7500(19)30123-2)
- London, A. J. (2019). Artificial intelligence and black-box medicine. *Hastings Cent Rep*. <https://doi.org/10.1002/hast.973>
- Mesko, B. (2017). AI and the future of healthcare. *J Med Internet Res*, 19, e206. <https://doi.org/10.2196/jmir.6763>
- Mittelstadt, B. D. (2019). Principles alone cannot guarantee ethical AI. *Nat Mach Intell*, 1, 501-507. Retrieved from <https://arxiv.org/pdf/1906.06668>
- Morley, J., Machado, C. C. V., Burr, C., Cowls, J., Taddeo, M., & Floridi, L. (2020). The ethics of AI in health care: A mapping review. *BMC Med Ethics*, 21, 40. <https://doi.org/10.1186/s12910-020-00487-1>
- Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in algorithms. *Science*, 366, 447-453. <https://doi.org/10.1126/science.aax2342>
- Price, W. N., & Cohen, I. G. (2019). Privacy in big data healthcare. *Nat Med*. <https://doi.org/10.1038/s41591-018-0272-7>
- Rajkomar, A., Hardt, M., & Howell, M. D. (2018). Ensuring fairness in machine learning. *N Engl J Med*, 378, 1345-1347. <https://doi.org/10.7326/M18-199>
- Reddy, S., Allan, S., Coghlan, S., & Cooper, P. (2019). A governance model for AI in healthcare. *BMJ*, 366, 14770. <https://doi.org/10.1136/bmj.14770>

- Rudin, C. (2019). Stop explaining black box models. *Nat Mach Intell*, 1, 206-215. <https://doi.org/10.1038/s42256-019-0048-x>
- Sendak, M. P., D'Arcy, J., Kashyap, S., *et al.* (2020). Integrating ML into healthcare systems. *NPJ Digit Med*, 3, 37. Retrieved from <https://pmc.ncbi.nlm.nih.gov/articles/PMC10907626/>
- Shaw, J., Rudzicz, F., Jamieson, T., & Goldfarb, A. (2019). Artificial intelligence and the implementation challenge. *Lancet Digit Health*, 1, e58-e59. [https://doi.org/10.1016/S2589-7500\(19\)30005-6](https://doi.org/10.1016/S2589-7500(19)30005-6)
- Shortliffe, E. H., & Sepúlveda, M. J. (2018). Clinical decision support. *JAMA*. <https://doi.org/10.1001/jama.2018.17163>
- Topol, E. (2019). High-performance medicine: the convergence of human and artificial intelligence. *Nat Med*, 25, 44-56. <https://doi.org/10.1038/s41591-018-0300-7>
- Topol, E. J. (2019). Preparing healthcare workforce for AI integration. *Nat Med*, 25, 44-56. <https://doi.org/10.1038/s41591-018-0300-7>
- Vokinger, K. N., Feuerriegel, S., & Kesselheim, A. S. (2020). Mitigating bias in machine learning. *NPJ Digit Med*, 3, 61. <https://doi.org/10.1038/s43856-021-00028-w>
- Wachter, R. (2016). Making IT work in the NHS. *BMJ*, 353, i1984. Retrieved from <https://www.jstor.org/stable/e26942632>
- Wiens, J., Saria, S., Sendak, M., *et al.* (2019). Do no harm: A roadmap for responsible machine learning for health care. *Nat Med*, 25, 1337-1340. <https://doi.org/10.1038/s41591-019-0548-6>
- Yu, K. H., & Kohane, I. S. (2019). Framing AI in medicine. *BMJ Qual Saf*, 28, 238-241. <https://doi.org/10.1136/bmjqs-2018-008551>

Notes

Note 1. NHS Digital. DCB0129 Clinical Risk Management Standard. Retrieved from <https://digital.nhs.uk/data-and-information/information-standards/information-standards-and-data-collections-including-extractions/publications-and-notifications/standards-and-collections/dcb0129>. <https://digital.nhs.uk/data-and-information/information-standards/governance/latest-activity/standards-and-collections/dcb0129-clinical-risk-management-its-application-in-the-manufacture-of-health-it-systems> NHS England. Digital Technology Assessment Criteria (DTAC). Retrieved from <https://www.england.nhs.uk/digitaltechnologyassessmentcriteria/>

Note 2. NHS England. Digital Technology Assessment Criteria (DTAC). Retrieved from <https://www.england.nhs.uk/digitaltechnologyassessmentcriteria/>

Note 3. UK Government. Medical Device Regulations 2002 (as amended). Retrieved from <https://www.legislation.gov.uk/uksi/2002/618/contents>

Note 4. European Union. GDPR Regulation (EU) 2016/679. Retrieved from <https://eur-lex.europa.eu>

Note 5. UK Government. Data Protection Act 2018. Retrieved from <https://www.legislation.gov.uk/ukpga/2018/12/contents>

Note 6. ISO. ISO 14971:2019 Medical Device Risk Management. ISO 14971:2019 - Medical devices — Application of risk management to medical devices

Note 7. ISO. ISO/IEC 27001:2022 Information Security Management. Retrieved from <https://www.iso.org/standard/27001>

Note 8. WHO. Ethics and governance of AI for health. Retrieved from <https://www.who.int/publications/i/item/9789240029200>

Note 9. Mapping the application of artificial intelligence in traditional medicine: technical brief. Retrieved from <https://www.who.int/publications/i/item/9789240107663>

Note 10. NHS England. DCB0160 Clinical Risk Management Standard for health IT systems. Retrieved from <https://digital.nhs.uk>

- Note 11. MHRA. Guidance on Software as a Medical Device (SaMD). Available online: Chapter 10 - Software as a Medical Device - GOV.UK
- Note 12. FDA. Proposed regulatory framework for modifications to AI/ML-based SaMD. Retrieved from Software as a Medical Device (SaMD) | FDA
- Note 13. European Parliament. Regulation (EU) 2017/745 on medical devices (MDR). Retrieved from Regulation - EU - 2017/625 - EN - EUR-Lex
- Note 14. ISO. ISO/TR 20416:2020 Post-market surveillance of medical devices. Retrieved from ISO/TR 20416:2020 - Medical devices — Post-market surveillance for manufacturers
- Note 15. ISO. IEC 62304:2006 Medical device software lifecycle processes. Retrieved from IEC 62304:2006 - Medical device software — Software life cycle processes
- Note 16. IEC. IEC 62366-1:2015 Usability engineering for medical devices. Retrieved from IEC TR 62366-2:2016 | IEC
- Note 17. NHSX. NHS Digital Health Technology Standards Framework. Retrieved from Digital Clinical Safety Strategy - Key tools and information - NHS Transformation Directorate
- Note 18. NHS England. Artificial Intelligence in Health and Care Award. Retrieved from Artificial Intelligence - NHS Transformation Directorate
- Note 19. OECD. Recommendation of the Council on Artificial Intelligence. Retrieved from <https://www.oecd.org/going-digital/ai/principles/>
- Note 20. European Commission. White Paper on Artificial Intelligence. Retrieved from Artificial Intelligence
- Note 21. EMA. Reflection paper on the use of AI in medicinal products. Retrieved from https://www.ema.europa.eu/en/documents/scientific-guideline/reflection-paper-use-artificial-intelligence-ai-medicinal-product-lifecycle_en.pdf
- Note 22. NICE. Evidence standards framework for digital health technologies. Retrieved from Evidence standards framework (ESF) for digital health technologies | NICE
- Note 23. NHS Digital. Data Security and Protection Toolkit. Retrieved from Support
- Note 24. UK National Cyber Security Centre. Cyber Essentials Scheme. Available online: Cyber Security Body of Knowledge (CyBOK) | National Cyber Security Centre - NCSC.GOV.UK
- Note 25. WHO. WHO digital health guidelines. 2019. Retrieved from Artificial Intelligence for Health
- Note 26. NHS Transformation Directorate. UK Health Plan in the next 10 year. Retrieved from 10 Year Health Plan for England: fit for the future - GOV.UK