European International Journal of Multidisciplinary Research and Management Studies (ISSN: 2750-8587)

Received: 18 January 2026 | Accepted: 12 February 2026 | Published: 28 February 2026

RESEARCH ARTICLE

Cloud-Based Data Governance and Predictive Analytics for Secure E-Healthcare Systems: Integrating Data Lineage, Multi-Cloud Architectures, And Artificial Intelligence for Digital Health Transformation

Amelia Robertson
Department of Information Systems and Digital Health, University of Edinburgh, United Kingdom

VOLUME: Vol.06 Issue02 2026

PAGE: 134-142

Abstract

The rapid digitization of healthcare systems has generated vast volumes of clinical, administrative, and operational data, necessitating advanced technological infrastructures capable of storing, processing, and analyzing such information securely and efficiently. Cloud computing has emerged as a transformative paradigm capable of supporting scalable health information systems, enabling interoperability, and facilitating advanced data analytics. Simultaneously, the integration of artificial intelligence and predictive analytics into healthcare environments has created opportunities for enhanced clinical decision-making, operational efficiency, and population health management. However, the convergence of cloud computing, predictive analytics, and sensitive healthcare data raises critical concerns related to data governance, security, compliance, and transparency. In particular, the concept of data lineage-the ability to track the origin, movement, and transformation of data across systems-has become increasingly significant for ensuring accountability, regulatory compliance, and trustworthy analytics in digital healthcare ecosystems.

This research develops a comprehensive theoretical framework examining the intersection of cloud computing infrastructures, predictive analytics, and data governance mechanisms in modern e-healthcare systems. Drawing upon interdisciplinary literature across cloud computing, healthcare informatics, financial data analytics, and information governance, the study explores how cloud-based architectures can support scalable healthcare information systems while maintaining secure and traceable data management processes. The research synthesizes existing literature on cloud storage mechanisms, multi-cloud data architectures, healthcare information systems, predictive analytics, and data lineage frameworks to construct a conceptual model for secure and intelligent healthcare data management.

Using a systematic literature synthesis methodology informed by PRISMA principles and information systems review methodologies, the study analyzes theoretical and empirical findings from cloud computing, health informatics, and data governance research. The findings demonstrate that integrated cloud architectures combined with robust data lineage frameworks can significantly enhance transparency, reliability, and regulatory compliance in healthcare analytics environments. Moreover, the integration of predictive analytics capabilities within cloud-based healthcare systems enables proactive clinical insights, improved risk management, and data-driven healthcare decision-making.

The study contributes to the growing body of knowledge on digital healthcare transformation by proposing a holistic approach to cloud-enabled healthcare analytics governance. The research highlights practical implications for healthcare organizations, policymakers, and technology developers seeking to implement secure, scalable, and intelligent digital health infrastructures. Future research directions include empirical validation of the proposed framework, evaluation of real-world healthcare cloud deployments, and further exploration of governance mechanisms for AI-driven healthcare analytics.

KEYWORDS

Cloud computing, e-healthcare systems, predictive analytics, data lineage, healthcare information systems, digital health transformation, data governance.

INTRODUCTION

The global healthcare sector has undergone a profound digital transformation over the past two decades, driven by the proliferation of electronic health records, digital medical imaging, telemedicine platforms, and health data analytics systems. These technological developments have fundamentally altered how healthcare organizations collect, store, analyze, and utilize information to support patient care, administrative operations, and strategic decision-making. The increasing digitization of healthcare processes has resulted in an unprecedented expansion of health-related data, encompassing clinical records, biomedical signals, genomic information, administrative records, and population health datasets (Agarwal et al., 2010).

The rapid growth of healthcare data presents both opportunities and challenges for healthcare institutions and policymakers. On one hand, the availability of large-scale health data provides opportunities for advanced analytics, predictive modeling, and evidence-based medical decision-making. On the other hand, the management of such vast and heterogeneous data sources requires sophisticated technological infrastructures capable of ensuring data availability, security, interoperability, and scalability (Payne et al., 2013).

Cloud computing has emerged as a transformative technology that addresses many of these challenges by offering flexible, scalable, and cost-effective computing resources delivered through network-based services. In the context of healthcare systems, cloud computing enables organizations to store large volumes of health data, support distributed healthcare applications, and facilitate real-time collaboration among medical professionals across geographic boundaries (Kuo, 2011). By shifting computing resources from localized infrastructures to centralized or distributed cloud environments, healthcare organizations can reduce operational costs while improving system performance and accessibility (Pechette, 2012).

The adoption of cloud computing in healthcare has accelerated in recent years due to the growing demand for integrated healthcare information systems capable of supporting complex medical workflows and data-intensive applications. Cloud-based healthcare systems enable interoperability between different healthcare providers, facilitate the sharing of medical information across institutions, and support advanced analytics tools that can generate actionable insights from clinical data (Griebel et al., 2015). Additionally, cloud infrastructures provide the computational power required for emerging technologies such as artificial intelligence and predictive analytics, which are increasingly being used to analyze large healthcare datasets and improve clinical decision-making (Challapalli, 2023).

Despite these advantages, the integration of cloud computing into healthcare environments raises significant concerns related to data security, privacy protection, and regulatory compliance. Healthcare data is among the most sensitive forms of personal information, and its unauthorized disclosure or misuse can have severe consequences for individuals and organizations. Consequently, healthcare institutions must ensure that cloud-based systems comply with strict regulatory frameworks governing the handling and protection of patient information (Tahir et al., 2020).

In addition to security concerns, the increasing reliance on data-driven decision-making in healthcare highlights the importance of transparency and traceability in data management processes. As healthcare organizations adopt advanced analytics and artificial intelligence technologies, it becomes essential to understand how data flows through complex information systems, how it is transformed during processing, and how analytical outputs are generated. This requirement has led to the growing importance of data lineage, which refers to the ability to track the origins, transformations, and movement of data throughout its lifecycle within information systems (Peuralinna, 2024).

Data lineage plays a critical role in ensuring the reliability and accountability of analytics processes in data-intensive environments. By providing detailed insights into how data is collected, processed, and utilized, data lineage frameworks enable organizations to verify the accuracy of analytical results, identify potential data quality issues, and ensure compliance with regulatory standards. In healthcare systems, where analytical decisions may directly influence patient outcomes, the ability to trace data provenance and transformations is particularly important.

At the same time, predictive analytics has become a central component of modern healthcare data analysis. Predictive analytics involves the use of statistical and machine learning techniques to identify patterns in historical data and generate forecasts or predictions about future events. In healthcare contexts, predictive analytics can be used to identify disease risks, predict hospital readmissions, optimize treatment plans, and improve healthcare resource allocation (Olaniyi et al., 2023). The integration of predictive analytics into cloud-based healthcare systems enables organizations to leverage large-scale data resources while utilizing powerful computational capabilities to generate real-time insights.

However, the deployment of predictive analytics within cloud-based healthcare environments introduces additional challenges related to data governance and transparency. The complexity of machine learning algorithms and the distributed nature of cloud infrastructures can make it difficult to trace how analytical results are generated and how data transformations influence predictive outcomes. This lack of transparency can undermine trust in data-driven decision-making processes and create obstacles for regulatory compliance.

The concept of data governance has therefore become increasingly important in the context of cloud-based healthcare analytics. Data governance refers to the policies, processes, and technologies used to ensure the quality, security, and proper use of organizational data assets. Effective data governance frameworks provide mechanisms for managing data access, maintaining data quality, and ensuring that data is used in accordance with regulatory and ethical guidelines (Rao, 2022).

Within healthcare environments, data governance must address a range of complex challenges, including patient privacy protection, interoperability between heterogeneous systems, and the integration of diverse data sources. The increasing use of cloud computing and predictive analytics further complicates these challenges by introducing new layers of technological complexity and potential security risks.

The convergence of cloud computing, predictive analytics, and data governance has therefore emerged as a critical area of research within the fields of health informatics and information systems. While numerous studies have examined individual aspects of these technologies, there remains a significant gap in the literature regarding the integration of cloud-based infrastructures, predictive analytics frameworks, and data lineage mechanisms within comprehensive healthcare data management systems.

Existing research on cloud computing in healthcare has primarily focused on evaluating the potential benefits and challenges of cloud adoption, including cost efficiency, scalability, and security concerns (Kuo et al., 2011). Other studies have examined the role of predictive analytics in improving healthcare decision-making and operational efficiency (Mullangi, 2017). Meanwhile, research on data lineage has largely focused on financial and data governance contexts, emphasizing the importance of traceability and transparency in complex data environments (Peuralinna, 2024).

Despite the growing relevance of these topics, relatively few studies have explored how cloud computing, predictive analytics, and data lineage can be integrated to support secure and transparent healthcare data ecosystems. The absence of such integrated frameworks represents a significant limitation for healthcare organizations seeking to implement advanced data analytics capabilities while maintaining robust governance and compliance mechanisms.

This research addresses this gap by developing a comprehensive conceptual framework for cloud-based healthcare analytics systems that incorporates predictive analytics capabilities and data lineage governance mechanisms. By synthesizing interdisciplinary research from healthcare informatics, cloud computing, data governance, and predictive analytics, the study seeks to provide a holistic perspective on how digital health infrastructures can be designed to support secure, scalable, and transparent data management.

The central objective of this research is to examine how cloud computing infrastructures can be integrated with predictive analytics and data lineage frameworks to support secure and trustworthy healthcare data ecosystems. Specifically, the study aims to achieve three primary goals. First, it seeks to analyze the technological foundations of cloud-based healthcare systems and their role in supporting digital health transformation. Second, it explores the integration of predictive analytics within cloud-enabled healthcare environments and its implications for healthcare decision-making. Third, it investigates the role of data lineage and governance mechanisms in ensuring transparency, reliability, and regulatory compliance in healthcare analytics systems.

By addressing these objectives, the research contributes to the growing body of knowledge on digital healthcare transformation and provides practical insights for healthcare organizations seeking to implement advanced data-driven technologies. The findings are expected to inform policymakers, healthcare administrators, and technology developers about the critical factors that influence the successful adoption of cloud-based analytics systems in healthcare environments.

The remainder of this study is structured around a comprehensive analysis of cloud computing infrastructures, predictive analytics applications, and data governance frameworks in healthcare contexts. The research methodology involves a systematic synthesis of interdisciplinary literature, followed by a conceptual analysis of how these technologies interact within integrated digital health ecosystems. Through this approach, the study aims to provide a detailed understanding of the opportunities and challenges associated with the adoption of cloud-based predictive analytics in healthcare systems.

METHODOLOGY

The methodological framework of this research is grounded in a systematic literature synthesis approach designed to integrate insights from multiple academic domains including health informatics, cloud computing, predictive analytics, data governance, and information systems research. The complexity of modern digital healthcare ecosystems requires an interdisciplinary methodological strategy capable of capturing the interactions among technological infrastructures, analytical techniques, and governance frameworks. Consequently, the research design combines systematic review principles, conceptual synthesis, and analytical interpretation to construct a comprehensive theoretical model explaining the integration of cloud computing and predictive analytics in secure healthcare information systems.

The study follows structured literature review principles informed by the PRISMA methodology, which provides standardized guidelines for identifying, selecting, and synthesizing relevant academic research. The PRISMA framework emphasizes transparency, reproducibility, and methodological rigor in literature reviews by requiring explicit documentation of literature search strategies, inclusion criteria, and analytical procedures (Moher et al., 2009). In the context of information systems research, systematic literature reviews play a crucial role in consolidating fragmented knowledge across disciplines and generating conceptual frameworks that guide future empirical investigations (Levy and Ellis, 2006).

The first stage of the research methodology involved the identification and collection of relevant scholarly literature addressing cloud computing in healthcare, predictive analytics, data lineage, healthcare information systems, and digital transformation in medical services. The references provided for this study served as the foundational dataset for the literature synthesis process. These references span multiple academic fields, including medical informatics, cloud computing technologies, financial analytics, and data governance, reflecting the interdisciplinary nature of the research problem.

The literature selection process focused on studies that address three core thematic areas relevant to the research objectives. The first thematic domain concerns cloud computing architectures and storage mechanisms within healthcare environments. This includes research exploring the technological infrastructure required to support cloud-based healthcare systems, the advantages and limitations of cloud deployment models, and the implications of cloud computing for healthcare data management. Studies examining virtualization technologies, multi-cloud architectures, and cloud-based healthcare information systems are particularly relevant in this domain (Tahir et al., 2020; Sadoughi and Erfannia, 2017).

The second thematic domain relates to predictive analytics and artificial intelligence applications within healthcare and related data-intensive sectors. Predictive analytics techniques enable organizations to extract meaningful insights from large datasets by identifying patterns, correlations, and trends that can inform decision-making processes. Although predictive analytics has been widely applied in financial markets and other industries, its integration into healthcare systems has gained increasing attention in recent years as healthcare organizations seek to leverage data-driven insights to improve patient outcomes and operational efficiency (Challapalli, 2023; Olaniyi et al., 2023).

The third thematic domain focuses on data governance, data lineage, and information management frameworks that support transparency and accountability in data-driven systems. As organizations increasingly rely on complex data pipelines and distributed computing infrastructures, understanding how data moves across systems and how it is transformed during processing becomes essential for ensuring data quality, regulatory compliance, and analytical reliability. Research on data lineage provides valuable insights into methods for tracking data provenance and maintaining transparency in large-scale information systems (Peuralinna, 2024; Rao, 2022).

After identifying the relevant literature within these thematic domains, the second stage of the methodology involved a detailed qualitative analysis of the selected studies. Rather than relying on quantitative meta-analysis techniques, the research adopts a narrative synthesis approach that emphasizes conceptual integration and theoretical interpretation. Narrative synthesis is particularly appropriate for interdisciplinary research contexts where studies employ diverse methodologies, theoretical frameworks, and empirical datasets. This approach enables researchers to identify common themes, theoretical relationships, and conceptual patterns across a wide range of scholarly sources.

The qualitative analysis involved examining the key findings, theoretical arguments, and methodological approaches presented in each study. Particular attention was given to identifying recurring concepts and analytical frameworks that could contribute to the development of an integrated model for cloud-based healthcare analytics governance. For example, several studies emphasize the importance of scalable computing infrastructures for managing large healthcare datasets, while others highlight the role of predictive analytics in enhancing clinical decision-making and healthcare resource allocation.

In addition to identifying thematic patterns, the analysis also explored the technological and organizational challenges associated with the adoption of cloud computing and predictive analytics in healthcare systems. These challenges include concerns related to data security, privacy protection, regulatory compliance, and interoperability among heterogeneous healthcare information systems. By examining how different studies address these challenges, the research aims to develop a comprehensive understanding of the factors that influence the successful implementation of digital healthcare technologies.

A key component of the methodological framework involves the conceptual integration of findings from the three thematic domains into a unified theoretical model. This process requires synthesizing insights from cloud computing research, predictive analytics literature, and data governance studies to identify the interrelationships among these technological and organizational components. The resulting conceptual framework seeks to explain how cloud-based infrastructures can support advanced healthcare analytics while ensuring secure and transparent data management practices.

The conceptual integration process also incorporates insights from related fields such as financial data analytics and enterprise data governance. Although these fields operate in different industry contexts, they face similar challenges related to data management, analytical transparency, and regulatory compliance. For instance, research on data lineage in financial systems provides valuable lessons for healthcare organizations seeking to track the flow of data through complex analytics pipelines (Peuralinna, 2024).

To enhance the analytical rigor of the study, the research also incorporates a systems-oriented perspective on healthcare information technology. A systems approach emphasizes the interconnected nature of technological infrastructures, organizational processes, and governance mechanisms within complex digital ecosystems. By viewing cloud-based healthcare systems as integrated socio-technical systems, the research can examine how technological innovations interact with institutional structures and regulatory frameworks to shape the adoption and impact of digital health technologies.

The systems perspective is particularly relevant for understanding the role of cloud computing in healthcare environments. Cloud-based infrastructures do not operate in isolation but rather function as components within broader healthcare information ecosystems that include electronic health record systems, clinical decision support tools, telemedicine platforms, and health data analytics applications. Consequently, the successful implementation of cloud computing in healthcare requires careful coordination among technological architectures, organizational policies, and governance frameworks.

The final stage of the methodology involved constructing a theoretical model that integrates cloud computing infrastructures, predictive analytics capabilities, and data governance mechanisms within a comprehensive healthcare data management framework. This model serves as the foundation for the analytical discussions presented in the results and discussion sections of the research. By synthesizing insights from interdisciplinary literature, the model aims to provide a structured understanding of how digital technologies can be integrated to support secure, scalable, and intelligent healthcare systems.

RESULTS

The synthesis of interdisciplinary literature revealed several key findings regarding the integration of cloud computing infrastructures, predictive analytics capabilities, and data governance mechanisms within modern healthcare information systems. The analysis demonstrates that cloud computing serves as a foundational technological platform that enables the storage, processing, and analysis of large-scale healthcare datasets while supporting distributed healthcare applications and collaborative medical practices.

One of the most significant findings emerging from the literature concerns the scalability advantages offered by cloud computing in healthcare environments. Traditional healthcare information systems often rely on locally hosted data centers and on-premise computing infrastructures that require substantial capital investments and ongoing maintenance costs. Such systems frequently encounter limitations in storage capacity, computational performance, and system flexibility, particularly when dealing with rapidly expanding healthcare datasets (Kuo, 2011).

Cloud computing addresses these limitations by providing on-demand access to computing resources, including storage capacity, processing power, and networking infrastructure. Healthcare organizations can dynamically scale their computing resources according to fluctuating demand levels, thereby avoiding the need for costly hardware investments and improving system efficiency. This scalability is particularly valuable for healthcare analytics applications that require high computational performance for processing large volumes of medical data.

The literature also highlights the role of cloud-based healthcare information systems in facilitating interoperability among healthcare providers. Interoperability refers to the ability of different information systems and organizations to exchange and interpret shared data effectively. In many healthcare environments, interoperability challenges arise due to the existence of heterogeneous systems developed by different vendors and implemented across various institutional contexts (Griebel et al., 2015).

Cloud-based platforms provide a centralized infrastructure that enables healthcare providers to access shared data repositories and standardized data services. By supporting common data formats and communication protocols, cloud platforms facilitate the integration of electronic health records, laboratory information systems, medical imaging systems, and other healthcare applications. This integration enables healthcare professionals to access comprehensive patient information and coordinate care across multiple institutions.

Another important finding concerns the growing role of predictive analytics within cloud-enabled healthcare systems. Predictive analytics technologies leverage machine learning algorithms and statistical modeling techniques to identify patterns within large datasets and generate forecasts about future events or outcomes. In healthcare contexts, predictive analytics can support a wide range of applications, including disease risk prediction, patient outcome forecasting, and healthcare resource optimization (Olaniyi et al., 2023).

Cloud computing infrastructures provide the computational resources necessary to support predictive analytics applications, particularly those involving complex machine learning algorithms and large-scale data processing tasks. By hosting analytics platforms in cloud environments, healthcare organizations can access advanced analytical tools without the need to maintain specialized computing hardware locally. This accessibility enables smaller healthcare institutions to adopt predictive analytics capabilities that would otherwise be beyond their technical or financial capacity.

The results of the literature synthesis also emphasize the importance of data governance frameworks in ensuring the reliability and security of cloud-based healthcare analytics systems. Data governance encompasses the policies, procedures, and technologies used to manage data assets throughout their lifecycle, including data collection, storage, processing, and dissemination. Effective data governance frameworks help organizations maintain data quality, protect sensitive information, and ensure compliance with regulatory requirements (Rao, 2022).

In the context of cloud-based healthcare systems, data governance becomes particularly complex due to the distributed nature of cloud infrastructures and the involvement of multiple stakeholders in data management processes. Healthcare data may be stored across multiple cloud servers, processed by various analytical applications, and accessed by diverse user groups including clinicians, researchers, and administrators. This complexity creates challenges for maintaining transparency and accountability in data usage.

The concept of data lineage emerges as a critical component of data governance in cloud-based analytics environments. Data lineage refers to the ability to trace the origin, movement, and transformation of data as it flows through information systems. By documenting how data is generated, modified, and utilized within analytical processes, data lineage frameworks enable organizations to verify the integrity of analytical outputs and identify potential sources of error or bias (Peuralinna, 2024).

The analysis indicates that integrating data lineage mechanisms into cloud-based healthcare systems can significantly enhance the transparency and reliability of predictive analytics applications. When healthcare organizations deploy machine learning models to generate clinical insights, it is essential to understand which datasets were used for model training, how data was preprocessed, and how analytical outputs were generated. Data lineage frameworks provide this level of transparency by maintaining detailed records of data transformations and analytical workflows.

Another important finding relates to the role of multi-cloud architectures in enhancing the resilience and security of healthcare data infrastructures. Multi-cloud strategies involve distributing data and applications across multiple cloud service providers rather than relying on a single provider. This approach reduces the risk of service disruptions and enhances data availability by ensuring that healthcare systems remain operational even if one cloud provider experiences technical difficulties (Hussain, 2012).

Multi-cloud architectures also support advanced data security mechanisms, including distributed data storage and encryption techniques that protect sensitive healthcare information from unauthorized access. Secure data deduplication methods further enhance storage efficiency and data protection by ensuring that redundant data copies are managed securely within cloud environments (Rashid et al., 2013).

The integration of predictive analytics within cloud-based healthcare systems also enables new forms of data-driven healthcare decision-making. By analyzing historical clinical data, predictive models can identify patterns associated with disease progression, treatment effectiveness, and patient outcomes. These insights enable healthcare providers to implement preventive interventions, personalize treatment plans, and allocate healthcare resources more effectively.

Overall, the results of the literature synthesis demonstrate that the integration of cloud computing, predictive analytics, and data lineage frameworks has the potential to transform healthcare data management and analytics practices. However, achieving these benefits requires careful attention to data governance, security, and regulatory compliance considerations.

DISCUSSION

The findings of this research highlight the transformative potential of integrating cloud computing infrastructures, predictive analytics capabilities, and data governance frameworks within modern healthcare information systems. The convergence of these technological and organizational elements represents a fundamental shift in how healthcare data is managed, analyzed, and utilized to support clinical decision-making and healthcare service delivery.

One of the most significant implications of the study relates to the role of cloud computing as an enabling infrastructure for digital healthcare transformation. Cloud platforms provide the scalability, flexibility, and computational capacity required to support large-scale healthcare data management and analytics applications. As healthcare systems continue to generate increasing volumes of digital data, traditional on-premise computing infrastructures are likely to become increasingly inadequate for meeting the demands of modern healthcare analytics environments.

The adoption of cloud-based healthcare information systems therefore represents a strategic imperative for healthcare organizations seeking to leverage data-driven insights and improve operational efficiency. However, the transition to cloud-based infrastructures must be accompanied by robust data governance frameworks that ensure the secure and ethical use of healthcare data.

Another important implication concerns the integration of predictive analytics into healthcare decision-making processes. Predictive analytics technologies offer significant opportunities for improving patient outcomes, optimizing healthcare resource allocation, and enhancing the efficiency of healthcare service delivery. By identifying patterns in historical data and generating forecasts about future events, predictive models enable healthcare providers to adopt proactive approaches to patient care and population health management.

Nevertheless, the deployment of predictive analytics within healthcare systems also raises important ethical and governance challenges. Machine learning algorithms may inadvertently introduce biases into clinical decision-making processes if they are trained on datasets that reflect historical inequities or data quality issues. Ensuring transparency and accountability in predictive analytics processes is therefore essential for maintaining trust in data-driven healthcare systems.

Data lineage frameworks play a critical role in addressing these challenges by providing detailed documentation of data sources, transformations, and analytical workflows. By enabling organizations to trace how analytical results are generated, data lineage mechanisms enhance the transparency and interpretability of predictive models. This transparency is particularly important in healthcare contexts where analytical insights may directly influence patient treatment decisions.

Despite the significant benefits associated with cloud-based healthcare analytics systems, the research also identifies several limitations and challenges that must be addressed in future implementations. Data security remains one of the most pressing concerns associated with cloud computing in healthcare environments. Healthcare organizations must ensure that cloud service providers implement robust security measures capable of protecting sensitive patient information from unauthorized access or cyber threats.

Regulatory compliance represents another important challenge for cloud-based healthcare systems. Healthcare data is subject to strict regulatory frameworks that govern how patient information can be collected, stored, and shared. Ensuring compliance with these regulations requires careful coordination among healthcare organizations, cloud service providers, and regulatory authorities.

Future research should focus on empirically validating the conceptual framework proposed in this study by examining real-world implementations of cloud-based healthcare analytics systems. Such research could involve case studies of healthcare organizations that have successfully integrated cloud computing, predictive analytics, and data governance frameworks into their digital infrastructures.

CONCLUSION

The digital transformation of healthcare systems has created unprecedented opportunities for leveraging advanced technologies to improve patient care, operational efficiency, and healthcare system sustainability. Cloud computing, predictive analytics, and data governance frameworks represent three critical components of modern digital healthcare infrastructures. When integrated effectively, these technologies have the potential to enable secure, scalable, and intelligent healthcare data ecosystems capable of supporting data-driven decision-making.

This research has developed a comprehensive conceptual framework for integrating cloud computing infrastructures, predictive analytics capabilities, and data lineage governance mechanisms within healthcare information systems. By synthesizing interdisciplinary literature from healthcare informatics, cloud computing, and data governance research, the study highlights the technological and organizational factors that influence the successful adoption of digital healthcare analytics systems.

The findings demonstrate that cloud-based architectures provide the scalability and computational power necessary for advanced healthcare analytics, while predictive analytics technologies enable the generation of actionable insights from large healthcare datasets. At the same time, data governance and data lineage frameworks are essential for ensuring transparency, reliability, and regulatory compliance in cloud-based healthcare environments.

As healthcare systems continue to evolve in response to technological innovation and increasing data availability, the integration of cloud computing and predictive analytics will play an increasingly important role in shaping the future of healthcare delivery. By adopting robust data governance frameworks and ensuring transparency in analytical processes, healthcare organizations can harness the full potential of digital technologies while maintaining trust, security, and ethical responsibility in healthcare data management.

REFERENCES

1. Agarwal, R., Gao, G.D., DesRoches, C., & Jha, A.K. The digital transformation of healthcare: Current status and the road ahead. Information Systems Research.
2. Challapalli, S. R. From Data to Decisions: Artificial Intelligence in Financial Markets. International Journal of Recent Technology and Engineering.
3. Chow, F., Muftu, A., & Shorter, R. Virtualization and cloud computing in dentistry. Journal of the Massachusetts Dental Society.
4. Griebel, L., Prokosch, H.U., Köpcke, F., Toddenroth, D., Christoph, J., Leb, I., et al. A scoping review of cloud computing in healthcare. BMC Medical Informatics and Decision Making.
5. Hussain, S. An efficient and economical multi-cloud storage in cloud computing. IET.
6. Judijanto, L., Devi, E. K., & Yusuf, S. Trends and evolution of data-driven financial management. International Journal of Recent Technology and Engineering.
7. Kuo, A.M. Opportunities and challenges of cloud computing to improve health care services. Journal of Medical Internet Research.
8. Levy, Y., & Ellis, T.J. A systems approach to conduct an effective literature review supporting information systems research. Information Science.
9. Mayfield, C.A., Gigler, M.E., Snapper, L., Jose, J., Tynan, J., Scott, V.C., et al. Using cloud-based open-source technology to evaluate community interventions. Journal of the American Medical Informatics Association.
10. Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement.
11. Olaniyi, O., Shah, N. H., Abalaka, A., & Olaniyi, F. Harnessing predictive analytics for strategic foresight.
12. Payne, T.H., Bates, D.W., Berner, E.S., Bernstam, E.V., Covvey, H.D., Frisse, M.E., et al. Healthcare information technology and economics. Journal of the American Medical Informatics Association.
13. Peuralinna, J. Data lineage in the financial sector. Aalto University.
14. Rao, K.V. Review of Data Lineage: Challenges, Tools, Techniques and Approaches. International Journal of Research and Analytical Reviews.
15. Rashid, F., Miri, A., & Woungang, I. Secure enterprise data deduplication in the cloud.
16. Sadoughi, F., & Erfannia, L. Health information system in a cloud computing context.
17. Tahir, A., Chen, F., Khan, H.U., Ming, Z., Ahmad, A., Nazir, S., et al. A systematic review on cloud storage mechanisms concerning e-healthcare systems. Sensors.
18. Varanasi, S. R. (2025b). HIPAA-AS-Code: Automated Audit Trails in AWS Sage Maker Pipelines. European Journal of Engineering and Technology Research, 10(5), 23–26. https://doi.org/10.24018/ejeng.2025.10.5.3287
19. Whaiduzzaman, M., Gani, A., Anuar, N.B., Shiraz, M., Haque, M.N., & Haque, I.T. Cloud service selection using multicriteria decision analysis. TheScientificWorldJournal.