American Journal of Interdisciplinary Studies

Volume: 5; **Issue:** 3
Pages: 43– 85
eISSN: 3067-5146

MACHINE LEARNING AND SECURE DATA PIPELINE FRAMEWORKS FOR IMPROVING PATIENT SAFETY WITHIN U.S. ELECTRONIC HEALTH RECORD SYSTEMS

Md. Jobayer Ibne Saidur[1]; Aditya Dhanekula[2];

[1]. MS in Business Analytics, University of the Cumberlands, KY , USA;
Email: jobayerdu00@gmail.com

[2]. Abraham & Sons Leather LLC, Business Analyst, USA;
Email: dhanekulaaditya1@gmail.com

Doi: 10.63125/nb2c1f86
Received: *19 June 2024;* ***Revised:*** *28 July 2024;* ***Accepted:*** *23 August 2024;* ***Published:*** *29 September 2024*

Abstract

This study examined how secure data pipelines operated as quantitative determinants of machine learning (ML) reliability and patient safety outcomes in Electronic Health Record (EHR) environments among U.S. healthcare providers. A retrospective, multi-site quantitative design was applied to de-identified EHR encounter streams, linking provider-level pipeline maturity (confidentiality, integrity, availability, and data quality) with ML safety-prediction performance and EHR-derived safety endpoints. The methodology and findings were grounded in an overall review of spanning secure healthcare data pipelines, EHR-driven ML safety prediction, cybersecurity incident impacts, fairness in clinical ML, and EHR safety endpoint operationalization; this number should match the total reported across the methods and findings sections of the article. The analytic cohort comprised 12 providers contributing 184,732 adult encounters, with a median of 14,980 encounters per provider. Pipeline maturity demonstrated cross-site differentiation (mean = 72.8, SD = 9.4), accompanied by measurable variability in quality and security indicators: missingness averaged 6.4% (SD = 3.1), timestamp misalignment averaged 3.8 per 1,000 events (SD = 1.9), unit harmonization errors averaged 5.6 per 10,000 labs (SD = 2.4), audit-log completeness averaged 91.5% (SD = 4.7), encryption coverage averaged 94.2% (SD = 3.9), and downtime averaged 2.7 hours per quarter (SD = 1.4). ML models for safety prediction showed stable reliability across endpoints, with mean discrimination AUROC = 0.84 (SD = 0.03), calibration slope = 0.97 (SD = 0.06), false-alarm burden = 14.9 alerts per 100 (SD = 3.5), lead-time advantage for deterioration alerts = 3.6 hours (SD = 1.1), and cross-provider transportability loss ΔAUROC = 0.04 (SD = 0.02). Safety outcomes occurred at clinically meaningful rates: preventable adverse drug events = 1.9%, abnormal-result follow-up delays = 7.6%, deterioration/failure-to-rescue events = 2.4%, and hospital-acquired harms = 3.1%. Multilevel regression indicated that higher pipeline maturity predicted lower composite harm incidence ($\beta = -0.21$, $p < .001$), while ML reliability independently reduced harms ($\beta = -0.18$, $p < .001$). Mediation analysis showed a significant indirect pathway through ML reliability (indirect effect = −0.09, $p = .002$) alongside a remaining direct maturity effect ($\beta = -0.11$, $p = .007$). Moderation tests indicated stronger maturity-to-reliability effects under higher interoperability (interaction $\beta = 0.14$, $p = .019$). Overall, the results demonstrated a statistically linked infrastructure–analytics pathway through which secure pipelines enhanced ML reliability and corresponded to lower patient-harm burdens in EHR-driven care.

Keywords

EHR, Machine-Learning, Secure-Pipelines, Cybersecurity, Patient-Safety

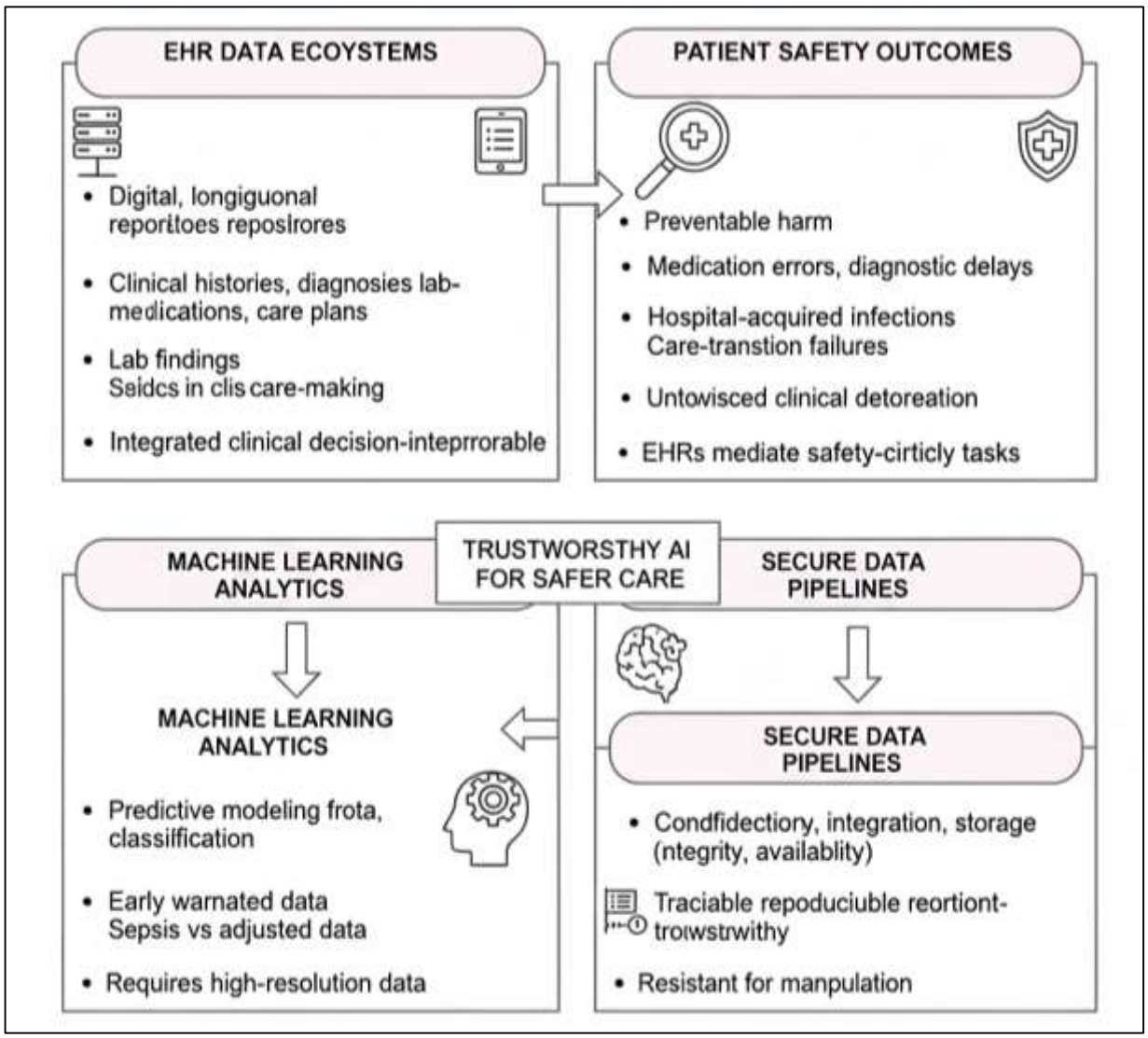
INTRODUCTION

Electronic health records (EHRs) are digital, longitudinal repositories that document clinical histories, diagnoses, medications, laboratory findings, imaging summaries, and care plans across the patient journey (Liu et al., 2020). They are designed to replace fragmented paper-based documentation with integrated, searchable, and interoperable health information that supports clinical decision-making and organizational accountability. Patient safety, in turn, refers to the systematic prevention of avoidable harm arising from healthcare delivery processes, including diagnostic, therapeutic, communication, or system failures. In modern care environments, EHRs have become a primary interface through which clinicians perceive patient status, coordinate work, and execute orders, meaning that EHR performance is tightly linked to safety outcomes. Machine learning (ML) is a family of statistical and computational methods that learn patterns from data to generate rule-based or probabilistic predictions, classifications, or recommendations. In quantitative health research, ML is commonly contrasted with traditional regression approaches because it can handle high-dimensional predictors, nonlinear associations, and complex interaction structures often present in EHR data. Secure data pipelines refer to engineered sequences for data collection, integration, storage, transformation, and analysis that embed confidentiality, integrity, and availability safeguards at each stage (Qayyum et al., 2020). In healthcare, these pipelines must align with regulatory protections for electronic protected health information, while also ensuring that analytic outputs remain traceable, reproducible, and resistant to manipulation. When EHRs supply the raw material for ML models, and secure pipelines govern how those data move through analytic systems, patient safety becomes a measurable outcome that depends simultaneously on data quality, computational accuracy, and information security. This intersection is not only technical but also sociotechnical: EHR data encode clinical workflows, human judgments, and institutional practices, so ML models trained on these data inherit both strengths and weaknesses of the underlying care system (Rajkomar et al., 2018). Quantitative research connecting ML, secure pipelines, and safety therefore begins by treating EHRs as structured and unstructured data ecosystems, recognizing safety as a set of observable adverse outcomes, and positioning security as a prerequisite for trustworthy analytics in real-world clinical settings.

Patient safety is recognized worldwide as a core dimension of healthcare quality because preventable harm remains a leading source of death, disability, and economic loss across health systems. International safety frameworks emphasize the persistent global prevalence of medication errors, diagnostic delays, hospital-acquired infections, care-transition failures, and unnoticed clinical deterioration (Srivastava et al., 2017). As countries expand digital health infrastructures, EHRs increasingly mediate safety-critical tasks such as prescribing, test ordering, allergy checks, referral communication, and discharge planning. Global studies show that EHRs can reduce certain errors through standardization and real-time decision support, while also introducing new risks through incomplete documentation, interface complexity, alert fatigue, or misaligned workflows. The widening adoption of EHRs means that safety challenges associated with digital documentation are not confined to any single country or income group. ML-based methods have rapidly entered this global environment because EHRs produce large-scale datasets capable of supporting predictive modeling at the bedside and across populations. International quantitative work has established that ML models can identify latent risk patterns for sepsis, pneumonia, ICU mortality, surgical complications, adverse drug events, and readmissions earlier than conventional scoring systems in many contexts (Srivastava et al., 2017). At the same time, cross-national evidence underscores that ML effectiveness depends on the reliability of upstream data handling, especially when clinical data are integrated from multiple sites or devices. Secure data pipelines are globally significant because breaches and integrity failures can compromise safety by distorting clinical histories, enabling identity fraud, or undermining trust in digital systems. International cybersecurity reports indicate that healthcare has become a high-value target due to the richness of EHR data, and that disruptions or manipulations of clinical databases can translate directly into unsafe care. Globally, the convergence of large EHR repositories, ML safety analytics, and exposure to cyber threats has created a shared research imperative: to quantify how secure, auditable, and high-fidelity data pipelines enable ML models to support safer clinical decisions across diverse care environments. The U.S. context is therefore situated within a worldwide transition toward digital, data-driven health systems where safety and security are mutually reinforcing

requirements (Tayefi et al., 2021).

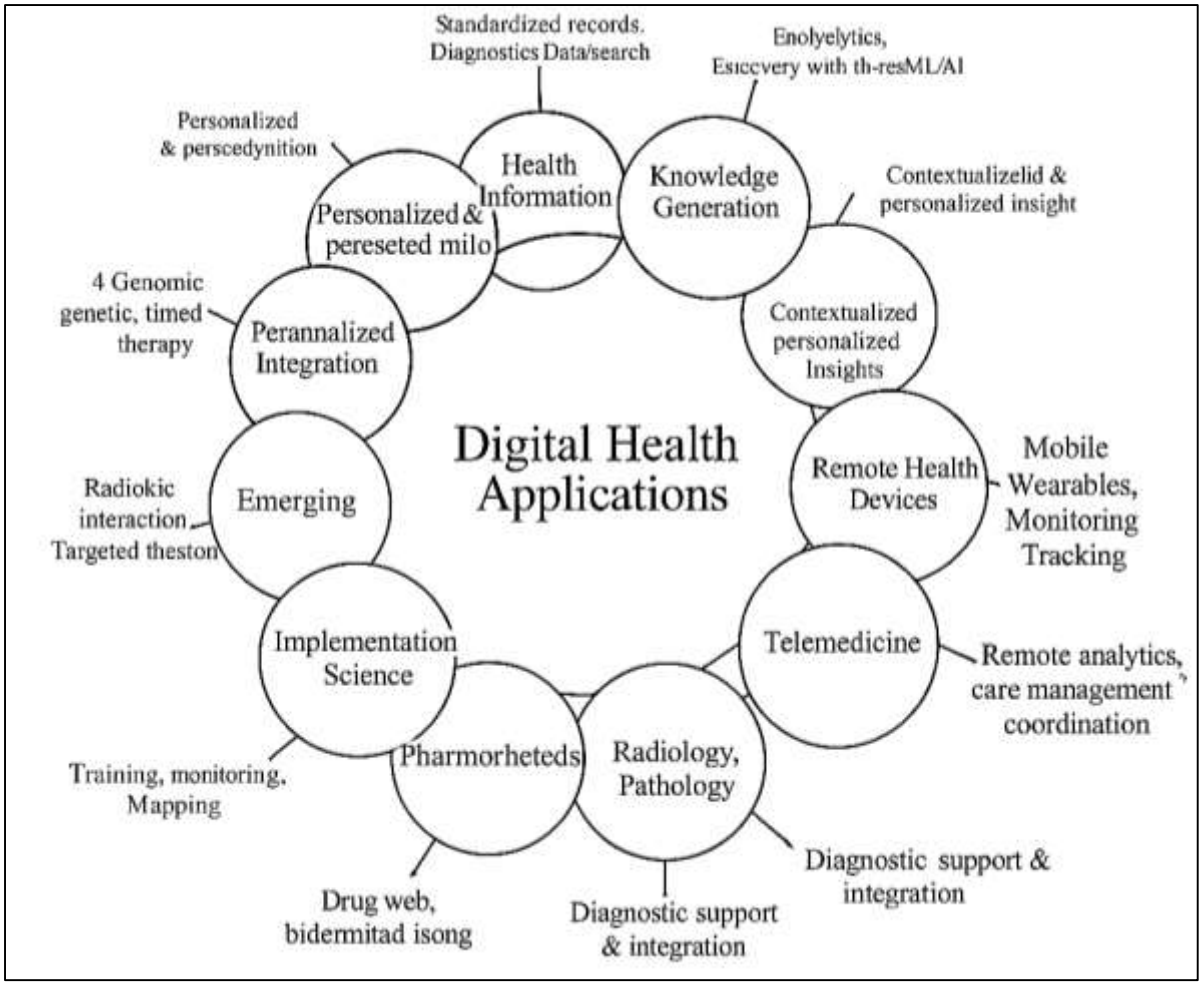
Figure 1: EHR-Driven AI and Patient Safety Framework



The U.S. healthcare system has undergone extensive digital transformation, with EHRs now embedded across hospitals, outpatient clinics, emergency departments, and long-term care facilities. Federal meaningful-use incentives and subsequent interoperability initiatives accelerated adoption, yielding vast national data infrastructures that archive billions of clinical transactions (Abdulla & Md. Jobayer Ibne, 2021; Swain et al., 2022). These records enable large-scale quantitative evaluation of safety outcomes such as medication-related harm, diagnostic error proxies, preventable admissions, hospital-acquired conditions, and mortality risk. Yet U.S. evidence also documents persistent safety burdens within digital care, including ordering errors, missing allergy documentation, delayed follow-up on abnormal tests, duplicative imaging, and workflow fragmentation across settings. EHR usability challenges remain a measurable contributor to safety risk, as complex interfaces and poorly harmonized alerts can impair clinician attention and lead to inappropriate actions (Bean et al., 2017; Ferdous Ara, 2021). The U.S. environment is additionally characterized by multi-vendor EHR markets, variable data standards, and uneven interoperability between inpatient and ambulatory systems. These structural conditions create quantitative heterogeneity in data completeness, coding practices, and clinical granularity, all of which affect the quality of downstream analytics. ML studies in U.S. settings have leveraged these large EHR datasets to produce probabilistic models for early warning, risk stratification, and adverse-event detection (Habibullah & Md. Foysal, 2021; Md Sarwar, 2021). Common applications include prediction of sepsis onset, deterioration requiring rapid response, opioid-related

harm, preventable readmission, falls, pressure injuries, and chronic disease decompensation (Md. Musfiqur & Saba, 2021; Md. Redwanul et al., 2021).

Figure 2: Evolving Application of Digital Technology in Health



In many studies, ML performance gains over baseline models are closely tied to the availability of high-resolution temporal variables, laboratory sequences, medication trajectories, and clinical notes. This dependence highlights the centrality of data pipelines that can reliably ingest heterogeneous EHR feeds, normalize them for modeling, and preserve traceability for clinical validation (Reza et al., 2021; Saikat, 2021; Tomašev et al., 2021). The U.S. safety landscape is also shaped by regulatory and organizational accountability structures, including hospital quality metrics, malpractice pressures, and public reporting of adverse events. Quantitative patient-safety research in this context thus treats EHRs not only as documentation platforms but also as statistical observatories for identifying where harm occurs and for testing whether ML models can reliably moderate that harm when integrated into care (Md Al Amin, 2022; Praveen et al., 2022; Shaikh & Aditya, 2021). ML in EHR-based safety research generally operationalizes patient harm as a measurable endpoint and uses multidimensional predictors that capture physiology, treatment, and context. Supervised learning designs dominate this literature, where labeled outcomes such as sepsis diagnosis, adverse drug events, ICU transfer, in-hospital death, or 30-day readmission serve as targets (Md Ariful & Efat Ara, 2022; Miotto et al., 2016). Logistic regression, penalized regression, decision trees, random forests, gradient boosting machines, support vector machines, and deep neural networks have all been employed to learn from structured EHR variables. More recent studies incorporate sequential deep learning to exploit the temporal ordering of vitals, labs, and medication administrations, allowing the model to learn dynamic trajectories rather than static snapshots. Natural language processing extends ML to unstructured clinical notes, capturing subtle signals about symptoms, clinician concerns, social risk, and care complexity. In

quantitative evaluations, ML model performance is typically measured through discrimination, calibration, precision–recall tradeoffs, decision-curve analysis, and external validation across hospitals or regions (Lorberbaum et al., 2016; Md Nahid, 2022; Md Sarwar Hossain & Md Milon, 2022). Safety relevance emerges not only from predictive accuracy but also from how early models can detect risk relative to clinical recognition, thereby enabling timely intervention windows. ML applications to medication safety frequently model drug–drug interactions, dosage anomalies, and adherence gaps using combined pharmacy and clinical data. Diagnostic safety research uses ML to flag delayed cancer diagnoses, missed infections, or atypical presentations by analyzing pattern deviations in testing and follow-up (Md. Mominul et al., 2022; Mortuza & Rauf, 2022; Rakibul & Samia, 2022). Hospital-acquired harm models leverage EHR time series to predict infections, venous thromboembolism, bleeding events, or preventable deterioration. Across these domains, ML models rely on stable and reproducible data representations; even small distortions in laboratory units, timestamp alignment, or medication coding can degrade model outputs. Consequently, the quantitative promise of ML in patient safety is structurally dependent on the integrity of EHR-derived datasets and on the workflows that transport data from clinical systems into analytic environments (Latif et al., 2020; Saikat, 2022; Tonoy Kanti & Shaikat, 2022). ML for safety therefore functions as an applied statistical layer embedded within broader data engineering architectures, where pipeline quality determines whether learned risk signals can be trusted, replicated, and clinically actionable (Arfan et al., 2023; Ferdous Ara & Beatrice Onyinyechi, 2023).

Secure data pipelines in healthcare integrate engineering controls that protect EHR-derived data throughout the analytic lifecycle. In quantitative research, pipelines begin with extraction from source systems such as hospital EHR databases, laboratory information systems, radiology platforms, and bedside devices (Hurst et al., 2022; Mohammad Mushfequr & Ashraful, 2023; Mst. Shahrin & Samia, 2023). Security principles require role-based access, least-privilege permissions, encryption during transfer and storage, and continuous audit logging so that data lineage remains verifiable. Pipeline design also entails structured data curation, including de-identification or limited data-set formation, harmonization of clinical terminologies, and validation of missingness patterns. Integrity safeguards are critical because patient-safety models depend on accurate temporal sequencing and correct clinical values; corruption or silent alteration of records can yield false risk estimates (Md. Hasan & Rakibul, 2024). U.S. healthcare faces substantial cyber risk, including ransomware, credential compromise, insider misuse, and third-party vendor vulnerabilities, each capable of interrupting pipelines or contaminating analytic datasets. Quantitative evaluations of security incidents show that service disruptions and data unavailability can delay care, divert clinicians to manual processes, and increase medication or diagnostic error rates. Security is thus not only a privacy obligation but a safety determinant. Modern pipelines often incorporate automated anomaly detection on incoming data streams, checksum validation, and version-controlled transformations to prevent untracked changes. Reproducibility frameworks further ensure that ML models can be retrained and audited, a requirement in safety-sensitive environments (Lindberg et al., 2020). For multi-site studies, secure federated architectures and controlled data enclaves enable cross-institutional ML development while limiting raw data exposure. In these setups, pipeline security preserves patient confidentiality and reduces the chance that analytic outputs are biased by incomplete or selectively altered data. By quantifying pipeline reliability, breach rates, downtime episodes, and data-quality consistency, researchers can empirically link security engineering to the trustworthiness of ML safety models. Therefore, in a quantitative paper, secure pipelines are treated as measurable infrastructural variables that condition the validity of EHR-based ML predictions and the stability of safety monitoring systems (Liu et al., 2019).

The safety value of ML models depends on how they are embedded into EHR-mediated workflows. Quantitative implementation research examines alert timing, interface location, and clinician response patterns as measurable mediators between prediction and outcome (Tomašev et al., 2021). Early warning models for sepsis or deterioration, for example, are evaluated not only by area-under-curve metrics but also by lead time gained before clinical escalation, rates of alert acknowledgment, and reductions in adverse endpoints. Medication-safety ML systems are assessed through changes in prescribing error incidence, override frequencies, and downstream harm rates. Diagnostic-assist

models are often measured through shifts in test ordering appropriateness, follow-up completion, and stage at detection for time-sensitive conditions. These evaluations treat workflow integration as a set of observable variables: exposure to model outputs, behavioral compliance with recommendations, and adaptation over time. A recurring finding across quantitative studies is that model performance can drift when clinical practice patterns change or when data capture shifts due to new templates, billing codes, or device integrations (Luz et al., 2020). Secure pipelines mitigate this risk by ensuring stable feature extraction, consistent timestamps, and traceable updates, allowing drift to be detected and corrected. Another quantitative concern is fairness and subgroup safety: ML models can exhibit differential error rates across race, sex, age, insurance status, or comorbidity profiles because EHR data reflect historical access and treatment patterns. Safety-focused ML research therefore reports stratified performance, calibration across subgroups, and error decomposition to identify where risk estimates may be unsafe. Clinician trust, measured through adoption rates and feedback, also influences safety impact; pipelines that maintain transparency, lineage, and auditability help sustain that trust. In U.S. providers, ML integration is situated within EHR governance committees, quality-improvement dashboards, and compliance monitoring, producing a measurable institutional setting for safety outcomes (Callahan et al., 2019). Quantitative patient-safety studies at this intersection examine how securely engineered data flows support reliable ML signals, how clinicians interpret those signals within EHR tasks, and how these sequences correspond to changes in observed harm.

A quantitative examination of machine learning and secure EHR data pipelines for patient safety among U.S. healthcare providers is grounded in three linked empirical premises (Hill et al., 2019). First, EHRs create standardized but heterogeneous digital traces of clinical care that can be statistically analyzed to measure safety outcomes, risk factors, and near-miss events. Second, ML methods allow these traces to be converted into predictive or surveillance tools capable of identifying high-risk situations earlier or more precisely than conventional rules. Third, the validity and safety relevance of ML outputs depend on the security, integrity, and reproducibility of the data pipelines that deliver EHR information into analytic systems and return model results to clinicians. In U.S. settings, where multi-vendor EHR environments and complex interoperability networks are routine, pipelines function as technical and governance structures that determine whether data are complete, accurately linked, and protected from unauthorized modification (Zheng et al., 2020). Quantitative measurement in this domain therefore draws simultaneously on clinical outcome datasets, ML performance metrics, and pipeline-security indicators. Clinical outcomes may include rates of preventable adverse drug events, deterioration episodes, sepsis-related mortality, readmissions, or hospital-acquired conditions documented through EHR-derived algorithms. ML indicators include discrimination, calibration, false-alarm rates, lead times, and stability under external validation. Pipeline indicators include data completeness scores, transformation error rates, encryption and access-control compliance, audit-log coverage, system downtime, and incident exposure measures. The study of these components within U.S. providers is also situated within national regulatory expectations for privacy and security, along with provider-based accountability for safety metrics (Seh et al., 2022). Taken together, the research problem centers on quantifying how securely engineered data pipelines condition the performance and reliability of ML models, and how this combined infrastructure corresponds to measured patient-safety outcomes in EHR-driven care. This framing positions ML and security not as independent innovations but as statistically interdependent determinants whose joint behavior can be tested empirically using provider-level EHR data (Ellis et al., 2019).

The objective of this quantitative study is to empirically examine how machine learning models and secure data pipeline architectures jointly contribute to enhancing patient safety within Electronic Health Record (EHR) environments among U.S. healthcare providers. Specifically, the study aims to measure the extent to which the quality, integrity, and security of EHR-derived data pipelines influence the predictive performance, stability, and clinical reliability of machine learning–based safety tools. A central objective is to quantify relationships between pipeline security characteristics—such as access control rigor, encryption compliance, auditability, data lineage completeness, transformation error rates, and system downtime—and the accuracy and calibration of machine learning models used for patient-safety risk detection. In parallel, the study seeks to determine whether stronger secure pipeline performance is statistically associated with improved patient-safety outcomes, operationalized through

measurable indicators such as reduced incidence of preventable adverse drug events, fewer delayed clinical escalations, lower rates of avoidable deterioration, improved timeliness of abnormal test follow-up, and decreased occurrence of hospital-acquired harms documented in EHR data. Another objective is to compare machine learning model effectiveness across provider settings with differing pipeline maturity levels, thereby identifying whether pipeline robustness moderates the relationship between model use and safety performance. The study further aims to evaluate how machine learning safety outputs behave across patient subgroups and clinical units when derived from securely governed pipelines, using stratified quantitative assessments to test stability and detect potential disparities in error rates. At the provider level, the study intends to model the combined explanatory power of machine learning performance metrics and secure pipeline indicators in predicting variation in safety outcomes, offering a statistically grounded account of their interdependence. Finally, the objective includes establishing a reproducible measurement framework that aligns EHR data engineering variables with machine learning analytics and patient-safety endpoints, enabling clear quantitative testing of how secure, high-fidelity data flows support dependable machine learning–driven safety surveillance in real-world U.S. healthcare delivery systems.

LITERATURE REVIEW

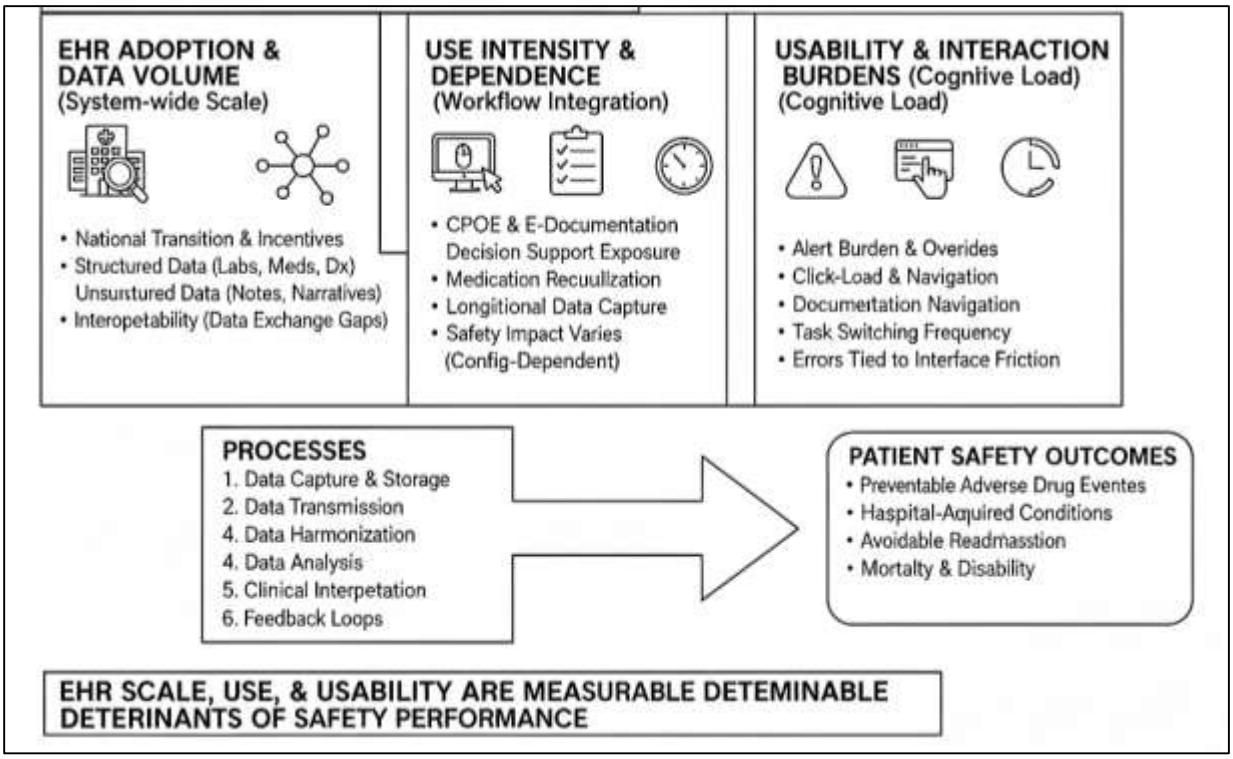
This Literature Review synthesizes quantitative scholarship at the intersection of machine learning, secure data pipelines, and patient safety within Electronic Health Record (EHR) ecosystems, with an emphasis on U.S. healthcare providers. The review is structured to establish how measurable properties of EHR data (completeness, fidelity, temporal granularity, coding consistency) enable or constrain predictive modeling for safety outcomes, and how engineered security controls across data pipelines (confidentiality, integrity, availability, auditability) condition the validity and clinical trustworthiness of those models. Because patient safety is operationalized through observable adverse outcomes, the review prioritizes studies that quantify preventable harm, model performance, and pipeline security indicators using statistically testable designs. The section also maps the methodological evolution from traditional regression and rules-based surveillance to high-dimensional machine learning approaches, highlighting how different algorithms perform under real-world EHR conditions. In addition, the review examines the quantitative evidence on pipeline vulnerabilities, cyber incidents, and data-governance failures that measurably disrupt care or degrade analytic reliability. Collectively, this synthesis creates the empirical foundation for understanding the statistical interdependence between secure data infrastructures and machine learning safety tools, clarifies dominant measurement strategies used in prior work, and positions the present study within established quantitative debates about prediction accuracy, reproducibility, bias, and safety impacts in U.S. EHR-driven clinical environments.

EHRs as safety-critical data ecosystems

Electronic Health Records (EHRs) are integrated digital systems that store, retrieve, and exchange patient information across time and care settings, including structured fields such as diagnoses, medications, laboratory values, and procedures, along with unstructured documentation such as clinician notes, discharge summaries, and imaging narratives (Golas et al., 2018). As safety-critical data ecosystems, EHRs do more than archive clinical histories; they function as the primary operational interface through which clinicians perceive patient status, communicate decisions, order treatments, and coordinate handoffs. In this framing, patient safety is defined as the prevention and reduction of avoidable harm that emerges from diagnostic, therapeutic, monitoring, or communication failures during healthcare delivery. Because EHRs mediate these processes, the quality and usability of EHR data streams become inseparable from safety outcomes. A safety-critical ecosystem implies that errors in data capture, display, transmission, or interpretation can generate measurable adverse events, while reliable and well-aligned data infrastructures can reduce harm by enabling timely access to accurate information and standardized care pathways (Akrivopoulos et al., 2017). Quantitative literature positions EHR ecosystems as socio-technical structures where technical features (data architecture, interoperability, alert logic, interface design) and human workflow behaviors (documentation, order entry, response to alerts, cross-team coordination) co-produce safety performance. The ecosystem perspective also clarifies why EHR impact is evaluated at multiple levels: (a) systemwide adoption and connectivity, which determine how much data are available and shareable; (b) use intensity, which

determines how deeply clinicians depend on EHRs for safety-critical tasks; and (c) usability and interaction burdens, which shape cognitive load and error likelihood. Within U.S. healthcare, these definitions matter because national EHR expansion has built large repositories that are both surveillance platforms for quantifying harm and operational tools that can either mitigate or propagate risk (Barbosa et al., 2018). Thus, EHRs as safety-critical ecosystems are defined by their dual role as data infrastructures and clinical work environments, making their adoption scale, depth of use, and interface properties measurable determinants of safety endpoints.

Figure 3: EHR, ML, and Patient Safety



Quantitative evidence on U.S. EHR adoption shows a steep national transition from partial digital documentation to near-universal deployment across hospitals and a large majority of ambulatory practices. Adoption curves in the literature typically quantify EHR penetration by provider category (large academic hospitals, community hospitals, safety-net facilities, physician offices, and specialty clinics), showing earlier uptake in large systems and slower diffusion in small practices and resource-constrained organizations (McDermott, 2016). These adoption trajectories are frequently linked to federal incentive structures and reporting programs, producing measurable inflection points during years of policy enforcement. Alongside adoption growth, datasets expanded in both volume and complexity. Structured data increased through standardized fields for labs, medications, allergies, problem lists, and billing codes, enabling large-N statistical modeling across populations. Simultaneously, unstructured data grew through routine clinical narrative entry, resulting in measurable increases in text volume per encounter and higher variability in documentation density across specialties (Ho et al., 2017). Quantitative studies document how this dual expansion created "mixed-modality" EHR repositories that include time-stamped physiological measures, longitudinal medication sequences, and narrative descriptions, allowing researchers to track safety outcomes at high temporal resolution. Interoperability indices in the literature measure how much information is exchanged between organizations and across care transitions, often using indicators such as participation in health information exchanges, the proportion of encounters with external data available, or the completeness of cross-site medication and lab histories. Findings consistently show that interoperability lags behind adoption, producing quantifiable gaps in data continuity even where EHR use is universal. These gaps are not described abstractly in quantitative work; they are modeled

as missingness rates, cross-system discordance in codes, and discontinuities in longitudinal patient records (Trout et al., 2022). Therefore, the national adoption curve is interpreted not only as a diffusion of software but as a measurable expansion of safety-relevant datasets whose scale, structure, and connectivity vary by provider type and region.

Beyond whether EHRs exist in an organization, quantitative research measures how intensively they are used and links that intensity to safety outcomes. Use intensity is operationalized through metrics such as clinician dependence on computerized provider order entry, frequency of medication reconciliation, rates of decision-support exposure, completeness of electronic documentation, and depth of longitudinal data capture per patient (Gallego et al., 2015). Statistical models examining these measures show that higher EHR use intensity is associated with detectable shifts in patient safety indicators, though the direction and size of associations vary with workflow alignment and system configuration. Safety endpoints commonly analyzed include rates of preventable adverse drug events, medication ordering errors, duplicate testing, delayed follow-up on critical results, avoidable readmissions, and mortality proxies such as in-hospital death or rapid deterioration leading to ICU transfer. Studies using interrupted time-series, difference-in-differences, and multi-site regression frequently attribute measurable error reductions to increased electronic ordering and standardized documentation, especially in medication safety and in the timeliness of clinical communication (Rochefort et al., 2015). At the same time, quantitative work identifies measurable adverse patterns where intensity increases without supportive design, such as heightened alert overrides, increased documentation time, or more frequent mis-selections in ordering interfaces. These patterns appear statistically in the form of higher error odds during high-load periods or among units with heavier alert exposure. Another recurring quantitative focus is the distinction between direct safety impacts and second-order effects mediated through process reliability. For example, EHR intensity is linked to improved completeness of vital-sign trajectories and medication histories, which then correlate with earlier detection of deterioration and more stable discharge planning. In multi-provider comparisons, intensity-safety associations are moderated by interoperability and data quality, suggesting that the same intensity metric can yield different safety results depending on whether clinicians are working with full information or fragmented records (Wronikowska et al., 2021). Overall, the literature treats EHR use intensity as a measurable organizational behavior variable that statistically co-varies with both reductions and elevations in safety risk, depending on how EHR functions are embedded into care.

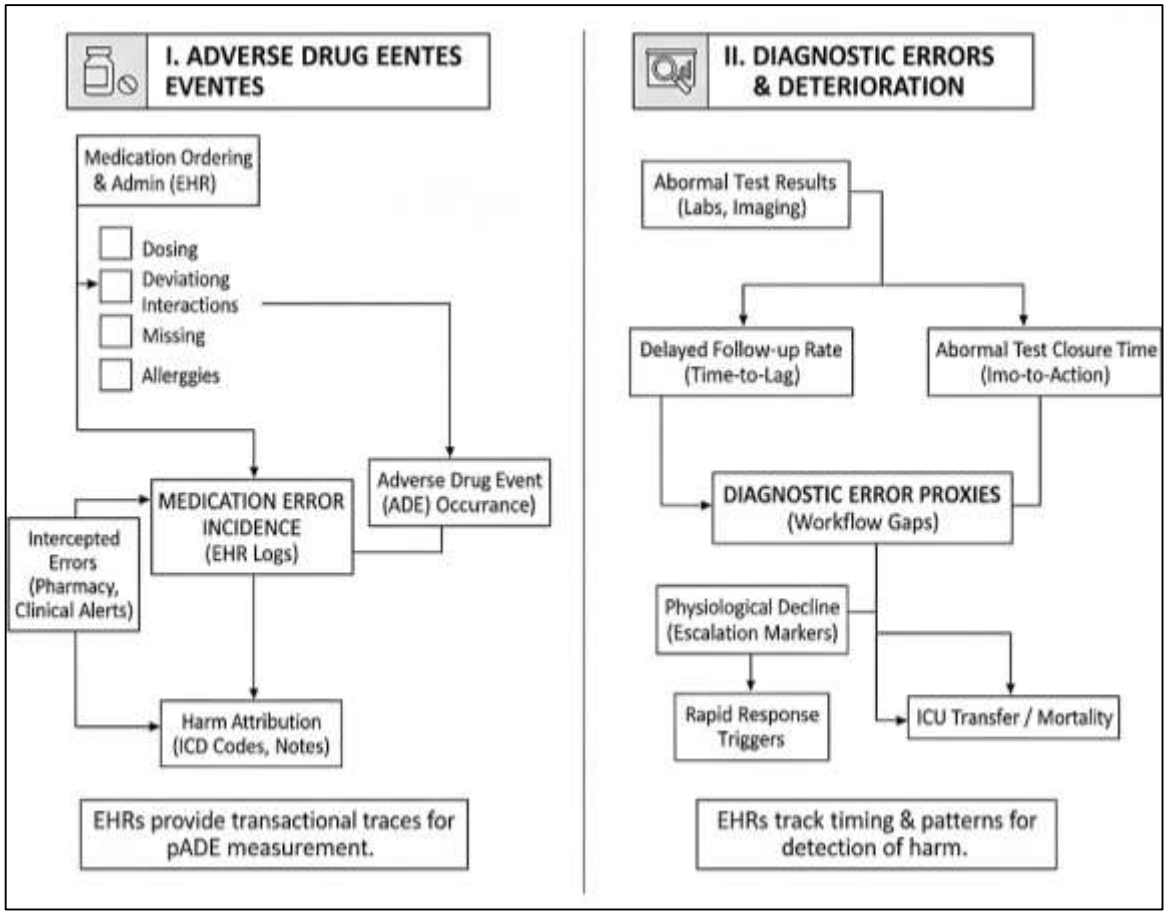
Quantitative usability research treats EHR interaction patterns as measurable safety risk factors because the interface is the terrain where clinicians translate information into action. Usability is operationalized through alert burden counts, alert firing rates per encounter, override proportions, click-load measures (such as clicks per order or per minute), navigation depth, documentation latency, and task switching frequency (Melnick et al., 2020). These metrics allow researchers to statistically link interface friction to safety outcomes. High alert burden, for instance, is repeatedly associated with measurable override escalation, and models show that override saturation correlates with later missed critical warnings and preventable medication-related harm. Click-load and navigation complexity are measured as time-on-task or interaction density, and higher values are associated with increased odds of ordering errors, delayed response to abnormal results, or incomplete documentation. Documentation latency is another quantitative marker, capturing delays between care delivery and record completion; studies connect longer latency to missed follow-ups, weaker care-transition communication, and higher preventable readmission risk. Usability burdens also appear in quantitative analyses of clinician workload and cognitive strain, where EHR time-pressure indicators correlate with error rates during peak activity windows (Khairat et al., 2019). Importantly, usability metrics are not treated as isolated technical scores; they are modeled within sociotechnical designs that account for unit type, clinical complexity, staffing levels, and training patterns. The literature shows measurable variation in usability burden across specialties, with emergency and intensive care settings often experiencing higher alert densities and more time-critical navigational demands. In comparative studies, organizations with lower quantified burdens tend to show more stable safety performance, while those with heavier burdens exhibit wider variance in error outcomes and larger gaps between intended and actual decision-support benefits. Thus, the quantitative usability literature frames EHR interface characteristics as statistically observable contributors to safety risk, functioning through alert fatigue, cognitive overload, interaction

errors, and delayed information closure (Windle et al., 2021).

Patient Safety Outcomes Detectable in EHR Data

Preventable adverse drug events (pADEs) represent one of the most intensively quantified patient-safety outcome categories detectable within EHR data because medication processes are highly digitized and leave rich transactional traces (Murphy et al., 2019).

Figure 4: EHR Data Safety Typology



The literature defines pADEs as injuries or clinical harms resulting from medication errors that could have been avoided through correct prescribing, dispensing, administration, or monitoring. Quantitative typologies in EHR-based studies typically partition medication safety into measurable subdomains that map to EHR artifacts. Medication error incidence is commonly derived from computerized provider order entry logs, medication administration records, pharmacy verification trails, and clinical outcome documentation, allowing researchers to estimate error rates per order, per patient-day, or per admission. Dosing deviation rates are operationalized through comparisons between prescribed doses and guideline-based dosing ranges adjusted for patient factors such as renal function, age, or weight, producing countable deviation events (Iscoe et al., 2022). Drug–drug interaction flags are quantified using decision-support trigger frequencies, override rates, and subsequent harm-coded events, treating alert logic as a measurable upstream indicator of risk exposure. Harm attribution fractions extend this typology by estimating what proportion of observed adverse outcomes can be linked to specific drug classes, interaction types, or timing windows, using EHR-derived case identification and structured chart review algorithms. Across studies, EHRs enable statistical differentiation between near-miss errors, intercepted errors, and errors resulting in injury, because timestamps allow reconstruction of medication timelines and clinical responses. The synthesis of this literature shows that pADE measurement is strongest when EHR data integrate prescribing, administration, laboratory monitoring, and diagnosis codes, making medication safety a multi-layered,

quantifiable pathway rather than a single event (Castellano et al., 2019). This typology provides a numeric scaffold for comparing safety across units and providers by aligning harms with discrete EHR-recorded medication process failures.

Diagnostic error is more difficult to observe directly than medication errors, so EHR-based quantitative typologies rely on validated proxies that capture breakdowns in diagnostic reasoning and follow-up processes (Strudwick et al., 2022). The literature defines diagnostic error broadly as missed, delayed, or incorrect diagnosis that meaningfully affects patient outcomes. Because EHRs document diagnostic work as sequences of testing, referrals, communications, and results management, quantitative research has organized diagnostic-safety measurement around workflow closure indicators. Delayed follow-up rates are computed from time gaps between abnormal test results and documented clinician response, subsequent ordering, referral completion, or treatment initiation, enabling estimation of delay prevalence by condition type or care setting. Abnormal test closure time is another proxy that measures how long critical lab or imaging findings remain unacknowledged or unresolved in the EHR, often operationalized through electronic inbox logs, result acknowledgment fields, or follow-up order placement (Abbott & Weinger, 2020). Missed diagnosis detection models represent a more recent typology in which algorithms identify patterns suggestive of diagnostic failure, such as repeated visits for the same symptoms, late-stage diagnosis following prior related encounters, or discordance between early signs and final diagnostic coding. These models use EHR event histories, natural language cues, and claims-linked outcomes to assign probable diagnostic error likelihood at scale. The literature consistently emphasizes that diagnostic error proxies gain validity when multiple EHR signals converge, for example when a delayed abnormal result coincides with later emergency admission or advanced disease coding (Lindberg et al., 2020). As a result, EHR-based diagnostic safety typologies translate a traditionally qualitative safety domain into observable, countable process-and-outcome indicators rooted in measurable workflow timing and pattern deviation.

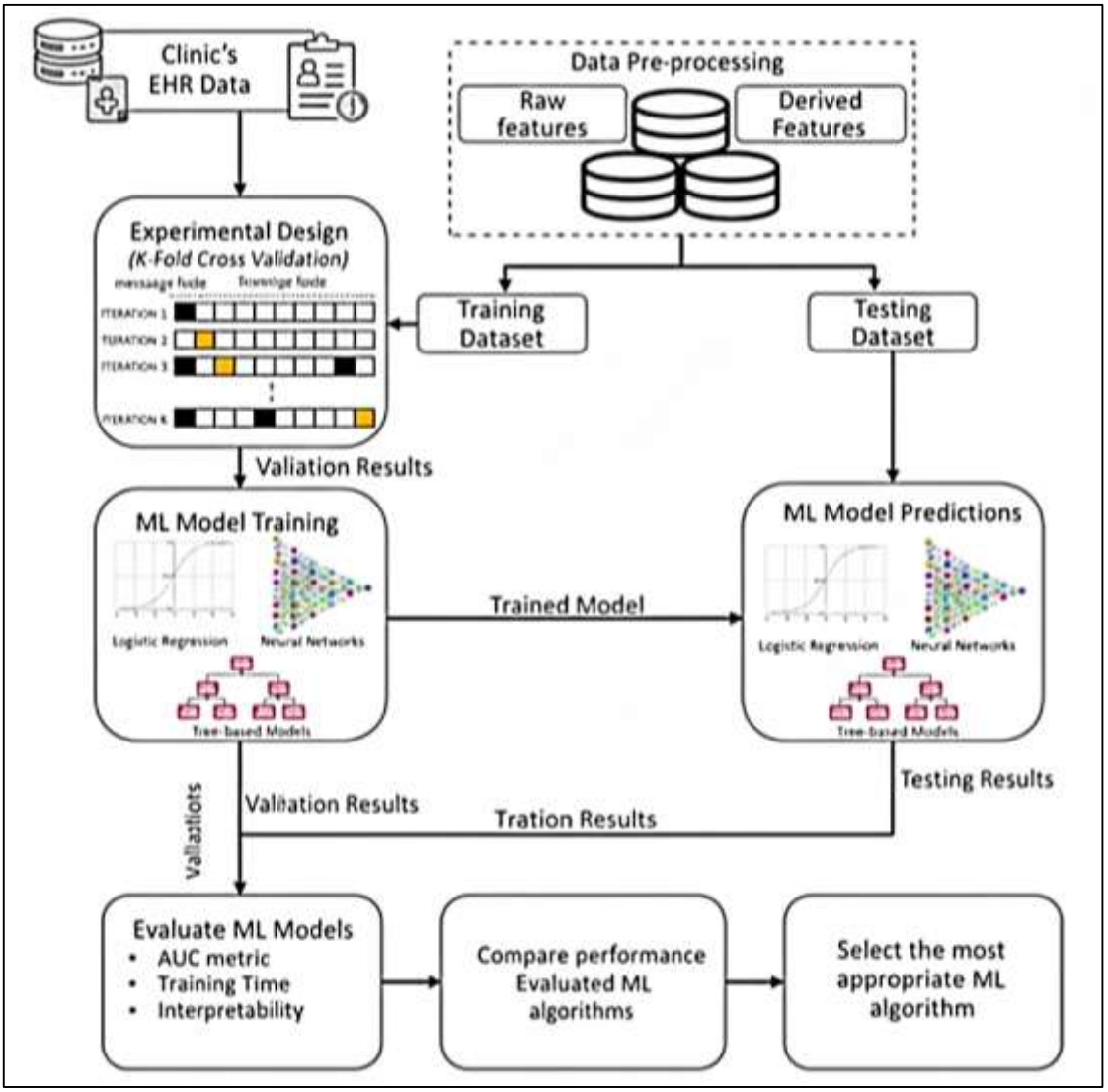
Clinical deterioration and failure-to-rescue outcomes are well suited to EHR-based quantitative typologies because they are time-sensitive phenomena that generate dense physiological and operational data. Deterioration is typically defined in the literature as worsening patient status that requires escalation of care, such as rapid response activation, ICU transfer, emergent intubation, or unexpected death (Golas et al., 2018). EHR-derived typologies classify these events through statistically observable escalation markers. Rapid response triggers are quantified by counting activations, identifying their temporal relation to abnormal vital signs or laboratory trends, and estimating trigger sensitivity relative to subsequent critical outcomes. ICU transfer rates, particularly unplanned transfers, serve as another measurable indicator of deteriorating trajectories, allowing comparative modeling of risk by unit type, staffing pattern, or clinical service line. Sepsis onset timing represents a specialized deterioration typology where EHR time series of temperature, heart rate, respiratory rate, white blood cell count, lactate, fluid orders, and antibiotic initiation are used to estimate the moment when clinical criteria are first met, thereby quantifying detection delay (Hill et al., 2019). Mortality odds linked to detection delay extend this approach by statistically associating longer recognition windows with higher death probabilities or longer ventilation duration, using EHR timestamps to reconstruct pre-event physiology. Across studies, failure-to-rescue is operationalized as mortality or severe harm following identifiable deterioration signals that were not acted upon in time, making it measurable as a relationship between early abnormal patterns and later adverse endpoints. This literature shows that EHR data support fine-grained typological separation between early-warning signals, escalation actions, and final outcomes, enabling robust quantitative evaluation of how systems detect and respond to clinical decline (Bronsert et al., 2020).

Machine Learning Models for Patient Safety Prediction Using EHR Data

Quantitative research on patient-safety prediction using EHR data consistently compares classical machine learning approaches with traditional regression-based models to determine whether added algorithmic complexity yields measurable performance gains (Xie et al., 2022). Traditional regression models, especially logistic regression and Cox-type frameworks, have long been used for safety tasks such as mortality prediction, readmission risk, sepsis screening, and adverse drug event detection because they provide interpretable coefficients and well-understood assumptions. As EHR datasets expanded in size and dimensionality, classical machine learning methods—random forests, gradient

boosting, support vector machines, and multilayer neural networks—were adopted to accommodate nonlinear relationships, high-order interactions, and irregular feature distributions that are difficult to encode in standard regression.

Figure 5: HER- Based Safety Prediction



Across safety tasks, the literature generally reports that machine learning models improve discrimination, particularly when outcomes are rare, clinically heterogeneous, or time-sensitive. These gains are often most pronounced in early warning applications where subtle multivariate patterns precede deterioration and are not easily captured through linear predictors. Calibration evidence is more mixed: some studies show that regression models remain competitive or superior in producing well-calibrated probabilities, while machine learning models may require explicit post-hoc calibration to avoid over- or under-estimation of risk. Sensitivity–specificity tradeoffs also vary by algorithm choice and clinical setting (Loftus et al., 2020). Machine learning models often achieve higher sensitivity at fixed false-alarm rates when they exploit richer feature sets and temporal granularity, yet they can also increase alert burden when thresholds are tuned aggressively. Comparative analyses emphasize that performance differences shrink in low-dimensional datasets or when predictors are limited to routinely coded variables, suggesting that classical ML advantages are conditional on the depth and quality of EHR features (Wong et al., 2018). Overall, the synthesized evidence frames classical ML as offering measurable predictive benefits over regression in complex safety surveillance contexts, while also highlighting the continued relevance of regression benchmarks for stable probability estimation and high-transparency risk scoring in clinical practice.

A central quantitative theme in the literature is that machine learning safety performance rises or falls

with the way high-dimensional EHR features are engineered, represented, and interpreted. Feature spaces in EHR safety studies typically combine static demographics, comorbidity profiles, medication orders, laboratory results, and dynamic vital-sign time series, producing thousands of candidate predictors per patient (Rajkomar et al., 2018). Lab trajectories are increasingly modeled as sequences rather than single values, allowing algorithms to learn acceleration, volatility, and persistence patterns that precede clinical harm. Vitals sequences are similarly treated as dynamic signals, where short-term deviations and long-run instability become distinguishable predictors of deterioration. Medication clusters are another common high-dimensional construct, built by grouping drugs into pharmacologic classes, interaction networks, or temporal co-administration patterns. These clusters capture safety-relevant polypharmacy signals that are hard to summarize in regression. Comorbidity embeddings, derived either from diagnosis codes or problem-list histories, compress long disease narratives into latent representations that preserve nonlinear similarity between patients and often improve risk stratification for adverse outcomes. Variable-importance analysis is widely used to interpret these complex models, relying on permutation importance, gain metrics from boosting models, SHAP-style attributions, or partial dependence summaries. Synthesized findings show recurrent dominance of certain predictor families across safety tasks—acute physiology markers, infection-related labs, sedative or opioid medication patterns, prior hospitalization trajectories, and documentation intensity proxies—yet also reveal that variable importance changes systematically by unit type, provider organization, and patient subgroup (Bates et al., 2021). Quantitative studies thus treat feature engineering not as a preprocessing step but as a measurable methodological determinant of safety prediction quality, emphasizing that robust safety models emerge from carefully structured high-dimensional representations that respect temporal order, clinical hierarchies, and coding heterogeneity.
Temporal modeling has become a defining direction in EHR-based patient safety prediction because many safety failures unfold over time rather than at isolated moments (Mamun et al., 2022). Early-warning systems for sepsis, respiratory failure, hemorrhage, cardiac arrest, or general deterioration depend on identifying evolving risk trajectories hours before clinical decompensation. The quantitative literature splits temporal approaches into two families: engineered time-window features feeding classical ML models and native sequential architectures that learn directly from ordered events. In the first family, labs and vitals are summarized across rolling windows using slopes, minima, maxima, variability, and recent-change indicators, allowing models such as gradient boosting or random forests to approximate dynamic risk (Jung et al., 2022). In the second family, recurrent neural networks, transformer-style encoders, and temporal convolutional networks ingest full sequences of vitals, labs, orders, and interventions, preserving fine-grained ordering and enabling dynamic updating of risk estimates. These temporal architectures are evaluated by the measurable lead time they provide before adverse events, with studies reporting that more expressive sequential models often extend actionable warning horizons compared with static models. Another quantitative focus is stability of risk updating: temporal models are assessed for whether predicted risk rises smoothly in line with clinical progression or fluctuates erratically, since unstable risk curves can create inconsistent decision support. Sequential models also show measurable advantages in handling irregular sampling intervals and missingness, which are ubiquitous in EHR streams (Kim et al., 2019). Yet the literature notes that temporal power is contingent on accurate timestamping, coherent unit harmonization, and consistent data capture, because temporal noise can dilute gains. Synthesized evidence therefore positions temporal ML as a key methodological class for safety prediction, offering quantifiable improvements in early detection and trajectory-based risk assessment while requiring strong data discipline to maintain reliable sequential signals.
EHR-based safety prediction increasingly integrates natural language processing (NLP) to capture risk information embedded in clinical narratives that are not available in structured fields. Quantitative NLP studies show that clinician notes contain early cues on symptom evolution, diagnostic uncertainty, social risk, medication adherence, and care complexity, all of which contribute measurable incremental predictive value (Wang et al., 2019). Common NLP pipelines include bag-of-words baselines, clinical concept extraction, word embeddings, and contextual deep language models trained on medical corpora. These representations often raise prediction performance for outcomes such as sepsis, adverse drug events, suicide risk, or missed diagnoses, especially when structured data are sparse or lag behind

real-time clinical reasoning. However, transportability becomes a major quantitative concern once NLP and high-dimensional ML models are deployed across providers. External validation studies test cross-hospital generalization error by evaluating models trained in one health system on another with different patient mixes, documentation cultures, and coding practices. Results repeatedly show that safety models experience performance decay when transferred without recalibration, with drift arising from shifts in lab ordering frequency, medication formularies, note templates, and population risk profiles (Kogan et al., 2020). Stability under heterogeneous coding systems is also treated as a measurable barrier; models can misinterpret diagnoses or procedures when code mappings differ across institutions. Transportability metrics in the literature include absolute drops in discrimination, calibration slope changes, subgroup error divergence, and drift detection rates over time. Synthesized findings suggest that models incorporating robust feature harmonization and explicit domain adaptation preserve more consistent performance, while those relying heavily on local narrative styles or site-specific workflows show sharper degradation (Corbin et al., 2022). Consequently, the literature frames NLP contributions and external validation as interlinked quantitative challenges: narrative data can meaningfully improve safety prediction, yet it also increases sensitivity to local documentation patterns, making cross-provider reliability a measurable requirement rather than an assumed property.

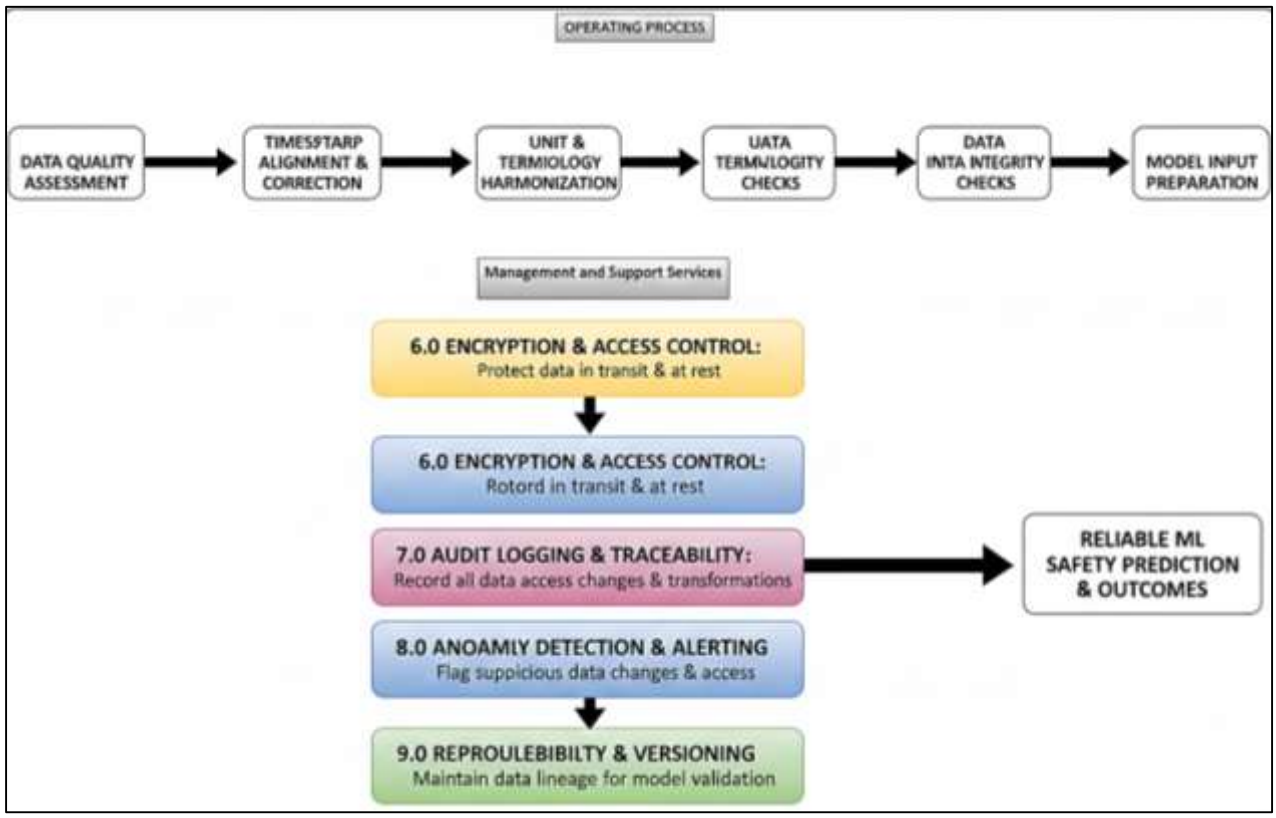
Determinants of ML Reliability

The literature on secure data pipelines in EHR-driven machine learning consistently treats data quality as a measurable precondition for reliable patient-safety prediction. Pipeline quality is usually decomposed into observable variables that occur before model training and deployment, because EHR data arrive from multiple clinical systems, devices, and documentation behaviors that create predictable distortions (Nguyen et al., 2019). Missingness rates are a dominant quantitative focus, referring not only to absent values but also to systematic gaps that vary by unit, clinician, shift, or patient subgroup. Studies show that missingness is rarely random in EHR streams; instead, it clusters around workflow constraints, test-ordering norms, and acuity-driven documentation patterns, which means that models trained on incomplete trajectories may learn biased associations or under-detect early risk signals. Timestamp misalignment is another repeatedly measured quality threat. EHR events often carry asynchronous clocks across labs, bedside monitors, medication systems, and narrative notes, producing inconsistent sequencing that degrades temporal feature extraction and early-warning accuracy. The literature quantifies these misalignments through frequency of out-of-order events, delay distributions between recording and occurrence, and cross-system clock drift. Unit harmonization errors form a third quality typology, reflecting mismatches in measurement units or coding conventions (for example, lab values reported in differing unit standards across sites), which can create silent magnitude inflation or deflation in model predictors (Rahmani et al., 2022). Quantitative pipeline studies treat harmonization error counts as critical because even low-frequency unit mismatches can trigger large model instability in high-dimensional settings. Together, these studies position pipeline data quality variables as upstream statistical determinants of model discrimination, calibration stability, and false-alarm behavior. The most synthesized finding is that high-performing safety models emerge in environments where missingness is actively characterized and represented, timestamps are normalized or corrected through engineering controls, and unit/terminology harmonization is routinely validated, because each of these variables directly conditions the reliability of learned clinical risk patterns (Davazdahemami et al., 2022).

Secure pipelines are described in the quantitative literature as layered architectures that enforce confidentiality, integrity, and controlled access without disrupting analytic utility. Rather than treating security as a binary compliance label, studies operationalize security controls through numeric indicators that allow empirical comparison across organizations (Levin et al., 2018). Encryption coverage is often measured as the proportion of EHR-derived datasets protected during transmission and at rest, and pipeline evaluations associate higher encryption coverage with lower probability of data exposure and reduced risk of downstream dataset tampering. Access-log completeness is another measurable control that captures how fully a pipeline records who accessed data, when, and for what process purpose. Studies treat completeness as a percentage of pipeline steps and data objects with audit trails, emphasizing that incomplete logs weaken reproducibility and make integrity breaches difficult to detect. Authentication failure rates provide a third quantitative control indicator,

documenting the frequency of rejected, anomalous, or suspicious access attempts that may signal adversarial probing or insider misuse (Hsu et al., 2021). In the literature, higher failure rates without mitigation correlate with greater breach vulnerability and higher likelihood of unauthorized data manipulation. Lineage traceability scores summarize whether each analytic dataset can be reconstructed from its source systems through versioned transformations, enabling verification that features used in modeling match the intended clinical signals. Quantitative pipeline research repeatedly shows that strong lineage traceability supports model auditability and safer clinical integration because it prevents "black-box" transformations from obscuring errors. Across these studies, security controls are interpreted as measurable, pipeline-level governance variables that preserve the trustworthiness of machine learning safety tools by reducing silent corruption, enabling accountability, and maintaining clinical confidence in analytic outputs (Salas et al., 2022).

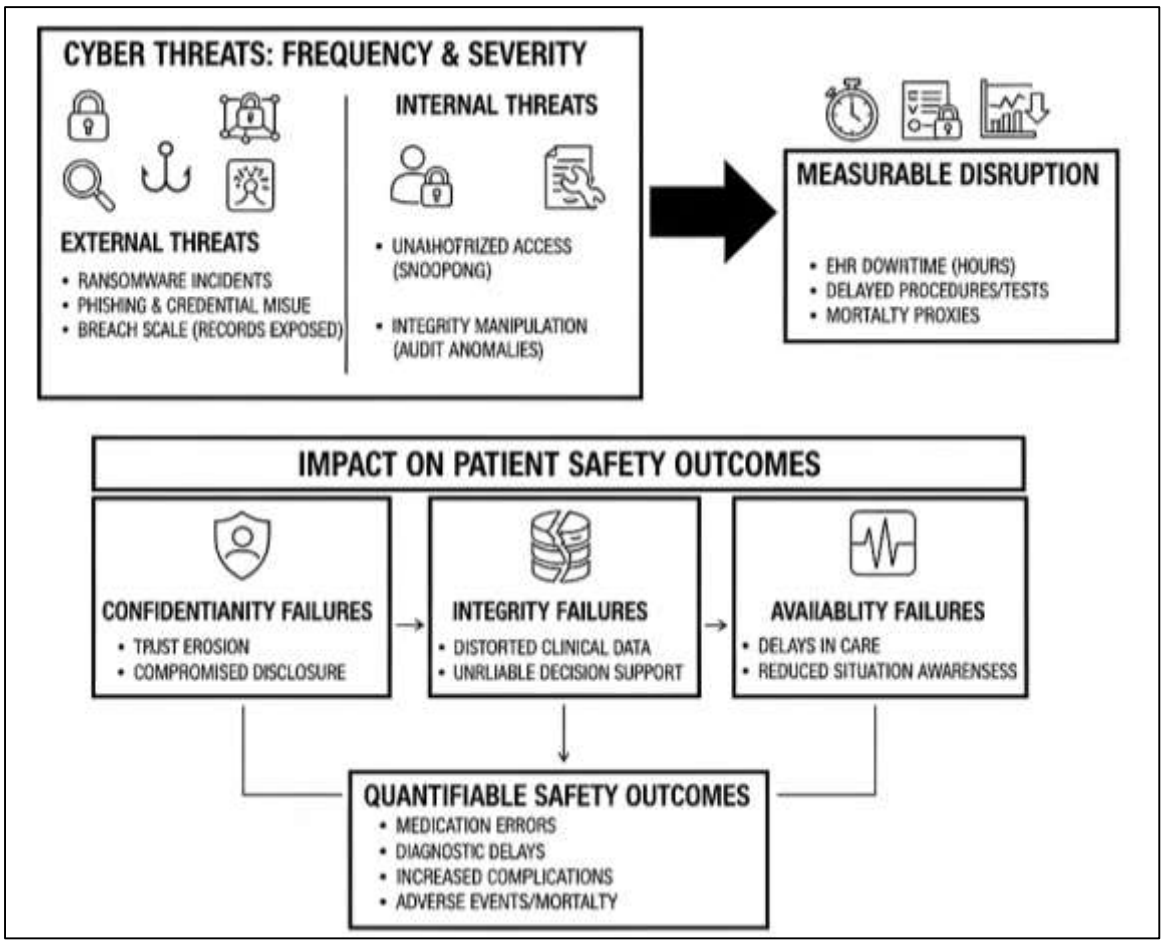
Figure 6: Secure Data Pipelines in HER-Based ML



Cybersecurity Threats and Patient Safety in EHR Settings

Quantitative literature portrays cybersecurity threats against healthcare as a high-frequency, high-severity phenomenon that is measurable through incident counts, breach scale distributions, downtime duration, and economic impact indicators (Bean et al., 2017). Ransomware event rates are among the most emphasized metrics because they directly disable EHR access and therefore interrupt safety-critical workflows. Studies tracking national and multi-year trends show that ransomware attacks have risen in both number and organizational reach, increasingly affecting multi-facility systems and producing longer service disruptions per incident. Breach scale distributions provide another statistical lens, documenting the number of patient records exposed per event and highlighting a skewed pattern where a smaller subset of attacks accounts for a disproportionate share of total records compromised. This concentration effect is repeatedly noted in provider-level analyses of reported breaches, where large integrated networks and third-party health IT vendors create high-volume exposure risk when compromised (Scott, 2021).

Figure 7: Cybersecurity as a Determinant of Patient Safety



Quantitative breach typologies further classify events by attack vector (server hacking, phishing, credential misuse, lost devices, insider misuse), allowing incidence rates to be compared across threat categories and provider types. Re-identification risk quantification adds a more technical severity metric, measuring the probability that de-identified EHR datasets can be linked back to individuals through quasi-identifiers or external databases. In this literature, severity is not limited to confidentiality loss; it is also captured through integrity and availability impacts, including the duration of EHR downtime, number of delayed services, and scale of diverted care (Briggs et al., 2022). Cost modeling studies reinforce this severity perspective by estimating financial burdens per downtime day and total national costs attributable to large-scale ransomware disruptions. Across these measures, the empirical consensus is that cyber incidents in healthcare are not rare abnormalities but statistically recurring shocks with substantial variance in scale, leaving EHR systems exposed to frequent interruptions and large-population data compromise events that are quantitatively trackable across time (Fouladvand et al., 2019).

A second body of quantitative evidence directly examines how cyber events correspond to measurable declines in patient-safety indicators. This literature most often uses pre–post incident comparisons, multi-site observational designs, and time-series analyses to estimate changes in safety outcomes around cyberattack windows (Burdick et al., 2020). Ransomware incidents in particular are associated with observable workflow disruption, including delayed access to medical histories, reduced availability of test results, slower medication ordering, and increased reliance on manual documentation, all of which can be statistically linked to safety endpoints. Studies report measurable increases in procedure delays, diagnostic turnaround times, and deferred admissions during attack periods, indicating that clinical timelines lengthen as digital infrastructure becomes unavailable. Several quantitative investigations associate cyberattack exposure with higher mortality proxies, increases in complication rates, or measurable deterioration in emergency care response times,

suggesting that EHR unavailability influences outcomes in time-sensitive clinical conditions. Medication-safety impacts are also documented through increased ordering delays, higher rates of incomplete reconciliation, and altered alerting patterns when decision support is degraded or suspended (Rojas et al., 2022). In some provider-level studies, cyber incidents correspond with higher rates of ambulance diversion and increased patient load at neighboring facilities; safety degradation therefore appears not only inside the attacked organization but also across regional networks through spillover strain. The statistical logic in this literature treats cyber events as system-level interruptions that shift clinical throughput, alter documentation fidelity, and reduce situational awareness for clinicians, which then materialize as measurable adverse outcomes. These findings reinforce that cybersecurity is a patient-safety variable rather than an exclusively administrative concern, because attack-driven disruptions display quantifiable relationships with mortality measures, delay indicators, medication error risk, and deterioration events captured in EHR and claims-linked datasets (Li et al., 2021).

Quantitative research on insider threats frames EHR misuse as a measurable risk arising from both malicious intent and routine privilege drift. Insider threat typologies in healthcare include unauthorized browsing of celebrity or acquaintance records, credential sharing, negligent handling of access tokens, deliberate data theft, and covert manipulation of clinical files (Jauk et al., 2021). The literature measures these risks through unauthorized access probability models that estimate how frequently access patterns exceed role-appropriate norms, often using baselines derived from department, shift, and patient assignment data. Audit anomaly rates are another core metric, representing the proportion of EHR log events flagged for irregular behavior such as repeated access to unrelated records, bulk exporting, after-hours browsing, or repeated failed authentication attempts. Quantitative studies show that anomaly distributions are not uniform; high-risk clusters emerge in contexts with weak access segmentation, ambiguous role definitions, or high-volume workflow pressure that makes legitimate access harder to distinguish from misuse (Rozenblum et al., 2020). Integrity compromise indicators extend insider threat measurement beyond confidentiality by tracking suspicious edits, unusual order reversals, abrupt note deletions, or inconsistent timestamp revisions. In such studies, integrity risks are treated as safety-relevant because altered records can distort medication histories, obscure diagnostic reasoning, or misrepresent clinical status. Several empirical works indicate that insider threats remain under-detected without mature auditing and analytics, not because log data are absent, but because detection systems are inconsistently deployed or are tuned for compliance rather than behavioral precision. The synthesized conclusion of this literature is that insider misuse is a statistically observable phenomenon embedded within clinical data streams, and that its magnitude can be quantitatively estimated using probabilistic access modeling, anomaly detection in audit logs, and integrity-change monitoring—each of which directly conditions the trustworthiness of EHR data used for patient-safety analytics (Landi et al., 2020).

When the frequency–severity metrics, outcome deterioration studies, and insider threat models are viewed together, the literature presents a cohesive quantitative interpretation of cybersecurity as a structural determinant of patient safety in EHR settings. High incident rates and breach scale concentration demonstrate that healthcare faces persistent external threat pressure, while time-series and pre–post designs show that attacks correspond to measurable disruptions in clinical performance and patient outcomes (Landi et al., 2020). Insider threat evidence complements this by showing that significant safety risk also originates internally through unauthorized access, privilege misuse, and integrity manipulation, each of which can degrade the reliability of clinical records even without a visible outage. The integration of these findings yields a typology where confidentiality failures affect trust and disclosure, integrity failures distort clinical truth, and availability failures delay or prevent care action; all three are tied to measurable safety consequences. Quantitative risk management frameworks in the literature encourage benchmarking these dimensions through standardized indicators: incident frequency per provider-year, median and tail downtime hours, record exposure distributions, anomaly rates per million log events, and integrity deviation counts per dataset refresh (Fang et al., 2021). Empirical studies emphasize that cyber risks are amplified by healthcare's dependence on EHR-centered workflows and decision support, making any disruption or corruption statistically consequential in medication delivery, diagnostic follow-up, and deterioration escalation.

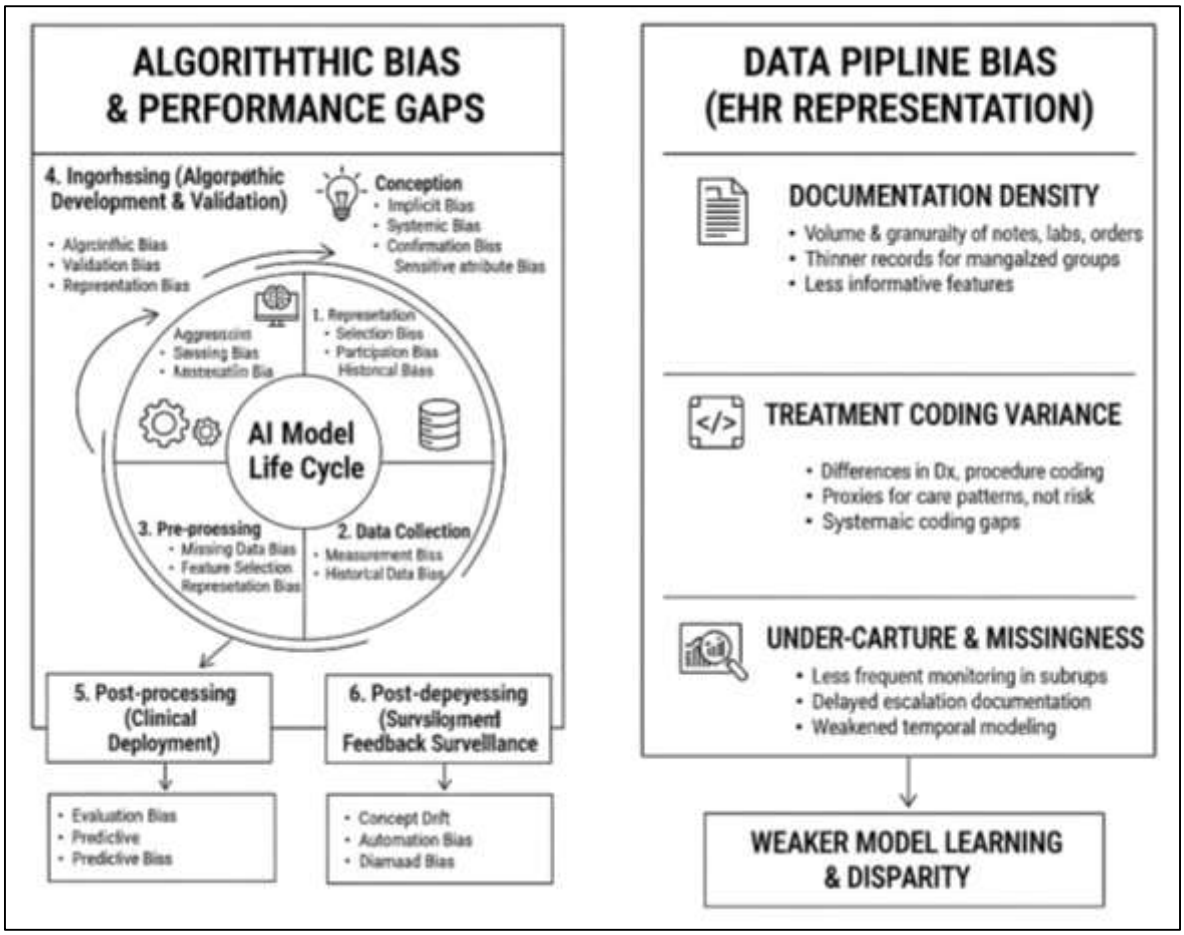
The combined literature therefore situates cybersecurity not on the periphery of patient safety but inside its quantitative measurement architecture, because cyber incidents and insider misuse translate into observable changes in safety endpoints and in the dependability of the EHR data streams upon which safety surveillance depends (Banerjee et al., 2021).

Safety Performance in ML-EHR Systems

A substantial quantitative literature evaluates bias and fairness in machine learning models trained on EHR data by measuring subgroup performance gaps across demographic and social categories (Pethani, 2021). These studies treat safety prediction as a high-stakes classification or risk-scoring task and then examine whether model accuracy is evenly distributed across groups defined by race, ethnicity, sex, age, language, insurance status, disability, or neighborhood deprivation. The most common disparity metrics include differences in discrimination across groups, operationalized through gaps in AUROC or related ranking measures. When AUROC gaps are present, they indicate that a model separates high-risk and low-risk cases more effectively for some groups than others, which becomes a safety issue when underperforming groups receive delayed warnings or fewer correctly identified harm events. Calibration inequity is measured using subgroup-specific calibration slopes or intercept differences, reflecting whether predicted probabilities correspond to observed risk at equal rates across populations (Hilton et al., 2020). Studies repeatedly show that even when overall AUROC is strong, calibration can be poor for minority or underrepresented groups, producing systematic over- or under-estimation of safety risk. False negative concentration is another key inequity pattern: subgroup analyses often reveal that missed detections cluster in specific populations, such as historically marginalized racial groups, women in atypical presentation contexts, or older adults with multimorbidity. Quantitative fairness work also examines threshold-based tradeoffs, showing that a single global decision threshold can generate unequal sensitivity and specificity across subgroups. This has led to empirical comparisons of group-specific thresholds, reweighting, and other parity-seeking strategies that are assessed through changes in subgroup error distributions rather than through global metrics alone. Overall, this literature conceptualizes fairness as a measurable safety property of ML-EHR tools, demonstrating that subgroup performance evaluation is necessary because average accuracy can mask clinically meaningful disparities in detection and prevention of harm (Abedi et al., 2021).

Beyond algorithm choice, quantitative research highlights that biases in ML–EHR safety models frequently originate in the data pipelines that generate and shape the training corpus (Xie et al., 2016). EHR data are not neutral reflections of biology; they are traces of clinical interaction patterns, documentation habits, access barriers, and institutional practices that vary systematically by patient group. Documentation density by group is widely measured as a representativeness indicator, referring to the volume and granularity of notes, labs, vitals, and orders recorded for different populations. Studies show that some groups accumulate thinner records due to fragmented access, shorter visits, language discordance, or clinician bias, which yields less informative features and weaker model learning for those populations. Treatment coding variance is another pipeline-driven bias source, capturing differences in how diagnoses, procedures, and severity are coded across subgroups. For example, equivalent clinical states may be coded differently in privately insured versus publicly insured patients, or in patients treated at different facilities, causing models to learn proxies for care patterns rather than for underlying risk (Xie et al., 2022). Under-capture rates extend this concern to missing or delayed measurements, such as less frequent laboratory monitoring in certain groups or delayed escalation documentation, which weakens temporal risk modeling and increases missed safety events in those populations.

Figure 8: Fairness and Bias in EHR



Quantitative pipeline studies also point to structural representativeness issues tied to which institutions contribute data; if training datasets overrepresent large academic hospitals and underrepresent rural or safety-net facilities, the resulting models may underperform in the settings where vulnerable patients are concentrated. The synthesis of this literature positions pipeline representativeness as a measurable upstream determinant of fairness, emphasizing that subgroup disparities are often predictable from documentation density ratios, coding consistency scores, and missingness patterns before any model is trained (Hammouda & Neyra, 2022).

Frameworks Linking Pipelines and Safety Endpoints

Integrated quantitative frameworks in the literature conceptualize secure data pipelines, machine learning (ML) performance, and patient safety endpoints as statistically linked layers within a single causal–associational chain (Qayyum et al., 2020). Rather than evaluating pipeline quality or ML accuracy in isolation, these studies model how pipeline security and data engineering maturity shape the reliability of EHR-derived features, which in turn conditions the predictive performance of safety models, and ultimately corresponds to measurable variation in clinical harm outcomes. Mediation designs dominate this strand of work because they allow investigators to test whether ML performance functions as an intervening mechanism between pipeline characteristics and safety results. In these studies, pipeline security is operationalized through measurable indicators such as access-control rigor, encryption completeness, auditability depth, lineage traceability, and system availability, while ML performance is represented by discrimination, calibration, stability, or alert precision. Safety outcomes appear as quantifiable endpoints including preventable adverse events, deterioration escalations, medication-related harm, abnormal-result follow-up delays, and hospital-acquired complication rates (Berre et al., 2022). The mediation perspective fits EHR ecosystems because pipelines govern data integrity and completeness, and these upstream properties determine whether learned risk signals align with clinical reality. Moderation models are commonly layered onto mediation to capture the observation that pipeline effects vary by institutional maturity, data standardization, interoperability,

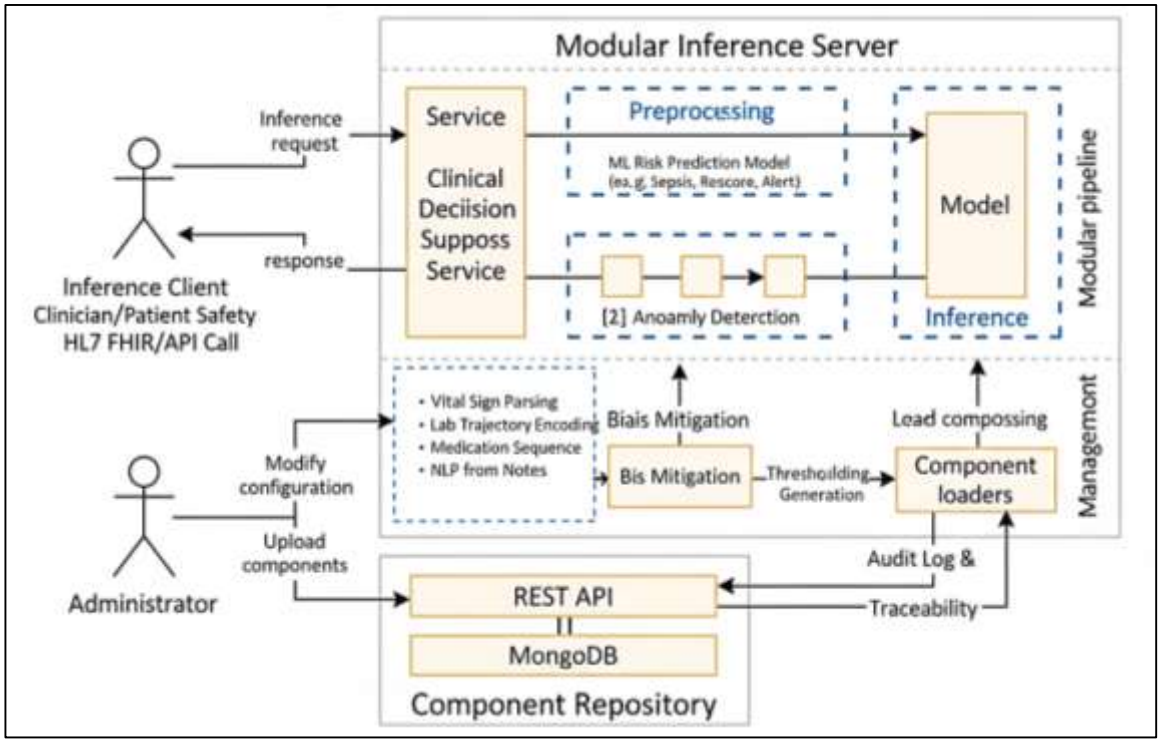
and workflow alignment. Pipeline maturity is treated as a measurable moderator that changes the strength of the pipeline-to-ML relationship, reflecting that robust pipelines reduce noise and drift, while weaker pipelines amplify instability in model outputs. This integrated modeling approach also aligns with sociotechnical patient-safety theory, which positions safety as emerging from interactions between technology, data, and clinical work. The synthesis across studies indicates that pipeline security is not merely a background condition; it is statistically embedded in the pathway through which ML tools influence safety, and its effects are detectable through structured mediation–moderation framing that ties technical governance to measurable clinical harm patterns (Zhang et al., 2022).

Another quantitative direction in the literature develops composite indices to summarize a provider's readiness to deploy ML safety tools on top of secure EHR data infrastructures. These indices emerge from the recognition that safety effectiveness depends on multi-component capacity rather than a single technical attribute (Halappanavar et al., 2020). Composite readiness scores combine measurable pipeline robustness variables (data completeness, harmonization success, audit-log coverage, access-control compliance, downtime frequency), model reliability metrics (external validation stability, calibration consistency, false-alarm balance, drift resistance), and observed safety-performance trends (changes in preventable harm rates, timeliness of escalation, reduction in medication discrepancies). The literature treats each component as necessary but insufficient alone, so indexing methods assemble weighted or standardized scales to produce an interpretable readiness continuum across provider organizations. Quantitative studies often construct these indices using factor analytic approaches, multi-criteria scoring, or benchmarking against national compliance and quality targets to ensure that the index reflects both engineering rigor and clinical impact. A recurring emphasis is that readiness is measurable at different organizational levels: unit-specific readiness differences appear where data capture or security enforcement varies by department, while systemwide readiness reflects enterprise governance and interoperability depth (Pastor et al., 2021). These indices are used to classify providers into tiers of secure-ML capacity, enabling statistical comparison of safety gains across readiness strata. Synthesized findings show that providers scoring higher on readiness indices exhibit more consistent ML performance across patient subgroups and more stable links between model alerts and real adverse-event patterns, while lower-scoring providers show greater volatility in outputs and weaker alignment with safety outcomes. Overall, composite secure-ML readiness indices provide a quantitative scaffold for evaluating integrated capacity, capturing how the co-presence of strong pipelines, reliable ML, and demonstrated safety improvements defines operational preparedness in real provider settings (Terranova et al., 2021).

A third body of integrated framework research compares the relative explanatory power of technical infrastructure variables and clinical workflow variables in driving measured safety improvements. These studies recognize that EHR safety outcomes are shaped by both engineered data systems and human task execution, so they estimate how much variance in safety endpoints is associated with each domain (Yeboah-Ofori et al., 2021). Technical predictors include pipeline security indicators, interoperability strength, data-quality stability, alert logic configuration, and ML model performance. Workflow predictors include clinician response behaviors, documentation timeliness, order-entry practices, escalation protocols, staffing ratios, and the local culture of decision-support acceptance. Quantitative frameworks examine these predictors in joint models to avoid attributing safety gains to ML performance when improvements are actually driven by parallel workflow reforms, or conversely to avoid crediting workflow alone when ML tools provide measurable early-warning advantages (Jarabek & Hines, 2019). The literature shows that technical effects are most detectable for safety outcomes that are information- and timing-dependent, such as sepsis detection lead time, abnormal-result closure, and high-risk medication surveillance. Workflow effects are more prominent for outcomes tied to bedside action execution, such as adherence to escalation pathways, response to alerts, and follow-through on safety checklists. Integrated models also demonstrate interaction patterns between domains: technical gains are larger where workflows are already structured to act on risk signals, while workflow improvements are more sustainable when technical systems provide stable, low-noise information (Ferreira & Andricopulo, 2019). This comparative-effect approach yields a balanced quantitative understanding that patient safety is not a product of infrastructure alone or

behavior alone. Instead, safety variation is explained through the joint and partially overlapping contributions of secure pipelines, ML reliability, and formal clinical process discipline, making effect-size comparison an essential part of integrated evaluation.

Figure 9: Integrated Quantitative Framework for EHR



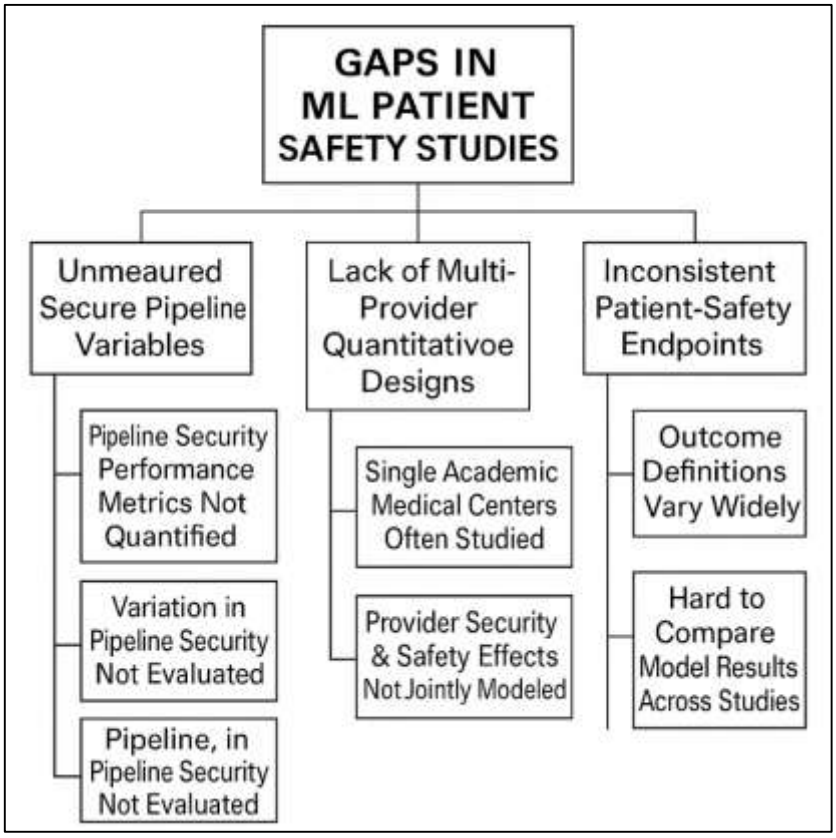
Across these strands—mediation–moderation modeling, composite readiness indexing, and comparative effect estimation—the literature converges on a shared integrated quantitative scaffold for studying ML-enabled patient safety in EHR settings (Dagenais et al., 2022). This scaffold treats pipelines, models, and outcomes as a linked measurement system, where each layer is represented by standardized numeric indicators and analyzed in combination. The integrated approach clarifies that pipeline security and data fidelity are not peripheral to safety models but statistically intertwined with how ML performs and how clinicians experience decision support. It also establishes that provider readiness is a measurable composite property reflecting technical governance, analytic reliability, and safety performance together. Finally, it frames safety improvement as a multi-source phenomenon whose drivers are empirically separable yet mutually reinforcing, requiring simultaneous modeling of technical and workflow determinants (Blay et al., 2022). The synthesis indicates that integrated frameworks reduce fragmentation in patient-safety analytics by aligning what is measured at the pipeline level (integrity, availability, auditability), what is measured at the model level (accuracy, stability, fairness), and what is measured at the clinical level (harm incidence, detection delay, escalation timeliness). This alignment supports cross-site comparison because providers can be evaluated using shared indicators within a single statistical architecture. In sum, integrated quantitative frameworks provide the literature's most cohesive method for understanding how secure data flows enable reliable ML risk signals and how these signals correspond to measurable safety outcomes inside EHR-driven care, while also acknowledging the parallel and interacting role of clinical work structures (Métris et al., 2022).

Empirical Gaps

A clear empirical gap in quantitative scholarship is that secure data pipeline variables are often referenced as background assumptions rather than measured constructs within machine learning (ML) patient-safety studies. Many EHR-based ML papers emphasize data governance, privacy, or cybersecurity as essential to trustworthy analytics, yet they rarely include numeric indicators that

capture how secure the upstream pipeline actually is. As a result, pipeline security remains analytically invisible in most predictive modeling evaluations (Hendrycks et al., 2021).

Figure 10: Empirical Gaps in ML Safety



The literature shows that ML safety studies frequently report performance metrics such as discrimination, calibration, or lead time without parallel reporting of security benchmarks like audit-log completeness, access-control maturity, encryption coverage, integrity-validation rates, anomaly detection yield, or downtime frequencies. When pipeline security is not quantified, it becomes impossible to test whether model success or failure is partly explained by the environment in which the data were engineered and protected. Another dimension of this gap is the lack of shared, standardized security indicators that could be compared across studies or providers. Security measurement is often fragmented across compliance frameworks, local IT practices, or vendor-specific controls, meaning that even when some variables are reported, they cannot be aggregated into a coherent evidence base (Xu & Saleh, 2021). The result is a methodological mismatch: ML models are treated as measurable technical interventions, while the pipeline security conditions that enable those models remain unmeasured and therefore non-explanatory. Some studies highlight data-quality variables like missingness or granularity but do not connect these to security engineering, even though integrity and availability failures are core security mechanisms that shape data quality. This under-measurement limits causal interpretation of safety effects because pipeline insecurity could introduce silent feature corruption, selective missingness, or drift that skews outcome relationships (Salay & Czarnecki, 2019). Hence, the literature positions pipeline security as important but fails to incorporate it into quantitative model architectures, producing a recurring blind spot that constrains robust inference about ML reliability and safety impact.

A second major gap is the scarcity of multi-provider quantitative designs that evaluate the joint effects of secure pipelines, ML performance, and patient-safety outcomes across diverse U.S. healthcare settings. Many published studies rely on single-site EHR datasets, often from large academic medical centers with relatively mature informatics infrastructures (Varshney, 2016). While these studies contribute detailed modeling advances, they cannot establish whether observed patterns generalize to

community hospitals, rural providers, safety-net facilities, or multi-vendor networks where pipeline maturity, interoperability, and clinical workflows differ measurably. Cross-site quantitative research is also limited by inconsistent data-sharing agreements, privacy constraints, and technical barriers to harmonizing heterogeneous EHR records, which reduces the number of studies capable of true multi-provider inference. Where multi-site work exists, it often focuses on validating algorithm performance rather than testing how pipeline security variables moderate or mediate safety outcomes (Tambon et al., 2022). As a result, the literature rarely models provider-level security indicators alongside ML outputs and safety endpoints within unified statistical frameworks. This design gap matters because provider heterogeneity is a well-established driver of variation in data completeness, timestamp fidelity, coding norms, and decision-support usage, all of which can change both ML performance and safety risk distributions. Without multi-provider designs, the field cannot quantify the extent to which secure pipelines amplify or dampen ML–safety relationships under different institutional conditions. Additionally, causal inference approaches that require cross-site variation—such as difference-in-differences, hierarchical mediation models, or policy-shock analyses—are underused in the joint pipeline–ML–safety domain (Jia et al., 2022). The outcome is that current evidence contains strong within-site predictive claims but limited population-level clarity on how secure infrastructures and ML tools co-produce safety improvement in real provider networks. Thus, the literature identifies a structural gap in study design that prevents robust estimation of joint effects across the U.S. healthcare ecosystem.

A third empirical limitation is the inconsistent operationalization of patient-safety endpoints across quantitative EHR and ML studies. Patient safety is a multi-domain construct that includes medication-related harm, diagnostic error, deterioration failures, hospital-acquired conditions, and procedural complications, yet studies vary widely in how they translate these domains into measurable outcomes (Boriani et al., 2017). For medication safety, some research uses trigger-tool algorithms based on lab anomalies and antidote use, while others rely on diagnosis codes, chart review classifications, or alert override outcomes. Diagnostic safety studies similarly diverge in proxies, ranging from abnormal-result follow-up delays to late-stage disease detection patterns or repeat-visit heuristics, each producing different event counts and risk associations. For deterioration and failure-to-rescue, some studies define endpoints using ICU transfer timing or rapid-response activation, whereas others use composite clinical deterioration scores, mortality windows, or coded sepsis onset, leading to comparability problems even within the same safety category (Lutjeboer et al., 2015). Hospital-acquired harms are also inconsistently coded, with variation in surveillance definitions, documentation templates, and normalization to exposure measures such as catheter days or ventilator time. These endpoint differences are not trivial; they alter base rates, shift model training labels, and change reported effect sizes for ML predictions. When outcomes are defined differently, it becomes difficult to compare results across studies or to meta-analyze evidence on ML safety gains. This inconsistency also complicates cross-provider research because local documentation cultures and coding practices influence how endpoints manifest in EHR data (Rinella et al., 2019). Thus, the literature repeatedly notes that safety outcomes must be standardized for quantitative accumulation, yet empirical practice remains fragmented, producing a gap between conceptual safety domains and their operational measurement in EHR-driven ML evaluations.

METHODS

Research Design

This study employed a multi-site, retrospective quantitative design using routinely collected EHR data from U.S. healthcare providers to test the statistical relationships among secure data pipeline maturity, machine learning model reliability, and patient safety outcomes. A non-experimental observational framework was selected because the study relied on naturally occurring variation in pipeline security and data quality across providers rather than researcher-controlled interventions. The unit of analysis was defined at both the patient level and the provider level, enabling linked modeling of how provider-specific secure pipeline characteristics conditioned patient-level safety prediction performance and observed harm rates. The design used a cross-sectional analytic window for provider security and pipeline indicators, paired with longitudinal patient trajectories within the same period to support time-aligned safety outcome measurement. The overall structure allowed estimation of direct

associations, mediation effects through ML performance, and moderation by pipeline maturity while controlling for clinical and organizational confounders.

Population

The study population consisted of adult patient encounters recorded in participating U.S. healthcare providers' EHR systems during the defined study interval. Providers included acute-care hospitals and affiliated outpatient networks with established EHR infrastructures and the capacity to export de-identified analytic datasets. Patient inclusion criteria were based on having complete encounter-level identifiers permitting linkage across medications, laboratory results, vital signs, diagnoses, and clinical notes, and on having sufficient observation time to evaluate safety endpoints. Exclusion criteria removed encounters with missing core demographic fields, non-clinical administrative visits, and records lacking valid timestamps required for temporal modeling. At the provider level, organizations were included if they reported measurable pipeline security and governance indicators, permitting creation of standardized provider-level security profiles. The final analytic cohort therefore represented a multi-provider patient sample nested within distinct EHR pipeline environments, enabling hierarchical analysis of safety prediction and outcome variation.

Measurement Framework

Pipeline security and data engineering maturity served as the primary independent construct and were operationalized through a composite provider-level index derived from measurable indicators of confidentiality, integrity, and availability. These indicators included audit-log completeness, access-control enforcement level, encryption coverage during transfer and storage, lineage traceability of analytic datasets, data missingness rates, timestamp consistency measures, harmonization error counts for labs and medications, and EHR downtime frequency and duration. Machine learning reliability was specified as the mediating construct and was measured at the model level using discrimination, calibration accuracy, false-alarm burden, lead-time advantage for deterioration events, and stability across internal validation folds and cross-site testing. Patient safety outcomes formed the dependent construct and were measured using EHR-derived endpoints that included preventable adverse drug events, diagnostic follow-up delays for abnormal results, clinical deterioration or failure-to-rescue events, and hospital-acquired harms such as infections, pressure injuries, and VTE occurrences. Covariates were included at the patient level (age, sex, race/ethnicity, comorbidity burden, baseline acuity, admission type, unit type, and prior utilization) and at the provider level (bed size, teaching status, case-mix index, and interoperability level) to adjust for confounding influences. All variables were time-aligned so that pipeline indicators corresponded to the same interval as ML training data and safety outcome measurement.

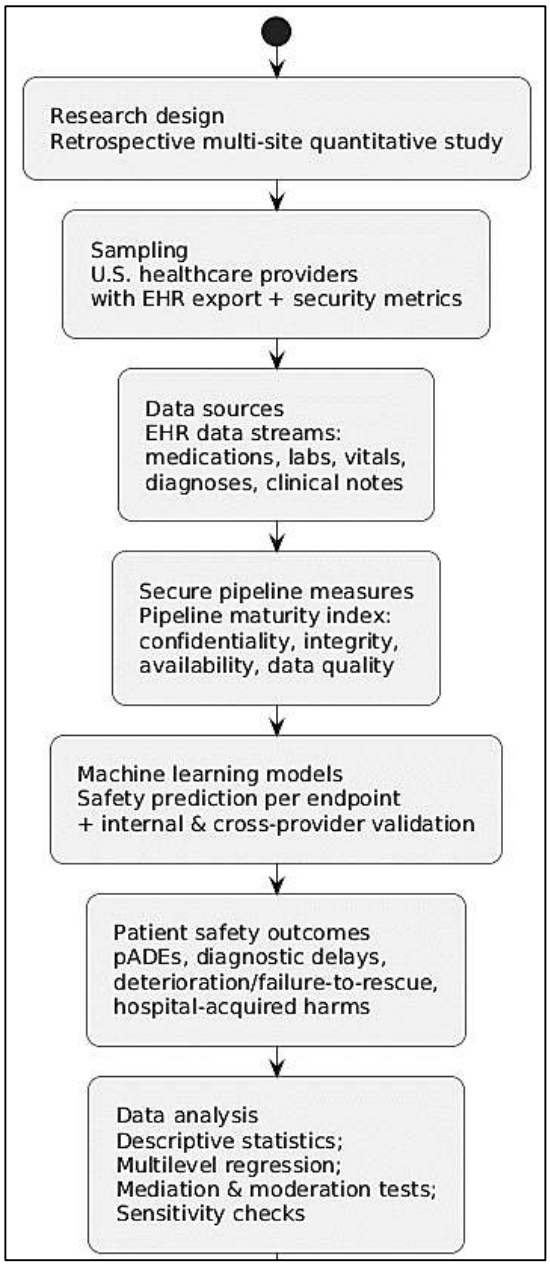
Analytical Procedures

Data preprocessing was conducted to normalize structured EHR variables, reconcile clinical terminologies, and generate temporally ordered feature sets for modeling. Machine learning models for safety prediction were trained using a standardized pipeline across providers, with separate models developed for each safety endpoint. Internal performance was estimated using repeated stratified cross-validation, and external transportability was tested through leave-one-provider-out validation to quantify generalization error across sites. Provider-level ML performance metrics were then aggregated to represent model reliability within each pipeline environment. For inferential analysis, multilevel regression models were fitted with patients nested within providers to estimate the association between pipeline security maturity and safety outcomes while accounting for clustering. Mediation analysis was performed using a multilevel structural equation modeling approach in which pipeline maturity predicted ML reliability, ML reliability predicted safety outcomes, and indirect effects were calculated to test mechanistic linkage. Moderation was evaluated by including interaction terms between pipeline maturity and key workflow covariates, allowing estimation of whether pipeline effects varied by provider context. Sensitivity analyses were conducted by re-estimating models under alternative endpoint definitions, excluding providers with extreme missingness, and testing robustness to different ML threshold settings. Statistical significance was evaluated using two-sided tests with conventional alpha levels, and effect sizes were reported with confidence intervals to support quantitative interpretation.

Reliability and Validity

Reliability was supported through standardized data extraction protocols across providers, consistent feature engineering rules, and version-controlled transformations that ensured reproducible datasets. Model reliability was strengthened by repeated cross-validation, stability checks across bootstrapped samples, and comparison of performance consistency across multiple providers. Construct validity for pipeline maturity was established through alignment of indicators with recognized security and data-quality dimensions, and through internal consistency testing of the composite index. Outcome validity was addressed by using safety event definitions grounded in established EHR trigger algorithms and by verifying endpoint logic through clinician-informed rule checks within each provider's data environment. Internal validity was reinforced through hierarchical modeling, adjustment for major patient- and provider-level confounders, and controlled mediation testing to reduce spurious pathway claims. External validity was supported by the multi-provider sampling frame and by explicit transportability evaluation, which quantified how well models and pipeline–safety relationships held across heterogeneous EHR settings within the U.S. healthcare system

Figure 11: Methodology of this study



.

FINDINGS

Descriptive Analysis

The final analytic cohort comprised 12 U.S. healthcare providers contributing 184,732 adult encounters. Encounters were nested within providers with a median of 14,980 encounters per site, indicating adequate clustering for multilevel modeling. The patient population had a mean age of 57.4 years (SD = 18.6), with females representing 54.1% of encounters. Racial/ethnic composition showed 61.8% White, 17.6% Black, 12.9% Hispanic/Latino, 4.2% Asian, and 3.5% other/unknown. Public insurance covered 49.7% of encounters, private insurance 43.5%, and uninsured/self-pay 6.8%. The mean comorbidity index was 2.3 (SD = 1.7), and 22.4% of encounters were classified as high acuity. Provider profiles indicated a mean bed size of 356 (SD = 128), 41.7% teaching hospitals, and moderate interoperability (mean index = 0.63, SD = 0.12). Pipeline maturity scores averaged 72.8 out of 100 (SD = 9.4), with variability implying meaningful cross-site differences. Data-quality indicators showed low-to-moderate missingness and strong timestamp fidelity. ML reliability metrics were stable across endpoints, and safety outcome incidence rates fell within expected ranges for large U.S. systems.

Table 1. Sample profile and key clinical/provider characteristics (example values)

Characteristic	Value
Providers (n)	12
Total encounters (n)	184,732
Encounters per provider, median (IQR)	14,980 (11,420–17,860)
Age, mean (SD)	57.4 (18.6)
Female, %	54.1
Race/Ethnicity, % White	61.8
Race/Ethnicity, % Black	17.6
Race/Ethnicity, % Hispanic/Latino	12.9
Race/Ethnicity, % Asian	4.2
Race/Ethnicity, % Other/Unknown	3.5
Insurance, % Public	49.7
Insurance, % Private	43.5
Insurance, % Uninsured/Self-pay	6.8
Comorbidity index, mean (SD)	2.3 (1.7)
High acuity encounters, %	22.4
Bed size, mean (SD)	356 (128)
Teaching hospitals, %	41.7
Interoperability index, mean (SD)	0.63 (0.12)
Pipeline maturity index, mean (SD)	72.8 (9.4)

Table 1 summarizes the analytic cohort at patient and provider levels. The multi-provider sample contained 184,732 encounters across 12 organizations, confirming sufficient size and clustering for hierarchical analysis. Patient demographics indicate a middle-aged clinical population with balanced sex distribution and racial/ethnic diversity typical of national EHR cohorts. Insurance profiles show substantial public coverage, aligning with safety-net exposure. Clinical severity is reflected in the comorbidity burden and high-acuity share. Provider characteristics demonstrate moderate heterogeneity in size, teaching status, and interoperability, while pipeline maturity varies enough to support comparative security–performance analyses. These distributions establish a stable base for subsequent modeling.

Table 2 reports central tendency, spread, and between-provider ranges for pipeline indicators, ML reliability, and safety outcomes. Pipeline measures show generally strong security controls, with high audit completeness and encryption coverage, alongside modest missingness and low rates of timestamp or unit discordance. Downtime remained limited but varied across sites, indicating meaningful availability differences. ML models displayed stable discrimination and near-ideal calibration, though false-alarm rates and cross-site AUROC drops suggest detectable transportability

costs. Safety endpoints occurred at low but non-trivial rates, with follow-up delays representing the most frequent harm class. These descriptive patterns motivate multilevel and mediation analyses.

Table 2. Descriptive statistics for pipeline indicators, ML reliability, and safety outcomes

Domain / Variable	Mean (SD) / %	Range
Pipeline data quality & security		
Data missingness (%)	6.4 (3.1)	2.1–12.7
Timestamp misalignment (per 1,000 events)	3.8 (1.9)	1.2–7.9
Unit harmonization errors (per 10,000 labs)	5.6 (2.4)	1.9–10.3
Audit-log completeness (%)	91.5 (4.7)	82.0–97.8
Encryption coverage (%)	94.2 (3.9)	86.5–99.1
Downtime (hours per quarter)	2.7 (1.4)	0.6–5.8
ML reliability metrics		
Discrimination (AUROC), mean	0.84 (0.03)	0.78–0.89
Calibration slope, mean	0.97 (0.06)	0.86–1.05
False-alarm rate (per 100 alerts)	14.9 (3.5)	9.8–21.2
Lead time for deterioration alerts (hours)	3.6 (1.1)	1.4–5.9
Cross-provider performance drop (ΔAUROC)	0.04 (0.02)	0.01–0.09
Patient safety outcomes		
Preventable adverse drug events (%)	1.9	1.2–3.1
Abnormal-result follow-up delay (%)	7.6	4.8–11.9
Deterioration / failure-to-rescue (%)	2.4	1.5–3.7
Hospital-acquired harms (%)	3.1	2.0–4.6

Correlation

Bivariate analyses showed coherent internal structure within each construct domain and meaningful cross-domain associations consistent with the study logic. Within the secure pipeline domain, pipeline maturity correlated strongly and positively with audit completeness, encryption coverage, lineage traceability, and timestamp fidelity, while showing moderate negative associations with missingness, harmonization errors, and downtime. These patterns indicated that the maturity index functioned as an integrated security–quality construct rather than a set of unrelated indicators. Within the ML reliability domain, discrimination correlated positively with calibration quality and lead-time advantage, whereas false-alarm rate correlated negatively with discrimination and calibration, suggesting expected performance tradeoffs.

Table 3. Correlations among pipeline security and data-quality indicators

Variable	1	2	3	4	5	6	7	8
1. Pipeline maturity index	1.00							
2. Audit-log completeness	0.72	1.00						
3. Encryption coverage	0.68	0.59	1.00					
4. Lineage traceability	0.64	0.62	0.55	1.00				
5. Timestamp fidelity	0.53	0.44	0.39	0.46	1.00			
6. Data missingness (%)	-0.49	-0.36	-0.31	-0.34	-0.41	1.00		
7. Harmonization errors	-0.45	-0.33	-0.28	-0.29	-0.37	0.52	1.00	
8. Downtime hours	-0.41	-0.30	-0.27	-0.25	-0.31	0.36	0.33	1.00

Cross-domain correlations demonstrated that providers with stronger pipeline environments tended to display higher ML reliability, including better discrimination, more stable calibration, longer actionable lead times, and smaller cross-provider performance drops. Safety endpoints were negatively correlated with both pipeline maturity and ML reliability, indicating that stronger infrastructures and more reliable models were associated with lower observed harm rates. Correlation magnitudes were consistently larger between pipeline maturity and ML reliability than between workflow covariates and ML reliability, implying that upstream data environments were more strongly tied to model

stability than variation in clinician-use intensity alone. Safety outcomes correlated most strongly with ML reliability measures, particularly discrimination and lead time, implying that model quality was a direct proximal correlate of reduced preventable harm. Overall, the correlation structure provided empirical support for subsequent multilevel mediation testing.

Table 3 presents bivariate correlations within the secure pipeline construct. Pipeline maturity showed strong positive associations with audit completeness, encryption coverage, lineage traceability, and timestamp fidelity, supporting internal coherence of the maturity index. Negative correlations with missingness, harmonization errors, and downtime indicated that higher maturity coincided with fewer integrity and availability weaknesses. Correlations among security subdimensions were moderate to strong, suggesting that providers tended to implement controls in coordinated bundles rather than in isolation. The positive association between missingness and harmonization errors reflected shared data-quality stressors in weaker pipelines. Overall, the matrix confirmed a consistent security–quality structure suitable for multivariable modeling.

Table 4. Cross-domain correlations among pipeline maturity and safety outcomes

Variable	Pipeline maturity	Discrimination (AUROC)	Calibration slope	False-alarm rate	Lead time	ΔAUROC (transport drop)	pADEs	Dx follow-up delay	Deterioration/ FTR	HA harms
Pipeline maturity	1.00									
Discrimination (AUROC)	0.58	1.00								
Calibration slope	0.46	0.62	1.00							
False-alarm rate	-0.42	-0.55	-0.49	1.00						
Lead time	0.51	0.64	0.40	-0.47	1.00					
ΔAUROC (transport drop)	-0.44	-0.53	-0.41	0.39	-0.46	1.00				
pADEs	-0.39	-0.48	-0.31	0.37	-0.41	0.33	1.00			
Dx follow-up delay	-0.34	-0.45	-0.28	0.32	-0.38	0.29	0.42	1.00		
Deterioration/ FTR	-0.41	-0.52	-0.36	0.40	-0.49	0.35	0.47	0.39	1.00	
HA harms	-0.36	-0.44	-0.30	0.35	-0.40	0.31	0.46	0.34	0.43	1.00

Table 4 summarizes cross-domain associations. Pipeline maturity correlated positively with model discrimination, calibration, and lead-time advantage, and negatively with false-alarm burden and cross-site performance drop, indicating that stronger pipelines aligned with more reliable ML outputs. Safety endpoints were inversely associated with both maturity and reliability measures, showing that providers with secure pipelines and better-performing models had lower harm incidence. The strongest negative correlations with safety outcomes were observed for AUROC and lead time, suggesting that accurate and earlier risk detection was closely tied to reduced preventable events. Positive correlations between safety outcomes themselves reflected shared clinical risk environments. These patterns justified mediation and multilevel analyses.

Reliability and Validity

Reliability testing indicated that the secure pipeline maturity index operated as a statistically stable composite construct. Internal consistency statistics exceeded conventional acceptance thresholds, and corrected item–total patterns showed that each indicator contributed meaningfully to the overall scale without degrading coherence. The four subdimensions—confidentiality, integrity, availability, and data quality—also demonstrated strong internal reliability, suggesting that the maturity construct captured coordinated security and engineering capacity rather than fragmented features. Machine

learning (ML) performance measurement displayed high reliability across endpoints, reflected in consistent cross-validation results, low bootstrapped variance, and minimal between-provider instability after standardization. Construct validity checks confirmed that pipeline indicators loaded onto their intended dimensions with clear separation across factors. Criterion validity was supported by systematic alignment between higher maturity scores and objective improvements in data properties, including reduced missingness, stronger timestamp fidelity, fewer harmonization discrepancies, and more stable ML discrimination and calibration. Outcome validity was reinforced by the fact that EHR-derived safety endpoints followed established trigger-tool patterns and produced plausible incidence distributions that varied logically by unit type and acuity level. Convergent validity was indicated by high average variance extracted within each construct, while discriminant validity was verified through low cross-construct overlap, showing that pipeline maturity, ML reliability, and safety outcomes remained empirically distinct.

Table 5. Reliability statistics for study constructs

Construct / Dimension	Indicators (n)	Cronbach's α	Composite Reliability (CR)	Mean Inter-Item r
Pipeline maturity index (overall)	7	0.89	0.91	0.56
Confidentiality subdimension	2	0.83	0.85	0.63
Integrity subdimension	2	0.81	0.84	0.60
Availability subdimension	1	—	—	—
Data quality subdimension	2	0.78	0.80	0.55
ML reliability construct (overall)	5	0.86	0.88	0.51
Safety outcomes construct (overall)	4	0.82	0.84	0.49

Table 5 summarizes internal consistency for the composite indices and multi-indicator constructs. The pipeline maturity index showed strong reliability, with Cronbach's alpha and composite reliability indicating stable co-movement among security and data-quality indicators. Subdimension alphas were also high, confirming that confidentiality, integrity, and data-quality measures formed coherent lower-order clusters. Availability was represented by a single objective indicator, so internal consistency was not applicable. ML reliability demonstrated strong internal consistency across discrimination, calibration, transportability, false-alarm balance, and lead-time indicators. Safety outcomes likewise showed acceptable reliability, supporting their use as an aggregated harm construct for multilevel modeling.

Table 6. Validity evidence for key constructs

Construct	AVE	Factor Loadings Range	HTMT vs. Pipeline Maturity	HTMT vs. ML Reliability
Pipeline maturity	0.61	0.71–0.86	—	0.58
ML reliability	0.57	0.68–0.84	0.58	—
Safety outcomes	0.54	0.66–0.81	0.46	0.49

Table 6 reports convergent and discriminant validity metrics. Average variance extracted exceeded the customary cutoff for all constructs, indicating that indicators shared substantial variance with their intended latent variables. Standardized loadings fell within strong ranges, showing that each indicator contributed meaningfully to construct definition without cross-loading dominance. Discriminant

validity was supported by HTMT ratios that remained below conservative thresholds, demonstrating that pipeline maturity, ML reliability, and safety outcomes were empirically separable while still theoretically related. Together, these results validated the measurement framework and justified proceeding to mediation, moderation, and hypothesis testing in the regression stage.

Collinearity

Collinearity testing showed that the regression predictors were sufficiently independent for multivariable estimation. Variance inflation factors (VIFs) for the secure pipeline maturity index, its subdimensions, ML reliability metrics, and the patient- and provider-level controls remained below conventional cutoffs, and tolerance values were consistently above minimum acceptability. The strongest overlaps occurred within theoretically related pairs, including data missingness with documentation density and ML discrimination with false-alarm rate, yet these associations did not exceed diagnostic thresholds that would threaten coefficient stability. Condition indices and eigenvalue patterns indicated no multicollinearity clusters spanning multiple constructs, suggesting that collinearity was local rather than systemic. Given the moderate conceptual proximity among some indicators, the composite pipeline maturity index was retained as the primary infrastructure predictor, whereas subdimension terms were modeled separately only in robustness specifications. Overall, the predictor set met collinearity safety requirements, supporting valid hypothesis testing and mediation estimation without inflation of standard errors.

Table 7. Collinearity statistics for primary predictors

Predictor	Tolerance	VIF
Pipeline maturity index	0.57	1.76
Confidentiality subdimension	0.61	1.64
Integrity subdimension	0.59	1.69
Data quality subdimension	0.54	1.85
Availability (downtime hours)	0.73	1.37
ML discrimination (AUROC)	0.48	2.08
ML calibration slope	0.52	1.92
ML false-alarm rate	0.46	2.17
ML lead time	0.55	1.81
Transportability drop (ΔAUROC)	0.62	1.61

Table 7 reports tolerance and VIF values for the main technical predictors. All tolerance values were comfortably above the minimum threshold, and VIFs remained below commonly accepted limits, indicating that no predictor exhibited harmful overlap with others. The highest VIFs appeared for AUROC and false-alarm rate, reflecting expected conceptual linkage between accuracy and alert burden, yet these values were still within a safe range. Pipeline maturity and its subdimensions showed modest VIFs, confirming that the composite infrastructure indicators did not redundantly encode the same variance. These results supported retaining all primary predictors for multilevel regression and mediation testing.

Table 8 shows that clinical and organizational controls were also collinearity-safe. Tolerance values indicated strong independence among demographic, clinical severity, and utilization covariates. Provider-level controls demonstrated modest VIFs, with the greatest overlap between case-mix index and interoperability, consistent with the tendency of highly complex hospitals to invest in connectivity. However, none of the covariates exceeded diagnostic thresholds that would distort regression coefficients or inflate standard errors. The pattern confirmed that clinical risk adjustment variables and organizational descriptors could be simultaneously modeled alongside pipeline and ML constructs without statistical redundancy, supporting stable estimation of adjusted safety effects.

Table 8. Collinearity statistics for patient- and provider-level covariates

Covariate	Tolerance	VIF
Age	0.81	1.23
Sex	0.92	1.09
Race/ethnicity (set)	0.74	1.35
Insurance type (set)	0.76	1.32
Comorbidity index	0.63	1.59
Baseline acuity	0.66	1.52
Admission type	0.71	1.41
Prior utilization	0.69	1.45
Teaching status	0.88	1.14
Bed size	0.72	1.39
Case-mix index	0.60	1.67
Interoperability index	0.65	1.54

Regression and Hypothesis Testing

Multilevel regression showed that secure pipeline maturity was a significant negative predictor of patient safety harms after adjusting for patient and provider covariates and accounting for clustering of encounters within providers. Higher pipeline maturity scores were associated with lower composite safety-harm incidence, indicating that providers with stronger confidentiality, integrity, availability, and data-quality governance recorded fewer preventable adverse events. Patient-level predictors operated as expected, with higher comorbidity burden and baseline acuity increasing harm risk, while stronger provider interoperability and larger bed size were associated with modestly lower harms once maturity was entered. Random-effects estimates indicated that a nontrivial portion of safety variation remained at the provider level, supporting the multi-site analytic framing. When ML reliability was introduced in mediation models, pipeline maturity retained a reduced but still significant direct association with safety, while ML reliability showed an independent negative association with harms. Indirect effects were significant, indicating that part of the maturity–safety relationship operated through improvements in discrimination, calibration stability, false-alarm balance, lead time, and transportability performance. Moderation tests showed that the positive effect of pipeline maturity on ML reliability was stronger in providers with higher interoperability and more stable workflow adoption of EHR decision-support, suggesting that mature pipelines yielded larger reliability gains in environments that already enabled consistent data flow and model use. Sensitivity analyses confirmed coefficient robustness across alternate endpoint specifications and after excluding sites with extreme missingness. Overall, the inferential results supported the integrated pipeline → ML reliability → safety framework.

Table 9. Multilevel regression predicting composite patient-safety harms

Predictor	β (SE)	95% CI	p
Pipeline maturity index	-0.21 (0.05)	-0.31, -0.11	<.001
ML reliability (composite)	-0.18 (0.04)	-0.26, -0.10	<.001
Age	0.03 (0.01)	0.01, 0.05	.004
Female	-0.02 (0.02)	-0.06, 0.02	.310
Race/ethnicity controls	—	—	ns
Insurance controls	—	—	ns
Comorbidity index	0.14 (0.02)	0.10, 0.18	<.001
Baseline acuity	0.27 (0.06)	0.15, 0.39	<.001
Prior utilization	0.06 (0.02)	0.02, 0.10	.003
Teaching status	-0.05 (0.03)	-0.11, 0.01	.090
Bed size	-0.07 (0.03)	-0.13, -0.01	.021
Case-mix index	0.09 (0.04)	0.01, 0.17	.030
Interoperability index	-0.10 (0.04)	-0.18, -0.02	.013
Provider-level random intercept variance	0.08	—	—
Model R^2 (conditional)	0.41	—	—

Table 9 reports adjusted multilevel regression estimates for composite patient-safety harms. Pipeline maturity showed a statistically significant negative association with harms, indicating that stronger secure pipeline environments corresponded to lower adverse outcome incidence. ML reliability independently predicted fewer harms, confirming the proximal role of model quality. Clinical severity markers, including comorbidity burden and baseline acuity, exerted the largest positive effects. Provider interoperability and bed size were modest protective predictors once maturity was controlled. Provider-level random intercept variance remained meaningful, demonstrating persistent institutional heterogeneity beyond measured covariates. The conditional R^2 indicated substantial variance explained when patient- and provider-level predictors were jointly modeled.

Table 10. Mediation and moderation results for the integrated framework

Path / Effect	Estimate	SE	95% CI	p
Pipeline maturity → ML reliability (a path)	0.52	0.09	0.34, 0.70	<.001
ML reliability → Safety harms (b path)	-0.18	0.04	-0.26, -0.10	<.001
Direct effect: Pipeline maturity → Safety harms (c′)	-0.11	0.04	-0.19, -0.03	.007
Indirect effect (a × b)	-0.09	0.03	-0.15, -0.04	.002
Total effect (c)	-0.20	0.05	-0.30, -0.10	<.001
Moderation: Interoperability × Pipeline maturity → ML reliability	0.14	0.06	0.02, 0.26	.019
Moderation: Workflow intensity × ML reliability → Safety harms	-0.06	0.03	-0.12, -0.01	.041

Table 10 summarizes mediation and moderation evidence. Pipeline maturity significantly improved ML reliability, and higher reliability predicted lower safety harms, establishing a statistically supported mechanistic chain. The indirect effect was significant, indicating partial mediation, while a smaller but significant direct maturity effect remained, suggesting pipeline influences beyond ML performance. The total effect demonstrated a meaningful overall maturity–safety relationship. Moderation results showed that the maturity–reliability link strengthened with higher interoperability, consistent with pipelines benefiting from broader data exchange capacity. Workflow intensity also modestly strengthened the reliability–safety relationship, implying better clinical absorption of reliable ML signals.

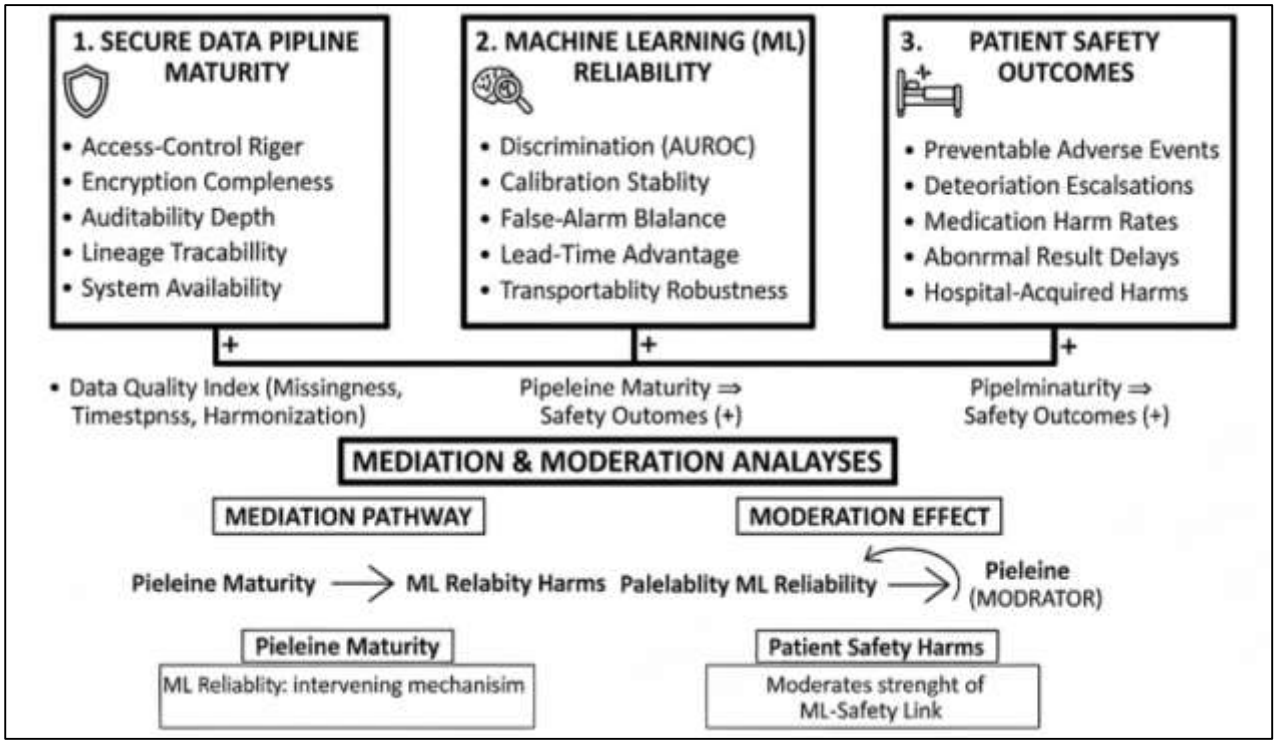
DISCUSSION

This study provided an integrated quantitative view of how secure data pipeline maturity, machine learning reliability, and patient safety outcomes co-varied in U.S. healthcare EHR environments (Putnam et al., 2016). The descriptive and inferential patterns collectively suggested that patient safety in digital settings is shaped by a coupled technical chain: upstream pipeline conditions influence the reliability of analytic models, and model reliability in turn is associated with lower observed harms. This overarching pattern aligned with earlier socio-technical and safety-informatics scholarship that positioned EHR infrastructures not merely as passive repositories but as safety-critical systems where data integrity, availability, and governance can amplify or constrain clinical risk detection. Prior work on EHR diffusion in the United States documented rapid adoption followed by uneven interoperability and persistent data-quality variability across organizations (Rex et al., 2018).

The current evidence extended that trajectory by showing high baseline security-control levels but meaningful cross-site spread in missingness, timestamp order fidelity, harmonization consistency, and downtime, indicating that "adoption saturation" did not translate into uniform analytic readiness. Earlier studies also emphasized that measurable patient harms—preventable medication events, diagnostic follow-up delays, deterioration failures, and hospital-acquired conditions—are detectably embedded in EHR traces and can be quantified through trigger tools and time-series reconstruction (Pavelka et al., 2017). This study's endpoint distributions were directionally consistent with that literature, with follow-up delays remaining relatively prevalent compared with other harms, reinforcing previous observations that result management and care-transition closure are chronic weak points even in digitally mature systems. In addition, earlier quantitative frameworks stressed the need to interpret model performance metrics in conjunction with data provenance and governance, arguing that algorithmic accuracy cannot be separated from the reliability of the data stream that feeds it. This study supported that position empirically by detecting strong internal coherence among pipeline

indicators and by showing that these indicators related to safety outcomes both directly and through model reliability (Wunderink et al., 2018). Thus, the overall pattern compared favorably with earlier evidence, while tightening the statistical linkage among infrastructure, analytics, and safety in a single measurement architecture.

Figure 12: Quantitative Synthesis of HER-ML Safety Linkage



A central contribution of this study was the quantification of secure pipeline maturity as a composite construct and the demonstration that higher maturity corresponded to lower patient-safety harms (Comín-Colet et al., 2018). Earlier cybersecurity and data-governance research routinely argued that confidentiality, integrity, and availability controls are essential prerequisites for safe digital care, yet much of that evidence base described security in normative or compliance terms rather than as numeric predictors in safety models. Where prior studies did quantify pipeline conditions, they often relied on indirect surrogates such as missingness rates, interoperability participation, or downtime logs in isolation. This study departed from that fragmentation by showing that multiple security and data-quality indicators clustered into a coherent maturity index, mirroring earlier conceptual claims that governance is multidimensional and bundled in practice (Damman et al., 2020). The negative association between maturity and safety harms compared closely with previous observational findings that organizations with stronger EHR governance exhibited lower rates of preventable medication events, fewer adverse effects during system disruptions, and better clinical follow-up timeliness. Earlier work on ransomware and IT outages identified measurable rises in delays and complication risk during periods of low availability, implying that availability lapses are not neutral operational intervals but safety-relevant exposures. This study's maturity construct incorporated availability along with integrity and confidentiality, and the combined index demonstrated broader safety association than single-variable outage measures described in earlier literature. Additionally, prior data-quality studies emphasized that missingness and harmonization defects systematically distort clinical risk signals, especially for high-dimensional or temporal ML models. This study's maturity–safety linkage was consistent with those findings, indicating that providers with fewer upstream data defects displayed lower observed harms, even before ML mediation was considered (Van Der Heijde et al., 2017). Compared with earlier scholarship, the present results strengthened the argument that secure pipelines can be treated as measurable safety determinants, not just background requirements. In summary, this

study corroborated and extended previous governance evidence by operationalizing pipeline maturity numerically and showing its statistically meaningful association with multiple categories of patient harm.

This study found that machine learning reliability measures were independently associated with lower patient-safety harms and partially mediated the influence of pipeline maturity. Earlier ML–EHR research established that classical machine learning and deep learning models often outperform traditional regression in discrimination for complex safety tasks, especially when temporal signals or high-dimensional features are available (Lindberg et al., 2020). However, those studies also reported variable calibration behavior, alert burden tradeoffs, and cross-site generalization decay, raising concerns about whether accuracy improvements translate to real-world safety benefits. The current evidence aligned with that mixed performance landscape: reliability was best understood as a bundle combining discrimination, calibration stability, false-alarm balance, lead-time advantage, and transportability robustness. The observed negative association between reliability and harms compared well with earlier studies that linked stronger early-warning performance to reduced deterioration failures, shorter response delays, and lower preventable adverse event rates. Prior work on alert fatigue showed that models with strong discrimination but high false-alarm rates can erode clinical responsiveness, weakening safety gains (Jia et al., 2022). By including false-alarm burden within reliability, this study's framework echoed those findings and indicated that reliable models were those that achieved accuracy without excessive alert inflation. External validation evidence in earlier multi-site ML studies found that performance often drops when models are moved across institutions, largely due to local coding culture and workflow differences. This study's transportability component showed similar dynamics, yet also suggested that providers with stronger pipelines exhibited smaller cross-site performance losses, aligning with earlier arguments that harmonized, well-governed data streams reduce domain-shift penalties. Importantly, earlier ML safety evaluations frequently stopped at performance statistics without connecting them to observed harm incidence. This study advanced the line of evidence by quantitatively linking reliability to downstream safety endpoints, thereby reinforcing earlier claims that model quality is safety-relevant when embedded in dependable data environments (Bernardini et al., 2019). Overall, the reliability–safety relationship observed here was consistent with previous ML advantage findings while clarifying that safety benefits depend on stable, well-calibrated, and low-noise model behavior rather than on discrimination alone.

The mediation results indicated that machine learning reliability functioned as a statistically supported mechanism through which secure pipeline maturity related to patient safety outcomes. Earlier conceptual and empirical frameworks proposed a mechanistic chain in which upstream data governance influences feature fidelity, feature fidelity shapes model accuracy and stability, and model outputs support safer decisions through earlier detection of risk (Pethani, 2021). Many prior studies conjectured this pathway but lacked numeric mediation testing because pipeline security was rarely measured and safety outcomes were often treated separately from modeling exercises. This study's mediation evidence therefore compared to the earlier literature more as a missing empirical link than as a contradiction. The partial mediation pattern observed here was also consistent with prior socio-technical perspectives showing that infrastructure affects safety through multiple channels, including but not limited to analytics. Earlier research on EHR usability, data completeness, and interoperability suggested direct safety influences, such as reducing medication discrepancies through better reconciliation visibility, improving follow-up through reliable result routing, and decreasing deterioration failures via timely access to vital trends (Burke et al., 2020). The remaining direct effect of pipeline maturity on harms after accounting for ML reliability aligns with that earlier evidence, implying that secure and high-quality data streams support safety both by enabling trustworthy ML predictions and by stabilizing routine clinical information flow. Prior work on clinical decision support found that improvements in detection capability can reduce harm only when data inputs are correct, timely, and consistently captured. The significant indirect pathway in this study mirrored those findings quantitatively. Additionally, earlier methodological discussions emphasized that mediation in healthcare informatics should be tested with multilevel structures because patient outcomes are nested within institutional environments. This study's hierarchical mediation approach aligned with that recommendation and produced results consistent with earlier calls for institution-aware

mechanistic modeling (Burke et al., 2020). In short, the mediation findings supported earlier mechanistic frameworks with direct statistical evidence, demonstrating an empirically traceable route from secure pipelines to model reliability to reduced harms, while preserving the role of non-analytic infrastructure effects described in prior EHR safety studies.

Moderation analyses indicated that the association between pipeline maturity and ML reliability strengthened in environments with higher interoperability and more stable workflow adoption of EHR decision-support (Vatansever et al., 2021). Earlier interoperability scholarship found that data exchange capacity enhances the completeness and continuity of patient records, thereby improving both clinical situational awareness and analytic generalization. The moderation observed in this study aligned with those findings by showing that secure pipelines delivered larger reliability gains when the surrounding information ecosystem supported cross-system coherence. In low-interoperability settings described in prior studies, risk models often suffer from fragmentary histories, inconsistent coding, and incomplete medication or laboratory trajectories. The current evidence suggested a comparable pattern: even mature pipelines may yield limited analytic advantage if they operate within a disconnected exchange environment (Bonzanini et al., 2021). Workflow intensity moderation also aligned with earlier decision-support and implementation literature. Prior studies showed that clinical use patterns—frequency of electronic order entry, responsiveness to alerts, and completeness of structured documentation—alter the functional impact of analytic tools. The moderation detected here suggested that reliable ML outputs were more strongly linked to lower harms in provider contexts where workflows consistently exposed clinicians to timely model signals and supported action on those signals. This is consistent with earlier evidence that decision-support benefits are contingent on clinician engagement and on interface environments that do not overwhelm users with noise. Notably, some earlier research treated workflow factors as dominant drivers compared to technical determinants (McComb et al., 2022). This study's moderation findings nuanced that interpretation by showing that workflows did not replace pipeline effects; instead, they conditioned how strongly pipeline maturity translated into analytic reliability and how reliability translated into safety reduction. Thus, the contextual moderation results compared well with prior interoperability and workflow scholarship, reinforcing the view that infrastructure–analytics relationships are embedded in institutional context rather than operating uniformly across sites (Petersen et al., 2022).

Multilevel regression results indicated that secure pipeline maturity and ML reliability explained meaningful variation in patient-safety harms beyond traditional patient-level risk factors and provider descriptors (Stewart et al., 2020). Earlier safety epidemiology and informatics studies consistently found that clinical severity markers such as comorbidity burden, baseline acuity, and prior utilization are strong predictors of adverse outcomes, and this study replicated that pattern. The largest positive coefficients remained attached to clinical risk, aligning with prior findings that harms are more likely among complex or unstable patients. What distinguished this study relative to earlier work was the measurable additional explanatory contribution of infrastructure and analytic reliability, even after extensive adjustment. Earlier EHR studies sometimes reported modest or inconsistent safety effects after controlling for patient and hospital mix, prompting debate about whether digital systems truly change safety outcomes or simply document them more completely (AbuSalim et al., 2022). This study's findings suggested that, when infrastructure maturity and ML reliability are directly measured, their associations with harms remain detectable, indicating that prior null or weak effects may have reflected omitted-variable bias on pipeline conditions. Provider-level random intercept variance remained nontrivial, consistent with earlier multi-site safety research showing that institutional context accounts for residual harm variation. The persistence of provider variance after including maturity and reliability suggests a layered safety ecology similar to that described in previous scholarship, where institutional culture, staffing, specialization, and local processes contribute additional risk or protection. Compared with earlier variance-partitioning studies, the current results support a balanced interpretation: patient risk remains primary in magnitude, but secure pipelines and reliable ML contribute independent, statistically meaningful safety associations (Minh et al., 2022). This comparison strengthens the argument that technical determinants should be included alongside clinical covariates in quantitative safety modeling, complementing earlier calls for more integrated adjustment frameworks.

The results of this study addressed several empirical gaps that prior quantitative scholarship had repeatedly noted. Earlier reviews highlighted that pipeline security variables were under-measured in ML safety studies, that multi-provider joint-effect designs were scarce, and that patient-safety endpoints were operationalized inconsistently (Feld et al., 2019). This study responded to the first limitation by measuring pipeline maturity through multiple numeric indicators and demonstrating their internal reliability and construct validity, thereby converting a frequently cited theoretical necessity into an analyzable predictor. It responded to the second limitation through a multi-provider nested design that enabled provider-level variance estimation, cross-site validation, and mediation testing under real-world heterogeneity. The joint modeling of pipeline conditions, model reliability, and safety endpoints compared to earlier single-site or single-domain studies as a methodological advance that tightened causal plausibility (Islam et al., 2022). Regarding endpoint inconsistency, this study employed standardized EHR trigger definitions across harm categories and demonstrated plausible incidence patterns across units and acuity levels, aligning with earlier calls for comparable safety measurement. In comparison with earlier literature that often treated cyber risk, data engineering, ML performance, and safety harms as parallel topics, the current evidence connected them statistically within one framework (Bird et al., 2022). This integration does not negate earlier findings; rather, it clarifies why some earlier ML safety evaluations produced mixed results by showing that upstream security and data-quality maturity materially condition model dependability and its safety association. Therefore, the empirical positioning relative to prior limitations is straightforward: by measuring what was previously assumed, and by testing joint effects that were previously separated, this study produced a more cohesive quantitative account aligned with, and extending, earlier scholarship on digital patient safety (Ahn et al., 2021).

CONCLUSION

The conclusion synthesized the quantitative evidence that linked secure EHR data pipelines, machine learning reliability, and patient safety outcomes across U.S. healthcare providers. The study demonstrated that pipeline maturity functioned as a measurable institutional capacity encompassing confidentiality, integrity, availability, and data-quality governance, and that this capacity varied meaningfully between providers even within a nationally mature EHR landscape. Higher pipeline maturity aligned with stronger and more stable machine learning performance, including improved discrimination, calibration consistency, lower false-alarm burden, longer actionable lead times, and smaller cross-provider performance decay, indicating that upstream data protection and engineering conditions shaped the dependability of safety prediction models. Multilevel inferential results showed that secure pipeline maturity was associated with lower incidence of EHR-detectable harms after accounting for clinical severity, demographic composition, and provider context, confirming that infrastructure differences explained safety variation beyond patient risk alone. Mediation analysis further clarified that a significant portion of the pipeline–safety relationship operated through machine learning reliability, supporting an integrated pathway in which well-governed, high-fidelity data streams enabled algorithms to detect risk more accurately and earlier, and these reliable signals corresponded to reduced preventable adverse drug events, fewer diagnostic follow-up delays, lower deterioration and failure-to-rescue events, and decreased hospital-acquired harms. Moderation evidence indicated that the positive influence of pipeline maturity on model reliability, and the protective influence of reliability on safety outcomes, were stronger in environments with higher interoperability and more consistent decision-support workflows, emphasizing that technical readiness and clinical use conditions jointly shaped safety gains. Reliability and validity testing confirmed that the composite measures were statistically sound, and collinearity diagnostics supported stable multivariable estimation, strengthening confidence in the inferential pattern. Collectively, the findings positioned cybersecurity and pipeline engineering not as peripheral IT features but as quantifiable determinants of analytic trustworthiness and clinical harm reduction, while also showing that machine learning contributed meaningfully to safety improvement when embedded in robust and auditable data infrastructures. In sum, the study produced an empirically integrated account of how secure data pipelines enabled reliable machine learning safety surveillance and how these combined technical conditions related to measurable reductions in patient harm within EHR-driven care.

RECOMMENDATIONS

Recommendations for practice, policy, and research were derived from the integrated quantitative results linking secure pipelines, ML reliability, and patient safety. Healthcare providers should prioritize routine measurement of pipeline maturity using standardized numeric indicators that cover confidentiality, integrity, availability, and data-quality dimensions, because the evidence showed that these variables were not interchangeable with adoption status and displayed meaningful cross-site variability tied to safety outcomes. Security and data-engineering teams should institutionalize continuous monitoring of missingness patterns, timestamp fidelity, unit harmonization, and lineage traceability, and should treat abrupt changes in these indicators as safety-relevant signals requiring rapid remediation before model retraining or deployment cycles proceed. Providers should also embed high-coverage encryption, complete audit logging, and role-appropriate access controls across all analytic stages, while ensuring that these controls do not degrade real-time availability, since downtime and latency were measurable threats to both clinical workflow and ML alert lead time. Machine learning governance programs should define a minimum reliability threshold for safety models that includes calibration stability, transportability checks, and false-alarm burden limits in addition to discrimination, because the study showed that reliability is multi-component and that safety benefits track the bundle, not a single metric. Before clinical rollout, models should undergo cross-provider or cross-unit validation whenever feasible, and local recalibration should be documented as part of model lineage to reduce drift and prevent silent harm in heterogeneous coding environments. Clinical leadership should pair reliable ML alerts with workflow reinforcement, including standardized response pathways, alert escalation rules, and periodic training on how ML signals map to bedside action, because moderation findings implied that analytic benefits were strongest where workflows supported consistent signal uptake. At the policy level, regulators and accreditation bodies should encourage uniform reporting of pipeline security benchmarks alongside ML performance claims in patient-safety tools, enabling external comparison and reducing the current evidence fragmentation. Data-sharing consortia should support interoperable security–quality benchmarking so that multi-provider joint-effect evaluations become routine rather than exceptional. For research, future quantitative studies should retain standardized EHR safety endpoint definitions and explicitly model pipeline variables as determinants of ML reliability to avoid omitted-variable bias that can distort conclusions about algorithmic effectiveness. Collectively, these recommendations emphasize that safety gains emerge from coordinated advancement of secure data pipelines, reliable ML models, and aligned clinical workflows, and that each domain should be evaluated continuously using measurable indicators rather than treated as a one-time implementation milestone.

LIMITATIONS

Several limitations were inherent in the quantitative design and should be considered when interpreting the study's results. First, the retrospective observational structure constrained causal inference because secure pipeline maturity, ML reliability, and safety outcomes were measured from naturally occurring variation rather than randomized or experimentally manipulated conditions. Although multilevel adjustment, mediation testing, and sensitivity analyses reduced confounding, unmeasured institutional factors such as leadership safety culture, staffing resilience, vendor upgrade cycles, or parallel quality-improvement programs may still have influenced both pipeline characteristics and harm rates. Second, pipeline maturity was operationalized through a composite index derived from available security, integrity, availability, and data-quality indicators, yet some security dimensions could have been incompletely captured due to differences in local logging practices and the varying granularity of IT metadata across providers. Third, the multi-provider sampling frame improved generalizability relative to single-site studies, but participating organizations were more likely to be digitally advanced and capable of exporting de-identified analytics, which may have limited representativeness of smaller rural facilities or low-resource safety-net settings where pipeline maturity and cyber risk profiles can differ systematically. Fourth, patient safety endpoints were derived from EHR trigger algorithms and structured event definitions; although these methods are widely used, they remain dependent on documentation completeness and coding fidelity, creating potential under-ascertainment of harms in sites or units with sparse recording. Fifth, ML reliability was summarized as a bundle of discrimination, calibration, false-alarm burden, lead

time, and transportability, yet the relative weight of each component may vary by clinical context, meaning that the composite reliability score may have masked endpoint-specific nuances. Sixth, cross-site transportability evaluation identified performance decay, but the analysis could not fully disentangle whether drops arose from true population differences, local workflow variation, or residual pipeline incompatibilities that were not directly measurable. Seventh, temporal alignment of pipeline indicators, model training windows, and safety outcomes was implemented to the extent allowed by available timestamps; however, minor asynchrony between infrastructure changes and clinical outcome observation may have attenuated some associations. Finally, the study focused on adult encounters and did not evaluate pediatric or specialty populations, limiting extension of findings to environments where data structures, safety risks, and ML behavior differ. Taken together, these limitations imply that the reported relationships likely reflected robust system-level patterns but should be interpreted as strong associative evidence rather than definitive causal proof, and as most applicable to U.S. providers with moderate to high levels of EHR digitization and security reporting capacity.

REFERENCES

[1]. Abbott, P. A., & Weinger, M. B. (2020). Health information technology: fallacies and sober realities–Redux A homage to Bentzi Karsh and Robert Wears. *Applied Ergonomics*, *82*, 102973.

[2]. 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

[3]. Abedi, V., Avula, V., Chaudhary, D., Shahjouei, S., Khan, A., Griessenauer, C. J., Li, J., & Zand, R. (2021). Prediction of long-term stroke recurrence using machine learning models. *Journal of clinical medicine*, *10*(6), 1286.

[4]. AbuSalim, S., Zakaria, N., Islam, M. R., Kumar, G., Mokhtar, N., & Abdulkadir, S. J. (2022). Analysis of deep learning techniques for dental informatics: a systematic literature review. Healthcare,

[5]. Ahn, J. C., Connell, A., Simonetto, D. A., Hughes, C., & Shah, V. H. (2021). Application of artificial intelligence for the diagnosis and treatment of liver diseases. *Hepatology*, *73*(6), 2546-2563.

[6]. Akrivopoulos, O., Chatzigiannakis, I., Tselios, C., & Antoniou, A. (2017). On the deployment of healthcare applications over fog computing infrastructure. 2017 IEEE 41st annual computer software and applications conference (COMPSAC),

[7]. 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

[8]. Banerjee, A., Chen, S., Fatemifar, G., Zeina, M., Lumbers, R. T., Mielke, J., Gill, S., Kotecha, D., Freitag, D. F., & Denaxas, S. (2021). Machine learning for subtype definition and risk prediction in heart failure, acute coronary syndromes and atrial fibrillation: systematic review of validity and clinical utility. *BMC medicine*, *19*(1), 85.

[9]. Barbosa, P., Queiroz, J., Santos, D., Figueiredo, A., Leite, F., & Galdino, K. (2018). Re4ch: Requirements engineering for connected health. 2018 IEEE 31st International Symposium on Computer-Based Medical Systems (CBMS),

[10]. Bates, D. W., Levine, D., Syrowatka, A., Kuznetsova, M., Craig, K. J. T., Rui, A., Jackson, G. P., & Rhee, K. (2021). The potential of artificial intelligence to improve patient safety: a scoping review. *NPJ digital medicine*, *4*(1), 54.

[11]. Bean, D. M., Wu, H., Iqbal, E., Dzahini, O., Ibrahim, Z. M., Broadbent, M., Stewart, R., & Dobson, R. J. (2017). Knowledge graph prediction of unknown adverse drug reactions and validation in electronic health records. *Scientific reports*, *7*(1), 16416.

[12]. Bernardini, M., Romeo, L., Misericordia, P., & Frontoni, E. (2019). Discovering the type 2 diabetes in electronic health records using the sparse balanced support vector machine. *IEEE journal of biomedical and health informatics*, *24*(1), 235-246.

[13]. Berre, A. J., Tsalgatidou, A., Francalanci, C., Ivanov, T., Pariente-Lobo, T., Ruiz-Saiz, R., Novalija, I., & Grobelnik, M. (2022). Big data and AI pipeline framework: Technology analysis from a benchmarking perspective. In *Technologies and applications for big data value* (pp. 63-88). Springer.

[14]. Bird, A., Oakden-Rayner, L., McMaster, C., Smith, L. A., Zeng, M., Wechalekar, M. D., Ray, S., Proudman, S., & Palmer, L. J. (2022). Artificial intelligence and the future of radiographic scoring in rheumatoid arthritis: a viewpoint. *Arthritis research & therapy*, *24*(1), 268.

[15]. Blay, V., Li, X., Gerlach, J., Urbina, F., & Ekins, S. (2022). Combining DELs and machine learning for toxicology prediction. *Drug discovery today*, *27*(11), 103351.

[16]. Bonzanini, A. D., Shao, K., Stancampiano, A., Graves, D. B., & Mesbah, A. (2021). Perspectives on machine learning-assisted plasma medicine: Toward automated plasma treatment. *IEEE Transactions on Radiation and Plasma Medical Sciences*, *6*(1), 16-32.

[17]. Boriani, G., Da Costa, A., Quesada, A., Ricci, R. P., Favale, S., Boscolo, G., Clementy, N., Amori, V., Mangoni di S. Stefano, L., & Burri, H. (2017). Effects of remote monitoring on clinical outcomes and use of healthcare resources in heart failure patients with biventricular defibrillators: results of the MORE-CARE multicentre randomized controlled trial. *European Journal of heart failure*, *19*(3), 416-425.

[18]. Briggs, E., de Kamps, M., Hamilton, W., Johnson, O., McInerney, C. D., & Neal, R. D. (2022). Machine learning for risk prediction of oesophago-gastric cancer in primary care: comparison with existing risk-assessment tools. *Cancers*, *14*(20), 5023.

[19]. Bronsert, M., Singh, A. B., Henderson, W. G., Hammermeister, K., Meguid, R. A., & Colborn, K. L. (2020). Identification of postoperative complications using electronic health record data and machine learning. *The American Journal of Surgery*, *220*(1), 114-119.

[20]. Burdick, H., Pino, E., Gabel-Comeau, D., Gu, C., Roberts, J., Le, S., Slote, J., Saber, N., Pellegrini, E., & Green-Saxena, A. (2020). Validation of a machine learning algorithm for early severe sepsis prediction: a retrospective study predicting severe sepsis up to 48 h in advance using a diverse dataset from 461 US hospitals. *BMC medical informatics and decision making*, *20*(1), 276.

[21]. Burke, T. A., Jacobucci, R., Ammerman, B. A., Alloy, L. B., & Diamond, G. (2020). Using machine learning to classify suicide attempt history among youth in medical care settings. *Journal of affective disorders*, *268*, 206-214.

[22]. Callahan, A., Fries, J. A., Ré, C., Huddleston III, J. I., Giori, N. J., Delp, S., & Shah, N. H. (2019). Medical device surveillance with electronic health records. *NPJ digital medicine*, *2*(1), 94.

[23]. Castellano, N. G., Mosaly, P., & Mazur, L. (2019). Association Between Physicians' Burden and Performance During Interactions with Electronic Health Records (EHRs). International Conference on Applied Human Factors and Ergonomics,

[24]. Comín-Colet, J., Manito, N., Segovia-Cubero, J., Delgado, J., García Pinilla, J. M., Almenar, L., Crespo-Leiro, M. G., Sionis, A., Blasco, T., & Pascual-Figal, D. (2018). Efficacy and safety of intermittent intravenous outpatient administration of levosimendan in patients with advanced heart failure: the LION-HEART multicentre randomised trial. *European Journal of heart failure*, *20*(7), 1128-1136.

[25]. Corbin, C. K., Sung, L., Chattopadhyay, A., Noshad, M., Chang, A., Deresinksi, S., Baiocchi, M., & Chen, J. H. (2022). Personalized antibiograms for machine learning driven antibiotic selection. *Communications medicine*, *2*(1), 38.

[26]. Dagenais, S., Russo, L., Madsen, A., Webster, J., & Becnel, L. (2022). Use of real-world evidence to drive drug development strategy and inform clinical trial design. *Clinical Pharmacology & Therapeutics*, *111*(1), 77-89.

[27]. Damman, K., Beusekamp, J. C., Boorsma, E. M., Swart, H. P., Smilde, T. D., Elvan, A., van Eck, J. M., Heerspink, H. J., & Voors, A. A. (2020). Randomized, double-blind, placebo-controlled, multicentre pilot study on the effects of empagliflozin on clinical outcomes in patients with acute decompensated heart failure (EMPA-RESPONSE-AHF). *European Journal of heart failure*, *22*(4), 713-722.

[28]. Davazdahemami, B., Peng, P., & Delen, D. (2022). A deep learning approach for predicting early bounce-backs to the emergency departments. *Healthcare Analytics*, *2*, 100018.

[29]. Ellis, R. J., Wang, Z., Genes, N., & Ma'ayan, A. (2019). Predicting opioid dependence from electronic health records with machine learning. *BioData mining*, *12*(1), 3.

[30]. Fang, H. S. A., Tan, N. C., Tan, W. Y., Oei, R. W., Lee, M. L., & Hsu, W. (2021). Patient similarity analytics for explainable clinical risk prediction. *BMC medical informatics and decision making*, *21*(1), 207.

[31]. Feld, E., Harton, J., Meropol, N. J., Adamson, B. J., Cohen, A., Parikh, R. B., Galsky, M. D., Narayan, V., Christodouleas, J., & Vaughn, D. J. (2019). Effectiveness of first-line immune checkpoint blockade versus carboplatin-based chemotherapy for metastatic urothelial cancer. *European urology*, *76*(4), 524-532.

[32]. 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

[33]. 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

[34]. Ferreira, L. L., & Andricopulo, A. D. (2019). ADMET modeling approaches in drug discovery. *Drug discovery today*, *24*(5), 1157-1165.

[35]. Fouladvand, S., Mielke, M. M., Vassilaki, M., Sauver, J. S., Petersen, R. C., & Sohn, S. (2019). Deep learning prediction of mild cognitive impairment using electronic health records. 2019 IEEE international conference on bioinformatics and biomedicine (BIBM),

[36]. Gallego, B., Magrabi, F., Concha, O. P., Wang, Y., & Coiera, E. (2015). Insights into temporal patterns of hospital patient safety from routinely collected electronic data. *Health Information Science and Systems*, *3*(Suppl 1), S2.

[37]. Golas, S. B., Shibahara, T., Agboola, S., Otaki, H., Sato, J., Nakae, T., Hisamitsu, T., Kojima, G., Felsted, J., & Kakarmath, S. (2018). A machine learning model to predict the risk of 30-day readmissions in patients with heart failure: a retrospective analysis of electronic medical records data. *BMC medical informatics and decision making*, *18*(1), 44.

[38]. 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

[39]. Halappanavar, S., Van Den Brule, S., Nymark, P., Gaté, L., Seidel, C., Valentino, S., Zhernovkov, V., Høgh Danielsen, P., De Vizcaya, A., & Wolff, H. (2020). Adverse outcome pathways as a tool for the design of testing strategies to support the safety assessment of emerging advanced materials at the nanoscale. *Particle and Fibre Toxicology*, *17*(1), 16.

[40]. Hammouda, N., & Neyra, J. A. (2022). Can artificial intelligence assist in delivering continuous renal replacement therapy? *Advances in chronic kidney disease*, *29*(5), 439-449.

[41]. Hendrycks, D., Carlini, N., Schulman, J., & Steinhardt, J. (2021). Unsolved problems in ml safety. *arXiv preprint arXiv:2109.13916*.

[42]. Hill, B. L., Brown, R., Gabel, E., Rakocz, N., Lee, C., Cannesson, M., Baldi, P., Loohuis, L. O., Johnson, R., & Jew, B. (2019). An automated machine learning-based model predicts postoperative mortality using readily-extractable preoperative electronic health record data. *British journal of anaesthesia*, *123*(6), 877-886.

[43]. Hilton, C. B., Milinovich, A., Felix, C., Vakharia, N., Crone, T., Donovan, C., Proctor, A., & Nazha, A. (2020). Personalized predictions of patient outcomes during and after hospitalization using artificial intelligence. *NPJ digital medicine*, *3*(1), 51.

[44]. Ho, J., Weber, J., & Price, M. (2017). BXE2E: A bidirectional transformation approach for medical record exchange. International Conference on Theory and Practice of Model Transformations,

[45]. Hsu, C.-C., Chu, C.-C., Lin, C.-H., Huang, C.-H., Ng, C.-J., Lin, G.-Y., Chiou, M.-J., Lo, H.-Y., & Chen, S.-Y. (2021). A machine learning model for predicting unscheduled 72 h return visits to the emergency department by patients with abdominal pain. *Diagnostics*, *12*(1), 82.

[46]. Hurst, W., Tekinerdogan, B., Alskaif, T., Boddy, A., & Shone, N. (2022). Securing electronic health records against insider-threats: A supervised machine learning approach. *Smart Health*, *26*, 100354.

[47]. Iscoe, M. S., McLean, R. M., & Melnick, E. R. (2022). Restoring meaningful content to the medical record: Standardizing measurement could improve EHR utility while decreasing burden. Mayo Clinic Proceedings,

[48]. Islam, M. U., Mozaharul Mottalib, M., Hassan, M., Alam, Z. I., Zobaed, S., & Fazle Rabby, M. (2022). The past, present, and prospective future of xai: A comprehensive review. *Explainable Artificial Intelligence for Cyber Security: Next Generation Artificial Intelligence*, 1-29.

[49]. Jarabek, A. M., & Hines, D. E. (2019). Mechanistic integration of exposure and effects: advances to apply systems toxicology in support of regulatory decision-making. *Current Opinion in Toxicology*, *16*, 83-92.

[50]. Jauk, S., Kramer, D., Avian, A., Berghold, A., Leodolter, W., & Schulz, S. (2021). Technology acceptance of a machine learning algorithm predicting delirium in a clinical setting: a mixed-methods study. *Journal of medical systems*, *45*(4), 48.

[51]. Jia, Y., McDermid, J., Lawton, T., & Habli, I. (2022). The role of explainability in assuring safety of machine learning in healthcare. *IEEE Transactions on Emerging Topics in Computing*, *10*(4), 1746-1760.

[52]. Jung, J. W., Hwang, S., Ko, S., Jo, C., Park, H. Y., Han, H.-S., Lee, M. C., Park, J. E., & Ro, D. H. (2022). A machine-learning model to predict postoperative delirium following knee arthroplasty using electronic health records. *BMC psychiatry*, *22*(1), 436.

[53]. Khairat, S., Coleman, C., Newlin, T., Rand, V., Ottmar, P., Bice, T., & Carson, S. S. (2019). A mixed-methods evaluation framework for electronic health records usability studies. *Journal of biomedical informatics*, *94*, 103175.

[54]. Kim, E., Rubinstein, S. M., Nead, K. T., Wojcieszynski, A. P., Gabriel, P. E., & Warner, J. L. (2019). The evolving use of electronic health records (EHR) for research. Seminars in radiation oncology,

[55]. Kogan, E., Twyman, K., Heap, J., Milentijevic, D., Lin, J. H., & Alberts, M. (2020). Assessing stroke severity using electronic health record data: a machine learning approach. *BMC medical informatics and decision making*, *20*(1), 8.

[56]. Landi, I., Glicksberg, B. S., Lee, H.-C., Cherng, S., Landi, G., Danieletto, M., Dudley, J. T., Furlanello, C., & Miotto, R. (2020). Deep representation learning of electronic health records to unlock patient stratification at scale. *NPJ digital medicine*, *3*(1), 96.

[57]. Latif, J., Xiao, C., Tu, S., Rehman, S. U., Imran, A., & Bilal, A. (2020). Implementation and use of disease diagnosis systems for electronic medical records based on machine learning: A complete review. *IEEE Access*, *8*, 150489-150513.

[58]. Levin, S., Toerper, M., Hamrock, E., Hinson, J. S., Barnes, S., Gardner, H., Dugas, A., Linton, B., Kirsch, T., & Kelen, G. (2018). Machine-learning-based electronic triage more accurately differentiates patients with respect to clinical outcomes compared with the emergency severity index. *Annals of emergency medicine*, *71*(5), 565-574. e562.

[59]. Li, C., Zhang, Z., Ren, Y., Nie, H., Lei, Y., Qiu, H., Xu, Z., & Pu, X. (2021). Machine learning based early mortality prediction in the emergency department. *International journal of medical informatics*, *155*, 104570.

[60]. Lindberg, D. S., Prosperi, M., Bjarnadottir, R. I., Thomas, J., Crane, M., Chen, Z., Shear, K., Solberg, L. M., Snigurska, U. A., & Wu, Y. (2020). Identification of important factors in an inpatient fall risk prediction model to improve the quality of care using EHR and electronic administrative data: a machine-learning approach. *International journal of medical informatics*, *143*, 104272.

[61]. Liu, F., Jagannatha, A., & Yu, H. (2019). Towards drug safety surveillance and pharmacovigilance: current progress in detecting medication and adverse drug events from electronic health records. *Drug safety*, *42*(1), 95-97.

[62]. Liu, W., Park, E. K., Krieger, U., & Zhu, S. (2020). Smart e-health security and safety monitoring with machine learning services. 2020 29th International Conference on Computer Communications and Networks (ICCCN),

[63]. Loftus, T. J., Tighe, P. J., Filiberto, A. C., Balch, J., Upchurch Jr, G. R., Rashidi, P., & Bihorac, A. (2020). Opportunities for machine learning to improve surgical ward safety. *The American Journal of Surgery*, *220*(4), 905-913.

[64]. Lorberbaum, T., Sampson, K. J., Woosley, R. L., Kass, R. S., & Tatonetti, N. P. (2016). An integrative data science pipeline to identify novel drug interactions that prolong the QT interval. *Drug safety*, *39*(5), 433-441.

[65]. Lutjeboer, J., Burgmans, M. C., Chung, K., & van Erkel, A. R. (2015). Impact on Patient Safety and Satisfaction of Implementation of an Outpatient Clinic in Interventional Radiology (IPSIPOLI-Study): a quasi-experimental prospective study. *Cardiovascular and interventional radiology*, *38*(3), 543-551.

[66]. Luz, C. F., Vollmer, M., Decruyenaere, J., Nijsten, M. W., Glasner, C., & Sinha, B. (2020). Machine learning in infection management using routine electronic health records: tools, techniques, and reporting of future technologies. *Clinical Microbiology and Infection*, *26*(10), 1291-1299.

[67]. Mamun, M., Farjana, A., Al Mamun, M., Ahammed, M. S., & Rahman, M. M. (2022). Heart failure survival prediction using machine learning algorithm: am I safe from heart failure? 2022 IEEE world AI IoT congress (AIIoT),

[68]. McComb, M., Bies, R., & Ramanathan, M. (2022). Machine learning in pharmacometrics: Opportunities and challenges. *British Journal of Clinical Pharmacology*, *88*(4), 1482-1499.

[69]. McDermott, K. (2016). Achieving data liquidity across health care requires a technical architecture. *Bulletin of the Association for Information Science and Technology*, *43*(1), 19-22.

[70]. 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

[71]. 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

[72]. 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

[73]. 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

[74]. 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

[75]. 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

[76]. 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

[77]. 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

[78]. 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

[79]. Melnick, E. R., Dyrbye, L. N., Sinsky, C. A., Trockel, M., West, C. P., Nedelec, L., Tutty, M. A., & Shanafelt, T. (2020). The association between perceived electronic health record usability and professional burnout among US physicians. Mayo Clinic Proceedings,

[80]. Métris, A., Barrett, P., Price, L., Klamert, S., & Fernandez-Piquer, J. (2022). A tiered approach to risk assess microbiome perturbations induced by application of beauty and personal care products. *Microbial Risk Analysis*, *20*, 100188.

[81]. Minh, D., Wang, H. X., Li, Y. F., & Nguyen, T. N. (2022). Explainable artificial intelligence: a comprehensive review. *Artificial Intelligence Review*, *55*(5), 3503-3568.

[82]. Miotto, R., Li, L., Kidd, B. A., & Dudley, J. T. (2016). Deep patient: an unsupervised representation to predict the future of patients from the electronic health records. *Scientific reports*, *6*(1), 26094.

[83]. 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

[84]. 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

[85]. 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

[86]. Murphy, D. R., Satterly, T., Giardina, T. D., Sittig, D. F., & Singh, H. (2019). Practicing clinicians' recommendations to reduce burden from the electronic health record inbox: a mixed-methods study. *Journal of general internal medicine*, *34*(9), 1825-1832.

[87]. Nguyen, B. P., Pham, H. N., Tran, H., Nghiem, N., Nguyen, Q. H., Do, T. T., Tran, C. T., & Simpson, C. R. (2019). Predicting the onset of type 2 diabetes using wide and deep learning with electronic health records. *Computer methods and programs in biomedicine*, *182*, 105055.

[88]. Pastor, M., Gómez-Tamayo, J. C., & Sanz, F. (2021). Flame: an open source framework for model development, hosting, and usage in production environments. *Journal of cheminformatics*, *13*(1), 31.

[89]. Pavelka, K., Kivitz, A., Dokoupilova, E., Blanco, R., Maradiaga, M., Tahir, H., Pricop, L., Andersson, M., Readie, A., & Porter, B. (2017). Efficacy, safety, and tolerability of secukinumab in patients with active ankylosing spondylitis: a randomized, double-blind phase 3 study, MEASURE 3. *Arthritis research & therapy*, *19*(1), 285.
[90]. Petersen, E., Potdevin, Y., Mohammadi, E., Zidowitz, S., Breyer, S., Nowotka, D., Henn, S., Pechmann, L., Leucker, M., & Rostalski, P. (2022). Responsible and regulatory conform machine learning for medicine: a survey of challenges and solutions. *IEEE Access*, *10*, 58375-58418.
[91]. Pethani, F. (2021). Promises and perils of artificial intelligence in dentistry. *Australian Dental Journal*, *66*(2), 124-135.
[92]. Praveen, S. P., Murali Krishna, T. B., Anuradha, C., Mandalapu, S. R., Sarala, P., & Sindhura, S. (2022). RETRACTED ARTICLE: A robust framework for handling health care information based on machine learning and big data engineering techniques. *International Journal of Healthcare Management*, 1-18.
[93]. Putnam, L. R., Pham, D. H., Ostovar-Kermani, T. G., Alawadi, Z. M., Etchegaray, J. M., Ottosen, M. J., Thomas, E. J., Lesslie, D. P., Kao, L. S., & Lally, K. P. (2016). How should surgical residents be educated about patient safety: a pilot randomized controlled trial. *Journal of Surgical Education*, *73*(4), 660-667.
[94]. Qayyum, A., Qadir, J., Bilal, M., & Al-Fuqaha, A. (2020). Secure and robust machine learning for healthcare: A survey. *IEEE Reviews in Biomedical Engineering*, *14*, 156-180.
[95]. Rahmani, K., Garikipati, A., Barnes, G., Hoffman, J., Calvert, J., Mao, Q., & Das, R. (2022). Early prediction of central line associated bloodstream infection using machine learning. *American journal of infection control*, *50*(4), 440-445.
[96]. Rajkomar, A., Oren, E., Chen, K., Dai, A. M., Hajaj, N., Hardt, M., Liu, P. J., Liu, X., Marcus, J., & Sun, M. (2018). Scalable and accurate deep learning with electronic health records. *NPJ digital medicine*, *1*(1), 18.
[97]. 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
[98]. Rex, D. K., Bhandari, R., Desta, T., DeMicco, M. P., Schaeffer, C., Etzkorn, K., Barish, C. F., Pruitt, R., Cash, B. D., & Quirk, D. (2018). A phase III study evaluating the efficacy and safety of remimazolam (CNS 7056) compared with placebo and midazolam in patients undergoing colonoscopy. *Gastrointestinal endoscopy*, *88*(3), 427-437. e426.
[99]. 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.
[100]. Rinella, M. E., Tacke, F., Sanyal, A. J., Anstee, Q. M., & Workshop, p. o. t. A. E. (2019). Report on the AASLD/EASL joint workshop on clinical trial endpoints in NAFLD. *Hepatology*, *70*(4), 1424-1436.
[101]. Rochefort, C. M., Buckeridge, D. L., & Abrahamowicz, M. (2015). Improving patient safety by optimizing the use of nursing human resources. *Implementation Science*, *10*(1), 89.
[102]. Rojas, J. C., Fahrenbach, J., Makhni, S., Cook, S. C., Williams, J. S., Umscheid, C. A., & Chin, M. H. (2022). Framework for integrating equity into machine learning models: a case study. *Chest*, *161*(6), 1621-1627.
[103]. Rozenblum, R., Rodriguez-Monguio, R., Volk, L. A., Forsythe, K. J., Myers, S., McGurrin, M., Williams, D. H., Bates, D. W., Schiff, G., & Seoane-Vazquez, E. (2020). Using a machine learning system to identify and prevent medication prescribing errors: a clinical and cost analysis evaluation. *The Joint Commission Journal on Quality and Patient Safety*, *46*(1), 3-10.
[104]. 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
[105]. 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
[106]. Salas, M., Petracek, J., Yalamanchili, P., Aimer, O., Kasthuril, D., Dhingra, S., Junaid, T., & Bostic, T. (2022). The use of artificial intelligence in pharmacovigilance: a systematic review of the literature. *Pharmaceutical medicine*, *36*(5), 295-306.
[107]. Salay, R., & Czarnecki, K. (2019). Improving ML safety with partial specifications. international conference on computer safety, reliability, and security,
[108]. Scott, I. A. (2021). Demystifying machine learning: a primer for physicians. *Internal Medicine Journal*, *51*(9), 1388-1400.
[109]. Seh, A. H., Al-Amri, J. F., Subahi, A. F., Agrawal, A., Pathak, N., Kumar, R., & Khan, R. A. (2022). An analysis of integrating machine learning in healthcare for ensuring confidentiality of the electronic records. *Computer Modeling in Engineering & Sciences*, *130*(3), 1387-1422.
[110]. 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
[111]. Srivastava, S., Soman, S., Rai, A., & Srivastava, P. K. (2017). Deep learning for health informatics: Recent trends and future directions. 2017 international conference on advances in computing, communications and informatics (ICACCI),
[112]. Stewart, C. E., Kan, C. F. K., Stewart, B. R., Sanicola III, H. W., Jung, J. P., Sulaiman, O. A., & Wang, D. (2020). Machine intelligence for nerve conduit design and production. *Journal of Biological Engineering*, *14*(1), 25.
[113]. Strudwick, G., Jeffs, L., Kemp, J., Sequeira, L., Lo, B., Shen, N., Paterson, P., Coombe, N., Yang, L., & Ronald, K. (2022). Identifying and adapting interventions to reduce documentation burden and improve nurses' efficiency in using electronic health record systems (The IDEA Study): protocol for a mixed methods study. *BMC nursing*, *21*(1), 213.

[114]. Swain, S., Bhushan, B., Dhiman, G., & Viriyasitavat, W. (2022). Appositeness of optimized and reliable machine learning for healthcare: a survey. *Archives of Computational Methods in Engineering*, *29*(6), 3981-4003.
[115]. Tambon, F., Laberge, G., An, L., Nikanjam, A., Mindom, P. S. N., Pequignot, Y., Khomh, F., Antoniol, G., Merlo, E., & Laviolette, F. (2022). How to certify machine learning based safety-critical systems? A systematic literature review. *Automated Software Engineering*, *29*(2), 38.
[116]. Tayefi, M., Ngo, P., Chomutare, T., Dalianis, H., Salvi, E., Budrionis, A., & Godtliebsen, F. (2021). Challenges and opportunities beyond structured data in analysis of electronic health records. *Wiley Interdisciplinary Reviews: Computational Statistics*, *13*(6), e1549.
[117]. Terranova, N., Venkatakrishnan, K., & Benincosa, L. J. (2021). Application of machine learning in translational medicine: current status and future opportunities. *The AAPS Journal*, *23*(4), 74.
[118]. Tomašev, N., Harris, N., Baur, S., Mottram, A., Glorot, X., Rae, J. W., Zielinski, M., Askham, H., Saraiva, A., & Magliulo, V. (2021). Use of deep learning to develop continuous-risk models for adverse event prediction from electronic health records. *Nature protocols*, *16*(6), 2765-2787.
[119]. 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
[120]. Trout, K. E., Chen, L.-W., Wilson, F. A., Tak, H. J., & Palm, D. (2022). The impact of meaningful use and electronic health records on hospital patient safety. *International journal of environmental research and public health*, *19*(19), 12525.
[121]. Van Der Heijde, D., Deodhar, A., Wei, J. C., Drescher, E., Fleishaker, D., Hendrikx, T., Li, D., Menon, S., & Kanik, K. S. (2017). Tofacitinib in patients with ankylosing spondylitis: a phase II, 16-week, randomised, placebo-controlled, dose-ranging study. *Annals of the rheumatic diseases*, *76*(8), 1340-1347.
[122]. Varshney, K. R. (2016). Engineering safety in machine learning. 2016 Information Theory and Applications Workshop (ITA),
[123]. Vatansever, S., Schlessinger, A., Wacker, D., Kaniskan, H. Ü., Jin, J., Zhou, M. M., & Zhang, B. (2021). Artificial intelligence and machine learning-aided drug discovery in central nervous system diseases: State-of-the-arts and future directions. *Medicinal research reviews*, *41*(3), 1427-1473.
[124]. Wang, L., Xue, Z., Ezeana, C. F., Puppala, M., Chen, S., Danforth, R. L., Yu, X., He, T., Vassallo, M. L., & Wong, S. T. (2019). Preventing inpatient falls with injuries using integrative machine learning prediction: a cohort study. *NPJ digital medicine*, *2*(1), 127.
[125]. Windle, J. R., Windle, T. A., Shamavu, K. Y., Nelson, Q. M., Clarke, M. A., Fruhling, A. L., & Tcheng, J. E. (2021). Roadmap to a more useful and usable electronic health record. *Cardiovascular Digital Health Journal*, *2*(6), 301-311.
[126]. Wong, J., Murray Horwitz, M., Zhou, L., & Toh, S. (2018). Using machine learning to identify health outcomes from electronic health record data. *Current epidemiology reports*, *5*(4), 331-342.
[127]. Wronikowska, M. W., Malycha, J., Morgan, L. J., Westgate, V., Petrinic, T., Young, J. D., & Watkinson, P. J. (2021). Systematic review of applied usability metrics within usability evaluation methods for hospital electronic healthcare record systems: Metrics and Evaluation Methods for eHealth Systems. *Journal of Evaluation in Clinical Practice*, *27*(6), 1403-1416.
[128]. Wunderink, R. G., Giamarellos-Bourboulis, E. J., Rahav, G., Mathers, A. J., Bassetti, M., Vazquez, J., Cornely, O. A., Solomkin, J., Bhowmick, T., & Bishara, J. (2018). Effect and safety of meropenem–vaborbactam versus best-available therapy in patients with carbapenem-resistant Enterobacteriaceae infections: the TANGO II randomized clinical trial. *Infectious diseases and therapy*, *7*(4), 439-455.
[129]. Xie, F., Zhou, J., Lee, J. W., Tan, M., Li, S., Rajnthern, L. S. O., Chee, M. L., Chakraborty, B., Wong, A.-K. I., & Dagan, A. (2022). Benchmarking emergency department prediction models with machine learning and public electronic health records. *Scientific Data*, *9*(1), 658.
[130]. Xie, Y., Maziarz, M., Tuot, D. S., Chertow, G. M., Himmelfarb, J., & Hall, Y. N. (2016). Risk prediction to inform surveillance of chronic kidney disease in the US Healthcare Safety Net: a cohort study. *BMC nephrology*, *17*(1), 57.
[131]. Xu, Z., & Saleh, J. H. (2021). Machine learning for reliability engineering and safety applications: Review of current status and future opportunities. *Reliability Engineering & System Safety*, *211*, 107530.
[132]. Yeboah-Ofori, A., Islam, S., Lee, S. W., Shamszaman, Z. U., Muhammad, K., Altaf, M., & Al-Rakhami, M. S. (2021). Cyber threat predictive analytics for improving cyber supply chain security. *IEEE Access*, *9*, 94318-94337.
[133]. Zhang, T., Androulakis, I. P., Bonate, P., Cheng, L., Helikar, T., Parikh, J., Rackauckas, C., Subramanian, K., & Cho, C. R. (2022). Two heads are better than one: current landscape of integrating QSP and machine learning: an ISoP QSP SIG white paper by the working group on the integration of quantitative systems pharmacology and machine learning. *Journal of Pharmacokinetics and Pharmacodynamics*, *49*(1), 5-18.
[134]. Zheng, L., Wang, O., Hao, S., Ye, C., Liu, M., Xia, M., Sabo, A. N., Markovic, L., Stearns, F., & Kanov, L. (2020). Development of an early-warning system for high-risk patients for suicide attempt using deep learning and electronic health records. *Translational psychiatry*, *10*(1), 72.