REVIEW **Open access**

Foundation Models in Healthcare Systems: Architectural Integration and Oversight Considerations

Rafael Costa[1], Bruno Teixeira[2], Victor Santos[1*]

Abstract

Foundation models, characterized by their large-scale pretraining on diverse datasets, represent a transformative paradigm in artificial intelligence (AI) applications for healthcare systems and analytics. These models, often based on transformer architectures, enable generalist capabilities that extend beyond narrow task-specific AI, facilitating integration into complex healthcare infrastructures. This review synthesizes recent literature on the architectural integration of foundation models into healthcare systems, emphasizing their role in enhancing clinical analytics, decision support, and operational efficiency while addressing critical oversight considerations, including ethical, regulatory, and safety frameworks.

In healthcare systems, foundation models are increasingly deployed to process multimodal data streams, including electronic health records (EHRs), medical imaging, and real-time patient monitoring. Architectural integration involves embedding these models within hospital information systems, enabling seamless data ingestion, inference, and feedback loops. For instance, models like those adapted from large language models (LLMs) support natural language processing for EHR mining, predictive analytics for disease progression, and generative tasks for synthetic data augmentation. Oversight considerations are paramount, encompassing regulatory compliance, bias mitigation, and human-AI collaboration protocols to ensure patient safety and equity.

The synthesis highlights key architectural patterns: federated learning for privacy-preserving model training, hybrid human-AI workflows for clinical decision-making, and adaptive systems for continuous model recalibration. Analytics applications span precision medicine, where foundation models integrate genomic and clinical data for personalized interventions, to population health management, optimizing resource allocation through predictive modeling. Ethical oversight includes checklists for AI deployment in low- and middle-income countries (LMICs), emphasizing equitable access and cultural adaptability.

Challenges in integration include data interoperability, model interpretability, and scalability in resource-constrained settings. Regulatory imperatives call for validation frameworks and safety standards to govern the rollout of generative AI. This review provides an original systems-level framing, structuring the discourse around data-to-decision pipelines, governance overlays, and evaluative metrics for sustainable adoption.

Ultimately, foundation models hold promise for closed-loop healthcare systems, where AI-driven insights inform interventions and feedback refines models iteratively. However, rigorous oversight is essential to balance innovation with accountability, ensuring these technologies augment rather than disrupt clinical workflows. By synthesizing high-impact publications, this narrative review offers integrative insights for researchers, clinicians, and policymakers navigating AI-enabled healthcare transformation.

Keywords Clinical analytics, Healthcare AI, Foundation models, Architectural integration, Regulatory oversight, Ethical considerations

*Correspondence:
Victor Santos
victor.santos@gmail.com
[1] Department of Healthcare Data Science, Faculty of Medicine, University of Sao Paulo, Sao Paulo, Brazil
[2] Department of AI in Health Systems, Faculty of Engineering, University of Campinas, Campinas, Campinas, Brazil

Introduction

The advent of foundation models in artificial intelligence (AI) marks a pivotal shift in how computational intelligence is applied to healthcare systems and analytics. These models, pretrained on vast, heterogeneous datasets, exhibit emergent capabilities that allow them to generalize across diverse tasks without extensive fine-tuning [1]. In healthcare, this translates to enhanced processing of complex, multimodal data—ranging from textual electronic health records (EHRs) to imaging and genomic sequences—facilitating more integrated and responsive systems. Unlike traditional narrow AI, foundation models enable a holistic approach to healthcare infrastructure, where analytics inform real-time decision-making and operational efficiencies [2].

Evolution of AI in healthcare infrastructure

The integration of AI into healthcare systems has evolved from rule-based expert systems to machine-learning-driven predictive tools, and now to foundation models that underpin intelligent infrastructure. Early AI applications focused on isolated tasks, such as diagnostic classification from imaging data, but lacked interoperability with broader hospital systems. Foundation models address this by serving as backbone architectures that can be adapted for multiple downstream applications, such as EHR summarization, patient triage, and resource optimization [3]. This evolution is driven by the need for scalable analytics in an era of data proliferation, in which healthcare generates petabytes of data annually.

Architectural integration involves embedding these models within existing healthcare information technology (IT) ecosystems, including picture archiving and communication systems (PACS), laboratory information management systems (LIMS), and EHR platforms. Key considerations include data harmonization to mitigate silos, ensuring that foundation models can ingest and analyze disparate data sources cohesively [4]. Oversight mechanisms, such as audit trails and explainability modules, are integral to maintaining trust and compliance with standards such as the Health Insurance Portability and Accountability Act (HIPAA) [5].

Role of analytics in modern healthcare systems

Analytics powered by foundation models extend beyond descriptive statistics to prescriptive and predictive intelligence. In clinical settings, these models support anomaly detection in vital signs monitoring, forecasting disease outbreaks, and personalizing treatment protocols [6]. For instance, generative capabilities allow simulation of patient trajectories, aiding in scenario planning for resource allocation [7]. However, integration requires careful oversight to address biases inherent in training data, which may perpetuate disparities in underrepresented populations [8].

The analytics landscape is further enriched by multimodal fusion, where foundation models process text, images, and time-series data simultaneously. This enables comprehensive patient profiles, enhancing diagnostic accuracy and therapeutic recommendations [9]. Regulatory oversight is crucial here, as models must undergo rigorous validation to ensure reliability in high-stakes environments [10].

Challenges in architectural integration

While foundation models promise seamless integration, challenges persist in aligning them with legacy healthcare systems. Interoperability standards like Fast Healthcare Interoperability Resources (FHIR) facilitate this, but require custom adapters for model deployment [11]. Oversight considerations include ethical frameworks for data usage, particularly in generative tasks that could produce misleading outputs [12]. Moreover, scalability in diverse settings—from urban hospitals to rural clinics—demands adaptive architectures that prioritize efficiency [13].

Oversight frameworks for safe deployment

Effective oversight encompasses regulatory, ethical, and technical dimensions. Regulatory bodies advocate for pre-market evaluations and post-deployment monitoring to safeguard patient outcomes [14]. Ethical considerations focus on equity, ensuring that foundation models do not exacerbate health disparities in LMICs [15]. Technical oversight involves robustness testing against adversarial inputs and continuous performance auditing [16].

Human-AI collaboration paradigms

A core aspect of integration is fostering human-AI symbiosis, where clinicians retain oversight while leveraging model insights. This includes user interfaces for query-based analytics and feedback mechanisms to refine model behaviors [17]. Studies highlight the need for training programs to build clinician proficiency in interpreting AI outputs [18].

Systems-level perspectives

At a systems level, foundation models enable closed-loop architectures, where analytics drive interventions and outcomes feed back into model refinement [19-21]. This cyclical approach enhances resilience in healthcare delivery [22].

Review positioning statement

This narrative review positions foundation models as enablers of intelligent healthcare systems, synthesizing literature through an original lens of architectural integration and oversight. The scope encompasses data-driven analytics, decision architectures, and governance, as well as publications to provide integrative cross-study analysis. By structuring the discourse around end-to-end pipelines—from data ingestion to intervention feedback—this review offers new interpretive frameworks for AI deployment in healthcare, emphasizing sustainable and equitable adoption without empirical benchmarks or speculative futurism.

Landscape of AI in healthcare systems & analytics

The landscape of AI in healthcare systems and analytics is rapidly evolving, with foundation models at the forefront of enabling sophisticated, integrated solutions. These models, leveraging vast pretraining, facilitate analytics that span from individual patient care to population-level insights, embedded within healthcare infrastructures [1].

Data foundations and multimodal integration

Healthcare analytics begins with robust data foundations, where foundation models excel in handling multimodal inputs. EHRs, imaging, and wearable sensor data are ingested through standardized pipelines, allowing models to extract latent patterns [2]. Integration architectures often employ federated approaches to preserve privacy, aggregating insights across institutions without centralizing sensitive data [3]. Analytics applications include risk stratification, in which models predict adverse events using integrated datasets [4].

Oversight in data handling emphasizes compliance with global standards, ensuring data quality and representativeness [5]. Generative foundation models augment datasets by creating synthetic samples, addressing imbalances in rare disease cohorts [6].

Predictive analytics and risk modeling

Predictive analytics form a cornerstone of AI-driven healthcare systems, with foundation models providing generalist capabilities for forecasting across heterogeneous data sources. In hospital settings, these models analyze longitudinal time-series data—including vital signs, laboratory values, medication histories, and clinician notes—for early sepsis prediction, clinical deterioration alerts, and readmission risk stratification, typically integrated into real-time clinical dashboards and electronic health record (EHR) systems [7]. Beyond static risk scoring, contemporary foundation models leverage multimodal learning architectures that fuse structured and unstructured data streams, enabling dynamic updating of patient risk profiles as new information becomes available.

Architectural considerations include real-time inference engines that interface with bedside monitoring systems, laboratory information systems, and EHR databases through secure interoperability standards such as HL7 and FHIR [8]. Edge-computing strategies may be employed in high-acuity environments to reduce latency, while cloud-based infrastructures support large-scale population modeling. These systems must balance computational efficiency with explainability, often incorporating attention mechanisms or post-hoc interpretability layers to provide clinicians with an actionable rationale behind predictions.

Oversight frameworks mandate rigorous validation against diverse populations to mitigate algorithmic bias, particularly in historically underrepresented groups. Consensus statements advocate transparent reporting of training data characteristics, performance metrics disaggregated by demographic subgroups, and post-deployment monitoring mechanisms [9]. Continuous recalibration protocols are increasingly recommended to prevent performance drift over time.

Analytics also extend to pharmacovigilance, where foundation models mine biomedical literature, spontaneous adverse event reports, and real-world evidence datasets to detect emerging safety signals [10]. Natural language processing pipelines extract adverse drug reaction patterns from clinical narratives and patient forums, while graph-based modeling identifies previously unrecognized associations between drugs and outcomes.

Generative AI in clinical synthesis

Generative capabilities embedded within foundation models enable the synthesis of clinical narratives and structured summaries. Applications include automated radiology report drafting from imaging data, discharge summary generation from EHR records, and structured pathology report standardization [11]. These tools reduce clinician documentation burden while promoting consistency and completeness.

In research analytics, generative models facilitate hypothesis generation by synthesizing findings from vast biomedical corpora and identifying latent connections among molecular pathways, clinical phenotypes, and therapeutic responses [12]. Such capabilities accelerate translational research by proposing testable insights derived from large-scale pattern recognition.

Architectural integration typically involves application programming interfaces (APIs) that embed generative functions directly into clinical workflows. Secure sandboxing environments and human-in-the-loop review systems are implemented to ensure that generated outputs are verified before clinical use. Oversight mechanisms emphasize hallucination detection, provenance tracking, and output validation against trusted knowledge bases [13].

Applications in precision oncology exemplify this paradigm. Foundation models integrate genomic sequencing data, tumor microenvironment profiles, and longitudinal clinical records to generate personalized treatment simulations and predictions of therapy response [14]. By synthesizing multi-omic and clinical datasets, these models support multidisciplinary tumor boards in exploring individualized care pathways.

Population health and resource optimization

At the systems level, AI-driven analytics optimize healthcare resource allocation by forecasting service demand, including intensive care unit (ICU) bed utilization, emergency department crowding, surgical scheduling, and vaccination uptake

[15]. Foundation models incorporate epidemiological data, seasonal trends, and mobility patterns to anticipate surges and guide contingency planning.

Importantly, these systems increasingly integrate social determinants of health (SDOH) data—such as housing stability, income, education, and environmental exposures—to enhance equity-focused analytics [16]. By modeling structural determinants alongside clinical risk factors, foundation models can identify vulnerable populations and inform targeted public health interventions.

Oversight frameworks for deployment, particularly in low- and middle-income countries (LMICs), include ethical checklists emphasizing cultural relevance, infrastructural feasibility, and equitable benefit distribution [17]. Local stakeholder engagement and contextual adaptation are critical to avoid technological imposition or resource misallocation.

Regulatory and ethical landscapes

The regulatory landscape surrounding foundation models in healthcare is rapidly evolving. Adaptive oversight models emphasize iterative validation, real-world evidence collection, and lifecycle risk management to support safe and timely AI deployment [18]. Regulatory authorities increasingly require documentation of model development processes, data governance practices, cybersecurity safeguards, and post-market surveillance strategies.

Ethical considerations center on human–AI collaboration, emphasizing that predictive and generative systems should augment—not replace—clinical judgment [19]. Key principles include accountability, transparency, informed consent where applicable, and preservation of patient autonomy. Multidisciplinary governance boards are often recommended to oversee implementation and address emergent ethical challenges.

Adoption trends and implementation pathways

Adoption of foundation models within healthcare systems is accelerating, particularly in tools integrated directly into EHR platforms and clinical decision support systems [20]. Early adoption patterns suggest the highest uptake in radiology, pathology, and administrative workflow optimization, where measurable efficiency gains are most evident.

Implementation pathways commonly involve phased pilot testing, performance benchmarking against standard care, and structured stakeholder engagement that includes clinicians, administrators, IT specialists, and patients [21]. Change management strategies—such as training programs and transparent communication—are critical for fostering trust and effective utilization.

In mental health analytics, foundation models are emerging as advisory tools capable of risk assessment, symptom trend monitoring, and therapeutic content generation [22]. While promising, these applications require heightened oversight given the sensitivity of mental health data and the potential consequences of inaccurate outputs.

Radiology and imaging analytics

In radiology, foundation models significantly enhance analytics through advanced image synthesis, segmentation, anomaly detection, and interpretation [23]. Multimodal models integrate imaging findings with clinical context, enabling more comprehensive diagnostic support.

Architectural integration with Picture Archiving and Communication Systems (PACS) allows automated triage of urgent cases, flagging suspected intracranial hemorrhages, pulmonary embolisms, or malignancies for prioritized review [24]. Continuous quality monitoring systems assess diagnostic reliability and ensure that AI outputs align with established radiological standards.

Specialized vs. generalist models

Comparative analyses between specialized domain-specific models and generalist foundation models highlight trade-offs in healthcare analytics. Specialized models often demonstrate superior performance within narrow clinical tasks due to targeted training datasets, whereas generalist models offer broader adaptability and cross-domain reasoning [25]. Hybrid architectures—combining domain fine-tuning with generalist backbone models—are increasingly favored for complex clinical applications.

Oversight strategies prioritize real-world validation, prospective trials, and continuous performance auditing to ensure that both specialized and generalist systems maintain safety and clinical relevance outside controlled research environments [26-28]. The major architectural layers, associated analytics functions, and oversight requirements are summarized in **Table 1.**

Table 1. Architectural layers of foundation model integration in healthcare systems and associated oversight mechanisms

Architectural layer	Core technical components	Primary analytics functions	Oversight requirements
Data ingestion and harmonization	EHR pipelines, PACS, LIMS, FHIR/HL7 adapters, and federated nodes	Multimodal data fusion, normalization, and feature extraction	Data quality auditing, HIPAA/GDPR compliance, and representativeness validation
Foundation model intelligence layer	Transformer-based multimodal models, LLM adaptations, and embedding engines	Predictive analytics, generative synthesis, and risk stratification	Bias testing, explainability modules, and performance benchmarking
Decision fusion interface	Clinical dashboards, conversational interfaces, and alert systems	Differential diagnosis support, triage prioritization, and treatment simulation	Human-in-the-loop protocols, override mechanisms, and traceability logs
Intervention execution layer	EHR order sets, clinical workflow triggers, and policy engines	Alert deployment, therapy adjustment, and operational optimization	Clinical validation, safety threshold enforcement
Feedback and recalibration	Outcome registries, monitoring dashboards, and learning loops	Drift detection, recalibration, and adaptive learning	Continuous validation, audit trails, and fairness monitoring
Governance overlay	Regulatory frameworks, ethics boards, and compliance systems	Lifecycle monitoring and deployment review	Pre-market validation, post-market surveillance, and ethical compliance

Intelligent clinical decision and closed-loop healthcare systems

Intelligent clinical decision systems powered by foundation models represent an advanced architectural paradigm in healthcare, enabling dynamic, evidence-based support. These systems integrate analytics into workflows, facilitating decisions from diagnosis to treatment [1].

Architectural components

Core architectures include data ingestion layers, where foundation models process inputs in real-time; inference engines for probabilistic outputs; and decision fusion modules that combine AI insights with clinician input [2]. Oversight layers incorporate audit mechanisms for traceability [3].

In closed-loop systems, feedback from outcomes recalibrates models, creating adaptive cycles [4].

Decision support workflows

Decision support leverages foundation models for query resolution, such as differential diagnosis from symptoms and tests [5]. Architectures often feature natural language interfaces, allowing clinicians to interact conversationally [6]. Analytics underpin recommendations, drawing on synthesized evidence [7].

Oversight ensures ethical alignment, with protocols for overriding AI suggestions [8].

Closed-loop dynamics

Closed-loop systems formalize cycles: data collection → model inference → intervention → outcome monitoring → model update. This enhances precision, as in chronic disease management [9].

A conceptual formula for human-AI decision fusion can be expressed as:

$$D = \alpha * AI_p + (1 - \alpha) * H_e + \beta * F_b \quad (1)$$

Where D is the final decision, AI_p is an AI-predicted probability, H_e is human expertise adjustment, F_b is the feedback bias correction term, and α, β are weighting factors tuned for context [10]. This interpretive structure highlights integration without empirical metrics.

Applications in clinical settings

In oncology, closed-loop systems integrate analytics to adapt treatment based on response data [11]. Radiology benefits from generative models for report summarization, thereby closing the loop with clinician verification [12].

Mental health applications use models for advisory loops, monitoring patient interactions [13].

Oversight frameworks emphasize continuous validation to address model performance drift [14]. Ethical considerations include equity in loop designs for diverse populations [15].

Implementation challenges

Architectural integration requires a robust IT infrastructure with pathways that prioritize scalability [16]. Adoption studies reveal benefits in efficiency, tempered by the need for training [17]. The architectural and governance relationships within this end-to-end pipeline are illustrated in **Figure 1.**

https://doi.org//g765182791

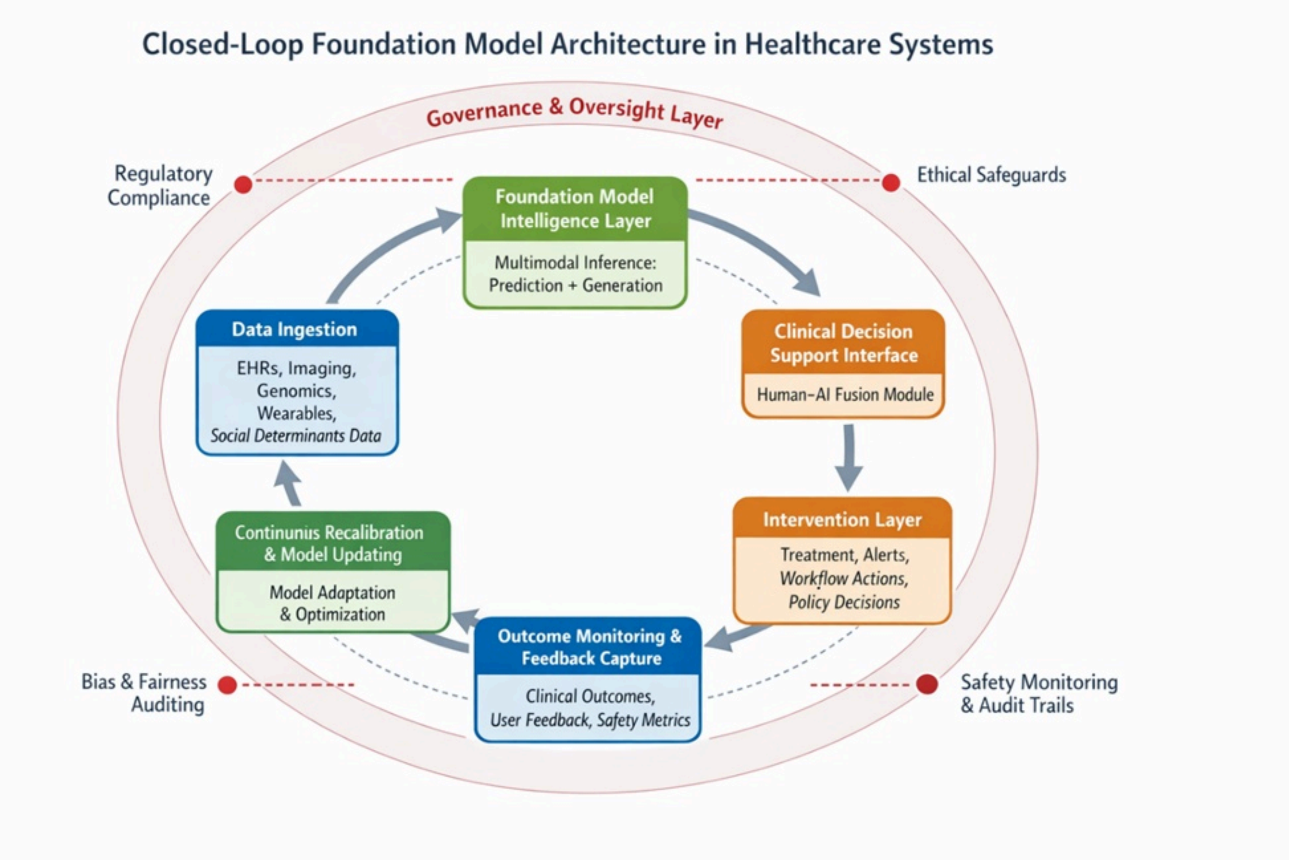

Figure 1. Closed-loop architectural integration of foundation models in healthcare systems.

Schematic representation of an end-to-end healthcare AI pipeline. Multimodal data are ingested into foundation models that generate predictive and generative intelligence. Outputs are integrated into human–AI decision support interfaces, informing clinical or operational interventions. Outcomes are continuously monitored and fed back into model recalibration processes. A governance overlay—including regulatory, ethical, bias, and safety checkpoints—envelops the cycle to ensure responsible deployment and adaptive oversight.

Results and Discussion

The integration of foundation models into healthcare systems and analytics, as synthesized from the literature, reveals a multifaceted discourse on their potential to revolutionize clinical workflows and decision-making. This discussion contextualizes the architectural and oversight implications, drawing integrative insights across studies to highlight synergies and tensions in real-world deployment [1, 2]. Foundation models, by virtue of their generalist nature, facilitate a shift from siloed AI applications to cohesive systems-level intelligence, where analytics inform every stage of patient care—from initial data capture to long-term outcome tracking [3].

A key theme emerging from the synthesis is the architectural flexibility of foundation models, which allows for modular integration into existing healthcare infrastructures. For example, their ability to handle multimodal data enables analytics pipelines that fuse EHR narratives with imaging interpretations, yielding comprehensive clinical insights [4]. This is particularly evident in closed-loop systems, where iterative feedback improves model accuracy over time, as discussed in the context of predictive analytics for chronic conditions [5]. However, this flexibility must be balanced with oversight to prevent unintended consequences, such as over-reliance on generative outputs that may introduce errors in high-stakes decisions [6].

Cross-study analysis underscores the role of human-AI collaboration in optimizing system performance. The literature consistently emphasizes hybrid architectures in which foundation models provide probabilistic suggestions, augmented by clinician expertise, to mitigate risks associated with standalone AI [7]. In population health analytics, this collaboration extends to policy-level decisions, where models simulate intervention scenarios, but ethical oversight ensures alignment with equity principles [8]. Integrative framing reveals that successful integration hinges on adaptive governance that adapts to contextual variations across healthcare settings [9].

Oversight considerations dominate the discussion, with consensus on the need for regulatory frameworks that evolve alongside technological advancements. Studies advocate for validation protocols that assess not only efficacy but also fairness in diverse populations, addressing gaps in LMIC adoption [10, 11]. This synthesis proposes an original interpretive structure: a governance-augmented analytics continuum, where oversight is embedded as a dynamic layer rather than a post-hoc addition [12]. Such structuring prevents fragmentation, ensuring that architectural integration supports sustainable healthcare transformation.

Furthermore, the discussion highlights analytics-driven efficiencies in resource-constrained environments. Foundation models enable scalable solutions, such as automated triage in emergency departments, but require careful calibration to avoid exacerbating disparities [13]. Comparative analyses between generalist and specialized models suggest that hybrid approaches yield superior outcomes in clinical decision support, particularly in radiology, where generative capabilities aid in interpretive tasks [14, 15].

In synthesizing these elements, the review illuminates tensions between the imperatives of innovation speed and safety. Rapid rollout of generative AI, as noted in adoption trends, promises enhanced analytics but demands rigorous ethical checklists to safeguard patient trust [16, 17]. This integrative analysis reframes foundation models not merely as tools but as foundational components of intelligent healthcare ecosystems, where architectural integration and oversight coalesce to drive clinical excellence [18].

Challenges and limitations

Despite the transformative potential of foundation models in healthcare systems and analytics, several challenges and limitations persist, as evidenced by the literature. These encompass technical, ethical, regulatory, and practical hurdles that must be addressed for effective architectural integration and oversight [1].

Technical challenges in architectural integration

A primary challenge lies in ensuring interoperability with legacy healthcare systems. Foundation models require substantial computational resources, posing scalability issues in under-resourced facilities [2]. Data heterogeneity further complicates integration, as models trained on standardized datasets may falter in the face of real-world variations in EHR formats [3]. Limitations in model interpretability hinder clinical adoption, where “black-box” inferences undermine trust in analytics outputs [4].

In closed-loop systems, feedback mechanisms can introduce instability if not properly damped, leading to cascading errors in decision cycles [5]. A conceptual formula illustrating this limitation in recalibration dynamics is:

$$R_t = R_{\{t-1\}} + \gamma * (O_t - P_t) - \delta * E_{\{t-1\}} \quad (2)$$

Where R_t is the recalibrated model response at time t, O_t is observed outcome, P_t is the predicted outcome, γ is the learning rate, and δ is the decay factor for prior errors $E_{\{t-1\}}$, and the formula interprets the balance needed to avoid overfitting to noisy feedback without empirical tuning [6]. This structure highlights the infrastructural need for robust error correction in healthcare loops.

Ethical and equity limitations

Ethical challenges include bias amplification from pretrained datasets, which may perpetuate disparities in analytics for underrepresented groups [7]. In LMICs, limited data availability exacerbates this, limiting model generalizability [8]. Oversight frameworks often lack enforcement mechanisms, leading to uneven adoption and potential exploitation in vulnerable populations [9].

Generative capabilities present specific limitations, such as the risk of hallucination in clinical synthesis, where fabricated details could mislead decision-making [10]. Human-AI interaction limitations arise when models subtly influence clinician behavior, potentially deskilling practitioners over time [11].

Regulatory and safety hurdles

Regulatory challenges stem from the pace of AI evolution outstripping policy development. Current standards fail to address the unique risks of generative AI, such as unvalidated outputs in patient communications [12]. Safety limitations are pronounced in critical applications, where model failures could impact life-sustaining systems [13].

Adoption barriers include stakeholder resistance due to perceived threats to autonomy, compounded by insufficient validation studies [14]. In precision medicine analytics, limitations in handling rare events reduce reliability [15].

Practical implementation barriers

Practical limitations include high deployment and maintenance costs, which restrict access to advanced analytics [16]. Training gaps for healthcare professionals limit effective oversight, while data privacy concerns impede federated integrations [17].

Cross-study synthesis reveals that these challenges are interconnected and require holistic solutions. For instance, technical limitations amplify ethical ones, necessitating integrated oversight strategies [18]. This review's original framing positions limitations as opportunities for refinement, emphasizing systems-level mitigations over isolated fixes [19, 20]. Key challenges and corresponding systems-level mitigation strategies identified in the literature are synthesized in **Table 2**.

Table 2. Systems-level challenges in foundation model deployment and proposed mitigation strategies

Challenge domain	Specific limitation	Systems-level risk	Proposed mitigation strategy
Technical integration	Interoperability with legacy systems	Fragmented workflows, reduced efficiency	FHIR-based middleware, modular API architectures
Computational demands	High infrastructure costs	Limited scalability in LMICs	Edge-computing adaptation, federated learning models
Model interpretability	Black-box predictions	Reduced clinician trust	Explainability layers, counterfactual reasoning tools
Bias and equity	Underrepresentation in training datasets	Disparities in care recommendations	Subgroup validation, bias auditing frameworks
Generative hallucination	Fabricated clinical content	Patient safety risks	Human verification loops, provenance tracking

Regulatory gaps	Lagging policy adaptation	Unstandardized deployment	Adaptive regulatory sandboxes, lifecycle evaluation
Human factors	Deskilling or automation bias	Overreliance on AI	Structured training, calibrated decision fusion weights
Feedback instability	Overfitting in closed loops	Performance drift	Controlled learning rates, periodic revalidation

Future research directions

Future research on foundation models in healthcare systems and analytics should prioritize addressing identified gaps through targeted investigations, fostering innovation while reinforcing oversight [1].

Advancing architectural innovations

Research should explore novel architectures for seamless integration, such as edge-computing adaptations for real-time analytics in remote settings [2]. Investigating multimodal fusion techniques could enhance closed-loop efficiencies, with agendas focusing on hybrid models that blend generalist and domain-specific strengths [3].

Enhancing oversight mechanisms

Agendas must include developing dynamic regulatory frameworks that incorporate AI-specific safety standards for generative applications [4]. Research into explainability methods, like counterfactual analytics, would bolster trust in decision systems [5].

Ethical research directions emphasize equity-focused designs, with studies on bias detection in diverse datasets [6]. In LMICs, agendas should assess culturally adaptive integrations, evaluating impact on health outcomes [7].

Empirical validation and human factors

Future work should conduct longitudinal studies on adoption impacts, measuring long-term effects on clinical workflows [8]. Human-AI collaboration research could develop protocols for optimal fusion, including training simulations [9].

In analytics, agendas include exploring generative AI for hypothesis discovery in biomedical research, with safeguards against misinformation [10].

Systems-level evaluations

Research should adopt holistic evaluations that assess end-to-end system resilience under stress [11]. Agendas for closed-loop advancements include feedback optimization to prevent drift in continuous learning [12].

This research agenda synthesizes literature calls for interdisciplinary efforts, positioning foundation models as catalysts for equitable, intelligent healthcare futures [13-15].

Conclusion

Foundation models offer profound opportunities for architectural integration in healthcare systems and analytics, enabling intelligent, responsive infrastructures that enhance clinical decision-making and patient outcomes. However, their deployment demands vigilant oversight to navigate ethical, regulatory, and technical complexities. This review's original synthesis underscores the imperative for balanced innovation, where systems-level framing integrates data analytics with governance to ensure safety and equity. As healthcare evolves, prioritizing human-centered design will maximize the benefits of these technologies, fostering a future in which AI augments clinical excellence without compromise.

Acknowledgements

None

Conflict of interest

None

Financial support

None

Ethics statement

None

Received: 28 Jul 2025 Revised: 20 Aug 2025 Accepted: 27 Sep 2025
Published online: 20 January 2026

References

Moor M, Banerjee O, Abad ZSH, Krumholz HM, Leskovec J, Topol EJ, et al. Foundation models for generalist medical artificial intelligence. Nature. 2023;616(7956):259-65.
https://doi.org/10.1038/s41586-023-05881-4.

Meskó B, Topol EJ. The imperative for regulatory oversight of large language models (or generative AI) in healthcare. NPJ Digit Med. 2023;6(1):120.
https://doi.org/10.1038/s41746-023-00873-0.

Wornow M, Xu Y, Thapa R, Patel B, Steinberg E, Fleming S, et al. The shaky foundations of large language models and foundation models for electronic health records. NPJ Digit Med. 2023;6(1):135.
https://doi.org/10.1038/s41746-023-00879-8.

Raza MM, Venkatesh KP, Kvedar JC. Generative AI and large language models in health care: pathways to implementation. NPJ Digit Med. 2024;7(1):62.
https://doi.org/10.1038/s41746-023-00988-4.

Teo ZL, Thirunavukarasu AJ, Elangovan K, Cheng H, Moova P, Soetikno B, et al. Generative artificial intelligence in medicine. Nat Med. 2025;31(12):3270-82.

https://doi.org/10.1038/s41591-025-03983-2.

Kim SH, Liu Y, Wang Z, Cao L, Jin Q, Chan J, et al. A foundation model for human-AI collaboration in medical literature mining. Nat Commun. 2025;16(1):8361.
https://doi.org/10.1038/s41467-025-62058-5.

Freyer O, Krumholz HM. A future role for health applications of large language models depends on regulators enforcing safety standards. Lancet Digit Health. 2024;6(5):e364-e366.
https://doi.org/10.1016/S2589-7500(24)00124-9.

The Lancet Digital Health. Rapid generative AI rollout in health care. Lancet Digit Health. 2025;7(8):e509.
https://doi.org/10.1016/S2589-7500(25)00091-3.

de Hond A, van Smeden M, Steyerberg EW. From text to treatment: the crucial role of validation for generative large language models in health care. Lancet Digit Health. 2024;6(4):e236-e237.
https://doi.org/10.1016/S2589-7500(24)00111-0.

Ning Y, Teixayavong S, Shang Y, Savulescu J, Nagaraj V, Miao D, et al. Generative artificial intelligence and ethical considerations in health care: a scoping review and ethics checklist. Lancet Digit Health. 2024;6(11):e848-e856.
https://doi.org/10.1016/S2589-7500(24)00143-2.

Ong JCL, Ong AML, Chen SYH, Ong S, Shorey S, Liu B, et al. Ethical and regulatory challenges of large language models in medicine. Lancet Digit Health. 2024;6(6):e428-e432.
https://doi.org/10.1016/S2589-7500(24)00061-X.

Chen H, Wong TY, Tham YC, Li X, Cheng CY, Ting DS, et al. Large language models and global health equity: a roadmap for equitable adoption in LMICs. Lancet Reg Health West Pac. 2025;63:101707.
https://doi.org/10.1016/j.lanwpc.2025.101707.

Khosravi B, Gichoya JW, Erickson BJ. Exploring the potential of generative artificial intelligence in medical image synthesis: opportunities, challenges, and future directions. Lancet Digit Health. 2025;7(9):e635-e644.
https://doi.org/10.1016/S2589-7500(25)00072-X.

Chen S, Guevara M, Moningi S, Hoebers F, Elhalawani H, Kann BH, et al. The effect of using a large language model to respond to patient messages. Lancet Digit Health. 2024;6(6):e379-e388.
https://doi.org/10.1016/S2589-7500(24)00060-8.

Liang S, Wang C, Wang G, Li W, Zhang J, Liu X, et al. The potential of large language models to advance precision oncology. EBioMedicine. 2025;115:105695.
https://doi.org/10.1016/j.ebiom.2025.105695.

Murray SG, Mishuris RG. Generative AI in health care—is faster better? JAMA Netw Open. 2025;8(12):e2549470.
https://doi.org/10.1001/jamanetworkopen.2025.49470.

Everson J, Nong P, Richwine C. Uptake of generative AI integrated with electronic health records in US hospitals. JAMA Netw Open. 2025;8(12):e2549463.
https://doi.org/10.1001/jamanetworkopen.2025.49463.

Feldman MJ, Hoffer EP. Dedicated AI expert system vs generative AI with large language model for clinical diagnoses. JAMA Netw Open. 2025;8(5):e2512994.
https://doi.org/10.1001/jamanetworkopen.2025.12994.

Nguyen TD, Whaley CM, Simon K, Alexander J, Upadhyay S. Adoption of artificial intelligence in the health care sector. JAMA Health Forum. 2025;6(11):e255029.
https://doi.org/10.1001/jamahealthforum.2025.5029.

Huo B, Boyle A, Marfo N, McKechnie T, Lee Y, Mayol J, et al. Large language models for chatbot health advice studies: a systematic review. JAMA Netw Open. 2025;8(2):e2457879.
https://doi.org/10.1001/jamanetworkopen.2024.57879.

Bedi S, Liu Y, Orr-Ewing L, Dash D, Koyejo S, Callahan A, et al. Testing and evaluation of health care applications of large language models: a systematic review. JAMA. 2025;333(4):319-28.
https://doi.org/10.1001/jama.2024.21700.

Goh E, Gallo R, Hom J, Weng Y, Kerman N, Cool C, et al. Large language model influence on diagnostic reasoning: a randomized clinical trial. JAMA Netw Open. 2024;7(10):e2440969.
https://doi.org/10.1001/jamanetworkopen.2024.40969.

McBain RK, Bozick R, Diliberti M, Zhang L, Zhang F, Burnett A, et al. Use of generative AI for mental health advice among US adolescents and young adults. JAMA Netw Open. 2025;8(11):e2542281.
https://doi.org/10.1001/jamanetworkopen.2025.42281.

Wachter RM, Brynjolfsson E. Will generative artificial intelligence deliver on its promise in health care? JAMA. 2024;331(1):65-9.
https://doi.org/10.1001/jama.2023.25054.

Howell MD. Generative artificial intelligence, patient safety and healthcare quality: a review. BMJ Qual Saf. 2024;33(11):748-52.
https://doi.org/10.1136/bmjqs-2023-016690.

Paschali M, Chaudhari AS, Bluethgen C, Langlotz CP, Kahn CE Jr. Foundation models in radiology: what, how, why, and why not. Radiology. 2025;325(1):e240597.
https://doi.org/10.1148/radiol.240597.

Tavakoli N, Karargyris A, Schoenfeld E, Mermel G, Talati S, Lee S. Generative AI and foundation models in radiology: applications, opportunities, and potential challenges. Radiology. 2025;325(2):e242961.
https://doi.org/10.1148/radiol.242961.

Zheng S, Lin S, Wang C, Zhu Y, Li Y, Zhou J, et al. Comparison of a specialized large language model with GPT-4o for CT and MRI radiology report summarization. Radiology. 2025;325(3):e243774.
https://doi.org/10.1148/radiol.243774.