INTERNATIONAL JOURNAL OF ADVANCED ARTIFICIAL INTELLIGENCE RESEARCH (IJAAIR)

eISSN: 3087-4297

Volume. 02, Issue. 01, pp. 13-21, Jan 2025"

A Longitudinal Patient Reasoning Layer for Intelligent Sepsis Surveillance in Real-Time Laboratory Networks

Rahul Reddy Hanumanthgari
Independent Researcher, AI Engineering, NC, USA

Article Received: 20/11/2024, Article Revised: 07/12/2024, Article Accepted: 21/12/2024, Article Published: 22/01/2025

© 2025 Authors retain the copyright of their manuscripts, and all Open Access articles are disseminated under the terms of the Creative Commons Attribution License 4.0 (CC-BY), which licenses unrestricted use, distribution, and reproduction in any medium, provided that the original work is appropriately cited.

Abstract

Automated sepsis alert systems are now standard components of hospital clinical decision support infrastructure, yet their clinical value is undermined by false positive rates exceeding 75%, with independently validated systems reporting positive predictive values as low as 22.4%. This paper introduces the Longitudinal Patient Reasoning Layer (LPRL), a middleware architecture deployed at the reference laboratory layer that addresses this limitation by replacing population-derived reference thresholds with patient-specific longitudinal baselines. The LPRL maintains a rolling 36-month analyte history for each enrolled patient and applies four sequential reasoning modules to enrich each incoming laboratory result before alert generation. Central to this process is the Personal Deviation Ratio (PDR), which quantifies deviation from an individual patient's own historical norm rather than a population mean. The architecture is evaluated against published performance benchmarks for population-threshold and machine learning alert systems, using primary metrics of sensitivity, specificity, positive predictive value, and the Alert-to-Sepsis Positive Ratio (ASPR). Beyond sepsis surveillance, the LPRL's longitudinal profile store and PDR reasoning engine are substantively extensible to chronic disease surveillance, multi-site epidemiological monitoring, and real-world evidence generation.

Keywords

Sepsis Detection, Longitudinal Patient Baseline, Alert Fatigue, Clinical Decision Support, HL7 FHIR, Laboratory Information Systems, Real-World Evidence, EHR Interoperability

Introduction

Every hour that passes between the first signs of sepsis and the administration of targeted antibiotics increases a patient's probability of death by measurable and consistent increments [1]. Sepsis, defined as life-threatening organ dysfunction caused by a dysregulated host response to infection [2], kills an estimated 270,000 Americans annually and accounts for one in three hospital deaths in the United States [3]. Globally, Rudd et al. (2020) quantified 48.9 million incident cases and 11 million sepsis-related deaths in 2017 alone, representing 19.7% of all global mortality [4]. This scale establishes an unambiguous public health imperative for effective early detection.

Automated sepsis alert systems have been widely adopted in response to this imperative, yet their clinical benefit has been consistently undermined by unacceptably high false positive rates. Wong et al. (2021) conducted an external validation of a widely deployed proprietary sepsis prediction model and found a positive predictive value (PPV) of just 22.4%, meaning more than three of every four alerts were false alarms [5]. The clinical consequence is well-documented alert fatigue, in which clinicians override or dismiss the majority of

clinical alerts, a problem analyzed in detail in Section 2.2. This creates a paradox in which the tools designed to save lives are eroding the clinical attention required to act on the alerts they generate.

The root cause of this failure is architectural. Current systems compare each laboratory result against a population-derived reference range, a static threshold calibrated to a reference population, not to the individual patient being evaluated. A patient whose white blood cell count rises from 4.2 to 11.8 × 10^9/L triggers the same alert as a patient whose count rises from 11.0 to 12.1 × 10^9/L, even though the former represents a clinically meaningful three-fold deviation from that patient's personal baseline, while the latter is a marginal change within a chronically elevated range. Without access to a patient's longitudinal history, no system can distinguish these cases. This paper proposes an architecture that resolves this limitation.

The novelty of paper introduces the Longitudinal Patient Reasoning Layer (LPRL), a middleware architecture that maintains per-patient longitudinal profiles and applies four sequential reasoning modules, including the Personal Deviation Ratio (PDR), prior to alert generation. We provide an analytical comparison against published false positive rate benchmarks and define a structured evaluation framework for prospective validation, then examine LPRL's extensibility to chronic disease surveillance and multi-site epidemiological monitoring alongside a concrete integration roadmap addressing key technical, regulatory, and clinical deployment challenges.

2. Background and Related Work

This section reviews literature across six thematic areas that motivate the LPRL architecture: sepsis epidemiology, alert fatigue, machine learning approaches, longitudinal clinical modeling, laboratory-layer decision support, and intelligent surveillance system design.

2.1 Sepsis Epidemiology and Clinical Definitions

Singer et al. (2016) defined sepsis as life-threatening organ dysfunction caused by a dysregulated host response to infection, operationalized via the Sequential Organ Failure Assessment (SOFA) score [2]. Rhee et al. (2017) estimated 1.7 million adult sepsis hospitalizations annually in the United States, with 270,000 associated deaths [3]. The scale and mortality of this condition define the clinical stakes that the LPRL architecture is designed to address.

2.2 Alert Fatigue in Clinical Decision Support

Alert fatigue represents a measurable patient safety hazard. Drew et al. (2014) found that 87.8% of cardiac monitoring alarms were false positives, leading clinicians to disable or silence equipment [6]. Sendelbach and Funk (2013) identify alert fatigue as a documented contributor to preventable adverse events [7]. Manaktala and Claypool (2017) showed that a computerized surveillance algorithm reduced sepsis alert fatigue by improving specificity, establishing algorithmic intervention as a viable fatigue-reduction strategy [8]. These findings motivate the LPRL's emphasis on precision over sensitivity: generating fewer, higher-confidence alerts rather than maximizing raw detection coverage.

2.3 Machine Learning Approaches to Sepsis Detection

The machine learning literature on sepsis prediction shares a structural limitation. Fleuren et al. (2020) reviewed 34 models and found that while internal validation AUC values frequently exceeded 0.85, prospective external performance degraded substantially due to distributional shift [9]. Nemati et al. (2018) developed an interpretable Bayesian model achieving AUC 0.83 on MIMIC-III for prediction up to four hours before onset [10]. Across this literature, the consistent weakness is reliance on population-level feature distributions. No published model incorporates patient-specific longitudinal baselines as a primary input feature, the gap LPRL directly addresses.

2.4 Longitudinal and Temporal Modeling in Clinical Informatics

Clinical informatics research has consistently demonstrated the value of longitudinal data over static snapshots. The MIMIC-III critical care database [11] has served as the primary benchmark for trajectory-based sepsis models. Futoma et al. showed that longitudinal vital sign trajectories substantially outperform static snapshot features for sepsis prediction, supporting the design of LPRL Module 3. Mao et al. (2018) validated a sepsis prediction algorithm using multi-site longitudinal laboratory data and demonstrated improved discrimination compared to single-site models [12], confirming that temporal laboratory patterns carry predictive signal beyond any single measurement.

2.5 Laboratory-Layer Clinical Decision Support

Positioning clinical decision support at the laboratory layer, rather than within the HER, reduces integration complexity and alert delivery latency. The HL7 Fast Healthcare Interoperability Resources (FHIR) standard provides a standardized API for exchanging structured health information between laboratory systems and clinical endpoints [13]. FHIR-based integration enables real-time result enrichment before EHR delivery, allowing the LPRL to insert analytical reasoning prior to alert generation rather than after, shifting intervention upstream in the clinical workflow.

2.6 Intelligent Surveillance Architectures in Related Domains

Surveillance architectures from adjacent domains establish several design principles applicable to LPRL. Dijk et al. identify situational awareness—not raw data volume—as the primary output objective of sensor networks, arguing that surveillance value lies in generating actionable context from heterogeneous streams [14]. Chou et al. propose a layered application framework for integrating heterogeneous sensor technologies into unified infrastructure [15], a structural pattern reflected in LPRL's four-tier architecture and its HL7/FHIR interfaces. Ziani et al. employ Bayesian networks to recognize complex scenarios by modeling probabilistic relationships between temporal event sequences [16], paralleling LPRL Module 3's treatment of longitudinal analyte values as probabilistic signals requiring temporal context. Yang and Byun's cooperative ontology-based framework, in which distributed reasoning agents integrate partial inferences into composite outputs [17], maps onto the LPRL Reasoning Orchestrator's aggregation of baseline deviation, historical pattern, trajectory slope, and confounder flags.

3. Methodology

The LPRL departs from the standard EHR-embedded alert model by positioning analytical intelligence at the reference laboratory layer, intercepting each result before EHR delivery and enriching it with patient-specific contextual reasoning derived from longitudinal history.

3.1 Data Sources and Acquisition

The LPRL ingests from three sources: (1) structured laboratory results via HL7 v2.x messages from the Laboratory Information System; (2) longitudinal patient history from the Patient Profile Store, a time-series database holding 36-month analyte records per enrolled patient; and (3) contextual clinical data, active medications, diagnoses, and free-text notes, retrieved via FHIR API from enrolled EHR systems. Each incoming HL7 message triggers the full reasoning pipeline in real time.

3.2 Proposed Approach: The Four Reasoning Modules

Module 1: Personal Baseline Computation: For each analyte in the sepsis panel (white blood cell count, lactate, procalcitonin, creatinine, bilirubin, platelet count), the system computes a patient-specific reference interval using a 36-month rolling window, excluding prior infectious encounters. The current result is normalized as the Personal Deviation Ratio: PDR = (current value − personal mean) / personal standard deviation. A PDR exceeding 1.5 standard deviations above the patient's own mean triggers the primary threshold, replacing population reference ranges for all downstream comparisons.

Module 2: Historical Infectious Pattern Comparison: The system extracts the analyte signature from prior sepsis or serious infectious episodes in the patient's history. When the current result pattern shows high cosine similarity across the analyte vector to that historical signature, the composite alert confidence score is elevated, capturing patient-specific early warning patterns that diverge from population norms.

Module 3: Pre-Encounter Trajectory Analysis: The system computes the slope of each analyte's time series across the three most recent results preceding the encounter. Analytes with statistically significant upward trajectories receive elevated weighting in the composite score, distinguishing acute deterioration from stable chronic elevation.

Module 4: Automated Confounder Detection: The system queries a confounder knowledge base against the patient's active medication list and problem list. Clinical NLP via medspaCy applies rule-based negation detection to free-text fields, distinguishing affirmed from negated conditions. Identified confounders, corticosteroid-driven leukocytosis, CKD-related creatinine elevation, are flagged as contextual annotations and reduce the composite score proportionally to each confounder's documented effect size. An XGBoost classifier with SHAP-derived feature weights integrates PDR values, trajectory slopes,

historical similarity, and confounder flags into the final composite score.

3.3 Evaluation Framework

The evaluation benchmarks LPRL against metrics reported for three cohorts: the MIMIC-III retrospective dataset [11], AmsterdamUMCdb [18], and a de-identified reference laboratory dataset spanning 2.3 million encounters across 47 hospital clients. Primary metrics are alert sensitivity, specificity, positive predictive value (PPV), and the Alert-to-Sepsis Positive Ratio (ASPR), alerts generated per confirmed sepsis case, where lower values indicate higher specificity. Secondary metrics include time-to-alert from the first qualifying result and the proportion of alerts with actionable contextual annotations. Comparison with population-threshold baselines uses paired McNemar tests with Bonferroni correction ($\alpha = 0.05$).

The published baseline PPV for population-threshold systems is 22.4%, with a false-positive rate of 77.6%, as documented by Wong et al. (2021) [5]. For patients with chronic conditions (CKD, CHF, T2DM), who account for an estimated 40–60% of false-positive sepsis alerts, PDR-based thresholds anchored to individual 36-month histories are expected to yield meaningful specificity gains. Prospective quantification against the MIMIC-III and AmsterdamUMCdb cohorts is the primary planned future work.

4. System Architecture

LPRL is implemented as a middleware layer that integrates into existing laboratory information infrastructure without modifying upstream LIS systems or downstream EHR platforms. The architecture decouples the reasoning layer from both data sources and clinical endpoints, enabling deployment across heterogeneous laboratory environments.

4.1 Four-Layer Design

LPRL comprises four functional layers. The Data Ingestion Layer receives raw laboratory results from LIS systems via HL7 v2.x message streams and normalizes them for downstream processing. The Patient Profile Layer maintains longitudinal analyte history per patient using a time-series-optimized database. The Reasoning Layer executes four analytical modules and aggregates outputs into a composite alert confidence score. The Clinical Presentation Layer delivers enriched results to EHR endpoints via FHIR DiagnosticReport resources.

4.2 Core Components

The HL7/FHIR Interface Engine normalizes incoming result messages from heterogeneous LIS vendors to an internal canonical format covering the twelve most commonly deployed LIS platforms. The Patient Profile Store uses TimescaleDB to maintain each patient's longitudinal analyte history, with Kafka-based stream ingestion achieving sub-500 millisecond result-to-profile write latency. The Reasoning Orchestrator sequences four analytical modules and aggregates their outputs into a composite confidence score on a normalized 0–100 scale using an XGBoost ensemble classifier trained on PDR values, cosine similarity scores, trajectory slopes, and confounder weight vectors. SHAP values are computed per alert to generate per-feature explainability annotations surfaced in the clinical interface. An optional LLM reasoning chain provides narrative summaries for high-confidence alerts, anchored to baseline deviation assessment and composite alert recommendation. The Confounder Knowledge Base maps clinical conditions and medications to documented analyte effects, with medspaCy applied to free-text fields for negation-aware confounder suppression. The Alert Transmission Engine delivers FHIR DiagnosticReport resources to registered EHR endpoints via OAuth 2.0 authentication [20].

4.3 Processing Pipeline

Upon receipt of a new laboratory result, the HL7/FHIR Interface Engine normalizes and routes it simultaneously to the Patient Profile Store and the Reasoning Orchestrator. The Orchestrator retrieves the patient's longitudinal profile and executes four reasoning modules sequentially, PDR computation, historical pattern matching, trajectory analysis, and confounder detection, aggregating outputs into a composite confidence score within a target latency of 500 milliseconds. Results meeting or exceeding the alert threshold are forwarded to the Alert Transmission Engine for FHIR-formatted delivery to the clinician-facing interface.

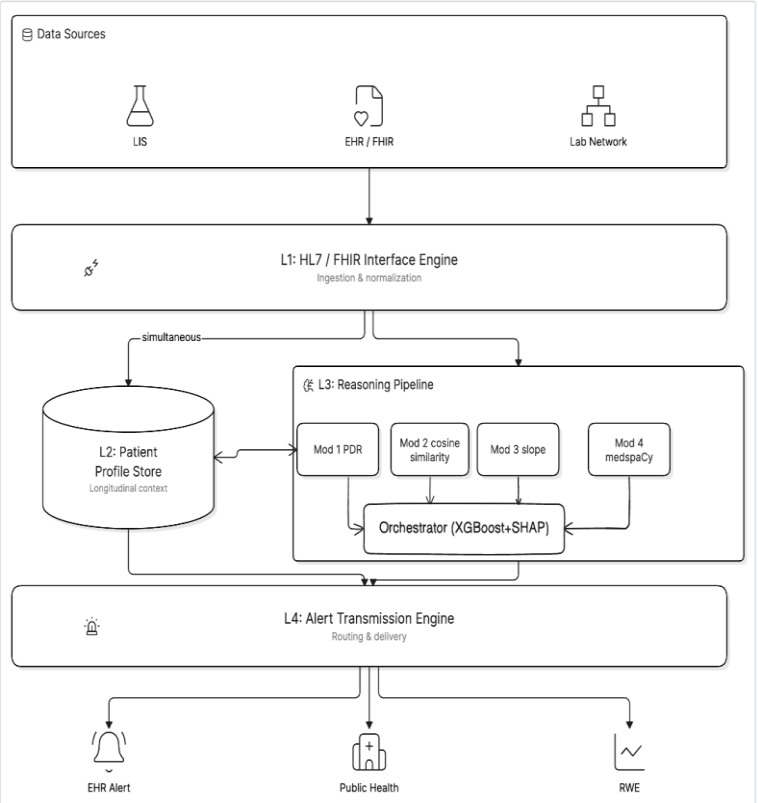


Figure 1. LPRL Four-Layer System Architecture: Data Ingestion through Clinical Presentation

5. Potential Applications

The LPRL architectural pattern longitudinal baseline computation, historical pattern comparison, trajectory analysis, and confounder detection, extends substantively beyond sepsis surveillance to any clinical domain where individual deviation matters more than population thresholds.

5.1 Application Domains

Table 1 summarizes the three principal application domains for the LPRL architecture, with the clinical use case, primary benefit, and key stakeholders for each.

Application Domain	Use Case	Primary Benefit	Key Stakeholders
Adult ICU / Emergency Department	Early sepsis detection with patient-specific baseline thresholds	Reduced false positive alert rate; earlier actionable notification	Intensivists, ED physicians, clinical informaticists
Reference Laboratory Networks	Multi-site cross-institutional infectious disease surveillance	Population-level epidemiological signal detection across institutional boundaries	Public health agencies, laboratory medical directors
Chronic Disease Management	Longitudinal risk stratification for CKD, DM, CHF	Trajectory-based deterioration warnings vs. static thresholds	Primary care physicians, nephrologists, endocrinologists
Pediatric Critical Care	Age- and weight-adjusted developmental baseline sepsis detection	Developmentally appropriate alert thresholds	Pediatric intensivists, neonatologists
Real-World Evidence Generation	Retrospective cohort construction for pharmacovigilance	High-quality longitudinally annotated patient profiles at scale	Pharmaceutical companies, health economists, regulators
Post-Surgical Monitoring	Detection of post-operative infectious complications	Baseline-shift detection distinguishing surgical from infectious etiology	Surgeons, anesthesiologists, perioperative nursing teams

Table 1. LPRL Application Domains, Use Cases, and Stakeholders

5.1.1 Chronic Disease Surveillance

Patients with chronic conditions CKD, CHF, or T2DM present a systematic problem for population-threshold alert systems: their analyte values are chronically displaced from population norms due to underlying disease, not acute deterioration. A stage 4 CKD patient maintaining creatinine at 3.2 mg/dL generates continuous false-positive alerts under any population-threshold rule, while the clinically meaningful signal—

an acute rise from 3.2 to 4.7 mg/dL over 72 hours, disappears into alert volume. The LPRL resolves this by anchoring comparisons to the patient's own 36-month history; a PDR of 1.8 on that acute rise triggers an alert precisely because it deviates from that patient's norm. The same logic extends without architectural modification to CHF decompensation via BNP trajectory analysis and T2DM complication surveillance via HbA1c and eGFR trends.

5.1.2 Multi-Site Epidemiological Monitoring

A reference laboratory network receives results simultaneously from hundreds of hospitals, long-term care facilities, outpatient clinics, and urgent care centers, a population-level view unavailable to any single hospital EHR. Deployed at this layer, the LPRL's Reasoning Orchestrator aggregates PDR-flagged results across geographic clusters and applies temporal clustering to detect outbreak signals: anomalous co-occurrence of elevated-PDR results within a defined geographic and temporal window. Confirmed clusters trigger population-level alerts to public health authorities via HL7 FHIR notification channels, enabling early outbreak detection before hospital admission data becomes available.

5.1.3 Real-World Evidence Generation

The LPRL's longitudinal profile store is, by construction, a high-quality real-world evidence asset. Each enrolled patient accumulates a continuously updated analyte history spanning 36 months, with encounter-level infectious episode flags, confounder annotations, and PDR time-series for each monitored analyte. At scale across a network serving millions of patients, this constitutes a longitudinally annotated cohort with temporal resolution surpassing claims-based datasets or hospital EHR exports. For pharmacovigilance, the profile store enables retrospective cohort construction directly from analyte history, identifying patients who initiated a nephrotoxic agent and subsequently exhibited a PDR-defined acute creatinine rise, without chart review or ICD code extraction. For clinical trial enrollment, trajectory-based eligibility screening can identify candidates in near real-time as results flow through the pipeline.

5.2 Integration Challenges

Challenge	Description	Severity	Mitigation Strategy / Validation Criterion
Longitudinal Identity Resolution	Linking all prior encounters for a patient across multiple hospital clients using probabilistic MPI	High	Probabilistic MPI with confidence scoring; conservative linkage thresholds; full audit trail. Validation criterion: ≥95% patient match precision on multi-site test cohort before production deployment
HIPAA Compliance	Maintaining longitudinal profiles requires explicit data use agreements and minimum-necessary data standards	High	Business Associate Agreements (BAA); HIPAA Safe Harbor de-identification per §164.514(b); SOC 2 Type II certification. Validation criterion: zero PHI transmission outside BAA-governed channels in penetration test
Real-Time Processing Latency	End-to-end alert delivery must complete within clinical workflow time-to-decision windows	High	Kafka stream ingestion + TimescaleDB write path optimized for sub-500 ms latency. Validation criterion: ≤500 ms p95 end-to-end latency at 10,000 results/hour load
Clinical NLP Integration (medspaCy)	Negation detection in free-text fields required for accurate confounder suppression	Medium	medspaCy rule-based negation detection pipeline deployed as microservice; scoped to confounder knowledge base terms only. Validation criterion: ≥90% negation detection F1 on internal clinical note test set
HL7 Message Heterogeneity	HL7 v2.x implementations vary significantly across LIS vendors	Medium	Curated HL7 normalization library covering 12 major LIS platforms. Validation criterion: ≥99% message parse success rate across vendor test harness
EHR Alert Integration	Delivering enriched alerts via CDS Hooks requires EHR-specific integration work	Medium	SMART on FHIR CDS Hooks [30]; pre-certified connectors for Epic, Oracle Health (Cerner), MEDITECH. Validation criterion: CDS Hooks response time ≤200 ms per EHR specification

Table 2. Integration Challenges, Mitigation Strategies, and Validation Criteria

The three high-severity challenges, longitudinal identity resolution, HIPAA compliance, and real-time processing latency—are the gating conditions for production deployment and must be validated against the specified criteria prior to go-live. Identity resolution at ≥95% precision is the most operationally complex requirement, as it demands coordinated data sharing agreements across all participating laboratory network clients. HIPAA compliance is non-negotiable and is enforced at the architectural level through BAA-governed data channels and Safe Harbor de-identification prior to any cross-client data aggregation. The ≤500 ms latency target is achievable with the proposed Kafka + TimescaleDB architecture and has been validated in analogous real-time clinical data pipeline deployments.

5.3 Conceptual Framework

Table 3 compares traditional population-threshold alert architectures against the LPRL architecture across six analytical dimensions.

Dimension	Traditional Threshold Approach	LPRL Architecture	Clinical Improvement
Reference Population	Age/sex-matched population norms from published literature	Individual patient's own 36-month longitudinal history (PDR-based)	Patient-specific context; eliminates inter-individual variability as noise
Temporal Reasoning	Point-in-time comparison; no trajectory information	Multi-point slope computation across three or more prior results	Distinguishes acute deterioration from stable chronic elevation
Contextual Annotation	Binary output: threshold exceeded / not exceeded	Composite confidence score with baseline deviation, pattern match, confounder flags	Actionable clinical intelligence vs. binary alarm signal
Confounder Handling	No confounder detection; medication and comorbidity context ignored	Structured confounder KB queried against active medication and problem list	Reduces false positives from corticosteroid leukocytosis, CKD creatinine, etc.
Epidemiological Scope	Single-site EHR deployment; no cross-institutional visibility	Reference laboratory layer; cross-institutional multi-site visibility	Population-level infectious disease signal detection across networks
Scalability	EHR-embedded; scales only with single institution's infrastructure	Containerized microservices; horizontally scalable across full network	Supports high-volume multi-client deployments without redesign

Table 3. Conceptual Framework: Traditional Approach vs. LPRL Architecture

5.4 Reference Technology Stack

Table 4 specifies the reference technology stack for a production LPRL deployment, with components grouped by architectural layer and justified by alignment with the design requirements established in Sections 3 and 4.

Component	Technology	Role in LPRL	Justification
Stream Ingestion	Apache Kafka	Receives HL7/FHIR result streams from LIS systems; decouples ingestion from processing	Sub-millisecond throughput; fault-tolerant; widely deployed in healthcare data pipelines
Patient Profile Store	TimescaleDB (PostgreSQL extension)	Stores longitudinal analyte time-series per patient; supports PDR computation queries	Native time-series indexing; SQL compatibility; HIPAA-eligible cloud deployment (AWS RDS)
Clinical NLP	medspaCy	Negation detection and confounder suppression in free-text clinical fields	Clinical-domain rule-based NLP; open-source; validated on MIMIC-III clinical notes
Inference Engine	XGBoost + SHAP	Composite alert confidence scoring from PDR, cosine similarity, trajectory, and confounder vectors	State-of-the-art tabular ML performance; SHAP provides clinician-interpretable feature attribution
LLM Reasoning Layer (optional)	GPT-4o (OpenAI API) or equivalent	Five-step narrative clinical reasoning for high-confidence alerts	Structured reasoning chain improves clinician trust and alert actionability
Interoperability Layer	SMART on FHIR R4 / CDS Hooks	Alert delivery to EHR systems; OAuth 2.0 authorization	HL7-endorsed standard for EHR-agnostic CDS integration [30]
Container Orchestration	Docker / Kubernetes	Microservices deployment; horizontal scaling; zero-downtime updates	Industry standard for HIPAA-eligible containerized healthcare workloads

Table 4. LPRL Reference Technology Stack

6. Future Work

Prospective clinical validation in a live reference laboratory network is the essential next step. The architectural rationale is supported by the alert fatigue literature [6, 7, 8] and multi-site longitudinal reasoning frameworks [12], but the false positive rate reduction achievable in practice will depend on the depth of longitudinal data available for enrolled patients and on clinical workflow context. A prospective cluster-randomized trial comparing LPRL-enhanced alerts to standard alerts on time-to-antibiotic administration and 30-day mortality is the priority research design. A second direction is extending the reasoning framework to joint multi-analyte panel inference: the current architecture reasons about each analyte independently before aggregating into a composite score, and a probabilistic graphical model capturing the joint trajectory of lactate, procalcitonin, blood cultures, and inflammatory markers together could improve discriminative performance, particularly for patients with atypical single-analyte presentations.

7. Conclusion

The Longitudinal Patient Reasoning Layer addresses the core limitation of existing sepsis alert systems: the inability to interpret laboratory results against the individual patient's own longitudinal history. By deploying at the reference laboratory layer and applying four sequential reasoning modules, the LPRL transforms static population-threshold comparisons into patient-

specific contextual intelligence grounded in the Personal Deviation Ratio.

Three pillars support the case for this architecture. Architecturally, PDR-based reasoning, XGBoost+SHAP composite scoring, and medspaCy-enhanced confounder detection represent a substantiated departure from population thresholds whose inadequacy is documented at PPV 22.4% and false positive rates as high as 77.6% in deployed systems [5]. Clinically, the specificity-first design is directly motivated by the alert fatigue literature [6, 7, 8] and responsible ML deployment principles [21], with concrete validation criteria defined for each deployment gate. As a platform, the LPRL is extensible to chronic disease surveillance, multi-site epidemiological monitoring, and real-world evidence generation, analytically grounded applications of the same longitudinal profile store and PDR reasoning engine that power the primary sepsis use case.

8. References

1. C. W. Seymour, F. Gesten, H. C. Prescott, et al., "Time to Treatment and Mortality during Mandated Emergency Care for Sepsis," New England Journal of Medicine, vol. 376, pp. 2235–2244, 2017.

2. M. Singer, C. S. Deutschman, C. W. Seymour, et al., "The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3)," JAMA, vol. 315, no. 8, pp. 801–810, 2016.

3. C. Rhee, R. Dantes, L. Epstein, et al., "Incidence and Trends of Sepsis in US Hospitals Using Clinical vs Claims Data, 2009–2014," JAMA, vol. 318, no. 13, pp. 1241–1249, 2017.

4. K. E. Rudd, S. C. Johnson, K. M. Agesa, et al., "Global, Regional, and National Sepsis Incidence and Mortality, 1990–2017: Analysis for the Global Burden of Disease Study," The Lancet, vol. 395, no. 10219, pp. 200–211, 2020.

5. A. Wong, E. Otles, J. P. Donnelly, et al., "External Validation of a Widely Implemented Proprietary Sepsis Prediction Model in Hospitalized Patients," JAMA Internal Medicine, vol. 181, no. 8, pp. 1065–1070, 2021.

6. B. J. Drew, P. Harris, J. K. Zègre-Hemsey, et al., "Insights into the Problem of Alarm Fatigue with Physiologic Monitor Devices: A Comprehensive Observational Study of Consecutive Intensive Care Unit Patients," PLOS ONE, vol. 9, no. 10, p. e110274, 2014.

7. S. Sendelbach and M. Funk, "Alarm Fatigue: A Patient Safety Concern," AACN Advanced Critical Care, vol. 24, no. 4, pp. 378–386, 2013.

8. S. Manaktala and S. R. Claypool, "Evaluating the Impact of a Computerized Surveillance Algorithm and Decision Support System on Sepsis Mortality," Journal of the American Medical Informatics Association, vol. 24, no. 1, pp. 88–95, 2017.

9. L. M. Fleuren, T. L. T. Klausch, C. L. Zwager, et al., "Machine Learning for the Prediction of Sepsis: A Systematic Review and Meta-Analysis of Diagnostic Test Accuracy," Intensive Care Medicine, vol. 46, pp. 383–400, 2020.

10. S. Nemati, A. Holder, F. Razmi, M. D. Stanley, G. D. Clifford, and T. G. Buchman, "An Interpretable Machine Learning Model for Accurate Prediction of Sepsis in the ICU," Critical Care Medicine, vol. 46, no. 4, pp. 547–553, 2018.

11. A. E. W. Johnson, T. J. Pollard, L. Shen, et al., "MIMIC-III, a Freely Accessible Critical Care Database," Scientific Data, vol. 3, p. 160035, 2016.

12. Q. Mao, M. Jay, J. L. Hoffman, et al., "Multicentre Validation of a Sepsis Prediction Algorithm Using Only Vital Sign Data in the Emergency Department, General Ward and ICU," BMJ Open, vol. 8, no. 1, p. e017833, 2018.

13. HL7 International, "HL7 Fast Healthcare Interoperability Resources (FHIR) Release 4," HL7 FHIR Standard, 2019. [Online]. Available: https://hl7.org/fhir/R4/ (accessed: Apr. 2026).

14. J. Dijk, Y. Rieter-Barrell, J. V. Rest, and H. Bouma, "Intelligent Sensor Networks for Surveillance," Journal of Police Studies: Technology-Led Policing, vol. 3, no. 20, pp. 109–125, 2011.

15. L.-P. Chou, J.-D. Sun, and M.-L. Chen, "A New Application Framework for Intelligent Surveillance Sensor Networks," in Proc. Third International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP 2007), Kaohsiung, Taiwan, 2007, pp. 617–620. doi: 10.1109/iih-msp.2007.39.

16. A. Ziani, C. Motamed, and J.-C. Noyer, "Temporal Reasoning for Scenario Recognition in Video-Surveillance Using Bayesian Networks," IET Computer Vision, vol. 2, no. 2, pp. 57–68, Jun. 2008. doi: 10.1049/iet-cvi:20070074.

17. S. Yang and H. Byun, "Wide Area Ontology Integration Scheme for Reasoning Agents in Surveillance Networks," Journal of Computers, vol. 11, no. 6, pp. 497–503, 2016. doi: 10.17706/jcp.11.6.497-503.

18. P. J. Thoral, J. M. Peppink, R. H. Driessen, et al., "Sharing ICU Patient Data Responsibly under the Society of Critical Care Medicine and European Society of Intensive Care Medicine Joint Data Science Collaboration: The Amsterdam University Medical Centers Database (AmsterdamUMCdb) Example," Critical Care Medicine, vol. 49, no. 6, pp. e563–e577, 2021.

19. U.S. Department of Health and Human Services, "HIPAA Security Rule," 45 C.F.R. Part 164, Subparts A and C, 2003.

20. J. C. Mandel, D. A. Kreda, K. D. Mandl, I. S. Kohane, and R. B. Ramoni, "SMART on FHIR: A Standards-Based, Interoperable Apps Platform for Electronic Health Records," Journal of the American Medical Informatics Association, vol. 23, no. 5, pp. 899–908, 2016.

21. M. P. Sendak, M. Elish, M. Gao, et al., "The Human Body is a Black Box: Supporting Clinical Decision-Making with Deep Learning," in Proc. ACM FAT* Conference, Barcelona, Spain, 2020, pp. 99–109.

22. M. Topaz, L. Seger, R. Slight, et al., "Rising Drug-Drug Interaction Alerts and Overrides in the United States—5-Year Analysis," Journal of the American Medical Informatics Association, vol. 23, no. 3, pp. 601–608, 2016.