World Journal of Advanced Engineering Technology and Sciences
eISSN: 2582-8266
Cross Ref DOI: 10.30574/wjaets
Journal homepage: https://wjaets.com/

(REVIEW ARTICLE)

Check for updates

Real-Time Patient Experience Monitoring in Healthcare: A Solution-Oriented Review of Lakehouse-Based Data Engineering Architectures

Fnu Mohammed Sirajuddin *

University of New Haven, Connecticut.

World Journal of Advanced Engineering Technology and Sciences, 2026, 18(02), 286-301

Publication history: Received on 10 January 2026; revised on 18 February 2026; accepted on 21 February 2026

Article DOI: https://doi.org/10.30574/wjaets.2026.18.2.0079

Abstract

The patient experience is the key indicator of health care quality and patient-centered care, whereas the traditional assessment instruments rely on the late feedback in the shape of a survey, which is not reflective of the care dynamics in real-time. The rapidly developing digital health systems, telemedicine platforms, and the associated connected patient engagement solutions have introduced the continued flood of experience data that has to be scalably and low-latency analytically processed. The paper presents the solution in the form of a solution-oriented review of the data engineering architectures on the lakeside patient experience monitoring issue in healthcare. The Lakehouse architectures address the issue of fragmentation and latency of legacy systems by uniting the scalability of data lakes with the transactional reliability, governance, and analytical performance of data warehouses. The proposed model enables real-time usage and consumption of heterogeneous signals of patient experience, e.g., electronic health record messages, comments on mobile applications, nurse call systems, and telehealth communication logs. The paper illustrates practical cases of emergency department wait-time analysis, inpatient responsiveness monitoring, and digital front-door experience optimization to indicate how lakehouse-oriented structures can help in the generation of continuous insight and response to proactive service interventions. In addition, the paper determines valuable governance, security, and compliance challenges, and the ways lakehouse platforms enabled scalable and auditable patient experience analytics and compliance-regulatory ones.

Keywords: Patient Experience Monitoring; Real-Time Healthcare Analytics; Lakehouse Architecture; Healthcare Data Engineering; Streaming Data Pipelines; Patient-Centered Care

1. Introduction

Patient experience has become one of the main pillars of healthcare quality to be used along with clinical outcomes, patient safety, and operational efficiency to enable truly patient-centered care provision [1]. It goes beyond quantifiable medical outcomes to include the perceptions, interpretations, and emotional reactions of patients to their experiences of healthcare systems along the entire spectrum of care. The interactions are in the form of the clinical, administrative, and digital points of contact, such as making appointments, registering, emergency department interactions, inpatient care responsiveness, discharge planning, billing communication, and follow-ups after the care delivery [2]. With every interaction, patient trust, perceived empathy, transparency, and confidence are built up, making patient experience a multidimensional concept, driven by both the behavior of human and technological performances. The quick digitalization of healthcare has tremendously extended the scale, fineness, and speed of information linked to experience. Current settings produce endless data through the electronic health records, patient portals, mobile health applications, telehealth systems, nurse call systems, remote monitoring technologies, and linked medical devices [3]. These sources yield heterogeneous data types, which include structured clinical attributes, semi-structured operational logs, unstructured free-text feedback, sensor readings, and data on behavioral interaction, providing more of a holistic

* Corresponding author: Fnu Mohammed Sirajuddin.

and near-real-time view of patient engagement and care pathway friction compared to more traditional survey-based methods. As a result, patient experience is no longer viewed as a fixed or retrospective outcome, which is measured only after the completion of the care delivery process but as a dynamic process that changes and heats up over the course of a patient journey and is influenced in real time by waiting times, delays in communication, gaps in care coordination, staffing responsiveness, and digital system usability [4]. Signals of experience can occur at any given time and in many cases necessitate instant visibility and action before it deteriorates into the levels of dissatisfaction, decreased treatment compliance, or worse clinical and operational results. The traditional methods of analytics, which are generally batch-based, siloed, and retrospective, however, are unable to embody the temporal, contextual, cross-system complexity of these ever-growing signals, especially when the incoming data is at high velocity via divergent platforms and organizational domains [5]. To overcome these constraints, more and more data engineering designs need to be able to ingest, process, and analyze patient experience information as it arrives. Here, the lakehouse-based data architectures have become an interesting basis of integrated and real-time patient experience analytics [6,7]. This intersection is particularly essential in patient experience monitoring, where standard clinical data has to be matched with semi-structured operational incidents and unstructured experience response subject to a severe privacy and regulatory policy [8], allowing timely responses, proactive actions, and ongoing enhancement of care quality.

Even though patient experience nowadays becomes a strategic priority, current monitoring practices are not responsive and analytical enough [1]. The majority of healthcare organizations are using post-encounter surveys and manual feedback systems that have high latency, low response rates, and are also vulnerable to recall bias and thus unsuitable for detecting the experience issues that are changing very fast [2,4]. The data on patient experience is additionally dispersed among siloed systems, whereas the traditional data warehouses and data lakes do not support high-velocity and unstructured data processing with sufficient governance and reliability [5,6]. Real-time analytics are further restricted by regulatory requirements of privacy, security, and auditability, which prevent relating signals of patient experience to operation and clinical workflows, or identifying emergent problems in a timely manner [8-10]. Consequently, continuous, real-time management of patient experience is increasingly becoming a change of direction brought about by present-day data engineering structures [3]. The lakehouse-oriented platforms offer an integrated platform to accept batch and streaming data ingestion into a controlled environment, decreasing latency and data duplication and enhancing the consistency of analytical processes [6,7]. This is made possible through their support of transactional guarantees, schema management, and advanced analytics to combine machine learning methods to perform sentiment analysis and detect anomalies, and allow the creation of proactive, compliant, and scalable patient experience monitoring in the healthcare environment [9,11,12].

The paper will give an in-depth, the lakehouse-based data engineering architectures for the real-time monitoring of patient experience in healthcare facilities. It looks at the architectural elements, data streams, and analysis ability needed to support ongoing experience evaluation throughout the patient experience [6]. The research also explains the practical applicability of the suggested solution using the real-world cases related to healthcare services, such as emergency care, inpatient services, and telehealth services [10]. Moreover, the issue of governance, security, and compliance peculiar to regulated healthcare settings is addressed [8]. The paper will combine theoretical foundations of architecture with empirical evidence to show that lakehouse-based solutions can help healthcare organizations to stop being fragmented, delayed in measuring experience, and integrate it into real-time experience management [12].

2. Background and evolution of patient experience analytics

The patient experience assessment development is also closely linked to the advancement of healthcare information systems and data analytics paradigm [13]. The innovative approaches during the initial years were premised primarily on the utilisation of qualitative feedback and standardized survey instruments to measure patient satisfaction, and the level of satisfaction was a high indicator but lacked operation in the process of care delivery [14]. One of the ways to support the organized analysis of experience indicators and clinical and operating data became the data warehouse-based analytics because the healthcare organizations started adopting electronic health records and enterprise reporting systems [15]. Such architectures, though, were mostly batch-based and ill-adapted to the dynamic and event-based nature of interactions with patients. The next trend researched the topic of data lake architectures as a scalability and data heterogeneity solution, particularly in response to the growing unstructured patient feedback and digital engagement mechanisms [16]. Even though data lakes facilitated increased flexibility in data storage, issues of administrative accountability, data quality, and reliability in data analysis remained a problem, limiting their effectiveness. Parallel work on streaming and real-time analytics demonstrated that it could be a low-latency monitoring in a healthcare environment, yet these systems were typically deployed individually and were not linked with the rest of the analytical and governance frameworks [11]. All these portray a disjointed architectural background of integrated paradigms that do not have a well-managed and monitored real-time, end-to-end patient experience tracking platform. This has left a gap of growing emphasis on convergent and hybrid architectures with the ability to include historical and real-time analytics and meet regulatory and compliance imperatives [17].

Table 1 Architectural Approaches for Patient Experience Monitoring

Architecture Paradigm	Primary Focus	Strengths	Key Limitations
Survey-Based Systems	Retrospective feedback	Standardization, ease of interpretation	Delayed insights, low response rates
Data Warehouses	Structured reporting	Strong consistency, regulatory compliance	Limited real-time and unstructured data support
Data Lakes	Scalable data storage	Flexibility, heterogeneous data handling	Weak governance, data reliability concerns
Streaming Architectures	Event-level processing	Low-latency detection	Limited historical analytics, integration complexity
Lakehouse Architectures	Unified analytics	Real-time + historical analytics with governance	Emerging maturity, architectural complexity

The evidence that can be summarized in Table 1 suggests that patient experience analytics architectures have a fundamental trade-off. Reliability and regulatory compliance systems are often not timely, whereas real-time responsive architectures are generally not governable and analytically consistent. To successfully monitor patient experience, it is essential to be capable of providing ongoing insight as well as processed, trusted, and auditable data simultaneously. Lakehouse architectures have come up as a converged model with the capability to meet this dual need by integrating both batch and streaming analytics into a single managed space. In contrast to previous hybrid systems that were based on loosely coupled systems, the lakehouse paradigm focuses on the architectural integrity, allowing the experience data to flow smoothly between the raw ingestion state and the analytical consumption state. This intersection is especially acute in healthcare environments, where patient experience indicators need to be converted into operational and clinical activity without violating the rigorous privacy, security, and auditability limitations.

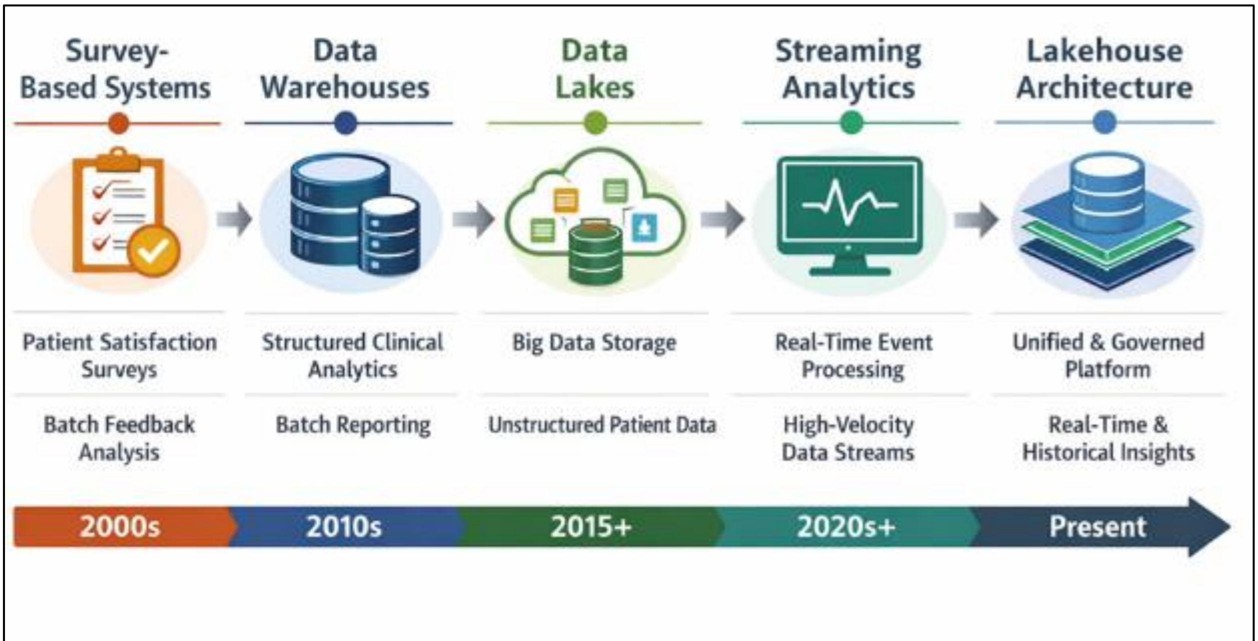

Figure 1 Evolution of Data Architectures for Patient Experience Monitoring

The evolution of architecture depicted in Figure 1 not only shows a transformation in the ways and means of the patient experience in the healthcare systems but also reflects the operating requirements of modern healthcare systems. Over time, the dependence of data and user interaction caused the weaknesses of the partitioned and batch processing styles of the architecture to surface. The shift towards lakehouse-type architectural designs is seen as not only a technological upgrade but also a change in analytics attitudes, from retrospective assessments to constant experience management. Lakehouse architectures integrate real-time data ingestion, historical analysis, and governance capabilities, thus providing a structural foundation for treating patient experience as a core operational metric rather than an auxiliary quality marker. This narrative considers the data engineering of lakehouse as a critical factor in the next-generation patient experience monitoring, thus preparing the ground for the next sections that will investigate the architectural components, the solution workflows, and the real-world healthcare applications more thoroughly.

Though the existing applications of patient experience analytics are disconnected from each other in terms of architectural approaches, data gathering techniques, and governance protocols, the existing applications have been alternating between the retrospective and survey applications on one side and the real-time and data-driven applications on the other hand. A reference model ensuring the combination of real-time patient experience indications, scalable analytics, and regulatory compliance on a single architecture is yet to be achieved. In order to address this gap, this paper suggests a reference architecture based on a paradigm of a lakehouse, which is explicitly designed to resolve the ongoing monitoring of patient experience. This proposed architecture is known as PX-LRA (Patient Experience Lakehouse Reference Architecture).

3. Proposed PX-LRA framework: a Lakehouse-based model for real-time patient experience analytics

To actualize the architectural concepts and analytical requirements as discovered in this study, the paper introduces the PX-LRA (Patient Experience Lakehouse Reference Architecture) conceptual framework that would potentially be utilized in supporting the continuous and real-time monitoring of patient experiences in controlled healthcare facilities. Although the structure is premised on a lakehouse foundation, it is mostly significant because it does not adhere to either traditional healthcare analytics directions or generic enterprise lakehouse designs, as it explicitly makes the patient experience a time-sensitive operational measure, rather than a retrospective quality measure.

3.1. Conceptual Overview of PX-LRA

The PX-LRA represents a model that outlines a comprehensive system of data engineering and analytics that combines heterogeneous patient experience signals with clinical and operational data streams, on a framework that relies on a lakehouse. Unlike in the traditional healthcare analytics systems, where the experience data is fragmented, post hoc, and delayed, PX-LRA is designed to facilitate the low-latency creation of insights, followed by historical analysis and enforcement of governance within one analytical system. Traditional analytics operations in healthcare are mostly retrospective and based on the delayed ETL operations, predetermined data warehouses, and post-encounter surveys to identify patient satisfaction and quality of service. The degradation of experience in these models is normally observed once the care delivery has been done, and timely intervention would be a challenging case. The PX-LRA abandons this paradigm because it has real-time ingest and analytics capability, which is directly associated with the temporal dimensions of care delivery. Therefore, patient experience is advanced as a downstream evaluative indicator to an operationally practical sign that can be made to aid the decision-making process as care is in progress.

3.2. Architectural Layers of PX-LRA

The PX-LRA framework consists of five interdependent architectural layers, each addressing a critical requirement for real-time patient experience analytics within regulated healthcare settings. PX-LRA architecture is a set of five layers, which are dependent on each other and considered one of the key requirements of real-time patient experience analytics in managed healthcare facilities.

3.2.1. Experience Signal Ingestion Layer

High-velocity, unstructured, and heterogeneous patient experience data can be found in various forms in this layer, such as electronic health records, nurse call systems, appointment and queue management systems, telehealth systems, mobile applications, and structured and unstructured patient feedback channels. It is able to deal with batch and streaming ingestion to overcome the limitations of the capabilities of the source systems. The ingestion layer provides the functionality of receiving the signals of experience with the lowest possible latency and maintaining the clinical and operational context.

3.2.2. Metadata Layer and Integrity Lakehouse Storage

All the read data is stored in a unified lakehouse environment with a mixture of large-scale storage of objects and transactional guarantees, schema control, and single-centralized metadata authorization. This layer brings a stable and dependable as well as traceable experience throughout the lifecycle of patient experience information. The PX-LRA lakehouse layer offers the same flexibility as the other data warehouses (no fixed schema), but the level of regulation and auditability is enforced, as compared to the data lakes, which are not regulated at all.

3.2.3. A superimposition of Real-time and Historical analytics

The real-time streaming analytics can be used with PX-LRA and historical analysis that is performed in the same analytics platform in a batch fashion. The layer allows descriptive and diagnostic, and machine learning-based pattern recognition to identify the experience degradation at an extremely early phase. The real-time and longitudinal analytics can enable healthcare organizations to address the issues that are arising in the moment, as well as evaluate the trends and performance of the system over time.

3.2.4. Governance, Security, and Compliance Layer

There is no second-rate matter with architects, which is governance. This layer includes fine-grain access control, encryption, data lineage, and audit trail so as to abide by the regulatory and institutional policies. PX-LRA allows experiencing sensitive data in real-time without jeopardizing privacy, accountability, and compliance by specially applying governance tools in the data engineering and analytics pipeline.

3.2.5. Experience Intelligence and Action Layer

The Experience Intelligence and Action layer converts analytical products into operationally actionable intelligence, which is in the form of dashboards, alerts, and decision-support interfaces. This level provides a clear correspondence of analytics of patient experience to the clinical and operational workflow in such a manner that interventions can then be executed in the real time as augmented/reduced staffing, increased communications, or re-engineering the workflow. The combination of action and analytics is a complete cycle that will start PX-LRA to feel the management rather than report.

3.3. Comparative Positioning of PX-LRA

Although PX-LRA is deployed on a lakehouse basis, its input does not lie in the embrace of lakehouse technology itself, but the redesigning of this architectural paradigm to support the specific task of real-time patient experience monitoring, in terms of the embodiment of its temporal, regulatory, and operational needs. The current healthcare analytics systems and generic lakehouse implementations do not explicitly consider patient experience as an operating signal in their architecture, which restricts their capabilities of timely intervention in the process of active care discharge. By identifying the differences between the design goals, analytical priorities, and governance assumptions of PX-LRA and these prevailing approaches, this section will locate PX-LRA in relation to them, which is not necessarily a technological difference. The comparison points out that PX-LRA is not intended as a general analytics back-end or a retrospective reporting platform, but as an operational intelligence platform that makes it possible to be aware of experience in real time and take action in controlled healthcare processes.

Table 2 Architectural Comparison of PX-LRA with Existing Analytics Approaches

Dimension	Traditional Healthcare Analytics	Generic Lakehouse Architectures	PX-LRA Framework
Analytical Orientation	Retrospective assessment	Use-case agnostic analytics	Real-time operational intelligence
Temporal Focus	Post-encounter analysis	Variable, not experience-driven	In-process experience monitoring
Patient Experience Treatment	Outcome metric	Optional data domain	Primary analytical construct
Data Processing Model	Batch-centric	Batch + streaming (generic)	Streaming-first with historical context

Governance Integration	External or downstream	Platform-level, generic	Embedded, experience-aware
Workflow Coupling	Weak or indirect	Minimal	Direct integration with care operations
Intervention Capability	Reactive	Analytical only	Proactive and closed-loop

3.4. Distinguishing Features of PX-LRA

The PX-LRA framework introduces several characteristics that distinguish it from existing healthcare analytics models and generic lakehouse platforms:

- Experience-Centric Design: Patient experience is modeled as a primary, time-sensitive operational signal rather than a secondary quality indicator.
- Unified Analytics Paradigm: Real-time and batch analytics coexist within a single governed architecture, eliminating fragmentation and reducing insight latency.
- Governance-by-Design: Compliance, auditability, and lineage tracking are embedded directly into the architecture instead of being enforced post hoc.
- Operational Responsiveness: Analytical outputs are tightly coupled with clinical and operational workflows, enabling proactive intervention during active care delivery.

3.5. Applicability and Extensibility

Even though PX-LRA is devoted to the real-time observation of patient experience, its architectural structure is specifically oriented to other healthcare analytics aspects that should develop an insight into time and in a highly controlled and monitored environment. The ease of combination with a single source of data ingestion, low-latency analytics, embedded governance, and workflow-intelligent intelligence into the framework permits its application to broader categories of ongoing healthcare analytics problems than experience measurement itself. PX-LRA will delimit an architecture of reuse by clearly decoupling itself as well as historic retrospective analytics pipelines and a generic lakehouse platform to support situational awareness and in-process decision-making. In this respect, the framework provides the researchers and healthcare organizations with a well-organized architectural platform on which they can transform the episodic and post hoc-oriented measurements into continuous and real-time experience and performance management of controlled clinical environments.

4. Lakehouse-oriented data engineering architectures in healthcare

Recently, lakehouse architecture has come to be a convergent data engineering paradigm in order to overcome the endemic fragmentation of historic healthcare analytics infrastructures [17, 18]. Lakehouse systems combine the scalability and openness of data lakes with the transactional reliability, as well as schema enforcement and query performance optimizations that were historically linked to enterprise data warehouses [6,19]. The relevance of this architectural synthesis concerns especially healthcare settings, where the analytical workloads are to maintain both structured clinical data, semi-structured operational and administrative data, and unstructured patient experience data in the form of free-text feedback or communication transcripts and digital interaction traces [20]. Lakehouse architectures allow both continuity-based analytics and longitudinal historical analysis to perform using the same platform without the need to move or replicate data across heterogeneous systems through redundant data movement or replication [11,21].

In contrast to previous hybrid architectures that have been built on the notion of weakly-crusting pipelines and multiple specialized platforms, lakehouse-based architecture has been mentioned as an integrated and unified infrastructures that share storage, metadata management, and data processing capabilities [19]. This centralization eases the data pipelines and simplifies the architectural complexity, which, as previously mentioned, research has cited as a substantial hindrance to scalable healthcare analytics [17]. Moreover, some of the improvements in analytical consistency, decreased data latency, and better governance are mentioned as the significant advantages of the lakehouse paradigm [18, 21, 22].

Figure 2 depicts architectural building blocks that are indicative of the general trends of consolidated, governance-sensitive healthcare analytics. The lakehouse architectures make it easy to combine real-time and historical information and are suitable to allow longitudinal analysis without affecting the low-latency access of new signals. The latter use is especially useful in context-sensitive analysis tasks like patient experience monitoring, in which operational processes, clinical processes, and digital interactions overlap throughout the healthcare enterprise. Lakehouse-based designs with mechanisms of governance, such as metadata-based management, fine-grained access control, and extensive auditability, can be aligned to the strict regulatory and compliance standards that are typical of a healthcare setting. Instead of describing the lakehouse as a patient experience analytics solution that was designed on purpose, it should be considered a data engineering model that can be used to achieve a number of different analytical goals in one, consistent, and controlled system.

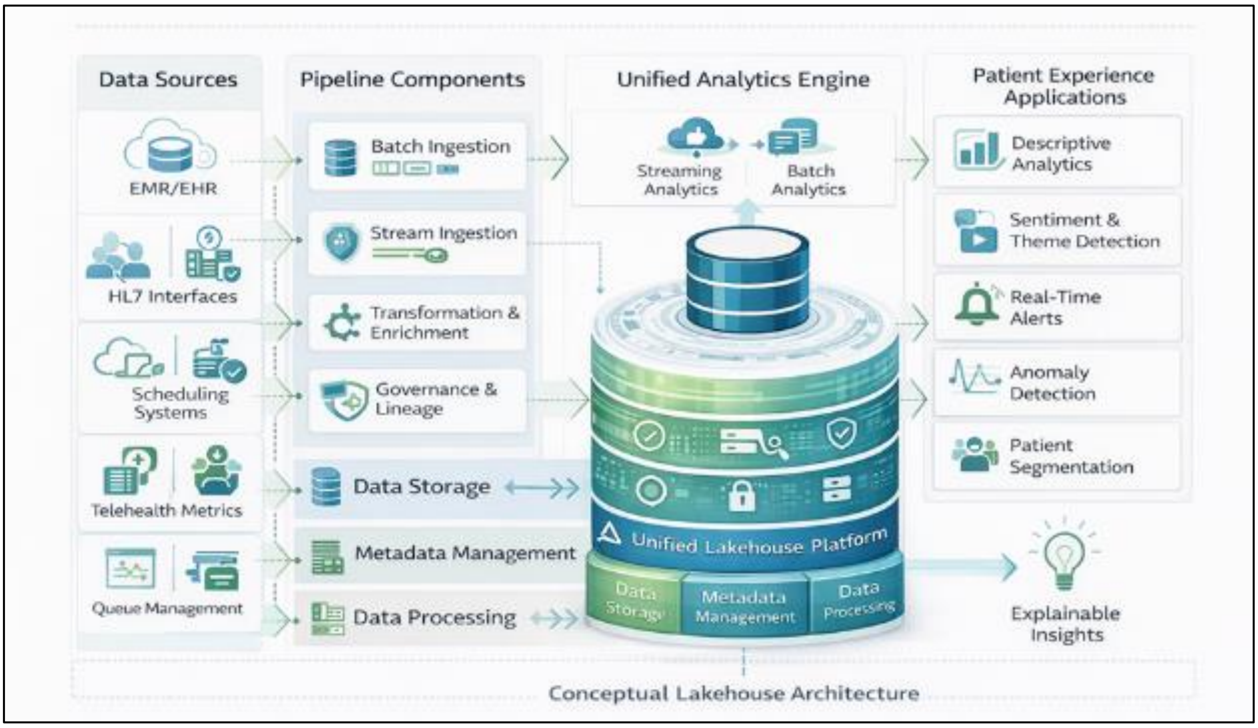

Figure 2 Conceptual View of Lakehouse Architecture in Healthcare Analytics

5. Analytical methods used in patient experience monitoring

The proposed patient experience analytics study presents the obvious methodological sequence and shows the process of maturing healthcare data ecosystems and an increase in appreciation of patient experience as a multidimensional and dynamic construct [23]. Initial empirical studies were mostly based on descriptive analytics and were based on summary statistics, frequency distribution, and composite scores of satisfaction obtained by using standardized surveys [14]. These methods were mostly retrospective and cross-sectional, in the sense that they provided top-level images of patient experience but without explanatory power or a time-sensitive nature. With the rise of electronic health records, patient portals, and digital feedback channels, analytical attention shifted towards unstructured sources of data, and text analytics methods and natural language processing methods started to be applied in order to make sense of free-text comments, grievances, and narrative feedback [24]. This development allowed more contextualization of patient emotion and disclosed experiential aspects that the use of numerical scoring alone could easily cover up. Increasingly more recent studies synthesized in this work are focused on real-time and near-real-time analytical paradigms, such as stream processing, event-driven analytics, and continuous monitoring frameworks, which are aimed at capturing signatures of experience as they arrive in care delivery and not once care delivery has occurred [11], [25]. In this framework, machine learning techniques are discussed as the means of discovering latent patterns, grouping patients according to their experiences, and identifying the exceptions or anomalies in experience paths, especially in data-heavy conditions like emergency departments, outpatient clinics, and telehealth environments [9].

Analytical processes are best understood within broader socio-technical frameworks, where performance depends on data integration architectures, data quality oversight, interoperability, and governance mechanisms. Rather than treating analytics as standalone computational tools, they are increasingly viewed as embedded capabilities shaped by organizational processes and operational constraints. A recurring perspective is that the primary objective of patient experience analytics is not predictive accuracy or optimized scoring alone, but timely, explainable, and context-aware interpretation of patient experience signals. This approach enables healthcare organizations to connect patient perceptions with underlying clinical workflows, resource availability, and service delivery dynamics in support of responsive and patient-centered decision-making.

Table 3 Analytical Methods Commonly Reported in Patient Experience

Analytical Category	Specific Techniques	Data Sources Commonly Analyzed	Analytical Objective	Typical Outputs	Limitations	PX-LRA Distinction
Descriptive Analytics	Aggregation, frequency analysis, trend analysis	Patient satisfaction surveys, feedback forms	Summarize historical experience levels	Experience scores, satisfaction indices, trend summaries	Retrospective focus, delayed insights	PX-LRA supports continuous descriptive views with streaming refresh, enabling a near-real-time experience baselining
Inferential Analytics	Correlation analysis, regression models	Operational metrics, clinical workflow data	Identify associations between experience and care processes	Statistical relationships, coefficients	Static assumptions, limited actionability	PX-LRA embeds inference within an operational context, enabling rapid validation against live process data
Text Mining	Keyword extraction, topic modeling	Free-text feedback, complaints, reviews	Identify recurring experience themes	Topic clusters, term frequencies	Noise sensitivity, weak context retention	PX-LRA combines text mining with event context and metadata, improving interpretability and relevance
Sentiment Analysis	Lexicon-based and ML-based scoring	Patient narratives, call-center transcripts	Quantify emotional tone	Sentiment scores, polarity trends	Context loss, sarcasm ambiguity	PX-LRA augments sentiment with temporal and situational signals, reducing misclassification risk
Event Stream Analytics	Window aggregation, event correlation	Digital interactions, system logs	Monitor experience signals in real time	Alerts, live dashboards	High integration complexity	PX-LRA natively integrates streaming and batch pipelines, reducing architectural fragmentation
Machine Learning Models	Clustering, classification, anomaly detection	Multimodal experience and operational data	Detect latent patterns and deviations	Patient segments, anomaly flags	Explainability and bias issues	PX-LRA enforces governed ML pipelines with lineage and auditability, improving trust and adoption

Temporal Analytics	Time-series analysis, sequence modeling	Longitudinal interaction data	Track experience evolution	Experience trajectories, change points	Requires consistent time alignment	PX-LRA ensures schema consistency and temporal integrity across streaming and historical data
Hybrid Analytical Pipelines	Statistical + ML workflows	Integrated clinical, operational, and PX data	Contextual interpretation of experience	Enriched dashboards, decision signals	Governance and orchestration complexity	PX-LRA unifies pipelines under a single governance, metadata, and access model, enabling scalable deployment

6. Real-time use cases and evidence-based scenarios

Real-time patient experience stands as a current quality of current healthcare frameworks, particularly in high-throughput and digitally mediated care locations [26]. Unlike the traditional retrospective models of assessment that use delayed surveys and episodic reports, the real-time use cases are based on real-time capture and processes of the experiential signals during the process of delivery of care by the patients. They are events related to operations, communication patterns, and patient-reported inputs that are utilised in order to provide situational awareness during the care delivery process and not in the post-care delivery process [27].

One of the examples of the real-time patient experience use cases demonstrates that with the assistance of the existing data engineering and analytics platform, the experience data may be directly incorporated into clinical and operational scenarios [23,28]. Patient experience is not a value in isolation, but in such framing a moving process, which varies as care paths and interactions with services. These plans are not only applicable to the implementation of experimental strategies; a trend of such plans as repeated approaches to the implementation of analytical strategies is emerging in the settings of emergency care, inpatient service, outpatient management of workflows, and virtual care platforms [10,29]. All these data mean that the experience-based risks could be identified earlier, the variability of service delivery could be reduced, and more patient-centered decisions could be made through ensuring the opportunity to act in response to the new operational and clinical changes in time with the help of real-time analytics [12]. Some of the common instances of application and evidence application cases are listed in the sub-sections below and are often reported in the research and industry applications.

6.1. Emergency Department Experience Monitoring

It is common to find emergency departments mentioned as a critical location of real-time patient experience analytics because of the large volumes of patients, unpredictable workflows, and the sensitivity of delays. Poor patient experience in emergency care setups has always been related to long waiting times, late triage, and poor communication. On-the-fly tracking of the queue length, triage, and care progression events will help to identify the manifestation of the experience degradation in its early phases before it manifests itself in the negative satisfaction ratings or official complaints. The operational measures are frequently compared to the real-time patient feedback based on kiosks, mobile messages, or short in-visit surveys in order to put dissatisfaction into context and determine its root causes. Real-time monitoring of congestion trends and delay in response associated with it can permit informed operational responses, like dynamic staffing or workflow redesign, to enhance perceived quality of care in situations of high demand [30].

6.2. Inpatient Care Responsiveness and Communication

Real-time observation of the indicators of responsiveness and communication in inpatient care conditions is the leading factor in defining patient experience. Some of the most often monitored indicators are the activity of the nurse call system, the response time, the number of visits to the patient and staff, and the time lapse between reception of the service request and its solution. These operation signals are usually studied in conjunction with patient feedback, which is received with the help of bedside devices or mobile applications. Delays are the main factors leading to negative patient assessments of quality of care and staff attention, as well as inconsistent patterns of responses or even a breakdown of communication. Real-time analytics can be used to identify bottlenecks in the service (understaffed shifts or workflow inefficiencies) before they become a problem, allowing for action before dissatisfaction sets in. Close connection with enhanced continuity of care, enhanced trust and transparency in hospital stay is highly correlated with this continuous monitoring capability [31].

6.3. Outpatient and Ambulatory Care Experience

In outpatient and ambulatory care environments, workflows are high-throughput and heavily scheduled, and patient experience monitoring in real time has grown to become more significant. There are myriad streams of experiential data produced by digital check-in solutions, appointment scheduling, and patient portals that capture patient interactions at various points of care. In such settings, real-time monitoring of appointment attendance, check-in delays, the time of consultation commencement, and follow-ups is also instrumental towards the development of patient experience. Late or poor coordination that is usually overlooked by retrospective reporting may be identified in real-time analysis, and corrective measures undertaken in a timely manner [32].

6.4. Telehealth and Digital Front-Door Experience

The rapidly growing telehealth services have paid more attention to digital patient experience monitoring. The digital experience is a very important determinant of patient satisfaction in remote care delivery models. Some of the real-time indicators that are usually measured are session latency, the stability of the connection, waiting time in virtual environments, and sentiment based on chatbot or messaging interactions. Slow connections, hiccups, or breakdowns in electronic communication may have a great impact on the perception of the patients concerning the availability of care and the quality of the service. It is required to continuously monitor these indicators to ensure continuity of care, reduce patient frustrations, and maintain trust in virtual healthcare services. The criterion of smooth and reliable digital experiences is thus one of the major ones when it comes to patient engagement and the acceptance of telehealth as a practical alternative to face-to-face care.

6.5. Post-Discharge and Continuity-of-Care Experience

Continuity of care has also been shown to be important in patient satisfaction and long-term patient outcomes through the application of real-time monitoring of patient experience throughout the post-discharge phase. The experience of the patients does not end at discharge, and it also includes the continuum of following up with the patients, monitoring their recovery, and receiving post-care support services. Slowness in responding to post-discharge problems, like problems in managing symptoms, adhering to medicine, or following-up care, is also highly correlated with patient dissatisfaction and preventable hospital readmissions. Through real-time analytics, it is now possible to be able to detect emerging issues at the earliest stage of the recovery process, and the care teams can step in ahead of complications before they arise.

7. Performance and architectural benefits

Comparisons of traditional warehouse-based solutions, disjointed tool-specific analytics solutions, and more integrated solutions in terms of architecture show that significant improvement has been made in various dimensions of operation in the event of adoption of unified architectures [30, 31]. The legacy systems, which are often based on batch processing, are often associated with fixed schemas and repeated patterns of data that are often related to increased ingestion delays, slower insight generation processes, and inconsistent system-to-system analytical outcomes [15], [32]. Conversely, unified and lakehouse-based architectures allow the handling of both structured and unstructured patient experience data on the same platform and minimize the overhead of data movement as well as synchronization errors [18], [33]. The end-to-end ingestion latency, query execution time, elasticity managed at peak interaction volumes, and system resilience in the face of delayed or incomplete data streams are the typical metrics used to measure performance in healthcare settings [21].

These are particularly successful in real-time and on-the-fly implementations of patient experience where a timely feedback loop renders operative the applicability of experiential cues to operations [26]. In addition to the computational performance, architecture characteristics that relate to the governance and manageability, such as metadata management, routine access control, tracing of the data source, and processes of auditable data transformation, are essential in supporting credible analytics [22,34]. These characteristics combine with analytic confidence and repeatability, which are essential in controlled health care environments where compliance is restricted, as well as privacy and quality assurance policies [8]. Reduced coordination between pipelines and reduced architectural complexity further permits greater analytical iteration and more definite implementation of sophisticated algorithms like experience tracking based on machine learning [23]. Collectively, these findings demonstrate that architectural design choice may be a direct cause of the extent to which real-time patient experience analytics can be made scalable, reliable, and of use as a foundation of operational and clinical decision-making that is centrally based on patient needs [12].

Table 4 Performance and Architectural Benefits of Lakehouse-Oriented Approaches in Healthcare

Dimension	Reported Benefit	Evidence	Relevance to Patient Experience Monitoring
Data Ingestion Latency	Reduced end-to-end latency	A transition from batch-hour delays to near-real-time processing	Enables timely detection of experience degradation
Analytical Throughput	High concurrent query support	Unified storage supports multiple analytical workloads simultaneously	Supports clinical, operational, and experience analytics together
Scalability	Elastic scaling across data volumes	Object storage-based designs handle growing experience data streams	Accommodates increasing digital interaction data
Data Consistency	ACID-compliant transactions	Improved reliability compared to traditional data lakes	Enhances trust in experience metrics
Pipeline Complexity	Simplified data pipelines	Reduced system sprawl and redundant ETL processes	Improves maintainability and operational efficiency
Governance and Auditability	Centralized metadata and lineage	Automated tracking of data access and transformations	Supports compliance and accountability
Cost Efficiency	Lower infrastructure redundancy	A single platform reduces the need for multiple analytics systems	Enables sustainable long-term deployment
Analytical Flexibility	Support for batch and streaming analytics	Unified platform adapts to diverse experience use cases	Facilitates both historical and real-time insights

The provided performance and architecture benefits, as outlined in Table 3, explain the reason why the concept of lakehouse-oriented design is becoming one of the topics in the area of healthcare analytics as a potentially appropriate framework for real-time monitoring of the patient experience. These architectures offer system-level enhancements and are not independent performance enhancements, and are extended to data ingestion, processing, governance, and analytical consumption. The benefits are particularly efficient in such situations when patient experience data is supposed to be evaluated with clinical and operational data under stringent regulations. Lakehouse-oriented architectures enable healthcare organizations to be more operational in exercising their patient experience insight by reducing latency, improving reliability, and simplifying data management. The conclusion helps to support the assumption that the issue of architectural convergence is critical to the creation of patient experience analytics based on historic reporting to continuous, scalable, and reliable analytical practice.

8. Governance, security, and compliance considerations

There are three basic dimensions that governance of patient experience analytics is based on, such as security, privacy, and regulatory compliance, which are all prerequisites in healthcare settings [35]. Patient experience data is often shared with secured health information, clinical operations, and contacts, which are subject to rigorous privacy and security rules [8], [36]. Discontinuous data architectures also result in a further complication of governance as sensitive data is spread to different systems, with varying access controls and auditing processes [32], [37]. By contrast, centralized policy implementation, uniform metadata operations, and complete data provenance are all made possible by integrated data platforms [22], [33]. Governance is not limited to access control only, and it includes data quality, consistency, and lifecycle management, which are directly related to the credibility and reliability of patient experience analytics [35], [38]. With the volume of data, velocity, and diversity on the rise, governance can no longer be a post hoc operation, but rather a continuous one that has to be enforced. This fact explains why the necessity to have architectures that internalize the governance controls at the very data engineering layer itself [21], [34].

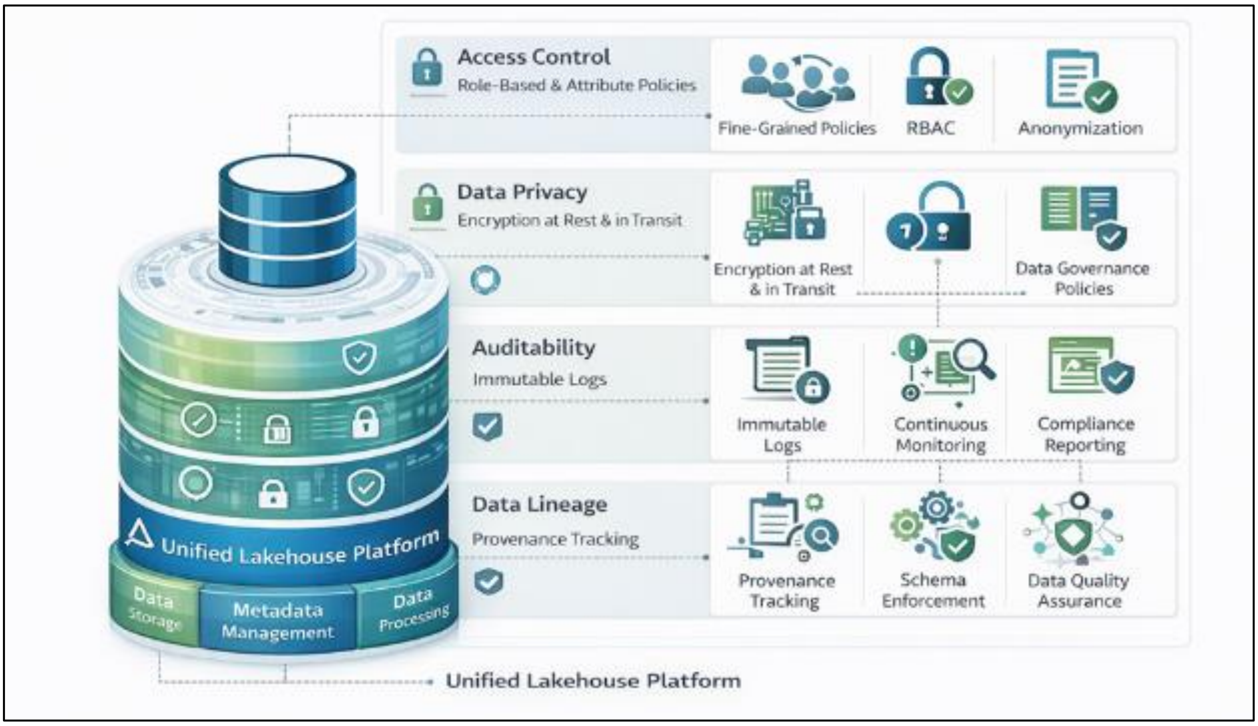

Figure 3 Governance, Security, and Compliance Layers in Lakehouse-Based Healthcare Analytics

The common design patterns in healthcare analytics design deployed in the form of governance layers are shown in Figure 3. The compliance requirements, which typically include the defense of patient rights, privacy, and regulatory reporting requirements, are usually met by centralized metadata catalogs, role access control, and unalterable audit trails. Both data encryption at rest and data encryption in transit are generally considered basic security requirements, and a fine-grained access control policy is significant in order to minimize the disclosure of sensitive data in an operating environment. Better attention is also paid to tracking data provenance and data lineage, particularly in real-time analytics pipelines receiving and storing signals of patient experience. All these features enable health care organizations to demonstrate responsibility, inquire about anomalies, and respond to compliance audits without interrupting the analytical workflow. Governance-sensitive architectures are offered as a primary source of scalable and credible real-time analytics instead of perceiving compliance as something external.

Table 5 Governance, Security, and Compliance Considerations Reported in Healthcare Analytics

Governance Dimension	Key Mechanism	Description in	Impact on Patient Experience Analytics
Access Control	Role-based and attribute-based policies	Restricts data visibility based on user roles	Prevents unauthorized access to sensitive PX data
Data Privacy	Encryption and anonymization	Protects patient-identifiable information	Enables compliant analysis of experience signals
Auditability	Immutable logs and lineage tracking	Records data access and transformations	Supports regulatory audits and accountability
Metadata Management	Centralized data catalogs	Maintains schema and data definitions	Improves the interpretability of experience metrics
Data Quality Governance	Validation and schema enforcement	Ensures consistency across data streams	Enhances the reliability of real-time insights
Regulatory Compliance	Policy-driven controls	Aligns analytics with healthcare regulations	Reduces compliance risk in real-time monitoring
Security Monitoring	Continuous access and activity logging	Detects anomalous data usage	Protects analytics infrastructure integrity

9. Challenges and open research gaps

Despite a growing interest in real-time monitoring of patient experiences and the growing popularity of data architectures based on lakehouses, a number of unanswered questions still restrict the ability to implement them at a large scale and with high effectiveness in healthcare environments [39]. These issues cut across technical, analytical, organizational, and regulatory levels, all of which are complex in nature as patient experience analytics becomes a part and parcel of daily care processes [40]. Even though integrated real-time analytics systems have shown technical effectiveness, there are ongoing limitations which are associated with data heterogeneity, interpretability of the analytical results, interoperability with traditional systems, as well as scalability under stringent governance and compliance criteria [17], [41]. Notably, such issues can be traced to both architectural and computational limitations; however, they are often a result when clinical practices, regulatory requirements, data management, and the changing needs of patient-centered analytics come into conflict [35], [42]. With patient experience becoming a more important factor in the assessment of care quality, reimbursement plans, and institutional responsibility in systems of value-based care, attention to these gaps is an urgent and unceasing concern [1], [26].

9.1. Data Quality and Signals of Patient Experience Subjectivity.

The data of patient experience is always problematic in terms of its quality, reliability, and interpretation. Patient experience indicators are subjective, unlike clinical measurements, and also subject to individual expectations, emotional conditions, cultural values, and situational context. Noise, ambiguity, and incomplete information are likely to impact real-time data sources like free-text feedback, call center transcripts, mobile application interactions, and digital communication logs. There is also inconsistency in data capturing processes, different rates of patient engagement, and differences in digital access, which further entail bias and sparsity in such datasets. These make real-time analytics difficult because they are needed when quick interpretation is needed, even when the inputs are uncertain or imperfect. The issue of data accuracy, completeness, and representativeness of the patient experience information has continued to be a serious unresolved concern, especially when the information is utilized to make operational or clinical decisions.

9.2. Explainability and Interpretability of Analytical Models.

The interpretation and transparency of models continue to be the primary issues of concern because machine learning and advanced analytics are increasingly used to monitor patient experience. Although predictive and pattern-discovery models are useful in determining the trends related to experience, the logic behind these models may be hard to comprehend for clinical, operational, and administrative stakeholders. This unaccountability is especially worrisome in a controlled medical setting, where accountability, traceability, and trust are key players. In situations where the outputs of an analytical process cannot be interpreted, the insights might be poorly received and will not contribute to the decision-making process. Consequently, there has been an acute unresolved problem of the successful compromise between analytical complexity and human explainability with respect to real-time systems, where an explanation as well as a decision will be required in a short period of time.

9.3. Compatibility and Interference with Existing Systems.

The nature of healthcare information ecosystems is fragmented in nature since these systems involve a number of disjointed systems, which have been developed independently and have been streamlined to fit a specific clinical or administrative purpose. It is always a challenge to integrate patient experience data into the electronic health records, operational systems, scheduling, organizing tools, and digital engagement tools. Semantic structures, differences in data standards, communications, and frequency of updates will make it impossible to share data flawlessly and conduct real-time analytics. Moreover, the old-fashioned clinical operations, which are usually highly controlled and risk-averse, should not be disturbed by the integration activities. Interoperable, standard, and low-latency integration between the old systems and the new digital systems is one of the recurrent problems of healthcare analytics.

9.4. Scalability and Compliance When using Real-Time Environments.

The architecture of lakehouses is designed to be scalable, but the incessant observation of patient experience poses unremitting ingestion, processing, and analytical loads, which it must be capable of performing with stringent compliance imperatives. Operationally complex maintenance of consistent governance, access controls, auditability, and data lineage becomes operationally hard, particularly in multi-hospital networks or geographically distributed healthcare systems. One of the continuous worries is that the improvement of the performance of the given system cannot be conducted at the expense of compliance with the regulatory requirements and the security of the data. When assessing the potential of real-time analytics architecture to operate in terms of high analytical performance and effective compliance, it is necessary to be more attentive in the future when the volumes of information, their velocity, and the complexity of analytical procedures are increasing continuously.

10. Discussion and implications for healthcare systems

The study examines existing literature and practices to identify the role of a lakehouse-based data engineering architecture in healthcare institutions for tracking patient experiences in real time. It is also in the results that a broader-based transformation of healthcare analytics to a more continuous, more data-driven experience management is highlighted. By enabling healthcare systems to combine experiential data with clinical and operational context by aggregating different indicators of the patient experience in a centralized and controlled analytical environment, Lakehouse-based architectures enable healthcare systems to process the experiential data effectively. This integrated methodology makes the trial of the unified approach not to be considered as a single dimension quality measure but a dynamic measure, which is closely linked to the care delivery processes and system performances. The research findings, in terms of organizational implications, include the fact that the healthcare systems, which can process the data in real-time, are in a better position to identify any risks that the experience poses and prevent them in good time. Real-time information regarding patient contact, communication delays, and delay points in the service process may be utilized to make operational changes to improve coordination of care and patient satisfaction. In addition to that, experience data and clinical workflow analysis enable making decisions more holistically, which contributes to preserving patient-centered care models and value-based healthcare objectives.

At the systems level, this paper will demonstrate that architectural convergence is critical towards enabling scalable, resilient, and regulation-conscious analytics in modern healthcare systems. The findings indicate that a data engineering architecture that is based on lakehouses has one technical foundation that can be used to provide governance, security, auditability, and high-performance analytics at the same time. Lakehouse platform removes the presence of fragmentation and inconsistency that characterize multi-system analytics environments by managing metadata within the same paradigm of architecture. The convergence is particularly relevant, as healthcare organizations are becoming more and more reliant on the high-velocity data on patient experience generated by digital health platforms, telemedicine, and operational systems. The study also indicates that the architectural abilities are especially relevant to large and dispersed healthcare institutions that are operating in multiple facilities, care locations, and geographical units. During such environments, having a unified set of data definitions, access controls, and compliance policies should be implemented to gain the reliability of the analytics and organizational responsibility across various systems. Lakehouse-based architectures can be used to compare, aggregate, and interpret the insights about patient experience across institutional boundaries without violating regulatory imperatives by supporting interoperability and standardized governance.

11. Conclusion

The use of lakehouse-based data engineering architecture as the key tool to enable real-time patient experience monitoring is discussed in the paper as one of the essentials in contemporary healthcare. The overlap of the available knowledge and evidence-based situations indicates the limitations of the traditional and retrospective approaches to patient experience assessment, and the contemporary care delivery with its dynamic and event-based nature, particularly. Patient experience data are flowing in real-time, at both clinical and operational and digital touch points, and require the application of scalable data structures that support ingestion of low-latency data, integrated analytics, and controlled data management. Architectures proposed in this paper based on lakeshouses address significant weaknesses of traditional data warehouses, data lakes, and standalone streaming systems proposed in this paper, by implementing both real-time and historical analytics within the same framework. The Architecture convergence enhances data reliability, scalability, and analytical consistency and satisfies regulatory and compliance requirements in the highly regulated healthcare environments. Integrated data structures offer real-time insight development and proactive reactions to real-time monitored operational circumstances in real-time care settings such as emergency services, inpatient units, telehealth systems, and post-discharge follow-up procedures. In addition to performance benefits, lakehouse-centered architecture improves the functionality of governance, data pipelines, and the integration of patient experience analytics and organizational decision-making. In the meantime, the existing problems related to data quality, interpretability, interoperability, and scalable compliance are the issues that should be considered further. Overall, the paper has discovered a robust, scalable foundation of lakehouse-based data engineering to facilitate real-time patient experience health and achieve a shift from fragmented and delayed measures to more responsive, patient-centered, and data-driven healthcare systems based on continuous, integrated, and compliant analytics.

References

[1] Committee on Quality of Health Care in America. (2001). *Crossing the quality chasm: a new health system for the 21st century*. National Academies Press.

[2] Quigley, D., Qureshi, N., Rybowski, L., Shaller, D., Edgman-Levitan, S., Cleary, P. D., ... & Hays, R. D. (2022). Summary of the 2020 AHRQ research meeting on 'advancing methods of implementing and evaluating patient experience improvement using consumer assessment of healthcare providers and systems (CAHPS®) surveys'. *Expert review of pharmacoeconomics & outcomes research*, *22*(6), 883-890.

[3] Aminabee, S. (2024). The future of healthcare and patient-centric care: Digital innovations, trends, and predictions. In *Emerging technologies for health literacy and medical practice* (pp. 240-262). IGI Global Scientific Publishing.

[4] Ahmed, F., Burt, J., & Roland, M. (2014). Measuring patient experience: concepts and methods. *The Patient-Patient-Centered Outcomes Research*, *7*(3), 235-241.

[5] Inmon, W. H. (2005). *Building the data warehouse*. John wiley & sons.

[6] Armbrust, M., Ghodsi, A., Xin, R., & Zaharia, M. (2021, January). Lakehouse: a new generation of open platforms that unify data warehousing and advanced analytics. In *Proceedings of CIDR* (Vol. 8, p. 28).

[7] Hai, R., Koutras, C., Quix, C., & Jarke, M. (2023). Data lakes: A survey of functions and systems. *IEEE Transactions on Knowledge and Data Engineering*, *35*(12), 12571-12590.

[8] Moore, W., & Frye, S. (2019). Review of HIPAA, part 1: history, protected health information, and privacy and security rules. *Journal of Nuclear Medicine Technology*, *47*(4), 269-272.

[9] Rajkomar, A., Dean, J., & Kohane, I. (2019). Machine learning in medicine. *New England Journal of Medicine*, *380*(14), 1347-1358.

[10] Morgan, M. W., Salzman, J. G., LeFevere, R. C., Thomas, A. J., & Isenberger, K. M. (2015). Demographic, operational, and healthcare utilization factors associated with emergency department patient satisfaction. *Western Journal of Emergency Medicine*, *16*(4), 516.

[11] Akidau, T., Bradshaw, R., Chambers, C., Chernyak, S., Fernández-Moctezuma, R. J., Lax, R., ... & Whittle, S. (2015). The dataflow model: a practical approach to balancing correctness, latency, and cost in massive-scale, unbounded, out-of-order data processing. *Proceedings of the VLDB Endowment*, *8*(12), 1792-1803.

[12] Overmann, K. M., Wu, D. T., Xu, C. T., Bindhu, S. S., & Barrick, L. (2021). Real-time locating systems to improve healthcare delivery: A systematic review. *Journal of the American Medical Informatics Association*, *28*(6), 1308-1317.

[13] Manary, M. P., Boulding, W., Staelin, R., & Glickman, S. W. (2013). The patient experience and health outcomes. *New England Journal of Medicine*, *368*(3), 201-203.

[14] Sitzia, J., & Wood, N. (1997). Patient satisfaction: a review of issues and concepts. *Social science & medicine*, *45*(12), 1829-1843.

[15] Inmon, W. H., Strauss, D., & Neushloss, G. (2010). *DW 2.0: The architecture for the next generation of data warehousing*. Elsevier.

[16] Wu, C., Buyya, R., & Ramamohanarao, K. (2016). Big data analytics, machine learning cloud computing. *arXiv preprint arXiv:1601.03115*.

[17] Hussain, F., Nauman, M., Alghuried, A., Alhudhaif, A., & Akhtar, N. (2023). Leveraging big data analytics for enhanced clinical decision-making in healthcare. *IEEE Access*, *11*, 127817-127836.

[18] Mazumdar, D., Hughes, J., & Onofre, J. B. (2023). The data lakehouse: Data warehousing and more. *arXiv preprint arXiv:2310.08697*.

[19] Kaoudi, Z., & Quiané-Ruiz, J. A. (2022). Unified data analytics: state-of-the-art and open problems. *Proceedings of the VLDB Endowment*, *15*(12), 3778-3781.

[20] Batko, K., & Ślęzak, A. (2022). The use of Big Data Analytics in healthcare. *Journal of Big Data*, *9*(1), 3.

[21] Akidau, T., Chernyak, S., & Lax, R. (2018). *Streaming systems: the what, where, when, and how of large-scale data processing*. " O'Reilly Media, Inc."

[22] Pingos, M. (2025). *Modern Data Storage Architectures for Managing Big Data: The Role of Semantically Enrichment Mechanisms in Data Management and Security* (Doctoral dissertation).

[23] Bayne, L. E. (2018). Big data in neonatal health care: big reach, big reward?. *Critical Care Nursing Clinics*, *30*(4), 481-497.

[24] Greaves, F., Ramirez-Cano, D., Millett, C., Darzi, A., & Donaldson, L. (2013). Harnessing the cloud of patient experience: using social media to detect poor quality healthcare. *BMJ quality & safety*, *22*(3), 251-255.

[25] Machireddy, J. R. (2023). Optimizing healthcare resource allocation for operational efficiency and cost reduction using real-time analytics. *Nuvern Applied Science Reviews*, *7*(3), 12-33.

[26] Khanbhai, M., Flott, K., Darzi, A., & Mayer, E. (2019). Evaluating digital maturity and patient acceptability of real-time patient experience feedback systems: systematic review. *Journal of Medical Internet Research*, *21*(1), e9076.

[27] Tanaka, A., & Mahlangu, I. G. (2025). Intelligent Event-Driven Architectures for Real-Time Enterprise Decision-Making.

[28] Yang, X., Qi, X., & Zhou, X. (2023). Deep learning technologies for time series anomaly detection in healthcare: A review. *IEEE Access*, *11*, 117788-117799.

[29] Alsabe, A. A. A., & Alsharif, R. A. B. (2025). Digital Transformation in Business Administration: Analyzing the Impact of Emerging Technologies on Organizational Performance. *Middle East Journal of Pure and Applied Sciences (MEJPAS)*, 33-40.

[30] Analytics, M. (2016). The age of analytics: competing in a data-driven world. *McKinsey Global Institute Research*.

[31] Woo, J., Shin, S. J., Seo, W., & Meilanitasari, P. (2018). Developing a big data analytics platform for manufacturing systems: architecture, method, and implementation. *The International Journal of Advanced Manufacturing Technology*, *99*(9), 2193-2217.

[32] Buyya, R., Ramamohanarao, K., Leckie, C., Calheiros, R. N., Dastjerdi, A. V., & Versteeg, S. (2015, December). Big data analytics-enhanced cloud computing: Challenges, architectural elements, and future directions. In *2015, IEEE 21st International Conference on Parallel and Distributed Systems (ICPADS)* (pp. 75-84). IEEE.

[33] Zaharia, M., Xin, R. S., Wendell, P., Das, T., Armbrust, M., Dave, A., ... & Stoica, I. (2016). Apache Spark: a unified engine for big data processing. *Communications of the ACM*, *59*(11), 56-65.

[34] Sculley, D., Holt, G., Golovin, D., Davydov, E., Phillips, T., Ebner, D., ... & Dennison, D. (2015). Hidden technical debt in machine learning systems. *Advances in neural information processing systems*, *28*.

[35] Khatri, V., & Brown, C. V. (2010). Designing data governance. *Communications of the ACM*, *53*(1), 148-152.

[36] Ulloa-Zamora, M., Barría-Huidobro, C., Sánchez-Rubio, M., & Galeazzi, L. (2025). Integral Security Pillars for Medical Devices: A Comprehensive Analysis. *Applied Sciences*, *15*(12), 6634.

[37] Benfeldt Nielsen, O. (2017). A comprehensive review of data governance literature.

[38] Király, P. (2019). *Measuring metadata quality* (Doctoral dissertation, Georg-August-Universität Göttingen).

[39] Lee, C., Luo, Z., Ngiam, K. Y., Zhang, M., Zheng, K., Chen, G., ... & Yip, W. L. J. (2017). Big healthcare data analytics: Challenges and applications. *Handbook of large-scale distributed computing in smart healthcare*, 11-41.

[40] Topol, E. (2019). *Deep medicine: how artificial intelligence can make healthcare human again*. Hachette UK.

[41] Rumbold, J. M. M., & Pierscionek, B. (2017). The effect of the General Data Protection Regulation on medical research. *Journal of Medical Internet Research*, *19*(2), e47.

[42] Heath, J. (2013, June). A privacy framework for secondary use of medical data. In *2013, IEEE International Symposium on Technology and Society (ISTAS): Social Implications of Wearable Computing and Augmented Reality in Everyday Life* (pp. 174-179). IEEE.