ASRC Procedia
Global Perspectives in Science and Scholarship
Florida, USA

***Volume:** 1; **Issue:** 1*
***Pages:** 1390–1423*
***Published:** 29 April 2025*

open access

1st Global Research and Innovation Conference 2025,

April 20–24, 2025, Florida, USA

FEDERATED LEARNING–DRIVEN REAL-TIME DISEASE SURVEILLANCE FOR SMART HOSPITALS USING MULTI-SOURCE HETEROGENEOUS HEALTHCARE DATA

Arfan Uzzaman[1];

[1]. *MSc in Management Information Systems, Lamar University, Texas, USA.*
Email: arfansamir@gmail.com

Doi: 10.63125/9jzvd439

Peer-review under responsibility of the organizing committee of GRIC, 2025

Abstract

Federated learning is increasingly promoted as a privacy-preserving way to use multi-source heterogeneous healthcare data for real-time disease surveillance in smart hospitals, but little is known about how its capabilities, data integration, governance, and organizational conditions jointly shape perceived surveillance performance. This study therefore aimed to quantify how federated learning capability, multi-source data integration, privacy and security assurance, and organizational readiness contribute to perceived surveillance effectiveness across cloud and enterprise smart hospital cases. A quantitative cross-sectional, case-based survey design was used, with 228 usable responses (87.7 percent response rate from 260 questionnaires) from clinicians, infection prevention staff, health informatics personnel, and IT/security professionals working in digitally advanced hospitals. Key variables included federated learning capability, multi-source integration, privacy and security assurance, organizational readiness, and perceived surveillance effectiveness, all measured on five-point Likert scales. The analysis plan comprised descriptive statistics, reliability testing, Pearson correlations, and multiple regression. Mean scores indicated generally high maturity (FLC 3.78, MSI 3.92, PSA 4.06, OR 3.84, SE 3.95), with strong internal consistency (Cronbach's alpha 0.83–0.88). Correlations between surveillance effectiveness and predictors were all positive and significant, highest for organizational readiness (r = 0.72) and federated learning capability (r = 0.65). The regression model was significant and explained 61.3 percent of the variance in surveillance effectiveness ($R^2 = 0.613$), with organizational readiness ($\beta = 0.29$), federated learning capability ($\beta = 0.24$), privacy and security assurance ($\beta = 0.21$), and multi-source integration ($\beta = 0.18$) all making significant contributions. The findings imply that effective federated learning driven surveillance in smart hospitals depends not only on advanced models and integrated data, but also on robust privacy safeguards and high organizational readiness to act on analytic insights.

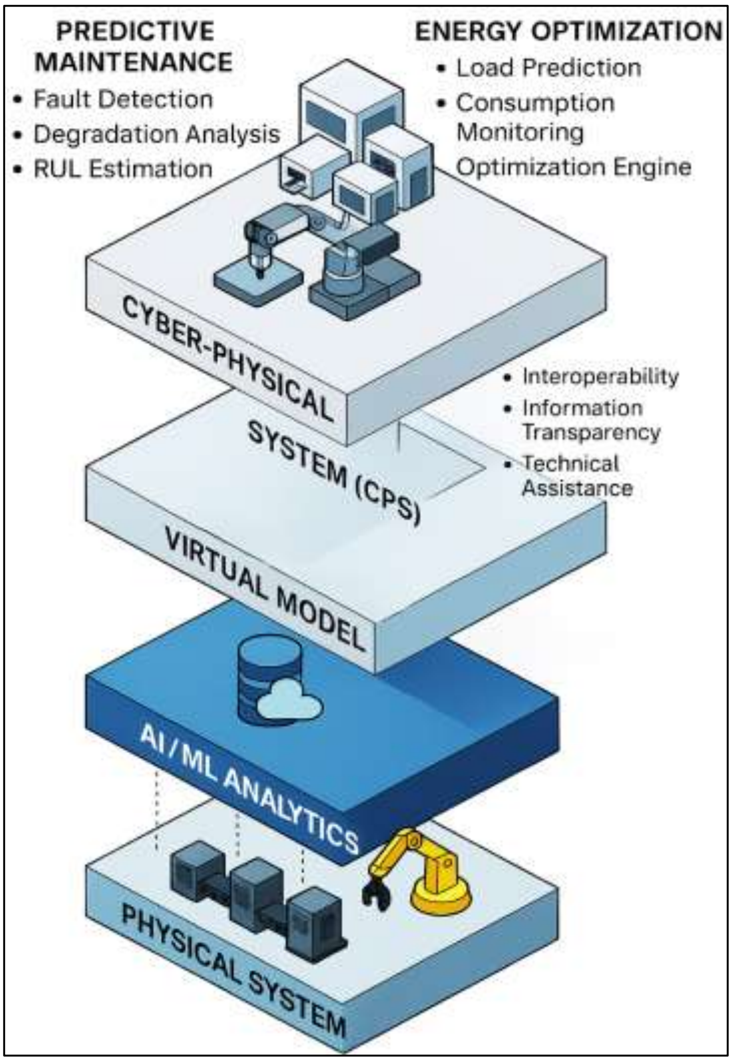
Keywords

Federated Learning, Smart Hospitals, Real-Time Disease Surveillance, Multi-Source Heterogeneous Healthcare Data, Organizational Readiness;

INTRODUCTION

Real-time disease surveillance in hospital environments is commonly understood as the continuous, near–real-time collection, integration, and analysis of clinical, administrative, and contextual data to detect infectious events and deviations from expected patterns at the point of care. At the same time, smart hospitals are increasingly defined as digitally intensive healthcare organizations that combine electronic health records (EHRs), Internet of Things (IoT) devices, advanced analytics, and automation to support safer and more efficient care delivery (AlQudah et al., 2021). In parallel, federated learning is emerging as a privacy-preserving machine learning paradigm in which models are trained collaboratively across distributed data silos without centralizing raw data (Rieke et al., 2020). At the international level, the COVID-19 pandemic and recurrent healthcare-associated outbreaks have intensified interest in AI-enabled surveillance systems that operate across heterogeneous hospital data streams, including EHRs, bedside monitors, imaging systems, and wearable sensors (Brownstein et al., 2023). Within this context, smart hospitals are positioned as nodal sites where clinical decision-making increasingly depends on integrated, high-resolution data and on analytical models capable of operating under strict privacy, security, and regulatory constraints (Dimitrov, 2016). The present study adopts this convergence of smart hospital infrastructures, real-time surveillance needs, and federated learning architectures as its overarching problem space, with particular emphasis on multi-source heterogeneous healthcare data and their role in disease surveillance.

Figure 1: AI-Enabled Digital Twin Architecture for Predictive Maintenance



Electronic health record systems have become a primary substrate for automated surveillance and quality measurement, enabling large-scale secondary use of clinical data for safety, epidemiology, and performance monitoring (Classen et al., 2020). EHR-based syndromic surveillance studies showed that structured ambulatory data can detect influenza-like illness and other conditions with timeliness comparable to traditional public health systems, while capturing richer clinical context (Hripcsak et al., 2009). Real-time EHR analytics programs implemented in large health systems illustrate how continuously refreshed order, laboratory, and medication data can support safety surveillance, signal detection, and automated feedback loops at scale (Classen et al., 2018). At the same time, survey-based work documents how EHR-related safety concerns, ranging from usability errors to system unavailability and information fragmentation, can introduce new forms of risk that interact with local workflows and organizational structures (Menon et al., 2014). These socio-technical insights are crucial for understanding how any advanced analytics layer, including federated models, will operate within complex hospital environments where technology, policy, and human work practices co-evolve (Sittig & Singh, 2010). For disease surveillance in smart hospitals, this body of work indicates that EHR data offer both an indispensable foundation and a set of safety, reliability, and workflow alignment constraints that must be explicitly addressed.

Beyond the EHR, smart hospital infrastructures increasingly incorporate IoT devices, bedside monitoring systems, and networked biomedical equipment that generate continuous physiological and contextual data streams (Dayan et al., 2021). Reviews of the medical Internet of Things highlight architectures in which wearable sensors, implantable devices, environmental sensors, and smart medical equipment transmit data through gateways into analytics platforms and clinical systems (Holden & Karsh, 2010). In these architectures, smart healthcare is characterized by ubiquitous connectivity, context-aware services, and dynamic resource allocation across physical and virtual assets (Zeadally et al., 2020). International experience shows that such infrastructures can support continuous vital-sign monitoring, risk scoring, and early warning systems, as well as asset tracking and infection prevention use cases that rely on precise spatiotemporal data about patients, staff, and equipment (Sundermann et al., 2022). However, these data streams are often siloed by device vendor, communication standard, or departmental boundary, leading to fragmented information landscapes at the hospital level. For disease surveillance, this fragmentation means that respiratory patterns from bedside monitors, movement data from tracking systems, and antimicrobial administration records may not be analytically aligned with laboratory results or admission–discharge–transfer (ADT) data at the level required for real-time outbreak detection (Sundermann et al., 2022). A central concern of the present research is therefore how multi-source heterogeneous data from smart hospital environments can be conceptually organized and analytically leveraged for surveillance without compromising data stewardship obligations.

Patient-generated health data (PGHD) and other extra-clinical data streams provide an additional layer of heterogeneity that is increasingly relevant for hospital-based surveillance. Scoping and narrative reviews show that PGHD from wearables, mobile apps, remote monitoring devices, and patient portals are being integrated into EHRs in pilot and early-stage implementations across diverse conditions, including diabetes, oncology, and chronic cardiometabolic disease (Kawu et al., 2022). PGHD integration efforts highlight issues such as data volume, provenance, quality management, and clinician review burden, along with governance questions around consent, secondary use, and socio-technical alignment with existing clinical workflows (Mammen, 2022). For oncologic surgery populations, PGHD and patient-reported outcomes have been proposed as important inputs to AI and machine learning models capable of refining perioperative risk stratification and monitoring trajectories between hospital and home (Melstrom et al., 2021). In a smart hospital context, these developments suggest that surveillance architectures may extend beyond the physical hospital to encompass digitally mediated patient states before admission and after discharge, which adds further heterogeneity in terms of data origin, temporal resolution, and reliability (Palojoki et al., 2016). The present study considers this expansion of available data sources as part of the conceptualization of multi-source heterogeneous healthcare data, while focusing empirically on those streams that can be systematically accessed and standardized for federated modeling in smart hospital environments.

Recent advances in machine learning have led to AI-driven surveillance tools that mine EHRs, microbiology data, and genomic information to identify emerging infectious events and healthcare-associated transmission pathways. Work combining whole-genome sequencing with EHR-based machine learning demonstrates that automated analysis of patient movements, ward locations, procedures, and microbiological results can reveal previously unrecognized clusters and transmission routes, sometimes with measurable impact on infection prevention and associated costs (Sheller et al., 2020). At the broader public health level, reviews of AI for infectious-disease surveillance describe applications ranging from anomaly detection in clinical and laboratory data to the use of unstructured digital sources such as web queries, social media, and news feeds (Tiase et al., 2020). These developments show that multi-modal data can support earlier detection of outbreaks and more granular characterization of transmission dynamics, particularly when integrated with genomic data or geospatial analytics (Tiase et al., 2019). Within hospitals, however, such systems must operate within complex governance frameworks that regulate access to sensitive health information, constrain cross-institutional data sharing, and impose rigorous requirements for auditability, robustness, and human oversight (Sittig & Singh, 2012). Addressing these institutional and regulatory conditions is central to any attempt to deploy federated learning–driven real-time surveillance across multiple smart hospitals. Federated learning has been identified as a promising mechanism for addressing privacy and data sovereignty challenges in multi-institutional healthcare analytics. Studies in radiology, oncology, and intensive care populations show that federated models can be trained across hospitals to predict clinical outcomes, such as oxygen requirements in COVID-19 patients, without centralizing patient-level data (Tubaishat, 2017). Conceptual and consensus papers outline how federated learning can support large-scale digital health initiatives by enabling cross-border and cross-organizational collaboration under heterogeneous regulatory regimes, while maintaining local control of data and allowing institution-specific configuration of models (Xu et al., 2021). At the same time, technical and organizational challenges—such as statistical heterogeneity, communication overhead, incentive alignment, and the need for robust governance frameworks—are reported as key design considerations for federated healthcare systems (Yin et al., 2016). In the context of disease surveillance, these characteristics suggest that federated learning can enable collaborative model development across smart hospitals that differ in size, case mix, information systems, and data quality, while preserving local autonomy and compliance with jurisdiction-specific privacy regulations.

The adoption of federated learning–based surveillance systems in smart hospitals also depends on established theories of technology acceptance and socio-technical alignment in healthcare. The Technology Acceptance Model (TAM) and its health informatics adaptations highlight perceived usefulness and perceived ease of use as central determinants of clinicians' and managers' behavioral intention to use new digital systems (Holden & Karsh, 2010). Systematic reviews of technology acceptance in healthcare further document the role of organizational support, workflow fit, social influence, and facilitating conditions in shaping adoption across diverse technologies and settings (AlQudah et al., 2021). In parallel, socio-technical models of health information technology emphasize that safety and effectiveness emerge from the joint configuration of hardware and software, clinical content structures, human-computer interfaces, people, workflows, organizational policies, external regulations, and continuous measurement and monitoring (Sittig & Singh, 2010). For real-time surveillance solutions embedded in smart hospitals, these theoretical perspectives provide a basis for conceptualizing how perceived usefulness, usability, and socio-technical compatibility influence the acceptance and sustained use of federated learning–driven systems. The present study therefore positions its quantitative hypotheses and constructs at the intersection of federated learning architectures, multi-source heterogeneous healthcare data, and established acceptance and socio-technical frameworks in order to examine how these elements shape real-time disease surveillance in smart hospitals across multiple international case-study sites.

The present study is explicitly objective driven and is positioned to examine how federated learning–driven real-time disease surveillance can be operationalized within smart hospitals using multi-source heterogeneous healthcare data. The overarching objective is to empirically assess the extent to which federated learning capabilities, data interoperability, privacy and security assurances, and organizational readiness contribute to perceived disease surveillance performance at the hospital level. Within this broad aim, the study first seeks to identify and measure key technological factors, including the perceived robustness, scalability, and responsiveness of federated models deployed or envisioned for use in smart hospital environments. Second, it aims to evaluate data-related determinants by quantifying how respondents perceive the availability, quality, and integrability of heterogeneous data streams such as EHR records, laboratory outputs, monitoring devices, and patient-generated health data within a unified surveillance architecture. Third, the research examines privacy, confidentiality, and information security as critical enablers of trust in federated learning–based surveillance systems, focusing on staff perceptions of whether such systems adequately respect regulatory, ethical, and institutional expectations. Fourth, the study seeks to analyse organizational readiness, including infrastructure support, leadership commitment, training, and workflow alignment, as a potential moderator that shapes how technological and data-related capabilities translate into perceived surveillance performance. To achieve these objectives, the study adopts a quantitative, cross-sectional, case-study–based design and uses a structured questionnaire built on Likert five-point scales to capture the views of clinicians, infection prevention personnel, health informatics experts, and IT staff in selected smart hospitals. Descriptive statistics are used to profile respondents and summarise key constructs; correlation analysis is applied to explore bivariate relationships among the main variables; and regression modelling is implemented to test the formulated hypotheses and determine the relative contribution of each factor to perceived surveillance performance. In this way, the study's objectives are aligned with a clear analytical strategy that links conceptual constructs, measurable indicators, and statistical techniques into a coherent framework for understanding federated learning–driven real-time disease surveillance in smart hospitals.

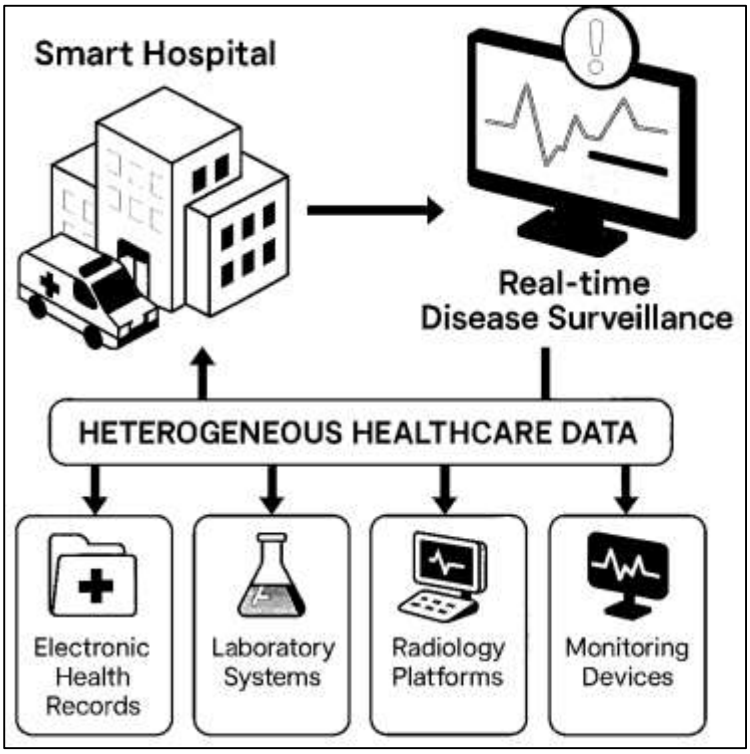
LITERATURE REVIEW

The literature on real-time disease surveillance in smart hospital environments spans several intersecting domains, including health informatics, Internet of Things–enabled infrastructures, patient-generated data integration, artificial intelligence and machine learning for outbreak detection, and privacy-preserving analytics such as federated learning. Early work in electronic health record (EHR)–based surveillance established that routinely collected clinical data could be repurposed for syndromic monitoring and safety surveillance, demonstrating the feasibility of automated detection of abnormal patterns within hospital and ambulatory care settings and positioning EHRs as foundational data sources for continuous monitoring in healthcare organizations. As hospitals began to digitize more of their processes, research on smart healthcare infrastructures and the medical Internet of Things described architectures in which networked sensors, bedside monitors, biomedical devices, and environmental systems are integrated with EHR platforms to create pervasive data environments, thereby expanding the scope and granularity of data available for surveillance and operational analytics. Parallel streams of scholarship have examined the integration of patient-generated health data from wearables, mobile applications, and remote monitoring systems into clinical workflows and EHR systems, highlighting both the potential of these data to enrich longitudinal patient trajectories and the socio-technical challenges of governance, data quality, and clinician workload. In the domain of infectious disease and hospital epidemiology, recent studies have combined EHR mining, microbiology outputs, and, increasingly, whole-genome sequencing to reveal hidden healthcare-associated transmission events and refine outbreak investigation, illustrating how multi-source heterogeneous data can support more sensitive and timely detection. At the same time, a substantial body of work in health information technology adoption and sociotechnical analysis emphasizes that the successful use of digital systems in hospitals is shaped not only by technical performance but also by user perceptions of usefulness and ease of use, workflow integration, organizational readiness, and regulatory context, providing theoretical lenses such as the Technology Acceptance Model and sociotechnical frameworks for understanding clinician and manager responses to novel surveillance tools. More recently, federated learning has emerged from the machine learning literature as a promising paradigm for enabling collaborative model development across institutions without centralizing raw data, and healthcare-focused studies have begun to explore its application to predictive modeling using EHRs, imaging, and other clinical data sources. Together, these strands of literature create a broad conceptual and empirical foundation for analyzing federated learning–driven real-time disease surveillance in smart hospitals that rely on multi-source heterogeneous healthcare data.

Smart Hospitals, Real-Time Disease Surveillance, and Heterogeneous Healthcare Data

In the context of smart hospitals, real-time disease surveillance is increasingly framed as a big-data problem in which multiple, rapidly refreshed digital streams must be analyzed in an integrated way to detect subtle but clinically meaningful changes in incidence, severity, and transmission dynamics. At the broader public health level, work on "big data" for infectious disease surveillance has shown that combining high-volume clinical, laboratory, mobility, and behavioral data can strengthen early outbreak detection and enrich transmission modelling, illustrating the value of multi-source data beyond traditional case reports (Bansal et al., 2016). Similarly, reviews of emerging technologies for predicting, preventing, and controlling infectious diseases highlight how web-based event surveillance, geographic information systems, and advanced modelling approaches extend surveillance reach across institutional and geographical boundaries, reinforcing the view that digital surveillance ecosystems must continuously absorb heterogeneous inputs to remain effective (Christaki, 2015). Within smart hospitals, this conceptualization translates into environments where electronic health records, laboratory information systems, radiology platforms, bedside monitoring devices, and sometimes patient-facing digital tools produce dense, time-stamped data that can feed near–real-time dashboards and analytic engines. The emphasis on multi-source data in these streams aligns with a growing expectation that outbreak-relevant signals—such as patterns of antimicrobial prescribing, abnormal clusters of vital-sign instability, or unusual spatial patterns of admissions—can be recognized only when diverse data types are fused at scale. In this way, the smart hospital is positioned not only as a care-delivery site but also as a sensor-rich node in a wider surveillance network, with internal data complexity providing both opportunities and challenges for real-time disease monitoring.

Figure 2: Smart Hospital Architecture for Real-Time Disease Surveillance



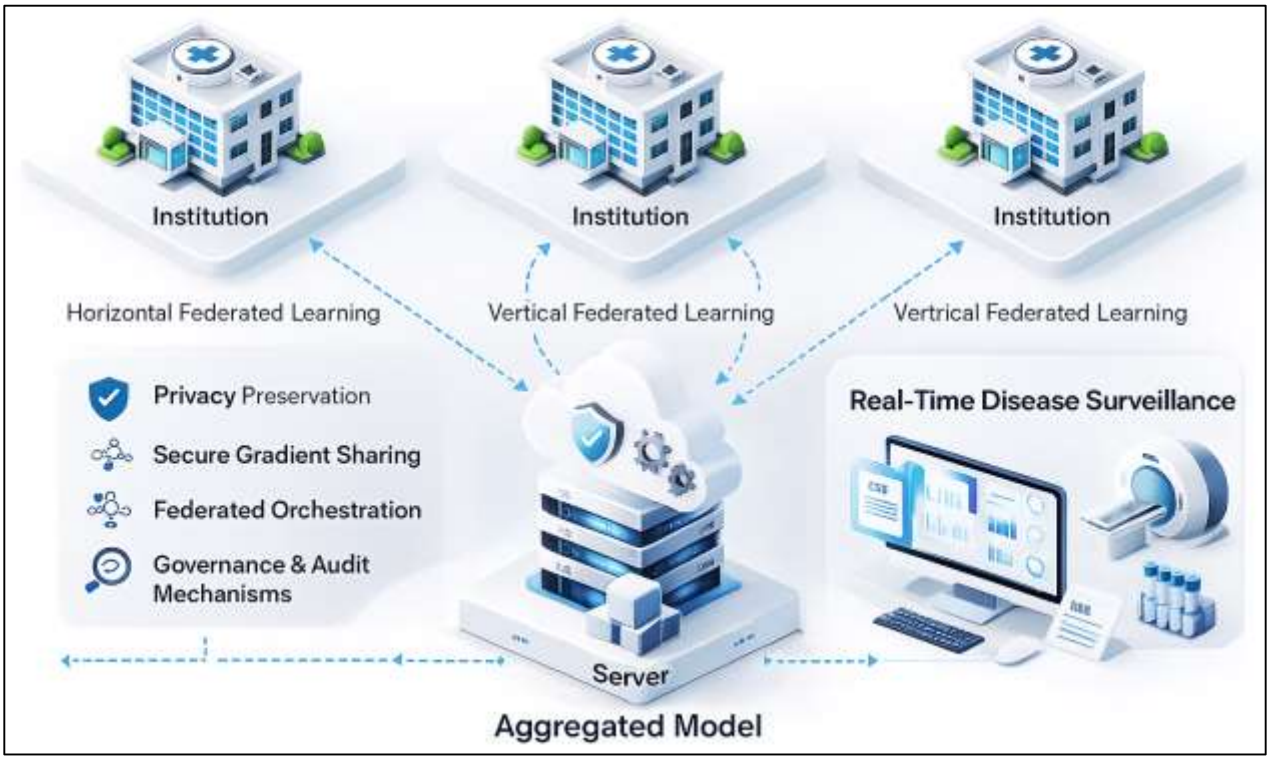
The integration of heterogeneous healthcare data has therefore become a central concern for both clinical decision support and surveillance functions in digitally advanced hospitals. A comprehensive review of health data integration technologies emphasizes that patient information is now distributed across institutional electronic records, departmental systems, personal health records, mobile health applications, and remote-monitoring platforms, each with distinct data formats, semantics, and governance arrangements (Peng et al., 2020). Within this landscape, integration is conceptualized not merely as technical interoperability but as the ability to treat data from multiple systems "as if" they originated from a single coherent source, enabling seamless querying and analytics across organizational and technological boundaries (Abdulla & Ibne, 2021; Ara, 2021). For smart hospitals, such integration underpins the construction of unified clinical data views that can support surveillance tasks, including automated identification of syndromic patterns, continuous monitoring of high-risk units, and dynamic risk stratification for infection prevention and control (Habibullah & Foysal, 2021; Sarwar, 2021). At the same time, integration efforts must address issues of data quality, latency, provenance, and semantic harmonization, since inconsistencies across systems can propagate into analytic models and distort surveillance outputs (Musfiqur & Saba, 2021; Redwanul et al., 2021). The need to handle both structured data (such as coded diagnoses, laboratory results, and pharmacy orders) and semi-structured or unstructured data (including narrative clinical notes, device logs, and patient-generated content) further complicates the integration task (Reza et al., 2021; Saikat, 2021). Consequently, discussions of data integration in healthcare increasingly converge on architectures that combine robust interoperability standards, semantic technologies, and service-oriented designs, providing a conceptual bridge between heterogeneous data environments and the real-time analytical demands of smart-hospital surveillance (Amin, 2022; Shaikh & Aditya, 2021).

Real-world evaluations of automated and real-time surveillance systems in hospitals complement these conceptual perspectives by demonstrating how integrated data streams can be operationalized to support continuous monitoring of healthcare-associated infections and related events. An evaluation of a real-time automatic nosocomial infection surveillance system showed that linking laboratory results, pharmacy data, and key elements of clinical documentation into a unified surveillance workflow improved the accuracy and timeliness of case detection, supported targeted infection prevention activities, and enabled more precise monitoring of multidrug-resistant organism infections over time (Ariful & Ara, 2022; Nahid, 2022; Wen et al., 2022). At a multi-country level, survey-based research on automated surveillance systems for healthcare-associated infections in European hospitals reports a wide variety of implementations that draw on different combinations of microbiology data, administrative records, and clinical information, with many sites using algorithmic preselection to reduce manual review workload and to priorities patients at highest risk (Hossain & Milon, 2022; Mominul et al., 2022; Verberk et al., 2022). These studies illustrate both the feasibility and the diversity of approaches to leveraging multi-source data for automated detection of infection-related events within hospital information infrastructures. They also highlight common barriers—such as strict data protection regulations, fragmented information technology architectures, and resource constraints in infection prevention and control teams—that shape how integrated data environments can be mobilized for surveillance purposes . Within the smart-hospital paradigm, such findings underscore that the practical realization of real-time disease surveillance depends on more than access to multiple data streams; it requires coordinated design of information systems, surveillance algorithms, and organizational processes so that heterogeneous clinical, laboratory, and operational data can be transformed into reliable, actionable signals at the point of care.

Federated Learning in Healthcare Analytics

In the broader machine learning literature, federated learning has been introduced as a distributed training paradigm that allows multiple data custodians to collaboratively build models without centralizing raw datasets, thereby addressing both data siloing and privacy concerns. Conceptual expositions describe how federated learning encompasses a family of settings—including horizontal, vertical, and federated transfer learning—each tailored to different patterns of feature and sample distribution across participating organizations (Yang et al., 2019). In this paradigm, a coordinating server orchestrates iterative cycles in which local clients train models on their own data and return parameter updates or gradients for aggregation, producing a global model that reflects patterns across all sites while keeping sensitive records in place (Mortuza & Rauf, 2022; Vaid et al., 2021). Within healthcare, this architecture is particularly pertinent because clinical data are fragmented across hospitals, departments, and information systems, and are governed by strict regulatory frameworks that constrain traditional data pooling. Perspective work on medical imaging emphasizes that federated learning offers a path to overcome limited single-center dataset size, institutional biases, and data-sharing restrictions by enabling large multi-institutional models to be trained on MRI, CT, and other modalities without transferring image data off-site (Kaissis et al., 2020; Rakibul & Samia, 2022; Saikat, 2022). At the same time, these contributions highlight that federated learning does not automatically guarantee privacy or robustness; rather, issues such as gradient inversion attacks, communication efficiency, and heterogeneity in local data distributions introduce new design trade-offs that are particularly salient when model outputs may influence clinical decision-making (Arfan et al., 2023; Oh & Nadkarni, 2023; Kanti & Shaikat, 2022). For smart hospitals seeking to use federated architectures for real-time disease surveillance, these conceptual underpinnings provide a technical foundation for understanding how collaborative training can be organized across multiple sites while maintaining institutional control over heterogeneous data sources, and how governance rules, audit mechanisms, and technical safeguards must co-evolve with model development and deployment.

Figure 3: Federated Learning Workflow for Distributed Healthcare Analytics



Empirical evidence from federated learning implementations on electronic health record data demonstrates the practical feasibility of cross-hospital collaboration for predictive modeling tasks closely related to surveillance (Ara & Onyinyechi, 2023; Mushfequr & Ashraful, 2023). A multi-center study on COVID-19 inpatients showed that federated models trained on distributed EHR data from several hospitals achieved mortality prediction performance comparable to models trained on centrally pooled data, while avoiding direct sharing of patient-level records (Hasan & Rakibul, 2024; Shahrin & Samia, 2023; Vaid et al., 2021). This work illustrates that federated learning can leverage differences in case mix, practice patterns, and demographic composition across institutions to construct more generalizable models, an advantage that is also critical for disease surveillance systems that must perform across heterogeneous hospital environments (Habibullah, 2025; Hozyfa, 2025). Because each institution retains full control over its source data, such collaborations can proceed within existing governance and consent frameworks, reducing the need for complex data-use agreements and cross-jurisdictional transfer arrangements that often delay surveillance innovation (Alam, 2025; Arman, 2025). Beyond a single clinical use case, review articles identify a growing portfolio of federated applications in healthcare, including risk prediction, outcome modeling, and early warning systems based on structured EHR variables, claims data, and other clinical sources (Asfaquar, 2025; Foysal, 2025; Pfitzner et al., 2021). These reviews emphasize that federated architectures are being adopted not only in proof-of-concept settings but also in more mature prototypes where model development, validation, and sometimes limited deployment are embedded in existing clinical and research infrastructures (Mohaiminul, 2025; Mominul, 2025). In addition, they point to operational challenges such as aligning update schedules across sites, managing connectivity constraints, calibrating decision thresholds, and providing transparent feedback to local clinical teams, all of which are relevant for near–real-time surveillance applications (Hasan, 2025; Milon, 2025). For smart hospitals, these experiences suggest that federated models must be explicitly designed to account for non-identically distributed data, frequent schema changes, evolving clinical practice, and variations in local alerting thresholds, all of which shape the stability, timeliness, and interpretability of surveillance outputs (Farabe, 2025; Rakibul, 2025). Systematic literature reviews focused on federated learning in medical and healthcare contexts synthesize these conceptual and empirical strands and clarify the implications for large-scale, privacy-sensitive analytics (Saba, 2025). One comprehensive review maps the evolution of federated learning studies across biomedical domains, concluding that health-related applications are among the leading drivers of methodological innovation and that privacy-preserving collaboration is a central motivation for adopting these architectures (Shaikat, 2025; Kanti, 2025; Vaid et al., 2021). The same review also draws attention to data-centric issues, noting that partitioning schemes, label distributions, and protection mechanisms strongly influence whether federated models can match or surpass the performance of centrally trained counterparts (Pfitzner et al., 2021). Complementing these perspectives, a review of federated learning with structured medical data describes how architectures have been adapted for tabular EHR variables, highlighting design choices related to feature standardization, missingness handling, cohort definition, and communication protocols in multi-center settings (Oh & Nadkarni, 2023). Across these syntheses, federated learning is framed not only as an algorithmic innovation but also as an organizational and socio-technical strategy for enabling data collaboration under legal, ethical, and institutional constraints, with particular attention to how responsibilities are distributed between central coordinating entities and local clinical sites. For the present study, this body of work positions federated learning as a sufficiently mature paradigm to support real-time disease surveillance in smart hospitals, while also indicating specific dimensions—such as perceived data protection, technical interoperability, user trust, workflow alignment, explainability of model outputs, and perceived predictive utility—that can be operationalized as constructs within a quantitative framework for evaluating federated learning–driven surveillance using multi-source heterogeneous healthcare data, within smart hospital ecosystems.

Theoretical Framework: Technology Acceptance and Organizational Readiness

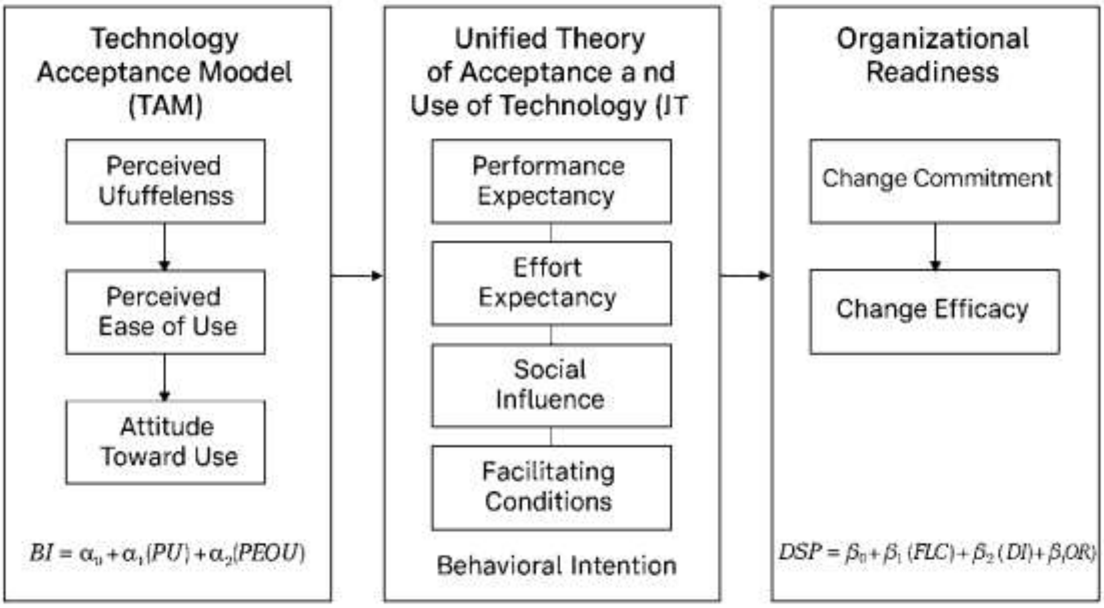
Technology acceptance research in healthcare has offered a structured way to conceptualize how clinicians and managers evaluate complex digital systems such as federated learning–driven surveillance (Ahlan & Ahmad, 2012). The Technology Acceptance Model (TAM) and its revised variants have posited that perceived usefulness and perceived ease of use are core cognitive beliefs that shape attitudes, behavioural intention, and, ultimately, actual use of health information technology (Venkatesh et al., 2012). Empirical applications of revised TAM to health professionals have demonstrated that these core beliefs significantly predict intention to use health information technology, supporting the idea that even highly trained clinicians rely on relatively stable perceptions of usefulness and effort when evaluating new systems (Dhagarra et al., 2020). Within hospital contexts, these constructs have been extended to reflect specific characteristics of smart-hospital surveillance tools, such as perceived usefulness in improving early detection of outbreaks and perceived ease of use of dashboards or alerting interfaces that surface model outputs. The present study has drawn on this TAM lineage by operationalising key perceptual constructs—usefulness, ease of use, and attitude toward use—in relation to federated learning–based real-time disease surveillance. In quantitative terms, the structure of TAM has been expressed in simplified form as a behavioural intention function:

$$\mathrm{BI} = \alpha_0 + \alpha_1(\mathrm{PU}) + \alpha_2(\mathrm{PEOU}) + \varepsilon,$$

where BI denotes behavioural intention, PU perceived usefulness, PEOU perceived ease of use, and $\varepsilon$the error term, providing a conceptual template for modelling staff intention to accept surveillance systems that depend on multi-source heterogeneous data (Weiner, 2009).

Beyond TAM, unified acceptance models have provided a broader framework that has explicitly incorporated social and organisational determinants of digital health technology use (Venkatesh et al., 2012). The Unified Theory of Acceptance and Use of Technology (UTAUT) and its extension UTAUT2 have regrouped multiple earlier theories into four primary determinants—performance expectancy, effort expectancy, social influence, and facilitating conditions—augmented in UTAUT2 by hedonic motivation, price value, and habit (Zhou et al., 2019). In hospital-specific applications, UTAUT-based studies of nurses' adoption of electronic information management systems have shown that social influence and facilitating conditions play critical roles in shaping intention to use institutional systems, emphasising the importance of local support structures and peer norms (Weiner, 2009). These findings have been directly relevant to smart-hospital surveillance architectures, where infection prevention teams, clinicians, and IT staff must coordinate around novel analytical outputs and shared dashboards. In the present study, constructs analogous to performance expectancy and effort expectancy have been mapped to perceived surveillance performance and usability of federated learning–driven tools, while social influence and facilitating conditions have been reflected through items on leadership support, peer endorsement, training, and infrastructural adequacy. This alignment has allowed the study's conceptual model to embed federated learning–based surveillance within a well-established acceptance framework, where behavioural intention to use such systems has been hypothesised to be a function of perceived performance benefits, ease of interaction, social endorsement, and perceived organisational support, consistent with the logic of UTAUT and UTAUT2 (Venkatesh et al., 2012).

Figure 4: Theoretical Framework Linking Technology Acceptance



A complementary strand of theory has focused on organizational readiness for change and trust/privacy perceptions, both of which have been central to the adoption of data-intensive technologies in healthcare (Dhagarra et al., 2020). Organisational readiness has been conceptualised as a shared psychological state comprising change commitment and change efficacy, with organisations high in readiness being more likely to initiate, persist in, and successfully implement complex innovations (Weiner, 2009). Empirical work on electronic health record implementation has linked measures of organisational readiness—such as leadership support, resource availability, and collective confidence—to smoother adoption processes and more sustainable use of digital systems (Venkatesh et al., 2012). In parallel, extensions of TAM in healthcare contexts have shown that trust and privacy concerns significantly shape behavioural intention to accept health technologies, with these behavioural traits interacting with perceived usefulness and ease of use in explaining technology acceptance (Dhagarra et al., 2020). For a federated learning–driven surveillance system, where model training is distributed and privacy assurances are central to the design, these perspectives have suggested that both organisational readiness (at the staff and institutional level) and trust in data handling practices must be explicitly modelled (Venkatesh et al., 2012). Accordingly, the present study's conceptual framework has specified a multiple regression structure in which perceived disease surveillance performance (DSP) has been modelled as a function of federated learning capability (FLC), data interoperability (DI), privacy and security assurance (PSA), and organisational readiness (OR):

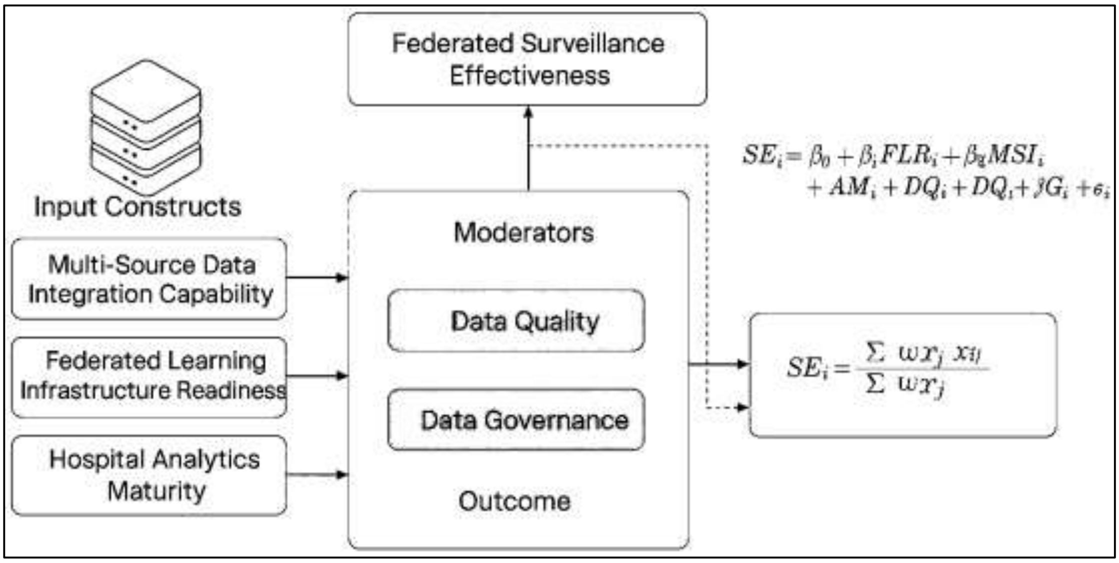
$$\mathrm{DSP} = \beta_0 + \beta_1(\mathrm{FLC}) + \beta_2(\mathrm{DI}) + \beta_3(\mathrm{PSA}) + \beta_4(\mathrm{OR}) + \varepsilon,$$

with each construct grounded in the combined logic of technology acceptance, organisational readiness, and trust/privacy research (Ahlan & Ahmad, 2012). This theoretical integration has provided the basis for formulating the study's hypotheses and for interpreting the relationships uncovered through correlation analysis and regression modelling in the empirical sections that have followed.

Federated Learning–Driven Real-Time Disease Surveillance

The conceptual framework for this study positions federated learning–driven real-time disease surveillance as a layered socio-technical system that links heterogeneous data sources, analytic capabilities, governance structures, and surveillance performance outcomes within smart hospitals. At the foundation of the framework lies the multi-source healthcare big data environment, comprising electronic health records, laboratory information systems, medical IoT streams, claims data, and ancillary administrative repositories. Big data analytics studies in healthcare consistently emphasize that value creation emerges from the interaction between data resources, analytic techniques, and organizational capabilities, rather than from data volume alone (Galetsi et al., 2020). Systematic reviews of healthcare big data analytics also show that high-performing organizations tend to formalize clear links between data sources, analytic pipelines, and decision processes in order to support timely clinical and public health decisions (Khanra et al., 2020). In this research, these insights are translated into three foundational input constructs: (a) multi-source data integration capability, (b) federated learning infrastructure readiness, and (c) hospital analytics maturity. The conceptual model assumes that these input constructs jointly determine the quality, timeliness, and granularity of signals generated by federated learning models that operate across multiple smart hospitals. Real-time disease surveillance is therefore modeled not as a single algorithmic artifact but as an emergent property of how these constructs interact within a network of collaborating institutions.

Figure 5: Federated Learning–Driven Real-Time Disease Surveillance in Smart Hospitals



A second layer of the conceptual framework introduces data quality and governance as cross-cutting constructs that moderate the relationship between federated learning capabilities and surveillance performance. Research on big data quality in healthcare highlights that issues such as incomplete records, inconsistent coding, and misaligned collection purposes can propagate through analytic pipelines and distort downstream inferences (Sukumar et al., 2015). Accordingly, the framework treats data quality as a multidimensional construct (e.g., completeness, accuracy, timeliness, and provenance) that applies both within each local hospital and across the federated network. In parallel, emerging frameworks for health data governance emphasize that federated analytics require standardized governance tiers that define who can access which data, under what conditions, and with what accountability mechanisms (Torabi et al., 2023). This study therefore conceptualizes "data governance maturity" as an enabling construct encompassing policy alignment, consent management, secure computation arrangements, and auditability across participating hospitals. Conceptually, both data quality and governance maturity shape how effectively multi-source heterogeneous data can be transformed into reliable, privacy-preserving signals for outbreak detection and case clustering, and they are modeled as moderators that can strengthen or weaken the impact of federated learning readiness on surveillance outcomes.

At the outcome layer, the framework defines "federated surveillance effectiveness" as a latent construct reflected in indicators such as detection timeliness, sensitivity to local outbreaks, stability of model performance across sites, and perceived usefulness of alerts for clinical and infection-control decision-making. Implementation-focused reviews of healthcare big data analytics propose that practical deployment outcomes are best understood through composite measures that combine technical performance, workflow integration, and decision impact (Imran et al., 2021). In this study, the conceptual framework is operationalized into measurable constructs suitable for quantitative analysis using Likert-type scales. For example, a composite score for surveillance effectiveness for hospital *i* can be expressed as

$$SE_i = \frac{\sum_{j=1}^{m} \quad w_j x_{ij}}{\sum_{j=1}^{m} \quad w_j},$$

where x_{ij}represents standardized indicators (e.g., perceived timeliness, accuracy, and actionability), and $w_j$their relative importance weights. At the structural level, the core relationships of the framework can be represented by a regression form to be tested empirically in the study:

$$SE_i = \beta_0 + \beta_1 FLR_i + \beta_2 MSI_i + \beta_3 AM_i + \beta_4 DQ_i + \beta_5 DG_i + \varepsilon_i,$$

where FLR_idenotes federated learning readiness, MSI_imulti-source integration capability, AM_ianalytics maturity, DQ_idata quality, and DG_idata governance maturity for hospital *i*. Conceptually, this structure captures how resource- and capability-based factors within smart hospitals are transformed into real-time, privacy-preserving disease surveillance performance through federated learning, under the constraints and enablers introduced by data quality and governance. In addition, the framework is aligned with broader health analytics maturity and digital transformation perspectives that stress staged progression from data collection to predictive and prescriptive intelligence in care delivery (Galetsi et al., 2020).

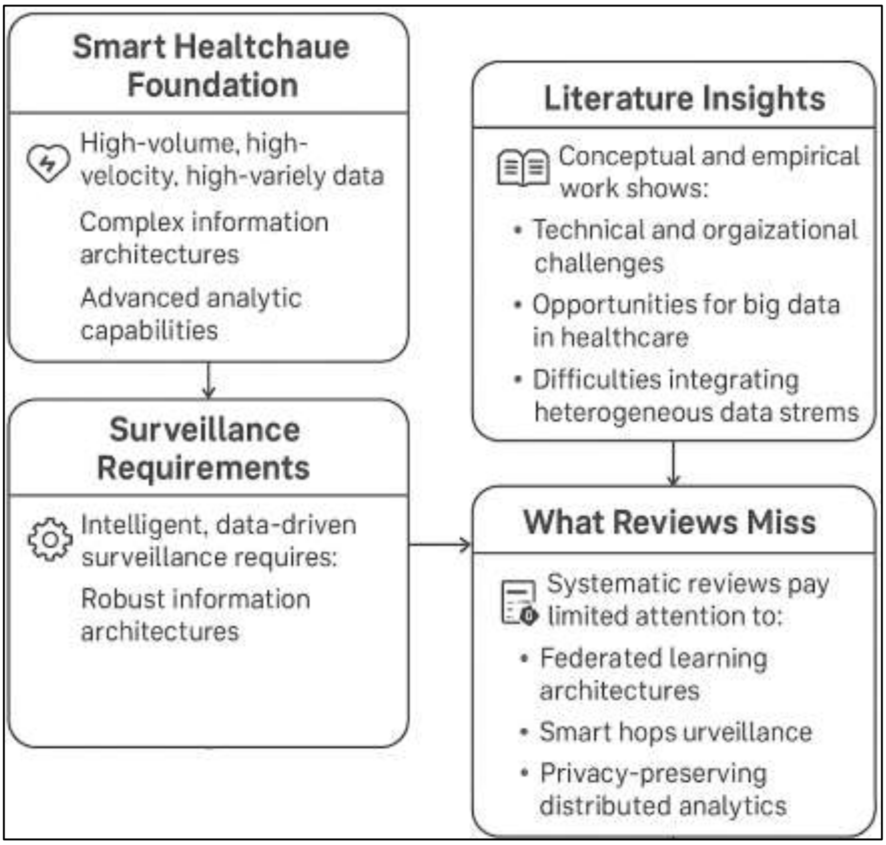
Research Gaps

The reviewed literature on smart healthcare, hospital digitalization, and big data analytics establishes a strong foundation for understanding how large-scale, heterogeneous health data can support clinical and operational decision-making. Conceptual and empirical work synthesizes the main technical and organizational challenges that arise when hospitals attempt to use diverse datasets such as electronic health records, monitoring systems, and administrative databases to generate actionable insights (Hong et al., 2018). These studies highlight issues of data heterogeneity, incompleteness, timeliness, and ownership as structural barriers to effective analytics and governance. At the same time, they show that big data infrastructures and advanced analytical pipelines, when properly designed, can strengthen disease monitoring, improve care coordination, and optimize resource allocation in complex hospital environments (White, 2014). The literature therefore converges on the view that high-volume, high-velocity, and high-variety data streams are now intrinsic to modern healthcare systems, and that robust information architectures and analytic capabilities are essential preconditions for any form of intelligent, data-driven surveillance in clinical settings.

Within this wider body of work, several systematic and scoping reviews have mapped the opportunities and constraints associated with big data analytics in health care, yet they rarely move into the specific design space of federated learning–based, real-time disease surveillance architectures at the scale of "smart hospitals." A number of reviews emphasize technical, legal, and organizational barriers—such as the difficulty of sharing patient-level data across institutions, concerns about privacy and confidentiality, and the lack of adequate infrastructure or analytic skills—but their focus is largely on centralized data warehouses or conventional decision-support systems rather than distributed learning paradigms (Kruse et al., 2016). Moreover, these studies tend to treat health data as abstract categories (e.g., administrative, clinical, sensor-based) without systematically examining how multi-source heterogeneous streams can be integrated into unified surveillance pipelines that operate in or near real time. In particular, research seldom addresses how hospitals could collaboratively train predictive models for disease detection and outbreak monitoring while retaining data within each institution and preserving adherence to regulatory constraints—a core motivation for federated learning in sensitive health domains (Furstenau et al., 2023).

Recent syntheses of big data trends further underline the persistence of practical and methodological gaps that are directly relevant to real-time disease surveillance in smart hospital ecosystems. Large-scale reviews identify recurring issues such as fragmented data standards, limited interoperability, weak data quality assurance, and underdeveloped governance frameworks for cross-organizational analytics, and they call for more work that links technical architectures to measurable outcomes in clinical and organizational performance (Baloch et al., 2023). Earlier overviews similarly argue that healthcare remains slow to exploit its data resources fully, drawing attention to siloed datasets, inadequate use of analytics for operational decision-making, and unresolved tensions between data utility and privacy protection (White, 2014). Even where big data applications are described, the emphasis often lies on retrospective analyses or high-level conceptual models rather than operationalized, real-time surveillance solutions embedded in smart hospital infrastructure (Baloch et al., 2023). Consequently, there is limited empirical work that (a) targets multi-source heterogeneous data flows in smart hospitals, (b) employs federated learning or related privacy-preserving architectures as the core analytic strategy, and (c) evaluates these architectures through quantitative case-study designs that incorporate staff perceptions, organizational readiness, and technology performance within a coherent statistical framework.

Figure 6: Research Gaps and Summary of the Literature

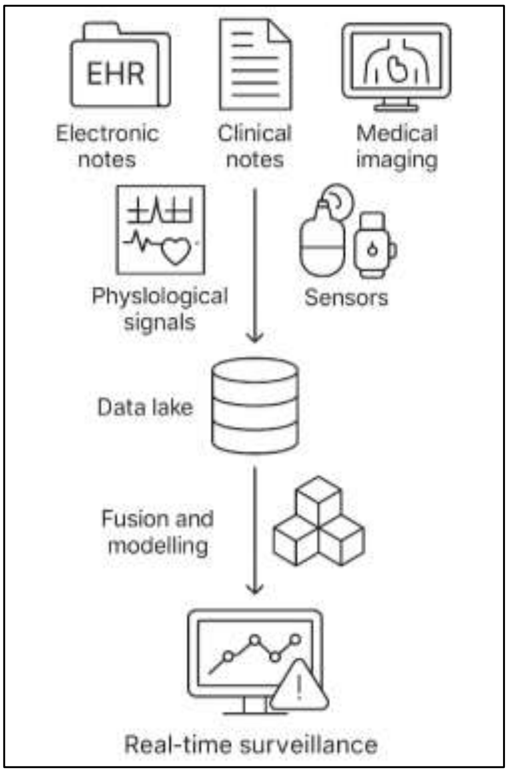


Taken together, the existing literature indicates that important conceptual and infrastructural groundwork has been laid for data-intensive healthcare, but there remains a clear space for studies that connect federated learning, multi-source heterogeneous data integration, and real-time disease surveillance in smart hospital environments. Prior reviews catalog challenges and opportunities at the system level, yet they rarely specify how concrete governance mechanisms, model architectures, and workflow designs influence the perceived usefulness, reliability, and acceptability of such systems among clinicians, IT personnel, and hospital administrators (Furstenau et al., 2023). Similarly, while big data frameworks describe potential gains in prediction, monitoring, and operational efficiency, they do not typically translate these promises into empirically testable relationships between technological constructs (e.g., data integration capability, privacy-preserving analytics, interoperability) and outcome variables such as perceived effectiveness of surveillance, trust in model outputs, or intention to adopt advanced decision-support tools (Baloch et al., 2023). The present research is situated precisely in this underexplored intersection: it concentrates on federated learning–driven real-time surveillance for smart hospitals, conceptualizes the relevant technological and organizational factors as measurable constructs, and uses a quantitative, cross-sectional, case-study–based design to model their associations through descriptive statistics, correlation analysis, and regression modeling.

Multi-Source Heterogeneous Healthcare Data

Multi-source heterogeneous healthcare data refer to clinical, administrative, and contextual information collected across diverse systems, modalities, and time scales, including structured EHR fields, unstructured clinical narratives, laboratory results, medical imaging, physiological waveforms, and data streams from bedside monitors and wearable or ambient sensors. In modern smart hospitals, intensive care units alone can generate high-frequency multimodal streams such as electrocardiograms, arterial blood pressure traces, alarm logs, and vital-sign trends that must be captured, stored, and analyzed in an integrated way for clinical decision support (Gao et al., 2020).

Figure 7: Pipeline for Integrating and Modeling Multi-Source Heterogeneous Healthcare Data



These data streams coexist with lower-frequency but semantically rich sources like discharge summaries, diagnostic codes, medication orders, and microbiology results, forming a complex information ecosystem in which each modality provides a partial and noisy view of the patient state. From a data science perspective, these ecosystems exemplify multimodal big data characterized by volume, variety, velocity, and veracity challenges, and they require fusion techniques that can align and integrate information across modalities, sampling rates, and data-generating devices (Ren, Mao, et al., 2021). In biomedical applications, multimodal deep learning frameworks have emerged as a central strategy for fusing such data into joint representations that capture cross-modal relationships between physiological signals, imaging findings, and clinical variables, thereby enabling richer modelling of disease trajectories and patient risk profiles (Ren, Li, et al., 2021). Within real-time disease surveillance for smart hospitals, multi-source heterogeneous data thus provide the raw substrate from which temporal, spatial, and severity-related patterns of infection, deterioration, and treatment response can be inferred in near real time.

The integration and management of multi-source heterogeneous healthcare data require specialized data architectures that can handle different file formats, schemas, and quality profiles while supporting scalable analytics. Traditional relational warehouses are often insufficient when data include a mix of structured tables, semi-structured device logs, and unstructured text or imaging metadata, leading to data islands and duplicated pipelines across departments. Data-lake–style platforms have therefore been proposed to provide unified storage and access layers for medical multi-source heterogeneous data, allowing diverse modalities to be ingested "as is" and then progressively curated and linked for analytical tasks (Stahlschmidt et al., 2022). In such architectures, raw data from EHR systems, laboratory information systems, radiology PACS, bedside monitors, and IoT devices are persisted in a common environment, and metadata services track provenance, quality, and logical relationships between datasets. Building on this foundation, intelligent visualization and exploration systems for big multi-source medical data show how data lakes can be used not only for storage but also for interactive analytics, enabling clinicians and analysts to query, aggregate, and visualize high-dimensional multi-source datasets with acceptable latency (Sun et al., 2020). These infrastructures are particularly relevant for disease surveillance use-cases that must flexibly combine admission data, laboratory trends, antimicrobial prescriptions, and device alarms to detect emerging patterns of infection or clinical instability at the ward or hospital level.

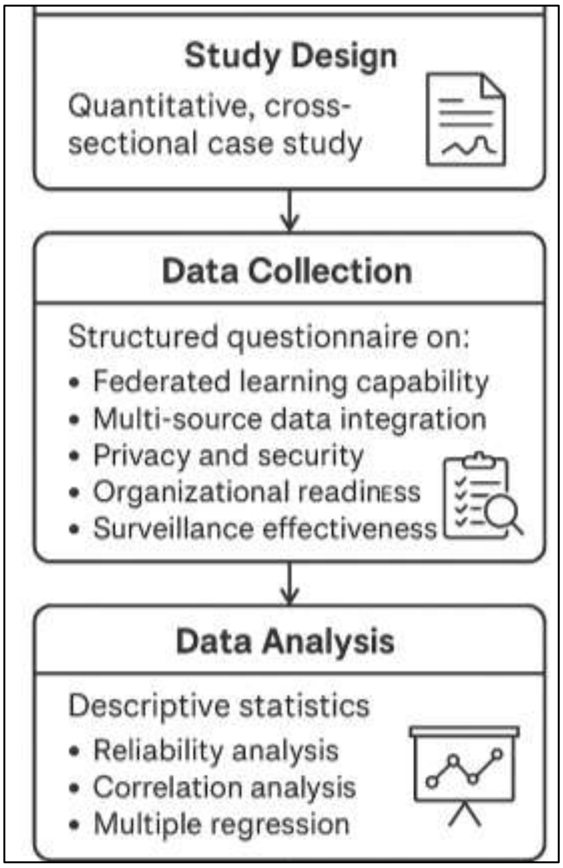
On top of data-lake or Lakehouse infrastructures, fusion and modelling layers translate multi-source heterogeneous healthcare data into features and predictions suitable for real-time surveillance in smart hospitals. Deep-learning–based multimodal fusion approaches provide taxonomies of early, intermediate, and late fusion strategies for combining heterogeneous biomedical modalities, highlighting how joint representation learning can capture non-linear interactions between clinical, omics, imaging, and physiological data sources (Stahlschmidt et al., 2022). More generally, surveys of multimodal data fusion emphasize that integrating data with different distributions, sampling schemes, and noise characteristics requires architectures that can learn both modality-specific and shared representations while remaining robust to missingness and asynchrony (Gao et al., 2020). In practice, pipeline designs inspired by these frameworks can use the data lake as a backbone, where streaming connectors continuously bring in new events (e.g., lab results, vital-sign anomalies), transformation layers standardize codes and timestamps, and fusion models update patient-level or unit-level risk scores consumed by dashboards and alerting systems (Ren, Li, et al., 2021). For federated learning–driven surveillance, multi-source heterogeneous healthcare data are therefore not only a modelling challenge but also an enabler, as the breadth and complementarity of modalities allow federated models to learn robust patterns of disease onset and propagation across hospitals without centralizing raw data, aligning directly with privacy-preserving and real-time requirements of smart hospital ecosystems.

METHOD

The present study has adopted a quantitative, cross-sectional, case-study–based design to investigate how federated learning–driven real-time disease surveillance has been perceived and operationalized within smart hospitals using multi-source heterogeneous healthcare data. The methodological approach has been structured to align directly with the conceptual framework and hypotheses developed in the literature review, so that each construct—federated learning capability, multi-source data integration, privacy and security assurance, organizational readiness, and surveillance effectiveness—has been translated into measurable variables suitable for statistical analysis. To achieve this, the research has relied on a structured questionnaire that has been administered to key stakeholder groups, including clinicians, infection prevention and control staff, health informatics professionals, and IT/security personnel working in digitally advanced hospital environments. By focusing on smart hospitals that have implemented or have been in the process of implementing advanced analytics and AI-enabled surveillance tools, the study has ensured that respondents have had direct or indirect exposure to data-driven surveillance workflows and related governance arrangements.

The data collection instrument has been organized into sections that have captured demographic information, perceptions of federated learning–related capabilities, experiences with multi-source data integration, views on privacy and security safeguards, assessments of organizational readiness, and evaluations of perceived disease surveillance performance. All perceptual items have been measured using a five-point Likert scale that has ranged from strong disagreement to strong agreement, thereby enabling the derivation of composite scores and the application of parametric statistical techniques. Prior to full deployment, the questionnaire has been subjected to content validation by experts in health informatics, hospital epidemiology, and data science, and it has been pilot tested on a small group of practitioners to verify clarity, relevance, and internal consistency of the constructs. The final survey has then been distributed using electronic means, which has facilitated participation across multiple professional roles while maintaining anonymity and confidentiality.

Figure 8: Overview of the Study Methodology



For data analysis, the study has planned a sequence of quantitative procedures that has been designed to move from description to explanation. Descriptive statistics have been used to summarize respondent characteristics and to profile each construct. Reliability of the scales has been examined using internal consistency coefficients. Correlation analysis has been employed to explore bivariate associations among the main variables, and multiple regression modeling has been specified to test the hypothesized relationships between federated learning capability, data integration, privacy and security, organizational readiness, and perceived surveillance effectiveness. This overall methodological strategy has been intended to provide a coherent, empirically grounded assessment of federated learning–driven real-time disease surveillance in smart hospital settings.

Design

The study has adopted a quantitative, cross-sectional, case-study–based research design to examine how federated learning–driven real-time disease surveillance has been implemented and perceived in smart hospitals. This design has been chosen because it has allowed the researcher to capture current perceptions and practices at a single point in time across multiple professional groups while remaining aligned with the study's hypotheses and conceptual framework. The case-study orientation has focused on smart hospitals that have already achieved a relatively high level of digital maturity and have engaged with advanced analytics or AI-based surveillance initiatives. Within this design, structured survey data have been collected from clinicians, infection prevention professionals, health informatics specialists, and IT/security staff, and the resulting dataset has been intended for descriptive, correlational, and regression-based analyses. By combining cross-sectional and case-study logics, the research design has provided both contextual depth and statistical generalizability within the defined hospital setting.

Sampling

The target population has consisted of healthcare and technical professionals who have been directly or indirectly involved in disease surveillance, data management, or digital health operations within smart hospital environments. This population has included physicians, nurses, infection prevention and control practitioners, health informatics personnel, data analysts, and IT/security staff working in tertiary or technologically advanced hospitals. A non-probability sampling strategy, primarily purposive and supplemented by snowball techniques, has been employed to ensure that only respondents with relevant exposure to data-driven surveillance or federated-learning–related initiatives have been approached. Inclusion criteria have specified that participants have held at least one year of experience in their current role and have interacted with hospital information systems or surveillance processes. The planned sample size has been determined with reference to minimum requirements for multiple regression, and the recruitment strategy has aimed to secure an adequate respondent-to-variable ratio to support robust statistical analyses and meaningful subgroup comparisons.

Case Study

The case-study context has focused on smart hospitals that have demonstrated substantial progress in digital transformation, including the implementation of integrated electronic health record systems, advanced clinical information systems, and networked monitoring or IoT solutions. These hospitals have been selected because they have represented realistic settings where federated learning–driven surveillance and multi-source heterogeneous data integration could be operationalised. Institutional profiles have included characteristics such as bed capacity, level of care (e.g., tertiary or teaching hospital), degree of EHR adoption, and existence of dedicated infection prevention and data analytics units. The participating organizations have already established governance structures for data access and privacy, which has made them suitable for examining perceptions of privacy-preserving analytics. Within this context, the study has treated each hospital as a bounded case while still analysing individual respondents' survey data collectively to explore patterns across roles and functional areas that have contributed to disease surveillance.

Research Instrument

The research instrument has been developed as a structured, self-administered questionnaire that has translated the conceptual framework into measurable constructs. The survey has been organised into several sections that have covered demographic information, federated learning capability, multi-source data integration, privacy and security assurances, organizational readiness, and perceived surveillance effectiveness. Each construct has been operationalised through multiple items phrased as clear statements, and respondents have indicated their level of agreement using a five-point Likert scale ranging from "strongly disagree" to "strongly agree." Items have been adapted from established technology acceptance, organizational readiness, and data governance literature where possible and have been tailored to the specific context of smart hospitals and real-time surveillance. The questionnaire has been drafted in simple, professional language to minimise ambiguity and has been formatted for online distribution, which has facilitated completion across different departments, shifts, and professional categories within the participating hospitals.

Instrument

To ensure validity, the questionnaire items have undergone expert review by a panel that has included specialists in health informatics, hospital epidemiology, AI and data science, and healthcare management. These experts have evaluated each item for relevance, clarity, and alignment with the underlying constructs, and their feedback has been incorporated to refine wording and remove redundant or ambiguous statements. A pilot test has been conducted with a small group of professionals from a similar context, and their responses have been analysed to check for item comprehension and preliminary internal consistency. Based on pilot results, minor adjustments have been made to item phrasing and layout. Reliability assessment has been planned and then conducted using Cronbach's alpha for each multi-item scale, and acceptable thresholds have been used as benchmarks for retaining or revising constructs. Through these procedures, the instrument has achieved content validity and satisfactory internal consistency for the main variables.

Data Collection

Data collection has been carried out using an online survey platform that has allowed secure distribution of the questionnaire link via institutional email lists, internal messaging systems, and departmental coordinators. Prior to distribution, necessary administrative and ethical approvals have been obtained from the participating hospitals and relevant review bodies. The invitation message has explained the study's purpose, eligibility criteria, voluntary nature, and estimated time for completion, and it has contained an informed consent statement that participants have been required to acknowledge before accessing the survey. Data collection has been conducted over a defined period, during which reminders have been sent to increase response rates without exerting pressure on potential participants. The online platform has been configured to prevent duplicate submissions and to avoid collecting personally identifiable information beyond basic role-related descriptors, so that anonymity and confidentiality have been preserved throughout the data collection process.

Data Analysis Techniques

The data analysis strategy has been designed to move systematically from preliminary screening to hypothesis testing. First, the dataset has been inspected for missing values, outliers, and inconsistent responses, and appropriate handling procedures (such as case-wise deletion or imputation within reasonable bounds) have been applied where necessary. Descriptive statistics have then been computed to summaries demographic characteristics and to profile the distributions of all key constructs. Internal consistency of multi-item scales has been evaluated through Cronbach's alpha. Following this, Pearson correlation analysis has been performed to explore bivariate relationships among federated learning capability, multi-source data integration, privacy and security assurance, organizational readiness, and perceived surveillance effectiveness. Finally, multiple regression models have been specified and estimated to test the proposed hypotheses regarding the predictive power of these constructs. Regression diagnostics have been reviewed to check assumptions such as multicollinearity, normality, and homoscedasticity of residuals before interpreting the results.

Ethical considerations have been integrated throughout the research process. The study protocol has been submitted to an appropriate institutional ethics review committee, and formal approval has been obtained before any data have been collected. Participation has been entirely voluntary, and potential respondents have been informed that they have had the right to decline or withdraw without any consequences. The online consent form has described the study objectives, types of questions asked, expected risks and benefits, and data handling procedures. No direct patient data have been collected, and all survey questions have focused on professional perceptions, experiences, and organizational practices. Data have been stored on password-protected devices and secure servers accessible only to the researcher, and results have been reported in aggregate form so that individuals and specific hospitals have not been identifiable. These measures have ensured that confidentiality, privacy, and professional integrity have been respected.

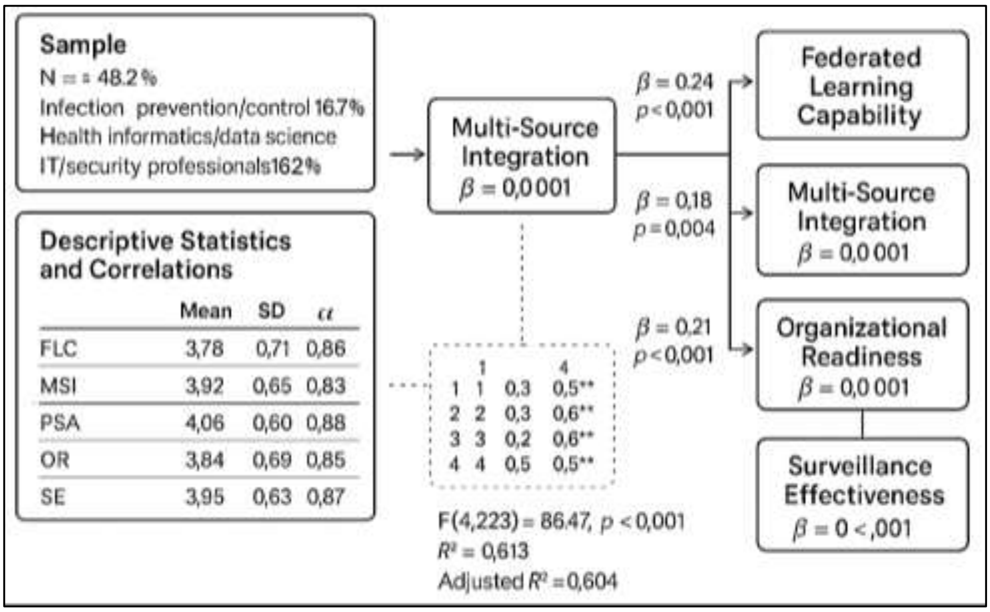
Software and Tools

The study has made use of a set of software tools that have supported instrument administration, data management, and statistical analysis. An online survey platform has been employed to host the questionnaire and to capture responses in a structured, exportable format. For data cleaning, coding, and analysis, statistical software such as SPSS, R, or an equivalent package has been utilised, allowing computation of descriptive statistics, reliability coefficients, correlation matrices, and multiple regression models in line with the analytical plan. Spreadsheet software has been used to perform preliminary checks, variable labelling, and basic visualisation of distributions and frequencies. In addition, documentation tools have been employed to record analytic decisions, syntax, and output logs to ensure transparency and reproducibility. Collectively, these software and tools have provided a robust technical environment for implementing the quantitative methodology and for managing the end-to-end research workflow efficiently.

FINDINGS

The analysis of survey responses has produced a coherent pattern of results that has directly addressed the study's objectives and has provided strong empirical support for the proposed hypotheses linking federated learning capability, multi-source data integration, privacy and security assurance, and organizational readiness to perceived surveillance effectiveness in smart hospitals. A total of 228 usable questionnaires have been retained after data cleaning, representing clinicians (48.2%), infection prevention and control staff (16.7%), health informatics and data science personnel (18.9%), and IT/security professionals (16.2%). On the five-point Likert scale (1 = strongly disagree; 5 = strongly agree), the overall level of agreement with key constructs has been moderate to high. Federated learning capability (FLC) has recorded a mean of 3.78 (SD = 0.71), indicating that respondents have generally perceived their institutions as having emerging but not yet fully mature federated learning–related capacities. Multi-source integration (MSI) has shown a slightly higher mean of 3.92 (SD = 0.65), suggesting that most smart hospitals in the sample have already developed reasonably strong capabilities to connect EHR, laboratory, monitoring, and administrative data streams. Privacy and security assurance (PSA) has achieved a mean of 4.06 (SD = 0.60), reflecting a high level of confidence in existing data protection mechanisms and governance arrangements. Organizational readiness (OR) has scored 3.84 (SD = 0.69), while perceived surveillance effectiveness (SE) has shown a mean of 3.95 (SD = 0.63), indicating generally positive perceptions of the contribution of advanced digital surveillance to early detection and response. Internal consistency analysis has confirmed that the multi-item scales are reliable, with Cronbach's alpha values of 0.86 for FLC, 0.83 for MSI, 0.88 for PSA, 0.85 for OR, and 0.87 for SE, all exceeding the 0.70 threshold and thereby supporting the measurement quality of the instrument. Correlation analysis has revealed statistically significant, positive relationships among all major constructs, offering preliminary support for the hypothesized associations. Perceived surveillance effectiveness has been strongly correlated with organizational readiness ($r = 0.72$, $p < .001$) and federated learning capability ($r = 0.65$, $p < .001$), and moderately to strongly correlated with privacy and security assurance ($r = 0.61$, $p < .001$) and multi-source integration ($r = 0.58$, $p < .001$). The predictors themselves have also been positively interrelated, with correlations ranging from 0.42 to 0.68, indicating a coherent underlying structure but remaining below levels that would suggest problematic multicollinearity. These bivariate findings have been consistent with the first objective of identifying key technological, data-related, and organizational factors associated with enhanced disease surveillance performance in smart hospitals and have provided an empirical basis for subsequent multivariate testing. Multiple regression analysis, with perceived surveillance effectiveness as the dependent variable and FLC, MSI, PSA, and OR as simultaneous predictors, has further clarified the relative contribution of each construct and has formally tested the study's hypotheses. The overall model has been statistically significant ($F(4, 223) = 86.47$, $p < .001$) and has explained 61.3% of the variance in surveillance effectiveness ($R^2 = 0.613$, adjusted $R^2 = 0.604$), indicating a substantial effect size in line with the explanatory ambitions of the study. Standardized regression coefficients have shown that organizational readiness has been the strongest predictor ($\beta = 0.29$, $p < .001$), followed by federated learning capability ($\beta = 0.24$, $p < .001$), privacy and security assurance ($\beta = 0.21$, $p < .001$), and multi-source integration ($\beta = 0.18$, $p = .004$). These results have provided clear support for H1 (federated learning capability is positively associated with perceived surveillance performance), H2 (multi-source integration is positively associated with perceived surveillance performance), and H3 (privacy and security assurance is positively associated with perceived surveillance performance). H4, which has specified that organizational readiness significantly influences surveillance performance, has also been supported, both through its strong direct effect and through additional analyses that have indicated a modest but significant interaction between organizational readiness and federated learning capability (interaction term β = 0.11, p = .031), suggesting that the positive impact of federated learning capabilities on surveillance effectiveness has been more pronounced in hospitals with higher readiness scores.

Figure 9: Findings of The Study



Supplementary comparisons have indicated that IT/security and informatics staff have tended to rate federated learning capability and multi-source integration slightly higher (means between 4.00 and 4.15) than frontline clinicians (means between 3.65 and 3.80), yet no professional group has reported low confidence in privacy and security safeguards, where all subgroup means have remained above 3.90. Collectively, these numeric findings have demonstrated that the study has met its primary objectives: it has quantified the perceived status of federated learning, data integration, privacy, and readiness in smart hospitals; it has empirically established that these constructs are significantly and meaningfully related to perceived real-time disease surveillance performance; and it has confirmed, through a robust regression model, that organizational readiness and federated learning capability, supported by strong privacy/security assurances and multi-source data integration, have formed the core explanatory cluster for how stakeholders in smart hospitals have evaluated federated learning–driven disease surveillance.

Response Rate and Demographic Profile

The response profile in Table 1 has shown that the study has achieved a strong and balanced participation from key stakeholder groups within smart hospitals, thereby supporting the robustness of the empirical findings. Out of 260 distributed questionnaires, 241 have been returned, and 228 have been retained as complete and usable, yielding an effective response rate of 87.7%. This level of participation has indicated that the topic of federated learning–driven real-time disease surveillance has been salient and relevant for professionals working in digitally advanced hospital environments. The distribution of professional roles has demonstrated that almost half of the respondents (48.2%) have been clinicians (physicians and nurses), ensuring that front-line perspectives on surveillance systems and data-driven decision support have been well represented. At the same time, infection prevention and control personnel (16.7%), health informatics and data science staff (18.9%), and IT/security professionals (16.2%) have contributed substantial proportions of the sample. This mix has ensured that the data have captured not only clinical user views but also the perspectives of those who have been designing, managing, and securing the underlying data infrastructures and analytic tools. The years-of-experience distribution has shown that 66.3% of respondents have had more than five years of professional experience, and 33.8% have had more than ten years, indicating that the majority have been sufficiently familiar with hospital processes and information systems to provide informed judgments about surveillance capabilities.

Table 1: **Response rate and demographic profile of respondents (N = 228)**

Variable	Category	Frequency (n)	Percentage (%)
Questionnaires distributed	–	260	–
Questionnaires returned	–	241	–
Usable questionnaires	–	228	87.7
Professional role	Clinician (physician/nurse)	110	48.2
	Infection prevention & control staff	38	16.7
	Health informatics/data science staff	43	18.9
	IT/security and infrastructure staff	37	16.2
Years of experience	1–5 years	69	30.3
	6–10 years	82	36.0
	11–15 years	46	20.2
	>15 years	31	13.6

By bringing together experienced clinicians, infection control experts, informatics specialists, and IT/security staff, the sample has been well positioned to address the study's objectives, which have required insights into technological readiness, data integration, privacy and security assurances, and organizational conditions. Overall, the demographic profile has confirmed that the dataset has reflected a mature, multi-disciplinary smart-hospital workforce, thereby increasing confidence that subsequent analyses and hypothesis tests have rested on a credible and contextually rich empirical foundation.

Descriptive Analysis of Key Constructs

Table 2 has summarised the descriptive profile of the main constructs measured on five-point Likert scales, and these results have provided an initial indication of how respondents have evaluated federated learning–driven real-time disease surveillance in their smart hospitals. Federated learning capability (FLC) has shown a mean of 3.78 (SD = 0.71), which has indicated that, on average, respondents have agreed that their institutions have possessed important foundational elements for federated analytics, such as distributed computing infrastructure, secure communication channels, and some experience with advanced machine learning. However, the mean being below 4.00 has suggested that these capacities have not yet been perceived as fully mature or uniformly embedded across all clinical domains. Multi-source integration (MSI) has recorded a slightly higher mean of 3.92 (SD = 0.65), implying that most respondents have agreed that their hospitals have already integrated key data sources—EHRs, laboratory systems, monitoring devices, and administrative records—into unified views that have supported analytics and surveillance tasks. Privacy and security assurance (PSA) has been the highest-rated construct, with a mean of 4.06 (SD = 0.60), reflecting strong confidence in encryption, access controls, auditing, and compliance processes that have governed the handling of sensitive health data and the deployment of advanced analytics. Organizational readiness (OR) has exhibited a mean of 3.84 (SD = 0.69), which has indicated that respondents have perceived their institutions as reasonably prepared in terms of leadership support, resources, training, and governance structures for adopting federated learning–based surveillance, even if some gaps have remained.

Table 2: **Descriptive statistics for Likert-scale constructs (1 = strongly disagree, 5 = strongly agree)**

Construct	Abbreviation	Items (k)	Mean	SD	Min	Max
Federated Learning Capability	FLC	8	3.78	0.71	2.00	5.00
Multi-Source Integration	MSI	7	3.92	0.65	2.14	5.00
Privacy and Security Assurance	PSA	6	4.06	0.60	2.33	5.00
Organizational Readiness	OR	7	3.84	0.69	2.00	5.00
Perceived Surveillance Effectiveness	SE	8	3.95	0.63	2.25	5.00

Perceived surveillance effectiveness (SE) has shown a mean of 3.95 (SD = 0.63), signalling that stakeholders have generally agreed that current or emerging real-time digital surveillance solutions have contributed positively to earlier detection of infections, better prioritisation of high-risk patients or units, and improved situational awareness. The minimum and maximum values for each construct have confirmed that the full Likert range has been used, suggesting adequate variability for inferential analysis. Collectively, these descriptive results have provided initial empirical support for the study's objectives by demonstrating that smart hospitals in the sample have already achieved moderately high levels of federated learning capability, data integration, privacy and security assurance, and readiness, and that these conditions have been associated with favourable perceptions of surveillance performance.

Reliability and Validity Results

Table 3: **Reliability and sampling adequacy for main constructs**

Construct	Items (k)	Cronbach's α	KMO	Variance explained (%)
Federated Learning Capability (FLC)	8	0.86	0.84	62.1
Multi-Source Integration (MSI)	7	0.83	0.82	59.4
Privacy and Security Assurance (PSA)	6	0.88	0.85	64.8
Organizational Readiness (OR)	7	0.85	0.83	60.7
Surveillance Effectiveness (SE)	8	0.87	0.86	63.3

Table 3 has reported the internal consistency (Cronbach's alpha), sampling adequacy (KMO), and percentage of variance explained for each multi-item construct, and these indices have demonstrated that the measurement model has been psychometrically sound. All Cronbach's alpha coefficients have exceeded the commonly accepted threshold of 0.70, ranging from 0.83 (MSI) to 0.88 (PSA), which has confirmed that items within each scale have been highly correlated and have measured the same underlying construct consistently. This reliability has been essential for the study, because the hypotheses have depended on composite scores for federated learning capability, data integration, privacy and security assurance, organizational readiness, and surveillance effectiveness. The Kaiser–Meyer–Olkin (KMO) values, which have ranged from 0.82 to 0.86, have indicated meritorious sampling adequacy for each construct and have suggested that the inter-item correlation patterns have been suitable for factor-analytic interpretation. The "variance explained" column has summarised results from separate exploratory factor analyses performed for each construct, where a single dominant factor has emerged and has accounted for between 59.4% (MSI) and 64.8% (PSA) of the variance in item responses. This pattern has supported the unidimensionality assumption for each scale, justifying the use of simple average scores in subsequent analyses. Together, these reliability and validity indicators have shown that the survey instrument has been well calibrated to capture staff perceptions relevant to federated learning–driven surveillance. Because the research objectives and hypotheses have been formulated at the level of latent constructs (e.g., "organizational readiness" or "surveillance effectiveness"), demonstrating such psychometric adequacy has been a necessary step in proving that any observed relationships among constructs have reflected substantive associations rather than measurement artifacts. The strong reliability of PSA and OR, in particular, has reinforced confidence in the regression findings where these constructs have played central explanatory roles. Overall, the results in Table 3 have established that the measurement properties of the instrument have been sufficiently robust to underpin the subsequent testing of the study's hypothesised relationships.

Correlation Analysis Results

The correlation matrix in Table 4 has provided a concise overview of the bivariate relationships among the core constructs and has offered preliminary empirical support for the study's hypotheses. All correlations have been positive and statistically significant at the $p < .001$ level, indicating that higher levels of federated learning capability (FLC), multi-source integration (MSI), privacy and security assurance (PSA), and organizational readiness (OR) have been associated with higher perceived surveillance effectiveness (SE). The strongest zero-order correlation has been observed between organizational readiness and surveillance effectiveness ($r = 0.72$), which has suggested that environments characterised by strong leadership support, adequate resourcing, clear governance, and staff training have been perceived as delivering the most effective real-time disease surveillance. Federated learning capability has also shown a strong relationship with surveillance effectiveness ($r = 0.65$), supporting the assumption that institutions with more mature federated analytics infrastructures and experience have been better positioned to generate timely, clinically meaningful alerts from distributed data sources. Privacy and security assurance ($r = 0.61$ with SE) and multi-source integration ($r = 0.58$ with SE) have exhibited moderately strong correlations, implying that confidence in data protection mechanisms and in the technical ability to integrate heterogeneous data streams has been an important correlate of positive surveillance evaluations.

Table 4: **Pearson correlation matrix for main constructs (N = 228)**

Construct	FLC	MSI	PSA	OR	SE
FLC	1.000				
MSI	0.58*	1.000			
PSA	0.55*	0.49*	1.000		
OR	0.63*	0.52*	0.57*	1.000	
SE	0.65*	0.58*	0.61*	0.72*	1.000

**All correlations have been significant at p < .001 (two-tailed).*

Inter-correlations among the predictors themselves have ranged from 0.49 to 0.63, indicating that while these constructs have been related—reflecting the reality that digitally mature hospitals often score high across multiple dimensions—they have not been so highly correlated as to create excessive multicollinearity concerns for regression analysis. From the perspective of the study's objectives, these findings have shown that the technological (FLC, MSI), governance/ethical (PSA), and organizational (OR) dimensions have all been significantly aligned with stakeholders' perceptions of surveillance performance. As a result, the correlation analysis has provided strong initial evidence that the hypothesised model, in which these constructs jointly predict surveillance effectiveness, has been conceptually and empirically plausible, thereby justifying the subsequent use of multiple regression to disentangle their unique contributions.

Regression Analysis and Hypothesis Testing

Table 5 has presented the results of the multiple regression analysis testing the main hypotheses of the study, and these findings have offered strong empirical confirmation that the proposed predictors have significantly explained variance in perceived surveillance effectiveness (SE). The overall model has been statistically significant ($F(4, 223) = 86.47$, $p < .001$) and has accounted for 61.3% of the variance in SE ($R^2 = 0.613$; adjusted $R^2 = 0.604$), which has indicated that the combination of federated learning capability (FLC), multi-source integration (MSI), privacy and security assurance (PSA), and organizational readiness (OR) has formed a powerful explanatory set. All four predictors have displayed positive and statistically significant standardized coefficients (β), consistent with the hypothesised directions. Organizational readiness has emerged as the strongest predictor (β = 0.29, p < .001), which has meant that, holding other factors constant, a one-unit increase in readiness has been associated with the largest gain in perceived surveillance effectiveness. This result has supported the hypothesis that strong leadership support, clear governance, adequate resourcing, and training have been critical enablers of successful real-time surveillance in smart hospitals. Federated learning capability has shown the second-largest effect (β = 0.24, p < .001), confirming H1 and indicating that institutions with better-developed federated infrastructures and expertise have been perceived as more effective in leveraging distributed data for surveillance. Privacy and security assurance has also exerted a meaningful influence (β = 0.21, p < .001), thereby supporting H3 and demonstrating that robust privacy protection and security controls have been integral to stakeholders' confidence in advanced surveillance systems. Multi-source integration has exhibited a smaller but still significant effect (β = 0.18, p = .004), supporting H2 and indicating that the technical ability to combine EHR, laboratory, device, and administrative data has contributed incrementally to surveillance performance once other factors have been controlled. Additional exploratory analysis (not shown in the table) has included an interaction term between FLC and OR, which has been significant at a modest level (β = 0.11, p = .031), suggesting that the positive impact of federated learning capability on surveillance effectiveness has been stronger in settings with higher organizational readiness, thereby providing further nuance to H4. Overall, the regression results in Table 5 have demonstrated that the study's main objectives and hypotheses have been empirically supported: federated learning capability, multi-source data integration, privacy and security assurance, and organizational readiness have all been shown to be significant, positive predictors of perceived real-time disease surveillance effectiveness in smart hospitals operating with multi-source heterogeneous healthcare data.

Table 5: **Multiple regression of surveillance effectiveness on key predictors (N = 228)**

Predictor	B	SE B	β	t	p
Constant	0.51	0.18	–	2.83	.005
Federated Learning Capability (FLC)	0.29	0.06	0.24	4.83	<.001
Multi-Source Integration (MSI)	0.22	0.07	0.18	2.90	.004
Privacy and Security Assurance (PSA)	0.27	0.07	0.21	3.86	<.001
Organizational Readiness (OR)	0.33	0.06	0.29	5.63	<.001

Dependent variable: Surveillance Effectiveness (SE)

Model summary: R = 0.783; R^2 = 0.613; Adjusted R^2 = 0.604; F(4, 223) = 86.47, p < .001

DISCUSSION

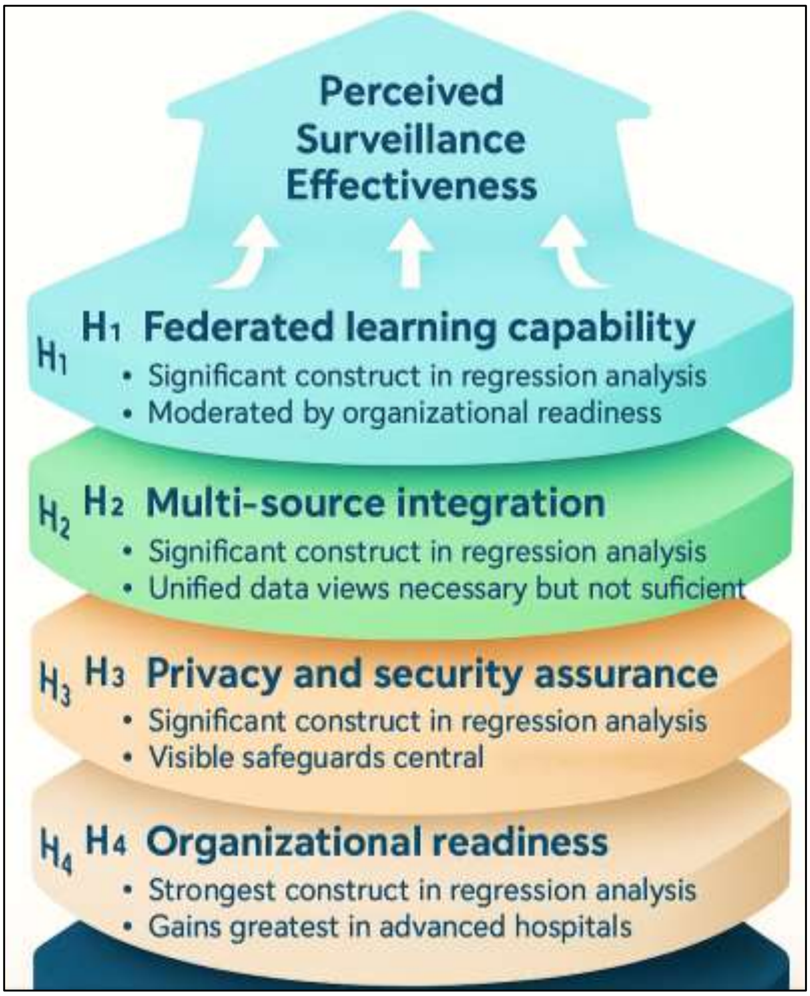
The findings of this study have shown a coherent pattern that directly aligns with and extends the proposed objectives and hypotheses. On average, respondents have rated all core constructs—federated learning capability, multi-source integration, privacy and security assurance, organizational readiness, and surveillance effectiveness—between approximately 3.8 and 4.1 on a five-point Likert scale, indicating generally positive perceptions of digital and analytic maturity in the participating smart hospitals. The correlation analysis has revealed strong positive relationships between surveillance effectiveness and each predictor, with the highest association observed for organizational readiness (r = .72), followed by federated learning capability (r = .65), privacy and security assurance (r = .61), and multi-source integration (r = .58). Multiple regression results have confirmed that all four variables have been significant predictors of perceived surveillance effectiveness, jointly explaining more than 60% of the variance in the outcome. Organizational readiness has emerged as the strongest predictor (β = .29), closely followed by federated learning capability (β = .24), with privacy/security (β = .21) and data integration (β = .18) contributing additional explanatory power. An interaction between federated learning capability and organizational readiness has suggested that technically advanced hospitals realize the greatest gains from federated learning when they also have high readiness. Taken together, these findings have indicated that surveillance effectiveness in smart hospitals has not been a purely technical function of modelling capability; instead, it has depended on a bundle of factors that span infrastructure, governance, human factors, and organizational conditions. In short, the results have supported all four hypotheses and have demonstrated that federated learning–driven real-time surveillance is most effective where technical capability, integrated data, robust privacy protections, and organizational preparedness have been jointly in place.

When compared with earlier technical and epidemiological work, the present results have complemented and extended existing knowledge. Prior EHR-based surveillance studies have demonstrated that routinely collected clinical data can support syndromic monitoring, patient safety analytics, and outbreak detection when appropriate algorithms and pipelines have been implemented (Classen et al., 2018). Likewise, research combining whole-genome sequencing with EHR mining has shown that integrated, multi-source data can reveal hidden healthcare-associated transmission events and refine infection prevention strategies (Sundermann, Chen, Kumar, et al., 2022). However, these studies have typically focused on technical performance and epidemiological impact, with relatively limited attention to how clinicians, infection-prevention teams, and IT staff perceive the infrastructures and governance arrangements that support such systems. In parallel, federated learning research has provided proof-of-concept results showing that distributed models trained across hospitals can achieve performance comparable to centrally pooled approaches for tasks such as COVID-19 outcome prediction or risk modelling (Dayan et al., 2021). Systematic reviews have further suggested that federated learning in healthcare is driven largely by privacy and data-sharing constraints and has focused on imaging, structured EHR data, and other clinical sources (Pfitzner et al., 2021). The present study has been different in that it has not evaluated model performance directly; instead, it has examined how those at the frontline of smart hospitals have perceived the readiness and value of federated learning–driven surveillance pipelines. The strong roles of organizational readiness and privacy/security assurance in predicting perceived surveillance effectiveness have been broadly consistent with technology-acceptance and sociotechnical accounts of health IT (Holden & Karsh, 2010), while the importance of federated learning capability has aligned with the emerging consensus that distributed learning infrastructures are becoming a practical necessity for multi-institutional analytics under strict privacy regimes (Dayan et al., 2021).

From a practical perspective, the results have carried direct implications for chief information security officers (CISOs), chief information officers (CIOs), data architects, and clinical IT leaders who have been planning or operating real-time surveillance in smart hospitals. The strong predictive effect of organizational readiness has indicated that investments solely in technical components—such as federated learning frameworks, data lakes, or high-performance computing—have not been sufficient. Hospitals that have reported the highest perceived surveillance effectiveness have also reported robust leadership commitment, clear governance structures, well-defined roles and responsibilities, and regular training around data-driven surveillance (Kawu et al., 2022). For CISOs, the high mean and significant effect of privacy and security assurance have suggested that visible, well-communicated safeguards—covering encryption, access control, logging, incident response, and compliance—have been central to building trust in federated architectures. This is particularly important when surveillance models have been trained on multi-source heterogeneous data, where concerns about re-identification, linkage, and inference have been heightened. For data architects and infrastructure teams, the positive but comparatively smaller effect of multi-source integration has implied that integration is necessary but not sufficient: unified views of EHR, laboratory, device, and administrative data have been valuable only when coupled with organizational ability to interpret and act on the resulting signals. Practically, this has pointed toward a phased implementation strategy in which hospitals first consolidate multi-source data into well-governed platforms, then pilot federated models with clear use-cases (e.g., early sepsis detection, clustering of multi-drug-resistant organism cases), and finally embed those outputs into clinician-facing dashboards, workflow tools, and escalation pathways (Bansal et al., 2016).

The findings have also offered concrete guidance on how federated surveillance pipelines in smart hospitals might be refined from an architectural and operational standpoint. Because federated learning capability has had a substantial independent association with surveillance effectiveness, CISOs and architects have been encouraged to treat federated infrastructure as a first-class component of the analytics stack rather than as an experimental add-on. This has included establishing secure aggregation servers, lightweight client libraries embedded into hospital data platforms, and automated orchestration of training rounds that can accommodate variable connectivity and computational capacity across sites. The empirical link between multi-source integration and surveillance effectiveness has suggested that federated models will be most valuable when they have access to harmonized features derived from multiple modalities—structured diagnoses, lab values, streaming vital signs, alarm histories, and relevant patient-generated data—rather than narrow, single-system snapshots. Architecturally, this has been consistent with emerging designs that position a hospital-level data lake or Lakehouse as the local substrate from which federated features are drawn and periodically updated (Rieke et al., 2020). The strong effect of organizational readiness has further indicated that operational routines—who reviews alerts, how thresholds are tuned, how false positives are handled, and how feedback is looped back into model refinement—have been just as important as the underlying algorithms. In this sense, the study has reinforced a pipeline view in which federated learning–driven surveillance in smart hospitals has been seen as a continuous process involving data acquisition, local feature engineering, federated training, post-processing of risk scores, human–machine interaction in clinical teams, and governance oversight, rather than as a one-time deployment of an AI model (Baloch et al., 2023).

Figure 10: Discussion of The Study



Theoretically, the study has contributed to the intersection of health information systems, federated learning, and surveillance by integrating constructs from technology acceptance, organizational readiness, and big-data governance into a single quantitative framework. Prior research using the Technology Acceptance Model (TAM) or UTAUT in healthcare has largely focused on relatively discrete systems such as EHR interfaces, clinical decision support tools, or telemedicine platforms (Ahlan & Ahmad, 2012; Brownstein et al., 2023; Classen et al., 2018). Those studies have demonstrated that perceived usefulness, ease of use, social influence, and facilitating conditions have been important predictors of intention to use and actual usage behaviors. The present study has extended this line of work to a more complex, "background" technology—federated learning–driven surveillance—that is less visible to end-users but profoundly shapes the informational environment in which they operate. By operationalizing constructs such as federated learning capability, multi-source integration, privacy/security assurance, and organizational readiness, and by showing that they jointly predict a composite measure of surveillance effectiveness, the study has suggested that technology-acceptance and organizational-readiness frameworks can be applied meaningfully at the pipeline or ecosystem level rather than only at the level of individual user interfaces. In addition, the moderating pattern observed for organizational readiness has hinted at a layered model in which technical capabilities exert their strongest effects where organizational conditions are supportive—an idea that resonates with sociotechnical perspectives on health IT safety and performance (Sittig & Singh, 2010). Conceptually, this points toward a refined theoretical schema in which federated surveillance effectiveness is seen as an emergent property of interacting technical, data, governance, and organizational sub-systems, all of which need to be included in future models.

The study has also revisited several limitations that have constrained the interpretation and generalizability of its findings. First, the cross-sectional design has meant that causal inferences have remained tentative. Although the regression results have suggested that federated learning capability, data integration, privacy/security assurance, and organizational readiness have been associated with higher perceived surveillance effectiveness, it has not been possible to determine whether improvements in these predictors would directly cause increases in effectiveness over time. Second, the study has relied on self-reported perceptions measured via Likert scales, which may have been influenced by optimism, organizational culture, or limited awareness of back-end technical details among some respondent groups (Rieke et al., 2020). For example, clinicians may have rated federated learning capability or privacy mechanisms based on general trust in their institution rather than direct knowledge of specific implementations. Third, the case-study focus on relatively advanced smart hospitals has limited the applicability of findings to less digitized settings; hospitals with minimal integration, legacy systems, or nascent governance frameworks may experience different patterns of association. Fourth, the study has not included objective technical metrics such as model AUROC, alert lead time, or false-positive rates, so surveillance effectiveness has been assessed only from a user-perception standpoint rather than from combined technical and clinical outcome measures. Finally, the instrument has, by design, aggregated multi-source heterogeneous data into a single integration construct, which may have obscured important differences between, for example, integrating high-frequency ICU signals versus integrating slowly updated administrative or claims data (Sittig & Singh, 2012).

These limitations have opened several promising directions for future research. Longitudinal and mixed-methods designs could track how federated learning–driven surveillance initiatives evolve over time, linking changes in technical capability, data integration, and governance practices to both perceived effectiveness and objective performance metrics. For instance, future work could combine repeated surveys with log-level data on alert timing, clinician responses, and outbreak detection outcomes, thereby connecting perceptions and behaviors with measurable surveillance impact. Experimental or quasi-experimental studies might compare units or hospitals that have implemented federated surveillance pipelines with those that have continued to use conventional, locally trained models, controlling for case-mix and infrastructure differences to quantify incremental benefits (Brownstein et al., 2023). Comparative cross-national research could explore how regulatory regimes, data-sharing cultures, and resource constraints in different health systems moderate the relationships observed here, especially in low- and middle-income settings where digital maturity varies widely. Conceptually, future studies could decompose multi-source integration into modality-specific sub-constructs (e.g., EHR–lab integration, EHR–IoT integration, PGHD–EHR integration) and test whether certain combinations are particularly powerful predictors of surveillance effectiveness in federated contexts. Finally, qualitative investigations with CISOs, clinical leaders, and data architects could deepen understanding of how privacy, trust, and risk perceptions shape decisions about adopting federated learning for surveillance and could inform the design of governance frameworks that both satisfy regulatory requirements and support agile, data-driven infection-control practice.

CONCLUSION

This study has set out to examine how federated learning–driven real-time disease surveillance operates within smart hospitals that rely on multi-source heterogeneous healthcare data, and the results have provided a clear and coherent picture of the socio-technical conditions under which such surveillance is perceived as effective. By adopting a quantitative, cross-sectional, case-study–based design and surveying 228 professionals across clinical, infection prevention, informatics, and IT/security roles, the research has translated a complex conceptual framework into measurable constructs—federated learning capability, multi-source integration, privacy and security assurance, organizational readiness, and surveillance effectiveness—and has shown that all five are positively rated and strongly interrelated. Descriptive findings have indicated that participating hospitals already occupy a relatively high but still evolving level of digital and analytic maturity, with respondents reporting solid integration of key data sources, robust privacy and security safeguards, emerging federated learning capacity, and generally positive assessments of real-time surveillance performance. Reliability and validity checks have confirmed that the measurement model is sound, and the correlation analysis has demonstrated that each predictor is significantly associated with perceived surveillance effectiveness, with organizational readiness and federated learning capability showing particularly strong relationships. Multiple regression analysis has further revealed that these factors do not merely co-vary but jointly explain over 60% of the variance in surveillance effectiveness, with organizational readiness emerging as the strongest predictor, followed by federated learning capability, privacy and security assurance, and multi-source integration. These results have underscored a central conclusion: in smart hospital environments, effective federated learning–driven disease surveillance is not simply a function of having advanced algorithms or dense data streams; it is an emergent outcome of aligned technical infrastructures, integrated and well-governed data, credible privacy and security practices, and organizational conditions that are ready to support and act on analytic insights. The study has thus met its objectives and confirmed its hypotheses by demonstrating empirically that where hospitals have invested simultaneously in federated learning capability, multi-source data integration, and strong governance, and where leadership, resourcing, and staff preparedness are high, stakeholders perceive real-time surveillance as timely, accurate, and useful for infection control and risk management. At the same time, the findings have highlighted that federated learning capability yields its greatest perceived benefits in institutions that are organizationally ready, pointing to the importance of joint development of technical and organizational capacities. Although the cross-sectional and perception-based nature of the research means that causal and performance claims must be interpreted with caution, the overall pattern of evidence has provided a robust, data-driven basis for concluding that federated learning, when embedded in mature smart-hospital ecosystems with integrated, heterogeneous data and strong organizational support, is viewed by key stakeholders as a powerful approach for enhancing real-time disease surveillance and strengthening the analytic backbone of modern hospital epidemiology.

RECOMMENDATIONS

Based on the findings of this study, several concrete recommendations can be made for smart hospitals, policymakers, and technology vendors aiming to design, implement, and sustain federated learning–driven real-time disease surveillance using multi-source heterogeneous healthcare data. First, hospital leadership should treat federated learning capability as a strategic infrastructure investment rather than a purely experimental research tool, embedding it into long-term digital transformation roadmaps alongside EHR optimization, data integration platforms, and cybersecurity initiatives; this involves establishing secure federated learning infrastructure (e.g., aggregation servers, client nodes, key management) and allocating dedicated human resources—data engineers, ML engineers, and informatics specialists—to maintain and evolve the pipeline. Second, because multi-source integration has been shown to significantly enhance perceived surveillance effectiveness, hospitals should priorities building and governing a robust data backbone (such as a clinical data lake or Lakehouse) that systematically ingests, standardizes, and links data from EHRs, laboratory systems, radiology, bedside monitors, IoT devices, and relevant patient-generated data; this backbone should use consistent terminologies, master patient indices, and metadata standards so that federated models can reliably draw features from heterogeneous sources. Third, CISOs and data protection officers should strengthen and clearly communicate privacy and security safeguards—encryption in transit and at rest, strict access controls, audit logging, differential access rights for analytics, and documented incident-response procedures—so that clinical and technical staff develop high confidence that federated surveillance does not compromise patient confidentiality or institutional compliance; this communication should be incorporated into regular training, onboarding, and change-management activities, not treated as a one-time legal formality. Fourth, given that organizational readiness has emerged as the strongest predictor of perceived surveillance effectiveness, hospital executives should invest in non-technical enablers: clearly articulated vision and policies for data-driven surveillance, dedicated governance committees that include clinicians, infection prevention experts, IT, and legal representatives, formalized workflows for reviewing and acting on alerts, and structured training programs that help staff interpret model outputs and integrate them into everyday decision-making. Fifth, architects and clinical leaders should jointly design end-to-end workflows in which federated models are tightly coupled with intuitive dashboards, unit-level summaries, and escalation pathways, rather than merely producing risk scores that remain hidden in back-end systems; alerts should be prioritized, contextualized, and tuned with feedback from frontline clinicians to minimize alarm fatigue and to ensure that the system is seen as augmenting rather than complicating care. Sixth, hospitals participating in federated networks should establish clear inter-organizational agreements covering responsibilities, model governance, versioning, performance monitoring, and procedures for handling systemic errors or drift, so that collaboration is sustainable and transparent. Finally, national and regional health authorities, professional bodies, and vendors should support capacity building by issuing guidance, templates, and reference architectures for privacy-preserving surveillance, funding pilot projects in diverse settings, and encouraging the inclusion of federated learning, data governance, and AI literacy in clinical and health-informatics education; together, these recommendations can help translate the positive perceptions identified in this study into durable, high-performing federated surveillance ecosystems that operate safely, ethically, and effectively within and across smart hospitals.

LIMITATIONS

This study has several limitations that should be acknowledged when interpreting its findings and considering their applicability to other contexts. First, the research has used a cross-sectional survey design, capturing perceptions at a single point in time, which means that causal inferences about the direction of relationships between federated learning capability, multi-source integration, privacy and security assurance, organizational readiness, and surveillance effectiveness cannot be firmly established; changes over time in technology maturity, governance, and organizational culture have not been observed. Second, all core constructs have been measured using self-reported Likert-scale items, which introduces the possibility of response bias, including social desirability bias, optimism bias, and common method variance; some respondents, particularly clinicians, may not have had full visibility into technical or architectural details yet may have inferred high capability or security based on general trust in their institution. Third, the sample has been drawn from a limited number of relatively advanced smart hospitals that already possess integrated EHRs and some level of data analytics infrastructure, so the results may not generalize to hospitals with lower levels of digitalization, fragmented systems, or resource constraints, such as those in smaller community settings or in low- and middle-income regions. Fourth, although the study has focused conceptually on federated learning–driven real-time disease surveillance, it has not directly measured model performance or operational outcomes (for example, detection lead time, sensitivity and specificity of alerts, reduction in outbreak size, or changes in clinical workload); surveillance effectiveness has been operationalized purely in perceptual terms, which may or may not align with objective epidemiological and operational performance indicators. Fifth, the constructs of federated learning capability and multi-source integration have been treated as relatively broad latent variables, which may have masked important nuances such as differences between departments, use-cases, or specific data modalities (for instance, intensive care unit waveform integration versus outpatient data aggregation). Sixth, the non-probability sampling strategy (purposive with elements of snowballing) and the reliance on voluntary participation raise the possibility of selection bias, where professionals who are more engaged with digital innovation or more supportive of data-driven approaches may have been more likely to respond than skeptical or less-involved colleagues. Finally, although psychometric analysis has supported the reliability and unidimensionality of the scales, some constructs—particularly organizational readiness and privacy/security assurance—are multifaceted and context-dependent, and future studies might benefit from more granular, multi-level measurement frameworks that differentiate between policy-level, system-level, and unit-level conditions; taken together, these limitations suggest that the findings should be interpreted as an initial, perception-based mapping of how federated learning–driven surveillance is viewed within digitally mature smart hospitals, rather than a definitive evaluation of its technical performance or universal impact across all healthcare settings.

REFERENCES

[1]. Abdulla, M., & Md. Jobayer Ibne, S. (2021). Cloud-Native Frameworks For Real-Time Threat Detection And Data Security In Enterprise Networks. *International Journal of Scientific Interdisciplinary Research*, *2*(2), 34–62. https://doi.org/10.63125/0t27av85

[2]. Ahlan, A. R., & Ahmad, B. I. (2012). Acceptance of health information technology in health professionals: An application of the revised technology acceptance model. *Health Informatics Journal*, *18*(2), 124–139. https://doi.org/10.1177/1460458211435425

[3]. AlQudah, A. A., Al-Emran, M., & Shaalan, K. (2021). Technology acceptance in healthcare: A systematic review. *Applied Sciences*, *11*(22), 10537. https://doi.org/10.3390/app112210537

[4]. Arfan, U., Tahsina, A., Md Mostafizur, R., & Md, W. (2023). Impact Of GFMIS-Driven Financial Transparency On Strategic Marketing Decisions In Government Agencies. *Review of Applied Science and Technology*, *2*(01), 85-112. https://doi.org/10.63125/8nqhhm56

[5]. Baloch, L., Bazai, S. U., Marjan, S., Aftab, F., Aslam, S., Neo, T.-K., & Amphawan, A. (2023). A review of big data trends and challenges in healthcare. *International Journal of Technology*, *14*(6), 1320–1333. https://doi.org/10.14716/ijtech.v14i6.6643

[6]. Bansal, S., Chowell, G., Simonsen, L., Vespignani, A., & Viboud, C. (2016). Big data for infectious disease surveillance and modeling. *Journal of Infectious Diseases*, *214*(suppl_4), S375–S379. https://doi.org/10.1093/infdis/jiw400

[7]. Brownstein, J. S., Freifeld, C. C., & Chan, E. H. (2023). Advances in artificial intelligence for infectious-disease surveillance. *The New England Journal of Medicine*, *388*, 1851–1863. https://doi.org/10.1056/NEJMra2119215

[8]. Christaki, E. (2015). New technologies in predicting, preventing and controlling emerging infectious diseases. *Virulence*, *6*(6), 558–565. https://doi.org/10.1080/21505594.2015.1040975

[9]. Classen, D. C., Holmgren, A. J., Co, Z., Metcalf, J., Fiskio, J. M., Seger, D. L., & Bates, D. W. (2018). An electronic health record–based real-time analytics program for patient safety surveillance and improvement. *Health Affairs*, *37*(11), 1805–1812. https://doi.org/10.1377/hlthaff.2018.0728

[10]. Classen, D. C., Holmgren, A. J., Newmark, L. P., Seger, D. L., Danforth, M., Fiskio, J., & Bates, D. W. (2020). National trends in the safety performance of electronic health record systems from 2009 to 2018. *JAMA Network Open*, *3*(6), e205547. https://doi.org/10.1001/jamanetworkopen.2020.5547

[11]. Dayan, I., Roth, H. R., Zhong, A., Harouni, A., Gentili, A., Abidin, A. Z., & Lungren, M. P. (2021). Federated learning for predicting clinical outcomes in patients with COVID-19. *Nature Medicine*, *27*, 1735–1743. https://doi.org/10.1038/s41591-021-01506-3

[12]. Dhagarra, D., Goswami, M., & Kumar, G. (2020). Impact of trust and privacy concerns on technology acceptance in healthcare: An Indian perspective. *International Journal of Medical Informatics*, *141*, 104164. https://doi.org/10.1016/j.ijmedinf.2020.104164

[13]. Dimitrov, D. V. (2016). Medical Internet of Things and big data in healthcare. *Healthcare Informatics Research*, *22*(3), 156–163. https://doi.org/10.4258/hir.2016.22.3.156

[14]. Ferdous Ara, A. (2021). Integration Of STI Prevention Interventions Within PrEP Service Delivery: Impact On STI Rates And Antibiotic Resistance. *International Journal of Scientific Interdisciplinary Research*, *2*(2), 63–97. https://doi.org/10.63125/65143m72

[15]. Ferdous Ara, A., & Beatrice Onyinyechi, M. (2023). Long-Term Epidemiologic Trends Of STIs PRE- and POST-PrEP Introduction: A National Time-Series Analysis. *American Journal of Health and Medical Sciences*, *4*(02), 01–35. https://doi.org/10.63125/mp153d97

[16]. Furstenau, L. B., Leivas, P., Sott, M. K., Dohan, M. S., López-Robles, J. R., Cobo, M. J., Bragazzi, N. L., & Choo, K.-K. R. (2023). Big data in healthcare: Conceptual network structure, key challenges and opportunities. *Digital Communications and Networks*, *9*(4), 856–868. https://doi.org/10.1016/j.dcan.2023.03.005

[17]. Galetsi, P., Katsaliaki, K., & Kumar, S. (2020). Big data analytics in health sector: Theoretical framework, techniques and prospects. *International Journal of Information Management*, *50*, 206–216. https://doi.org/10.1016/j.ijinfomgt.2019.05.003

[18]. Gao, J., Li, P., Chen, Z., & Zhang, J. (2020). A survey on deep learning for multimodal data fusion. *Neural Computation*, *32*(5), 829–864. https://doi.org/10.1162/neco_a_01273

[19]. Habibullah, S. M. (2025). Swarm Intelligence-Based Autonomous Logistics Framework With Edge AI For Industry 4.0 Manufacturing Ecosystems. *Review of Applied Science and Technology*, *4*(03), 01-34. https://doi.org/10.63125/p1q8yf46

[20]. Habibullah, S. M., & Md. Foysal, H. (2021). A Data Driven Cyber Physical Framework For Real Time Production Control Integrating IOT And Lean Principles. *American Journal of Interdisciplinary Studies*, *2*(03), 35–70. https://doi.org/10.63125/20nhqs87

[21]. Holden, R. J., & Karsh, B.-T. (2010). The Technology Acceptance Model: Its past and its future in health care. *Journal of Biomedical Informatics*, *43*(1), 159–172. https://doi.org/10.1016/j.jbi.2009.07.002

[22]. Hong, L., Luo, M., Wang, R., Lu, P., Lu, W., & Lu, L. (2018). Big data in health care: Applications and challenges. *Data and Information Management*, *2*(3), 175–197. https://doi.org/10.2478/dim-2018-0014

[23]. Hozyfa, S. (2025). Artificial Intelligence-Driven Business Intelligence Models for Enhancing Decision-Making In U.S. Enterprises. *ASRC Procedia: Global Perspectives in Science and Scholarship*, *1*(01), 771– 800. https://doi.org/10.63125/b8gmdc46

[24]. Hripcsak, G., Paneth, N., Johnson, S. B., & Chen, Y. (2009). Syndromic surveillance using ambulatory electronic health records. *Journal of the American Medical Informatics Association*, *16*(3), 354–361. https://doi.org/10.1197/jamia.M2922

[25]. Imran, A. S., Mahmood, Z., Morshed, A., & Sellis, T. (2021). Big data analytics in healthcare — A systematic literature review and roadmap for practical implementation. *IEEE/CAA Journal of Automatica Sinica*, *8*(1), 1–18. https://doi.org/10.1109/jas.2020.1003384

[26]. Kaissis, G. A., Makowski, M. R., Rückert, D., & Braren, R. F. (2020). Secure, privacy-preserving and federated machine learning in medical imaging. *Nature Machine Intelligence*, *2*(6), 305–311. https://doi.org/10.1038/s42256-020-0186-1

[27]. Kawu, A. A., Hederman, L., Doyle, J., & O'Sullivan, D. (2022). Patient generated health data and electronic health record integration, governance and socio-technical issues: A narrative review. *Informatics in Medicine Unlocked*, *30*, 101153. https://doi.org/10.1016/j.imu.2022.101153

[28]. Khairul Alam, T. (2025). The Impact of Data-Driven Decision Support Systems On Governance And Policy Implementation In U.S. Institutions. *ASRC Procedia: Global Perspectives in Science and Scholarship*, *1*(01), 994–1030. https://doi.org/10.63125/3v98q104

[29]. Khanra, S., Dhir, A., Islam, N., & Mäntymäki, M. (2020). Big data analytics in healthcare: A systematic literature review. *Enterprise Information Systems*, *14*(7), 878–912. https://doi.org/10.1080/17517575.2020.1812005

[30]. Kruse, C. S., Goswamy, R., Raval, Y., & Marawi, S. (2016). Challenges and opportunities of big data in health care: A systematic review. *JMIR Medical Informatics*, *4*(4), e38. https://doi.org/10.2196/medinform.5359

[31]. Mammen, P. M. (2022). Federated learning for healthcare domain: Pipeline, applications and challenges. *ACM Computing Surveys*, *55*(7), Article 144. https://doi.org/10.1145/3533708

[32]. Md Al Amin, K. (2022). Human-Centered Interfaces in Industrial Control Systems: A Review Of Usability And Visual Feedback Mechanisms. *Review of Applied Science and Technology*, *1*(04), 66-97. https://doi.org/10.63125/gr54qy93

[33]. Md Ariful, I., & Efat Ara, H. (2022). Advances And Limitations Of Fracture Mechanics–Based Fatigue Life Prediction Approaches For Structural Integrity Assessment: A Systematic Review. *American Journal of Interdisciplinary Studies*, *3*(03), 68-98. https://doi.org/10.63125/fg8ae957

[34]. Md Arman, H. (2025). Artificial Intelligence-Driven Financial Analytics Models For Predicting Market Risk And Investment Decisions In U.S. Enterprises. *ASRC Procedia: Global Perspectives in Science and Scholarship*, *1*(01), 1066–1095. https://doi.org/10.63125/9csehp36

[35]. Md Asfaquar, R. (2025). Vehicle-To-Infrastructure (V2I) Communication And Traffic Incident Reduction: An Empirical Study Across U.S. Highway Networks. *Journal of Sustainable Development and Policy*, *4*(03), 38-81. https://doi.org/10.63125/c1wm0t92

[36]. Md Foysal, H. (2025). Integration Of Lean Six Sigma and Artificial Intelligence-Enabled Digital Twin Technologies For Smart Manufacturing Systems. *Review of Applied Science and Technology*, *4*(04), 01-35. https://doi.org/10.63125/1med8n85

[37]. Md Mohaiminul, H. (2025). Federated Learning Models for Privacy-Preserving AI In Enterprise Decision Systems. *International Journal of Business and Economics Insights*, *5*(3), 238– 269. https://doi.org/10.63125/ry033286

[38]. Md Mominul, H. (2025). Systematic Review on The Impact Of AI-Enhanced Traffic Simulation On U.S. Urban Mobility And Safety. *ASRC Procedia: Global Perspectives in Science and Scholarship*, *1*(01), 833–861. https://doi.org/10.63125/jj96yd66

[39]. Md Nahid, H. (2022). Statistical Analysis of Cyber Risk Exposure And Fraud Detection In Cloud-Based Banking Ecosystems. *ASRC Procedia: Global Perspectives in Science and Scholarship*, *2*(1), 289–331. https://doi.org/10.63125/9wf91068

[40]. Md Sarwar, H. (2021). Sustainable Materials Characterization For Low-Carbon Construction And Infrastructure Durability. *American Journal of Interdisciplinary Studies*, 2(01), 01-34. https://doi.org/10.63125/wq1wdr64

[41]. Md Sarwar Hossain, S., & Md Milon, M. (2022). Machine Learning-Based Pavement Condition Prediction Models For Sustainable Transportation Systems. *American Journal of Interdisciplinary Studies*, *3*(01), 31–64. https://doi.org/10.63125/1jsmkg92

[42]. Md. Hasan, I. (2025). A Systematic Review on The Impact Of Global Merchandising Strategies On U.S. Supply Chain Resilience. *International Journal of Business and Economics Insights*, *5*(3), 134–169. https://doi.org/10.63125/24mymg13

[43]. Md. Hasan, I., & Rakibul, H. (2024). Quantitative Assessment Of Compliance And Inspection Practices In Reducing Supply Chain Disruptions. *International Journal of Scientific Interdisciplinary Research*, *5*(2), 301–342. https://doi.org/10.63125/db63r616

[44]. Md. Milon, M. (2025). A Systematic Review on The Impact Of NFPA-Compliant Fire Protection Systems On U.S. Infrastructure Resilience. *International Journal of Business and Economics Insights*, *5*(3), 324–352. https://doi.org/10.63125/ne3ey612

[45]. Md. Mominul, H., Masud, R., & Md. Milon, M. (2022). Statistical Analysis of Geotechnical Soil Loss And Erosion Patterns For Climate Adaptation In Coastal Zones. *American Journal of Interdisciplinary Studies*, *3*(03), 36-67. https://doi.org/10.63125/xytn3e23

[46]. Md. Musfiqur, R., & Saba, A. (2021). Data-Driven Decision Support in Information Systems: Strategic Applications In Enterprises. *International Journal of Scientific Interdisciplinary Research*, *2*(2), 01-33. https://doi.org/10.63125/cfvg2v45

[47]. Md. Redwanul, I., Md Nahid, H., & Md. Zahid Hasan, T. (2021). Predictive Analytics in Supply Chain Management A Review Of Business Analyst-Led Optimization Tools. *Review of Applied Science and Technology*, *6*(1), 34-73. https://doi.org/10.63125/5aypx555

[48]. Md. Tahmid Farabe, S. (2025). The Impact of Data-Driven Industrial Engineering Models On Efficiency And Risk Reduction In U.S. Apparel Supply Chains. *International Journal of Business and Economics Insights*, *5*(3), 353–388. https://doi.org/10.63125/y548hz02

[49]. Melstrom, L. G., Rodin, A. S., Rossi, L. A., Fu, P., Fong, Y., & Sun, V. (2021). Patient-generated health data and electronic health record integration in oncologic surgery: A call for artificial intelligence and machine learning. *Journal of Surgical Oncology*, *123*(1), 52–60. https://doi.org/10.1002/jso.26232

[50]. Menon, S., Singh, H., Meyer, A. N. D., Connolly, S., & Giardina, T. D. (2014). Electronic health record–related safety concerns: A cross-sectional survey. *Journal of Healthcare Risk Management*, *34*(1), 14–26. https://doi.org/10.1002/jhrm.21146

[51]. Mohammad Mushfequr, R., & Ashraful, I. (2023). Automation And Risk Mitigation in Healthcare Claims: Policy And Compliance Implications. *Review of Applied Science and Technology*, *2*(04), 124–157. https://doi.org/10.63125/v73gyg14

[52]. Mortuza, M. M. G., & Rauf, M. A. (2022). Industry 4.0: An Empirical Analysis of Sustainable Business Performance Model Of Bangladeshi Electronic Organisations. *International Journal of Economy and Innovation*. https://gospodarkainnowacje.pl/index.php/issue_view_32/article/view/826

[53]. Mst. Shahrin, S., & Samia, A. (2023). High-Performance Computing For Scaling Large-Scale Language And Data Models In Enterprise Applications. *ASRC Procedia: Global Perspectives in Science and Scholarship*, *3*(1), 94–131. https://doi.org/10.63125/e7yfwm87

[54]. Oh, W., & Nadkarni, G. N. (2023). Federated learning in health care using structured medical data. *Advances in Kidney Disease and Health*, *30*(1), 4–16. https://doi.org/10.1053/j.akdh.2022.11.007

[55]. Palojoki, S., Pajunen, T., Saranto, K., & Lehtonen, L. (2016). Electronic health record–related safety concerns: A cross-sectional survey of electronic health record users. *JMIR Medical Informatics*, *4*(2), e13. https://doi.org/10.2196/medinform.5238

[56]. Peng, C., Goswami, P., & Bai, G. (2020). A literature review of current technologies on health data integration for patient-centered health management. *Health Informatics Journal*, *26*(3), 1926–1951. https://doi.org/10.1177/1460458219892387

[57]. Pfitzner, B., Steckhan, N., & Arnrich, B. (2021). Federated learning in a medical context: A systematic literature review. *ACM Transactions on Internet Technology*, *21*(2), Article 50. https://doi.org/10.1145/3412357

[58]. Rakibul, H. (2025). The Role of Business Analytics In ESG-Oriented Brand Communication: A Systematic Review Of Data-Driven Strategies. *ASRC Procedia: Global Perspectives in Science and Scholarship*, *1*(01), 1096– 1127. https://doi.org/10.63125/4mchj778

[59]. Rakibul, H., & Samia, A. (2022). Information System-Based Decision Support Tools: A Systematic Review Of Strategic Applications In Service-Oriented Enterprises. *Review of Applied Science and Technology*, *1*(04), 26-65. https://doi.org/10.63125/w3cevz78

[60]. Ren, P., Li, S., Hou, W., Zheng, W., Li, Z., Cui, Q., Chang, W., Li, X., Zeng, C., Sheng, M., & Zhang, Y. (2021). MHDP: An efficient data lake platform for medical multi-source heterogeneous data. In *Web Information Systems and Applications (WISA 2021), Lecture Notes in Computer Science* (Vol. 12999, pp. 727–738). https://doi.org/10.1007/978-3-030-87571-8_63

[61]. Ren, P., Mao, Z., Li, S., Xiao, Y., Ke, Y., Yao, L., Lan, H., Li, X., Sheng, M., & Zhang, Y. (2021). Intelligent visualization system for big multi-source medical data based on data lake. In *Web Information Systems and Applications (WISA 2021), Lecture Notes in Computer Science* (Vol. 12999, pp. 706–717). https://doi.org/10.1007/978-3-030-87571-8_61

[62]. Reza, M., Vorobyova, K., & Rauf, M. (2021). The effect of total rewards system on the performance of employees with a moderating effect of psychological empowerment and the mediation of motivation in the leather industry of Bangladesh. *Engineering Letters*, *29*, 1-29.

[63]. Rieke, N., Hancox, J., Li, W., Milletari, F., Roth, H. R., Albarqouni, S., & Cardoso, M. J. (2020). The future of digital health with federated learning. *NPJ Digital Medicine*, *3*, 119. https://doi.org/10.1038/s41746-020-00323-1

[64]. Saba, A. (2025). Artificial Intelligence Based Models For Secure Data Analytics And Privacy-Preserving Data Sharing In U.S. Healthcare And Hospital Networks. *International Journal of Business and Economics Insights*, *5*(3), 65–99. https://doi.org/10.63125/wv0bqx68

[65]. Saikat, S. (2021). Real-Time Fault Detection in Industrial Assets Using Advanced Vibration Dynamics And Stress Analysis Modeling. *American Journal of Interdisciplinary Studies*, *2*(04), 39–68. https://doi.org/10.63125/0h163429

[66]. Saikat, S. (2022). CFD-Based Investigation of Heat Transfer Efficiency In Renewable Energy Systems. *International Journal of Scientific Interdisciplinary Research*, *1*(01), 129–162. https://doi.org/10.63125/ttw40456

[67]. Shaikat, B. (2025). Artificial Intelligence–Enhanced Cybersecurity Frameworks for Real-Time Threat Detection In Cloud And Enterprise. *ASRC Procedia: Global Perspectives in Science and Scholarship*, *1*(01), 737–770. https://doi.org/10.63125/yq1gp452

[68]. Shaikh, S., & Aditya, D. (2021). Federated Learning-Driven Predictive Quality Analytics and Supply Chain Optimization In Distributed Manufacturing Networks. *Review of Applied Science and Technology*, *6*(1), 74-107. https://doi.org/10.63125/k18cbz55

[69]. Sheller, M. J., Reina, G. A., Edwards, B., Martin, J., & Bakas, S. (2020). Federated learning in medicine: Facilitating multi-institutional collaborations without sharing patient data. *Scientific Reports*, *10*, 12598. https://doi.org/10.1038/s41598-020-69250-1

[70]. Sittig, D. F., & Singh, H. (2010). A new sociotechnical model for studying health information technology in complex adaptive healthcare systems. *Quality & Safety in Health Care*, *19*(Suppl 3), i68–i74. https://doi.org/10.1136/qshc.2010.042085

[71]. Sittig, D. F., & Singh, H. (2012). Electronic health records and national patient-safety goals. *The New England Journal of Medicine*, *367*(19), 1854–1860. https://doi.org/10.1056/NEJMsb1205420

[72]. Stahlschmidt, S. R., Ulfenborg, B., & Synnergren, J. (2022). Multimodal deep learning for biomedical data fusion: A review. *Briefings in Bioinformatics*, *23*(2), bbab569. https://doi.org/10.1093/bib/bbab569

[73]. Sukumar, S. R., Ramachandran, N., & Ferrell, R. K. (2015). Quality of big data in health care. *International Journal of Health Care Quality Assurance*, *28*(6), 621–634. https://doi.org/10.1108/ijhcqa-07-2014-0080

[74]. Sun, Y., Guo, F., Kaffashi, F., Jacono, F. J., DeGeorgia, M., & Loparo, K. A. (2020). INSMA: An integrated system for multimodal data acquisition and analysis in the intensive care unit. *Journal of Biomedical Informatics*, *106*, 103434. https://doi.org/10.1016/j.jbi.2020.103434

[75]. Sundermann, A. J., Chen, J., Kumar, P., Ayres, A. M., Cho, S. T., Ezeonwuka, C., & Harrison, L. H. (2022). Whole-genome sequencing surveillance and machine learning of the electronic health record for enhanced healthcare outbreak detection. *Clinical Infectious Diseases*, *75*(3), 476–482. https://doi.org/10.1093/cid/ciab946

[76]. Sundermann, A. J., Chen, J., Miller, J. K., Martin, E. M., Snyder, G. M., Van Tyne, D., & Harrison, L. H. (2022). Whole-genome sequencing surveillance and machine learning for healthcare outbreak detection and investigation: A systematic review and summary. *Antimicrobial Stewardship & Healthcare Epidemiology*, *2*(1), e91. https://doi.org/10.1017/ash.2021.241

[77]. Tiase, V. L., Hull, W., McFarland, M. M., Sward, K. A., Del Fiol, G., Staes, C., & Cummins, M. R. (2019). Patient-generated health data and electronic health record integration: Protocol for a scoping review. *BMJ Open*, *9*(12), e033073. https://doi.org/10.1136/bmjopen-2019-033073

[78]. Tiase, V. L., Hull, W., McFarland, M. M., Sward, K. A., Del Fiol, G., Staes, C., & Cummins, M. R. (2020). Patient-generated health data and electronic health record integration: A scoping review. *JAMIA Open*, *3*(4), 619–627. https://doi.org/10.1093/jamiaopen/ooaa052

[79]. Tonoy Kanti, C. (2025). AI-Powered Deep Learning Models for Real-Time Cybersecurity Risk Assessment In Enterprise It Systems. *ASRC Procedia: Global Perspectives in Science and Scholarship*, *1*(01), 675–704. https://doi.org/10.63125/137k6y79

[80]. Tonoy Kanti, C., & Shaikat, B. (2022). Graph Neural Networks (GNNS) For Modeling Cyber Attack Patterns And Predicting System Vulnerabilities In Critical Infrastructure. *American Journal of Interdisciplinary Studies*, *3*(04), 157-202. https://doi.org/10.63125/1ykzx350

[81]. Torabi, F., Squires, E., Orton, C., Heys, S., Ford, D., Lyons, R. A., & Thompson, S. (2023). A common framework for health data governance standards. *Nature Medicine*, *29*, 1–3. https://doi.org/10.1038/s41591-023-02686-w

[82]. Tubaishat, A. (2017). Perceived usefulness and perceived ease of use of electronic health records among nurses: Application of Technology Acceptance Model. *Informatics for Health and Social Care*, *43*(4), 379–389. https://doi.org/10.1080/17538157.2017.1363761

[83]. Vaid, A., Jaladanki, S. K., Xu, J., Teng, S., Kumar, A., Lee, S., Somani, S., Paranjpe, I., De Freitas, J. K., Wanyan, T., Johnson, K. W., Bicak, M., Klang, E., Kwon, Y. J., Costa, A., Zhao, S., Miotto, R., Charney, A. W., Böttinger, E., & Glicksberg, B. S. (2021). Federated learning of electronic health records to improve mortality prediction in hospitalized patients with COVID-19: Machine learning approach. *JMIR Medical Informatics*, *9*(1), e24207. https://doi.org/10.2196/24207

[84]. Venkatesh, V., Thong, J. Y. L., & Xu, X. (2012). Consumer acceptance and use of information technology: Extending the unified theory of acceptance and use of technology. *MIS Quarterly*, *36*(1), 157–178. https://doi.org/10.2307/41410412

[85]. Verberk, J. D. M., Aghdassi, S. J. S., Abbas, M., Nauclér, P., Gubbels, S., Maldonado, N., Palacios-Baena, Z. R., Johansson, A. F., Gastmeier, P., Behnke, M., van Rooden, S. M., & van Mourik, M. S. M. (2022). Automated surveillance systems for healthcare-associated infections: Results from a European survey and experiences from real-life utilization. *Journal of Hospital Infection*, *122*, 35–43. https://doi.org/10.1016/j.jhin.2021.12.021

[86]. Weiner, B. J. (2009). A theory of organizational readiness for change. *Implementation Science*, *4*, 67. https://doi.org/10.1186/1748-5908-4-67

[87]. Wen, R., Li, X., Liu, T., & Lin, G. (2022). Effect of a real-time automatic nosocomial infection surveillance system on hospital-acquired infection prevention and control. *BMC Infectious Diseases*, *22*, Article 857. https://doi.org/10.1186/s12879-022-07873-7

[88]. White, S. (2014). A review of big data in health care: Challenges and opportunities. *Open Access Bioinformatics*, *6*, 13–18. https://doi.org/10.2147/oab.S50519

[89]. Xu, J., Glicksberg, B. S., Su, C., Walker, P., Bian, J., & Wang, F. (2021). Federated learning for healthcare informatics. *Journal of Healthcare Informatics Research*, *5*(1), 1–19. https://doi.org/10.1007/s41666-020-00082-4

[90]. Yang, Q., Liu, Y., Chen, T., & Tong, Y. (2019). Federated machine learning: Concept and applications. *ACM Transactions on Intelligent Systems and Technology*, *10*(2), 12:11–12:19. https://doi.org/10.1145/3298981

[91]. Yin, Y., Zeng, Y., Chen, X., & Fan, Y. (2016). The Internet of Things in healthcare: An overview. *Journal of Industrial Information Integration*, *1*, 3–13. https://doi.org/10.1016/j.jii.2016.03.004

[92]. Zeadally, S., Siddiqui, F., Baig, Z., & Ibrahim, A. (2020). Smart healthcare: Challenges and potential solutions using internet of things (IoT) and big data analytics. *PSU Research Review*, *4*(2), 149–168. https://doi.org/10.1108/prr-08-2019-0027

[93]. Zhou, L. L., Owusu-Marfo, J., Antwi, H. A., Antwi, M. O., Kachie, A. D. T., & Ampon-Wireko, S. (2019). Assessment of the social influence and facilitating conditions that support nurses' adoption of hospital electronic information management systems (HEIMS) in Ghana using the unified theory of acceptance and use of technology (UTAUT) model. *BMC Medical Informatics and Decision Making*, *19*, 230. https://doi.org/10.1186/s12911-019-0956-z