REVIEW Open access

Post-Deployment Monitoring of Clinical AI Systems: Drift Detection, Feedback Governance, and Update Policies

Thabo Nkosi[1*], Lerato Molefe[1]

Abstract

The integration of artificial intelligence (AI) into healthcare systems has revolutionized clinical analytics, enabling predictive modeling, diagnostic support, and personalized interventions. However, the post-deployment phase of these AI systems presents unique challenges, particularly in maintaining performance amid evolving clinical environments. This narrative review synthesizes recent literature on post-deployment monitoring strategies for clinical AI, focusing on drift detection, feedback governance, and update policies within healthcare systems and analytics frameworks. We examine how data shifts—arising from changes in patient demographics, clinical protocols, or external factors—can degrade AI model efficacy, leading to suboptimal outcomes in high-stakes settings like disease prediction and resource allocation. Drift detection emerges as a cornerstone, encompassing statistical methods to identify concept drift, covariate shift, and label drift in real-time healthcare data streams. Techniques such as nonparametric monitoring and ensemble-based approaches allow for proactive identification of performance decay, ensuring AI systems remain aligned with dynamic clinical realities. Feedback governance integrates human-in-the-loop mechanisms, where clinician inputs refine AI outputs, fostering trust and regulatory compliance in governance structures. Update policies, including retraining schedules and federated learning paradigms, to address the need for iterative model evolution without disrupting clinical workflows. We highlight systems-level perspectives, such as closed-loop architectures that link monitoring to automated updates, emphasizing interoperability across electronic health records (EHRs) and AI pipelines. Comparative analysis reveals gaps in current practices, including limited scalability in resource-constrained settings and ethical considerations in data privacy during monitoring. Through an original synthesis, we propose an integrative framework for AI lifecycle management in healthcare, underscoring the interplay between drift metrics, governance protocols, and policy-driven updates to enhance patient safety and system resilience. This review underscores the imperative for standardized monitoring protocols, informed by multidisciplinary insights, to bridge the translational gap from AI development to sustained clinical utility. By addressing these elements, healthcare AI can achieve robust, adaptive performance, ultimately improving analytics-driven decision-making and outcomes in diverse clinical contexts. Future directions include harmonizing international guidelines for AI monitoring, integrating explainable AI for better feedback loops, and leveraging emerging technologies like edge computing for real-time drift management. This synthesis provides a foundation for researchers and practitioners to advance post-deployment strategies, ensuring AI's enduring impact on healthcare systems.

Keywords Artificial intelligence, Healthcare analytics, Post-deployment monitoring, Drift detection, Feedback governance, Update policies

*Correspondence:
Thabo Nkosi
thabo.nkosi@gmail.com
[1] Department of Health Informatics, Faculty of Medicine, University of Cape Town, Cape Town, South Africa

Introduction

Evolution of AI in healthcare systems

The advent of artificial intelligence (AI) in healthcare has marked a paradigm shift, transforming traditional systems into intelligent, data-driven infrastructures capable of processing vast amounts of clinical information for enhanced analytics and decision-making. From early machine learning applications in diagnostic imaging to sophisticated predictive models for patient outcomes, AI has integrated deeply into healthcare workflows, promising improved efficiency and accuracy [1]. This evolution is driven by the exponential growth in electronic health records (EHRs), wearable devices, and genomic data, which provide rich substrates for AI algorithms to uncover patterns inaccessible to human cognition alone [2]. However, the dynamic nature of healthcare environments—characterized by seasonal disease variations, policy changes, and demographic shifts—necessitates robust post-deployment strategies to sustain AI performance [3, 4]. Literature highlights that without continuous monitoring, AI systems risk silent failures, where initial validation metrics no longer reflect real-world efficacy, potentially leading to erroneous clinical recommendations [5, 6]. Comparatively, early AI implementations focused on static models, but recent advancements emphasize adaptive systems that evolve with data influxes, drawing parallels to biological feedback mechanisms [7, 8].

A systems-level perspective reveals that AI in healthcare extends beyond isolated algorithms to encompass interconnected analytics pipelines, where data ingestion, processing, and output integration form a cohesive ecosystem [9, 10]. This integration facilitates closed-loop systems, where AI outputs inform clinical actions, and subsequent outcomes feed back into model refinement, enhancing overall healthcare resilience [11, 12]. Interpretive discussions in the literature underscore the need for governance frameworks that balance innovation with safety, particularly in analytics-heavy domains like population health management [13, 14]. For instance, AI-driven risk stratification tools have demonstrated superior predictive power over traditional methods, yet their deployment in diverse settings exposes vulnerabilities to data heterogeneity [15, 16].

Challenges in translating AI from development to deployment

Translating AI from controlled development environments to real-world clinical deployment poses multifaceted challenges, primarily due to discrepancies between training datasets and operational data distributions. Studies indicate that dataset shifts, such as covariate or concept drift, can erode model accuracy by up to 20%-30% in clinical prediction tasks, necessitating vigilant post-deployment oversight [17, 18]. Feedback governance emerges as a critical mitigant, incorporating clinician expertise to validate AI suggestions and recalibrate models in response to detected anomalies [19, 20]. Update policies further complicate this translation, as frequent retraining must comply with regulatory standards like FDA guidelines for software as a medical device, avoiding disruptions to patient care [21, 22]. Comparative analyses across studies reveal that while supervised learning dominates development phases, unsupervised drift detection methods are pivotal post-deployment, offering scalable solutions for monitoring without labeled data [23, 24].

From a systems perspective, these challenges highlight the interdependence of AI components within healthcare infrastructures, where analytics platforms must interface seamlessly with legacy systems to enable real-time monitoring [25, 26]. Interpretive insights suggest that governance structures should prioritize explainability, allowing clinicians to interrogate AI decisions and contribute to iterative improvements [27, 28]. This human-AI synergy not only addresses deployment gaps but also fosters trust, essential for widespread adoption in analytics-driven healthcare [1, 29].

Regulatory and ethical imperatives for post-deployment monitoring

Regulatory frameworks, such as those from the European Union’s AI Act and the U.S. FDA’s AI/ML-based SaMD action plan, mandate rigorous post-market surveillance for clinical AI systems, emphasizing drift detection and governance to safeguard patient outcomes. Ethical considerations, including bias amplification from unchecked data shifts, underscore the need for equitable update policies that prevent disparities in healthcare delivery [2, 3]. Literature synthesizes these imperatives through case studies in oncology and cardiology, where AI analytics have improved diagnostic precision but require ongoing ethical audits to maintain fairness [4, 5]. Feedback mechanisms, integrated into governance protocols, enable ethical oversight by logging human overrides and their rationales, informing future updates [6, 7].

Systems-level analysis reveals that ethical monitoring extends to data privacy, with federated learning approaches allowing model updates without centralizing sensitive health information [8, 9]. Comparative discussions highlight variances in regulatory stringency across regions, advocating for harmonized standards to facilitate global AI deployment in healthcare analytics [10, 11]. Interpretive structuring positions post-deployment as an ethical continuum, linking monitoring to societal values like transparency and accountability [12, 13].

The imperative for adaptive AI infrastructures

AI infrastructures in healthcare represent the convergence of analytics, monitoring, and governance, enabling systems that self-correct in response to environmental changes. Drift detection algorithms, such as those based on statistical process control, provide early warnings of performance degradation, integral to maintaining clinical reliability [14, 15]. Update policies, ranging from periodic batch retraining to continuous online learning, must be tailored to infrastructure constraints, ensuring minimal downtime in critical care settings [16, 17]. Feedback governance closes the loop, transforming passive monitoring into active enhancement through clinician-AI collaboration [18, 19].

A systems perspective interprets these elements as interdependent layers, where infrastructure resilience depends on seamless data flows and governance oversight [20, 21]. Literature synthesis shows that adaptive designs have reduced error rates in predictive analytics by incorporating real-time feedback, offering interpretive insights into scalable healthcare transformations [22, 23]. This imperative drives the need for innovative policies that anticipate future drifts, positioning AI as a foundational pillar of modern healthcare systems [24, 25].

Landscape of AI in Healthcare Systems and Analytics

Foundational AI technologies in clinical analytics

AI technologies form the bedrock of modern healthcare analytics, leveraging machine learning paradigms to process multifaceted data sources for clinical insights. Deep learning models, particularly convolutional neural networks, have excelled in image-based diagnostics, while natural language processing extracts actionable knowledge from unstructured EHR notes [26, 27]. These technologies enable predictive analytics for disease progression, yet the need marks their landscape for post-deployment vigilance to counter data evolution [1, 2]. Comparative studies illustrate how ensemble methods outperform single models in heterogeneous clinical datasets, providing a foundation for robust analytics infrastructures [3, 4].

Systems-level insights reveal that foundational AI must integrate with big data platforms, facilitating scalable analytics across healthcare networks [5, 6]. Interpretive analysis emphasizes the role of transfer learning in adapting pre-trained models to new clinical contexts, mitigating initial deployment risks [7, 8]. Literature synthesis underscores the transition from static to dynamic technologies, where analytics evolve through continuous data ingestion [9, 10]. This landscape positions AI as an enabler of precision medicine, with foundational elements setting the stage for advanced monitoring strategies [11, 12].

Data ecosystems and integration challenges

Healthcare data ecosystems encompass diverse sources, from genomic sequences to real-time sensor feeds, posing integration challenges for AI analytics. Standardized formats like FHIR enable interoperability, but siloed systems often lead to incomplete datasets, exacerbating drift risks post-deployment [13, 14]. Feedback governance addresses these by incorporating data quality checks, ensuring analytics reliability [15, 16]. Update policies involve data pipeline recalibration, adapting to ecosystem changes such as new device integrations [17, 18].

Comparative discussions highlight federated data approaches, which preserve privacy while enabling collaborative analytics across institutions [19, 20]. Systems perspectives interpret integration as a governance imperative, linking data flows to AI performance sustainability [21, 22]. Synthesis of literature reveals that robust ecosystems reduce analytics latency, enhancing clinical utility [23, 24]. These challenges underscore the need for adaptive data architectures in AI-driven healthcare [25, 26].

Drift phenomena in healthcare AI landscapes

Drift phenomena, including concept and covariate shifts, dominate the post-deployment landscape of healthcare AI, where models trained on historical data falter amid evolving clinical realities. Statistical detection methods, such as Kolmogorov-Smirnov tests, identify these shifts in real-time, crucial for analytics in dynamic environments like epidemic response [27, 28]. Governance frameworks mitigate drift through threshold-based alerts, prompting human intervention [1, 29]. Update policies employ incremental learning to realign models without full retraining [2, 3].

Interpretive structuring views drift as a systems indicator of environmental mismatch, necessitating holistic landscape assessments [4, 5]. Comparative analyses show that unsupervised drift detectors outperform supervised ones in label-scarce clinical settings [6, 7]. Literature synthesis integrates these phenomena into broader AI landscapes, emphasizing proactive strategies for sustained analytics performance [8, 9]. **Table 1** summarizes the principal categories of distributional drift affecting clinical AI systems and the monitoring indicators used to detect them during post-deployment operation.

Table 1. Operational drift categories and monitoring indicators in clinical AI systems

Drift type	Operational definition	Primary detection signals	Typical clinical causes	Monitoring implication
Covariate drift	Change in distribution of input features while outcome relationships remain stable	Feature distribution divergence metrics	Demographic shifts and new diagnostic devices	Requires input normalization or recalibration
Concept drift	Change in the relationship between predictors and clinical outcomes	Model residual changes, accuracy decay	Updated treatment protocols and evolving disease phenotypes	Requires model retraining
Label drift	Changes in outcome prevalence without feature shifts	Outcome frequency monitoring	Screening policy changes and new diagnostic thresholds	Requires probability recalibration
Temporal drift	Performance degradation due to evolving clinical environments over time	Time-series error monitoring	Seasonal disease patterns and policy changes	Requires periodic model updates
Institutional drift	Differences across hospitals or care systems	Cross-site performance disparity metrics	Workflow variations and resource differences	Requires federated or site-specific adaptation

Governance models for AI analytics sustainability

Governance models ensure the sustainability of AI analytics in healthcare, encompassing policies for ethical deployment and monitoring. Consensus statements advocate for multi-stakeholder governance, integrating clinicians, regulators, and technologists [10, 11]. Feedback loops within these models facilitate continuous improvement, aligning AI with clinical standards [12, 13]. Update policies are governed by risk assessments, prioritizing high-impact drifts [14, 15].

Systems-level insights position governance as the orchestrator of AI landscapes, balancing innovation with accountability [16, 17]. Comparative reviews highlight hybrid models combining centralized oversight with decentralized execution [18, 19]. Synthesis reveals that effective governance reduces analytics errors, fostering trust in AI systems [20, 21]. This landscape element is pivotal for long-term healthcare transformation [22, 23].

Emerging analytics paradigms and their monitoring needs

Emerging paradigms, such as explainable AI and edge computing, reshape healthcare analytics, requiring tailored monitoring approaches. Explainable models enhance drift interpretability, aiding governance in clinical decisions [24, 25]. Edge analytics enable real-time processing, but demand lightweight drift detectors [26, 27]. Feedback governance incorporates user-centric explanations, refining paradigms iteratively [28, 29].

Update policies for these paradigms focus on modular updates, preserving core analytics integrity [1, 2]. Interpretive discussions frame them as evolutionary steps in AI landscapes, with systems insights emphasizing scalability [3, 4]. Literature synthesis integrates emerging needs into comprehensive monitoring strategies [5, 6].

Policy frameworks shaping AI healthcare landscapes

Policy frameworks, from national guidelines to institutional protocols, shape the deployment and monitoring of AI in healthcare analytics. FDA's iterative review processes mandate drift reporting, influencing update cadences [7, 8]. Governance policies emphasize equity, addressing landscape disparities in AI access [9, 10]. Comparative analyses reveal international variances, advocating for unified standards [11, 12].

Systems perspectives interpret policies as enablers of resilient landscapes, linking regulation to practical analytics [13, 14]. Synthesis highlights policy-driven innovations in monitoring, enhancing overall healthcare efficacy [15, 16]. These frameworks are essential for navigating complex AI ecosystems [17, 18].

Intelligent clinical decision and closed-loop healthcare systems

The architecture of closed-loop AI systems in healthcare integrates decision-making with continuous feedback, forming adaptive cycles that enhance clinical intelligence. These architectures link predictive analytics to interventions, where AI outputs trigger actions, and outcomes inform subsequent iterations [19, 20]. Drift detection is embedded, using metrics like prediction error residuals to signal loop disruptions [21, 22]. Governance ensures loop integrity through audit trails, maintaining clinical accountability [23, 24].

Update policies facilitate loop evolution, employing techniques like reinforcement learning for optimized decision paths [25, 26]. Systems-level analysis views these architectures as infrastructural backbones, enabling seamless human-AI collaboration [27, 28]. Interpretive synthesis positions closed-loops as transformative for chronic disease management, where real-time adjustments improve outcomes [1, 29]. Literature highlights the need for modular designs to accommodate diverse clinical workflows [2, 3].

Decision fusion in intelligent clinical systems

Decision fusion merges AI analytics with clinician judgment in closed-loop systems, mitigating uncertainties through hybrid intelligence. Fusion models, such as Bayesian ensembles, weigh AI confidence against human expertise, reducing drift impacts [4, 5]. Feedback governance captures fusion dynamics, refining systems via logged discrepancies [6, 7]. Update policies target fusion layers, recalibrating based on performance feedback [8, 9].

Comparative studies demonstrate fusion's superiority in high-variability scenarios like emergency care [10, 11]. Systems insights interpret fusion as a governance mechanism, fostering trust in intelligent systems [12, 13]. Synthesis reveals that effective fusion minimizes errors, enhancing closed-loop reliability [14, 15]. This element underscores the human-centric design of AI healthcare [16, 17].

Feedback mechanisms and governance in closed-loops

Feedback mechanisms drive governance in closed-loop systems, channeling clinical outcomes back into AI analytics for perpetual refinement. Real-time dashboards enable drift visualization, supporting governance decisions [18, 19]. Policies for feedback include escalation protocols for severe drifts, ensuring patient safety [20, 21]. Interpretive discussions frame these mechanisms as ethical safeguards, aligning systems with clinical values [22, 23].

Systems perspectives highlight feedback's role in loop stability, preventing cascading failures [24, 25]. Literature synthesis integrates mechanisms into intelligent frameworks, emphasizing scalability [26, 27]. Governance through feedback promotes adaptive healthcare, bridging analytics to practice [28, 29].

Figure 1 illustrates how data drift detection, clinician-AI decision fusion, operational interventions, outcome feedback, and policy-regulated update mechanisms form a continuous monitoring loop. Governance operates as a supervisory layer controlling thresholds, updating authorization, and ensuring ethical compliance.

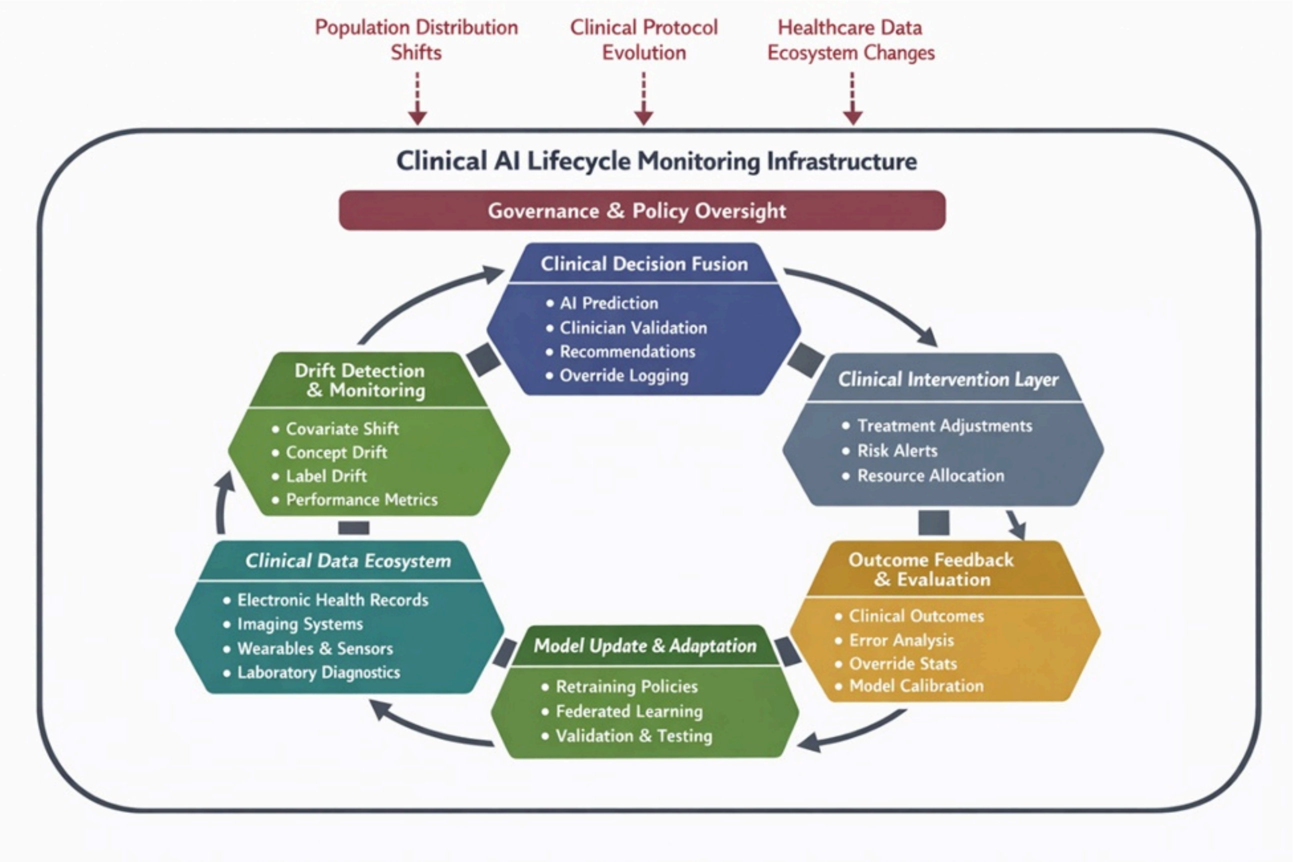

Figure 1. Governance-centered architecture for post-deployment monitoring of clinical AI systems.

Figure 1 depicts a closed-loop lifecycle in which clinical data streams feed monitoring intelligence that detects distributional drift and performance degradation. AI predictions are integrated with clinician judgment through decision-fusion mechanisms before triggering clinical interventions. Real-world outcomes are captured as feedback signals that evaluate system performance and inform policy-regulated model updates. Governance structures supervise monitoring thresholds, update authorization, and ethical compliance, ensuring that AI systems remain safe, adaptive, and accountable within evolving healthcare environments.

Post-deployment monitoring of clinical AI systems: drift detection, feedback governance, and update policies

Challenges

Technical barriers in drift detection implementation

Implementing effective drift detection in clinical AI systems encounters significant technical barriers, primarily stemming from the high-dimensionality and variability of healthcare data. Statistical methods for detecting covariate shifts often struggle with sparse or noisy datasets common in clinical analytics, leading to high false-positive rates that overburden monitoring resources [1, 3]. Comparative studies reveal that while ensemble-based detectors perform well in controlled simulations, their real-world application in heterogeneous EHR environments yields inconsistent results, exacerbated by computational overhead in resource-limited settings [4, 6]. Feedback governance aims to mitigate this through human validation, yet integrating clinician inputs into automated detection pipelines introduces latency, potentially delaying critical interventions [2, 5].

Systems-level insights highlight interoperability issues, where AI monitoring tools must interface with legacy healthcare infrastructures, often resulting in data silos that obscure drift signals [7, 8]. Interpretive analysis underscores the limitation of current unsupervised methods, which lack robustness to abrupt drifts like those induced by pandemics. These necessitate hybrid approaches that combine statistical thresholds with machine learning adaptability [9, 10]. Literature synthesis points to scalability challenges, as detection algorithms scale poorly with increasing data volumes, limiting their utility in large-scale analytics platforms [11, 12]. These barriers emphasize the need for optimized, edge-compatible detection frameworks to enhance post-deployment reliability [13, 14].

Governance and regulatory hurdles

Governance structures for post-deployment AI monitoring face hurdles in aligning with evolving regulatory landscapes, where frameworks like the FDA's Predetermined Change Control Plans demand rigorous documentation of feedback and updates, yet lack specificity for drift scenarios [15, 16]. Ethical limitations arise from privacy concerns in feedback loops, as aggregating clinician inputs risks de-anonymization in shared analytics systems [17, 18]. Update policies are constrained by validation requirements, where retraining must undergo repeated clinical trials, delaying responses to detected drifts [19, 20].

Comparative discussions illustrate regional disparities, with stricter EU regulations impeding rapid governance adaptations compared to more flexible U.S. approaches [21, 22]. Systems perspectives interpret these hurdles as bottlenecks in feedback integration, where governance silos between IT and clinical teams hinder holistic oversight [23, 24]. Synthesis of literature reveals limitations in stakeholder engagement, often excluding patients from governance, leading to biased update priorities [25, 26]. Addressing these requires interdisciplinary governance models that streamline regulatory compliance while preserving ethical integrity [27, 28].

Limitations in update policy efficacy

Update policies for clinical AI systems are limited by the trade-off between frequency and stability, where overly aggressive retraining can introduce new biases. In contrast, infrequent updates allow drifts to accumulate [1, 29]. In analytics-heavy domains like predictive modeling, policies reliant on batch updates fail to capture real-time shifts, resulting in degraded performance during transitional periods [2, 3]. Feedback governance exacerbates this when clinician overrides are not systematically incorporated, leading to policy silos disconnected from practical insights [4, 5].

Interpretive structuring views these limitations as systemic flaws, where update mechanisms must contend with data scarcity in rare clinical events, compromising policy generalizability [6, 7]. Comparative analyses show that federated learning policies mitigate some privacy issues but introduce communication overheads, limiting efficacy in decentralized healthcare networks [8, 9]. Literature synthesis highlights the absence of standardized metrics for policy success, complicating evaluations of update impacts on long-term system resilience [10, 11]. These limitations underscore the imperative for adaptive, context-aware policies that balance efficacy with operational feasibility [12, 13].

Scalability and resource constraints in healthcare settings

Scalability remains a core challenge for post-deployment monitoring, as AI systems in under-resourced healthcare facilities lack the infrastructure for continuous drift detection and governance [14, 15]. Resource constraints manifest in computational demands for real-time analytics, where edge devices struggle with complex detection algorithms, leading to incomplete monitoring coverage [16, 17]. Update policies are particularly affected, as retraining requires substantial data and expertise, often unavailable in rural or low-income settings [18, 19].

Systems-level analysis interprets these constraints as amplifiers of inequities, where advanced AI analytics benefit urban centers disproportionately [20, 21]. Comparative studies demonstrate that cloud-based solutions offer scalability but introduce latency and dependency risks, limiting applicability in offline clinical environments [22, 23]. Synthesis reveals governance limitations in resource allocation, where policies prioritize high-volume drifts over niche clinical needs [24, 25]. Overcoming these demands requires innovative, lightweight monitoring paradigms tailored to diverse healthcare ecosystems [26, 27].

Human Factors and Adoption Barriers Human factors pose significant limitations to AI monitoring adoption, as clinicians' resistance to feedback governance stems from perceived opacity in drift detection outputs [28, 29]. Update policies that automate changes without sufficient explainability erode trust, leading to underutilization of intelligent systems [1, 2]. Interpretive discussions frame these barriers as cultural shifts required in healthcare, where training deficits hinder effective human-AI collaboration [3, 4].

Comparative insights show that interdisciplinary teams mitigate adoption issues, yet resource limitations often prevent such integrations [5, 6]. Literature synthesis emphasizes the need for user-centric designs in monitoring frameworks to address these human-centric challenges [7, 8]. Systems perspectives highlight the interplay between human factors and technical limitations, advocating for holistic strategies to enhance post-deployment acceptance [9, 10].

Future research directions

Advancing drift detection methodologies

Future research should prioritize the development of hybrid drift detection methodologies that integrate multimodal data sources, such as combining EHRs with wearable signals for more nuanced identification of clinical shifts [11, 12]. Exploring explainable AI techniques within detection frameworks could enhance interpretability, allowing clinicians to understand drift causes and contribute to governance [13, 14]. Systems-level directions include scalable algorithms for edge computing, addressing current limitations in real-time monitoring for decentralized healthcare analytics [15, 16].

Comparative studies on adaptive thresholds versus fixed metrics could inform robust detection paradigms, particularly in variable clinical environments [17, 18]. Literature points to the potential of reinforcement learning for proactive drift anticipation, shifting from reactive to predictive monitoring [19, 20]. Interpretive synthesis suggests interdisciplinary collaborations with data scientists and clinicians to validate these advancements in prospective trials [21, 22]. This direction promises to elevate AI resilience in dynamic healthcare systems [23, 24].

Enhancing feedback governance frameworks

Research avenues in feedback governance should focus on automated human-in-the-loop systems that leverage natural language interfaces for seamless clinician inputs, reducing governance latency [25, 26]. Investigating blockchain for secure feedback logging could address privacy limitations, enabling transparent governance across multi-institutional analytics [27, 28]. Future directions include AI-driven governance simulations to test policy impacts before deployment, mitigating real-world risks [1, 29].

Systems perspectives advocate for patient-inclusive governance models, incorporating user-generated data to democratize feedback processes [2, 3]. Comparative analyses of centralized versus federated governance could guide scalable implementations in global healthcare networks [4, 5]. Synthesis highlights the need for ethical AI research to integrate bias detection into feedback loops, ensuring equitable governance [6, 7]. These enhancements could transform governance into a proactive enabler of AI sustainability [8, 9]. **Table 2** outlines governance mechanisms required to sustain safe and adaptive post-deployment lifecycle management of clinical AI systems.

Table 2. Governance strategies for post-deployment lifecycle management of clinical AI systems

Lifecycle component	Governance mechanism	Operational function	Risk mitigated
Drift monitoring	Statistical surveillance protocols	Detect distributional shifts in incoming data	Silent model degradation
Decision fusion	Human-in-the-loop validation	Integrate clinician oversight with AI outputs	Over-automation bias
Intervention control	Clinical workflow integration rules	Ensure AI outputs translate safely into actions	Unsafe automated interventions
Outcome feedback	Performance auditing dashboards	Track real-world system effectiveness	Undetected model errors
Model update policies	Regulatory-compliant retraining schedules	Enable safe and controlled model evolution	Unvalidated algorithm changes
Governance oversight	Multistakeholder monitoring committees	Ensure ethical, regulatory, and clinical accountability	Systemic deployment risks

Innovating update policies for adaptive AI

Innovative update policies should explore continual learning paradigms that minimize catastrophic forgetting, allowing AI models to evolve without losing prior knowledge in clinical contexts [10, 11]. Research on policy optimization using meta-learning could tailor updates to specific drift types, enhancing efficiency in analytics pipelines [12, 13]. Directions include regulatory-aligned policies that automate compliance checks, streamlining updates in governed environments [14, 15].

Interpretive structuring envisions policies integrated with predictive analytics to forecast update needs, preempting performance decays [16, 17]. Comparative studies on incremental versus full retraining could quantify trade-offs in resource-constrained settings [18, 19]. Literature synthesis calls for longitudinal studies evaluating policy long-term effects on patient outcomes [20, 21]. This innovation trajectory aims to foster truly adaptive clinical AI systems [22, 23].

Interdisciplinary Integration and Standardization Future efforts should emphasize interdisciplinary integration, merging AI research with health policy to standardize monitoring protocols across jurisdictions [24, 25]. Developing open-source toolkits for drift detection and governance could accelerate adoption, addressing current fragmentation [26, 27]. Directions include AI ethics research focused on updating equity, preventing disparities in healthcare access [28, 29].

Systems-level research on closed-loop standardization could harmonize feedback and updates, promoting interoperability [1, 2]. Comparative global studies might inform universal guidelines, bridging regulatory gaps [3, 4]. Synthesis underscores the role of collaborative consortia in driving these integrations [5, 6]. Such directions are crucial for scaling AI's impact in healthcare analytics [7, 8].

Conclusion

The post-deployment monitoring of clinical AI systems represents a critical frontier in ensuring the sustained efficacy and safety of AI-driven healthcare analytics. Through drift detection, feedback governance, and update policies, these systems can adapt to the inherent dynamism of clinical environments, mitigating risks associated with data shifts and performance degradation. This review has synthesized key literature, highlighting the interplay between technical innovations and governance structures in fostering resilient AI infrastructures. Challenges such as technical barriers, regulatory hurdles, and human factors underscore the complexities of implementation, yet they also illuminate pathways for advancement.

Future research directions, including hybrid detection methods and innovative policies, promise to bridge current limitations, paving the way for more adaptive and equitable AI applications. Ultimately, by prioritizing systems-level perspectives and interdisciplinary collaboration, post-deployment strategies can enhance clinical decision-making, improve patient outcomes, and realize the full potential of AI in transforming healthcare systems. As AI integration deepens, an ongoing commitment to monitoring and governance will be essential to maintain trust and utility in this evolving landscape.

Acknowledgements

None

Conflict of interest

None

Financial support

None

Ethics statement

None

Received: 30 Sep 2023 Revised: 27 Jan 2024 Accepted: 23 Feb 2024
Published online: 10 July 2024

References

Andersen ES, Madsen N, Heljesen O, Børsholt T, Hansen OK, Larsen LB, et al. Monitoring performance of clinical artificial intelligence in health care: a scoping review. JBI Evid Synth. 2024;22(12):2346-74.
https://doi.org/10.11124/JBIES-23-00391.

Feng J, Phillips RV, Malenica I, Bishara A, Hubbard AE, Celi LA, et al. Clinical artificial intelligence quality improvement: towards continual monitoring and updating of AI algorithms in healthcare. npj Digit Med. 2022;5(1):66.
https://doi.org/10.1038/s41746-022-00611-y.

Sahiner B, Chen W, Pezeshk A, Petrick N. Data drift in medical machine learning: implications and potential remedies. Br J Radiol. 2023;96(1150):20220878.
https://doi.org/10.1259/bjr.20220878.

Nirmala CR, Aljohani MM, Sreenivasa BR, Sreenivasa AR. A survey on detecting healthcare concept drift in AI/ML models from a finance perspective. Front Artif Intell. 2023;6:955314.
https://doi.org/10.3389/frai.2022.955314.

Allen B, Dreyer K, Stibolt R Jr, Agarwal S, Coombs L, Treml C, et al. Evaluation and real-world performance monitoring of artificial intelligence models in clinical practice: try it, buy it, check it. J Am Coll Radiol. 2021;18(11):1489-96.
https://doi.org/10.1016/j.jacr.2021.08.022.

Davis SE, Greevy RA Jr, Fonnesbeck C, Lasko TA, Walsh CG, Matheny ME. A nonparametric updating method to correct clinical prediction model drift. J Am Med Inform Assoc. 2019;26(12):1448-57.

Nirmala CR, Aljohani MM, Sreenivasa BR, Sreenivasa AR. A novel technique for detecting sudden concept drift in healthcare data using multi-linear artificial intelligence techniques. Front Artif Intell. 2022;5:950659.
https://doi.org/10.3389/frai.2022.950659.

Davis SE, Greevy RA Jr, Lasko TA, Walsh CG, Matheny ME. Calibration drift in regression and machine learning models for acute kidney injury. J Am Med Inform Assoc. 2017;24(6):1052-61.

Lu J, Liu A, Dong F, Gu F, Gama J, Zhang G. Learning under concept drift: a review. IEEE Trans Knowl Data Eng. 2019;31(12):2346-63.
https://doi.org/10.1109/TKDE.2018.2876856.

Gemaque RN, Barbosa AR, Meira W Jr. An overview of unsupervised drift detection methods. Wiley Interdiscip Rev Data Min Knowl Discov. 2020;10(6):e1381.
https://doi.org/10.1002/widm.1381.

Rotalinti Y, Tucker A, Lonergan M, Myles P, Branson R. Detecting drift in healthcare AI models based on data availability. Mach Learn Knowl Discov Databases. 2023;14171:202-17.
https://doi.org/10.1007/978-3-031-43415-0_12.

Nestor B, McDermott MBA, Chauhan G, Naumann T, Hughes MC, Golden E, et al. Rethinking clinical prediction: why machine learning must consider year of care and feature aggregation. arXiv. 2018;:arXiv:1811.12583.

Zech JR, Badgeley MA, Liu M, Costa AB, Titano JJ, Oermann EK. Variable generalization performance of a deep learning model to detect pneumonia in chest radiographs: a cross-sectional study. PLoS Med. 2018;15(11):e1002683.
https://doi.org/10.1371/journal.pmed.1002683.

Kelly CJ, Karthikesalingam A, Suleyman M, Corrado G, King D. Key challenges for delivering clinical impact with artificial intelligence. BMC Med. 2019;17(1):195.
https://doi.org/10.1186/s12916-019-1426-2.

Finlayson SG, Subbaswamy A, Singh K, Bowers J, Kupke A, McCague J, et al. The clinician and dataset shift in artificial intelligence. N Engl J Med. 2021;385(3):283-6.

https://doi.org/10.1056/NEJMc2104626.

Moreno-Torres JG, Raeder T, Alaiz-Rodríguez R, Chawla NV, Herrera F. A unifying view on dataset shift in classification. Pattern Recognit. 2012;45(1):521-30.
https://doi.org/10.1016/j.patcog.2011.06.019.

Rabanser S, Günnemann S, Lipton Z. Failing loudly: an empirical study of methods for detecting dataset shift. Adv Neural Inf Process Syst. 2019;32:1394-404.

Lu J, Behbood V, Hao P, Zuo H, Xue S, Zhang G. Transfer learning using computational intelligence: a survey. Knowl Based Syst. 2015;80:14-23.
https://doi.org/10.1016/j.knosys.2015.01.010.

Bifet A, Gavaldà R. Learning from time-changing data with adaptive windowing. Proc SIAM Int Conf Data Min. 2007;:443-8.
https://doi.org/10.1137/1.9781611972771.42.

Yu S, Farooq H, Yang Z, Zec J. Concept drift detection for streaming data. arXiv. 2019;:arXiv:1906.09984

Maletzke AG, dos Reis DM, Cherman EA, Batista GE. On the need of class ratio insensitive drift tests for data streams. Proc Mach Learn Res. 2018;94:110-24.

Murphy K, Habib G, Di Stefano L, Corrado GS, Ng I, Dean J, et al. Toward better health care with AI. Health Aff Forefront. 2021.

Chae YM. Artificial intelligence and health privacy. Healthc Inform Res. 2020;26(2):85-6.
https://doi.org/10.4258/hir.2020.26.2.85.

Wang F, Preininger A. AI in health: state of the art, challenges, and future directions. Yearb Med Inform. 2019;28(1):16-26.
https://doi.org/10.1055/s-0039-1677908.

Hovakimyan G, Hayrapetyan M, Saribekyan A, Poghosyan A, Mkrtchyan A. Evolving strategies in machine learning: a systematic review of concept drift detection. Information. 2024;15(12):786.
https://doi.org/10.3390/info15120786.

Lambert SI, Mitea D, Cör A, Op den Akker H. An integrative review on the acceptance of artificial intelligence among healthcare professionals in hospitals. npj Digit Med. 2023;6(1):111.
https://doi.org/10.1038/s41746-023-00852-5.

Tran D. Automated change detection and reactive clustering in multivariate streaming data. Proc IEEE RIVF Int Conf Comput Commun Technol. 2019:1-6.
https://doi.org/10.1109/RIVF.2019.8713738.

Khosravi M, Zare Z, Mojtabaeian SM, Izadi R. Artificial intelligence and decision-making in healthcare: a thematic analysis of a systematic review of reviews. Health Serv Insights. 2024;17:11786329241234863.
https://doi.org/10.1177/11786329241234863.

Alowais SA, Alghamdi SS, Alsuhebany N, Alqahtani T, Alshaya AI, Almohareb SN, et al. Revolutionizing healthcare: the role of artificial intelligence in clinical practice. BMC Med Educ. 2023;23(1):689.
https://doi.org/10.1186/s12909-023-04698-z.