SCI CENTRAL INTERNATIONAL PEER-REVIEWED JOURNALS

sci central

RESEARCH PAPER

State-Space Fusion of Physiologic, Imaging, and Molecular Data for Early Anastomotic Failure Detection

Rizky Pratama,[1] Fajar Nugroho,[2] and Dimas Saputra[3]

[1]Universitas Trunojoyo Madura, Department of Informatics Engineering, Raya Telang Street, PO Box 2, Kamal, Bangkalan 69162, Indonesia
[2]Universitas Sultan Ageng Tirtayasa, Department of Public Health, Jenderal Sudirman Street Km 3, Cilegon 42435, Indonesia
[3]Universitas Negeri Semarang, Department of Computer Science, Sekaran Campus, Gunungpati, Semarang 50229, Indonesia

Abstract

Clinical risk prediction in surgery is still dominated by structured variables that appear only after charting, aggregation, and manual abstraction. That timing constraint matters because physiologic deterioration often begins before a formal entry exists in the electronic health record. In perioperative care, the consequence is a recurrent mismatch between when a patient starts to deviate from a stable trajectory and when decision support systems are finally able to measure that deviation. This paper develops a technical framework for multimodal earlier-than-EHR risk signaling in perioperative anastomotic failure, a setting in which delays in recognition can change both salvage options and downstream morbidity. The proposed formulation treats earlier-than-EHR information not as a vague notion of temporal advantage, but as an estimable property of raw data streams whose acquisition time precedes structured chart availability. Continuous physiologic waveforms, intraoperative perfusion imaging, molecular compositional profiles, device telemetry, and sparse structured records are fused through a continuous-time latent state model linked to a hazard process for failure. The framework addresses asynchronous sampling, variable data quality, missing-not-at-random observations, label delay, and deployment constraints under finite alert budgets. It also distinguishes prediction performance from lead-time utility by defining metrics that reward accurate alarms only when they arrive early enough to matter clinically. The result is a unified research design for modeling, training, and evaluating pre-charting surgical risk signals. Rather than assuming that more modalities automatically improve prediction, the paper argues that the primary technical advantage of multimodal systems lies in recovering temporally proximal evidence before it is compressed into late EHR abstractions.

1. Introduction

Risk stratification in modern perioperative care is usually framed as a problem of extracting patterns from the electronic health record, yet the clinical process that produces the record is neither instantaneous nor lossless (1). A waveform is measured before it is summarized, a tissue image is captured before it is interpreted, a biologic specimen exists before its laboratory derivatives are written back into the chart, and clinician concern often emerges in speech, manipulation, or device behavior before it becomes codified as a diagnosis, order, or note. The structured EHR is therefore a delayed and compressed projection of a richer measurement process. For many complications that evolve over hours rather than seconds, that delay may appear tolerable. For perioperative anastomotic failure, however, the relevant question is not only whether prediction is possible, but whether a risk signal can be surfaced while the intervention set still differs from what would be done under routine surveillance. This paper takes that timing problem as primary and develops a technical framework for earlier-than-EHR risk detection in a multimodal setting.

The phrase earlier-than-EHR requires precision. It does not mean merely earlier than discharge coding, earlier than the final diagnosis, or earlier than catastrophic decompensation. It refers to signals that are acquired, measurable, and inferentially usable before they are translated into structured chart variables of the kind consumed by standard predictive pipelines. A perfusion video obtained during construction of an anastomosis is earlier than postoperative laboratory summaries. A raw arterial waveform with emerging vasomotor instability is earlier than the hourly vital-sign table that eventually records a simplified trajectory. A molecular profile that reflects impaired healing potential before incision is earlier than the postoperative note in which concern is documented. The central claim of this paper is that a clinically meaningful fraction of risk is encoded in these pre-charting modalities and that the right mathematical object for recovering it is not a static classifier on merged features, but a continuous-time latent hazard model with modality-specific reliability control.

This focus is motivated by a broader limitation of structured-data prediction. Even when structured EHR features are informative, they often become available only after the clinical state has already drifted. In the related context of postoperative leak prediction, Sadanandan (2024) emphasizes that structured records support meaningful risk stratification but also points toward modalities such as free-text signals, intraoperative perfusion video, and microbiome measurements whose timing can shift inference before or during the procedure (2). That observation is more general than the original application. It suggests that timing hierarchy across modalities is not incidental noise around a canonical EHR problem; it is a core design dimension of the problem itself (3). Once that is recognized, conventional feature engineering based on synchronized tabular snapshots becomes inadequate, because the data are neither synchronized nor naturally tabular at the point where they are most valuable.

Perioperative anastomotic failure provides a technically demanding and clinically consequential domain in which to formalize this idea. Failure is not a single instantaneous event in the mechanistic sense. It is an evolving process involving perfusion adequacy, tissue apposition, inflammation, microbial ecology, systemic response, and intervention timing. Some determinants are present long before surgery, including nutritional depletion, compositional microbiome signatures, and baseline vascular status. Others emerge intraoperatively through tissue handling, fluorescence wash-in dynamics, thermal injury, suture line geometry, and hemodynamic instability. Still others appear in the immediate postoperative period as waveform irregularity, microcirculatory failure, lactate kinetics, autonomic dysregulation, or changes in respiratory and vasopressor patterns. A prediction system that waits for these processes to be encoded into a small set of charted values effectively discards their most granular temporal content and may also miss the interval during which the choice set remains broad.

The problem is not solved by simply concatenating modalities. Multimodal fusion is difficult because different streams obey different geometries, noise models, and clocks. Waveforms are high-frequency and autocorrelated. Video is spatially structured, motion-contaminated, and conditionally informative only under adequate exposure or dye delivery. Molecular profiles are compositional and often sparse, with sequencing depth interacting with biological abundance. Device metadata may reflect clinician behavior as much as patient state. Structured chart variables are semantically useful but delayed and unevenly sampled. A coherent model must therefore answer several nontrivial questions simultaneously: which latent physiologic state explains the currently observed fragments, how much trust should be placed in each modality at each moment, how should data with different acquisition-to-chart delays be aligned, and how can training reward predictions that arrive early enough to matter rather than merely before the formal label?

This paper addresses those questions by proposing a continuous-time state-space hazard architecture for multimodal earlier-than-EHR risk inference. The core design principle is to treat each modality as a noisy view of a shared latent trajectory while preserving modality-specific timing and reliability (4). Instead of converting all streams into a common static table, the model maintains a latent state that evolves continuously, absorbs asynchronous evidence as it arrives, and drives a hazard process for anastomotic failure. The architecture is accompanied by a training objective that integrates censored time-to-event likelihood, acquisition-time lead utility, alignment regularization, and calibration control. The resulting framework is intended as a methodological contribution rather than a claim of immediate deployment readiness. Its purpose is to make the timing structure of multimodal perioperative data mathematically explicit and to show how earlier-than-EHR prediction can be evaluated on its own terms rather than as a minor variant of ordinary EHR risk scoring.

A second motivation is that the conventional distinction between preoperative and postoperative prediction is too coarse for many surgical problems. There is a continuous perioperative interval in which risk is being rewritten by events in the operating room and early recovery period. A model that only predicts before incision misses newly created danger. A model that only predicts after charted postoperative abnormalities misses the chance to intervene on subtle but evolving deviations. The proposed framework therefore uses a moving risk horizon over a continuous timeline anchored to acquisition events rather than documentation events. This timeline supports the idea of dynamic surveillance in which risk estimates are repeatedly updated as raw evidence accrues.

The remainder of the paper develops the technical components of this program. The next section formulates earlier-than-EHR prediction as an asynchronous, lead-sensitive time-to-event problem with delayed observation channels. Subsequent sections introduce multimodal representation geometry, the continuous-time hazard state model, composite training objectives under delayed and noisy labels, and an evaluation scheme that makes lead-time gain, calibration, and operational burden jointly visible. The final sections consider deployment as a constrained decision problem under alert budgets and discuss why earlier-than-EHR modeling should be interpreted as a temporal reorganization of surgical prediction rather than a simple increase in input dimensionality.

2. Problem Formulation

Consider a perioperative episode indexed by i, with surgery beginning at time $t = 0$ and with follow-up observed on a continuous interval $0, \tau_i$, where τ_i is the minimum of event time and censoring time. Let the adverse event of interest be anastomotic failure with event time T_i, understood clinically as the latent onset time of pathophysiologic failure or, when that is unobservable, the earliest decision-relevant manifestation of failure. The episode contains M modalities. For each modality $m \in \mathcal{M}$, observations arrive as an asynchronous sequence

$$\begin{aligned} \mathcal{O}_i^m &= \left\{ \left(a_{ij}^m, c_{ij}^m, y_{ij}^m, q_{ij}^m \right) : j = 1, \ldots, n_i^m \right\}, \\ \ell_{ij}^m &= c_{ij}^m - a_{ij}^m, \\ \mathcal{H}_i t &= \bigcup_{m \in \mathcal{M}} \left\{ \left(a_{ij}^m, c_{ij}^m, y_{ij}^m, q_{ij}^m \right) : a_{ij}^m \le t \right\}, \end{aligned} \tag{2.1}$$

where a_{ij}^m is acquisition time, c_{ij}^m is chart-availability time, y_{ij}^m is the raw observation, q_{ij}^m is a quality descriptor, and ℓ_{ij}^m is the acquisition-to-chart delay. Earlier-than-EHR evidence exists whenever $\ell_{ij}^m > 0$ and the observation can be incorporated at acquisition time rather than chart time. This definition is deliberately operational: it does not require perfect automation, only a measurable interval during which the signal is available outside the structured chart (5).

A standard EHR prediction pipeline observes the filtration $\mathcal{H}_i^{\mathrm{ehr}} t$ defined by replacing acquisition times with chart times. In contrast, the earlier-than-EHR pipeline observes $\mathcal{H}_i t$ directly. The difference between the two is not only a shift in timing but also a change in representation because raw signals often contain information that is destroyed by charting. The goal is

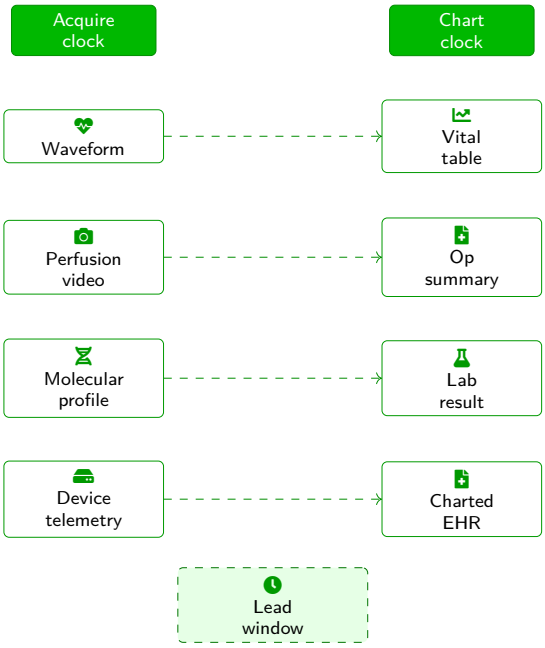

Figure 1. Timing hierarchy across perioperative modalities. Raw physiologic, imaging, molecular, and device streams exist on the acquisition clock before they are compressed into delayed chart abstractions, creating the earlier-than-EHR lead interval that the framework tries to exploit.

to estimate the dynamic risk

$$\begin{aligned} R_i t, \Delta &= \Pr\big(T_i \le t\ \Delta \mid T_i > t, \mathcal{H}_i t\big), \\ R_i^{\mathrm{ehr}} t, \Delta &= \Pr\big(T_i \le t\ \Delta \mid T_i > t, \mathcal{H}_i^{\mathrm{ehr}} t\big), \\ G_i t, \Delta &= R_i t, \Delta - R_i^{\mathrm{ehr}} t, \Delta, \end{aligned} \tag{2.2}$$

where Δ is the decision horizon. The object of interest is not merely large absolute risk, but clinically useful lead over the EHR-constrained estimator.

Because clinical action is thresholded, a more relevant quantity than G_i alone is the alarm time induced by a policy. For a threshold θ,

$$\begin{aligned} \tau_i^\star \theta &= \inf\{t \ge 0 : R_i t, \Delta \ge \theta\}, \\ \tau_{i,\mathrm{ehr}}^\star \theta &= \inf\{t \ge 0 : R_i^{\mathrm{ehr}} t, \Delta \ge \theta\}, \\ \Lambda_i \theta &= \tau_{i,\mathrm{ehr}}^\star \theta - \tau_i^\star \theta. \end{aligned} \tag{2.3}$$

The lead gain $\Lambda_i \theta$ captures the amount of clinically usable time won by incorporating earlier-than-EHR modalities. A model can achieve similar discrimination to an EHR-based comparator yet remain more valuable if it alerts substantially earlier at equivalent false-positive burden.

The event label itself is imperfect. Anastomotic failure is often recognized through downstream proxies such as reoperation, drainage, imaging, antibiotic escalation, or adjudicated diagnosis. Let $\tilde{T}_i$ denote the observed recognition time. In practice one often has $\tilde{T}_i \ge T_i$ with unknown lag. The learning problem is thus subject to delayed labeling. Rather than treating observed time as exact event onset, the model should permit a recognition kernel or interval-censoring representation. One can define a latent onset distribution conditional on recognition:

$$\begin{aligned} p T_i \mid \tilde{T}_i, \omega_i &\propto p \tilde{T}_i \mid T_i, \omega_i\, p T_i \mid \omega_i, \\ 0 \le T_i &\le \tilde{T}_i, \\ \omega_i &= \text{episode-level context}, \end{aligned} \tag{2.4}$$

where ω_i may include procedure type, surveillance intensity, and intervention propensity. This matters because apparent prediction delay can be confounded by label delay.

The modality set in the present framework is intentionally heterogeneous. Continuous waveforms include arterial pressure, photoplethysmography, ventilator flow, and capnography. Intraoperative imaging includes fluorescence perfusion sequences or ordinary video from which motion, color transport, and tissue handling can be inferred (6). Molecular signals include microbiome or metabolomic compositions measured preoperatively or early postoperatively. Structured chart variables include demographics, laboratory values, procedural metadata, and orders. Device and workflow telemetry include infusion pump activity, operating room timing traces, and interaction logs. Each modality has its own sampling density, cost structure, quality failures, and causal distance from the event process.

An important conceptual distinction arises between causally proximal and temporally early signals. A microbiome profile is temporally early but may be mechanistically distal. A perfusion image is both temporally early and causally proximal. A postoperative waveform abnormality may be temporally later than a molecular baseline but much closer to imminent failure. A useful model should not collapse these dimensions into a single notion of importance. Instead it should learn a latent state in which each modality contributes according to its time-varying relevance, reliability, and relation to the hazard process.

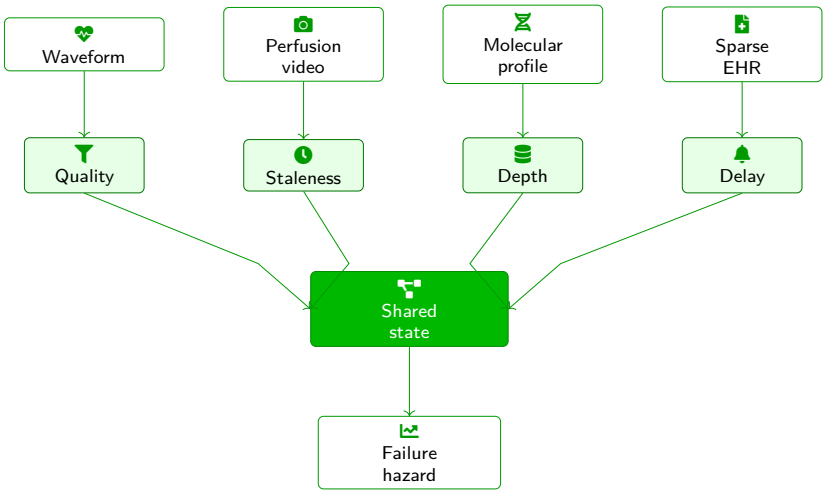

Figure 2. Reliability-aware multimodal fusion. Each stream is encoded locally, filtered by modality-specific quality or delay cues, and then projected into a shared latent state whose evolution supports dynamic hazard estimation rather than a single static risk score.

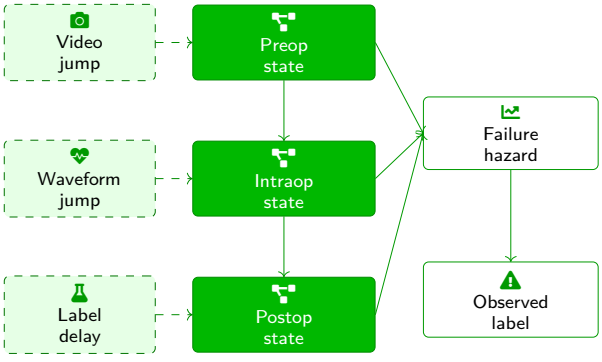

Figure 3. Continuous-time latent hazard dynamics. The hidden perioperative state evolves across clinical phases, receives asynchronous information jumps from arriving modalities, and drives a hazard process whose observed recognition label can appear later than the underlying deterioration.

The problem also differs from ordinary multimodal classification because missingness is often informative and resource-driven. A perfusion video is available only when a specific technology is used. Sequencing data may be absent because a clinician did not judge the case sufficiently unusual to order testing or because the institution lacks the assay. Hemodynamic waveforms may be high fidelity in an intensive monitoring environment and low fidelity otherwise. Structured records are often missing precisely because the patient appears stable. Missing-not-at-random patterns therefore contain information about both physiology and workflow. Discarding that structure by simple imputation can distort inference.

A final element of the formulation is operational utility (7). Let $A_i t$ denote an alarm action at time t. The benefit of an alarm depends on how much decision leverage remains. An alert five minutes before a clinically obvious decompensation may carry little value if the same response would have occurred without the model. By contrast, an alert during an otherwise routine postoperative interval or while the patient is still in the operating room may alter the diagnostic, technical, or monitoring plan. The objective function should therefore reward not only correctness but also clinically meaningful earliness. That requirement drives the architecture developed below: a continuous-time latent process, explicit acquisition-time handling, and losses that encode lead-sensitive risk rather than static discrimination alone.

3. Multimodal Representation Geometry

A central challenge in earlier-than-EHR inference is that modalities do not live in the same representational space. Waveforms are functions over time, imaging is tensor-valued over space and time, molecular assays are compositional vectors on the simplex, and structured records are sparse symbolic events. A meaningful fusion layer cannot assume that these objects are naturally commensurate. Instead the model must construct modality-specific encoders that preserve local geometry while producing state estimates that can be projected onto a shared latent manifold. The purpose of the common latent space is not aesthetic symmetry; it is to support clinically coherent state updates under asynchronous evidence.

For each modality m, define an encoder E_m that maps a raw observation window into a feature vector and an uncertainty

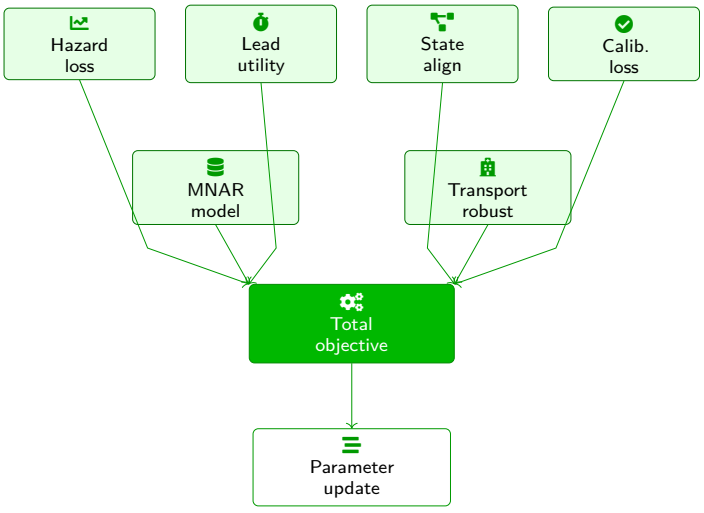

Figure 4. Composite learning objective. Training combines censored hazard likelihood, lead-sensitive reward, latent alignment, calibration control, missingness handling, and transport robustness so that early alerts are favored only when they remain stable and clinically credible.

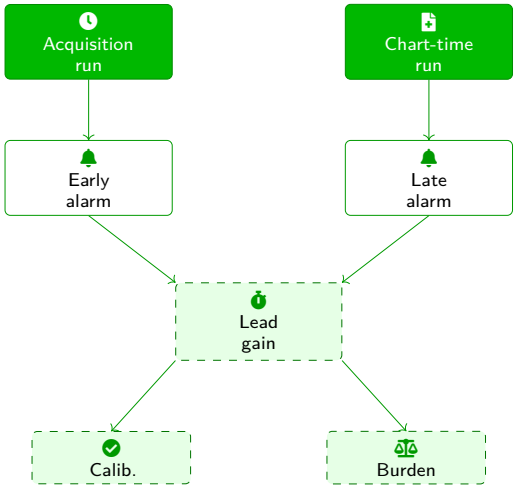

Figure 5. Clock-based evaluation. The same architecture is run once with acquisition-time evidence and once under artificial chart-time delay, so measured gains can be attributed to timing advantage rather than to richer representation alone.

descriptor:

$$\begin{aligned}\left(e_t^m, u_t^m\right) &= E_m\left(Y^m_{t-h_m,t}, Q^m_{t-h_m,t}\right),\\ e_t^m &\in \mathbb{R}^{d_m},\\ u_t^m &\in \mathbb{S}^{d_m},\end{aligned} \tag{3.1}$$

where $Y^m_{t-h_m,t}$ is the recent observation window, $Q^m_{t-h_m,t}$ is the corresponding quality stream, and u_t^m is a positive semidefinite uncertainty object. For waveforms, E_m should be sensitive to morphology, transient instability, and variability across scales. For video, it should retain transport dynamics, boundary confidence, dye arrival behavior, and motion quality. For molecular signals, the encoder should operate in a geometry that respects compositional structure rather than Euclidean abundance directly.

Microbiome and metabolomic compositions motivate a brief geometric detour. Let $p_t \in \Delta^{K-1}$ denote a composition with simplex coordinates summing to one. Direct Euclidean operations on p_t are ill-posed because absolute scale is not identifiable and negative covariance is induced by closure. A log-ratio embedding offers a more stable basis: (8)

$$\begin{aligned}r_t &= H\log p_t,\\ H\mathbf{1} &= 0,\\ r_t &\in \mathbb{R}^{K-1},\end{aligned} \tag{3.2}$$

where H is a contrast matrix. The resulting coordinates can then be embedded alongside other modalities while preserving relative abundance structure. This is especially useful when molecular signatures reflect broad ecological states associated with impaired healing rather than single taxa of interest.

To place modality-specific embeddings into a shared state space, introduce projection operators $P_m \in \mathbb{R}^{d\times d_m}$ and define

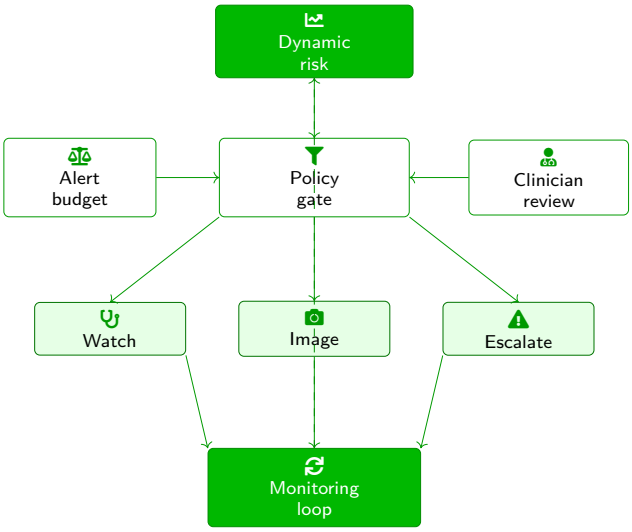

Figure 6. Clinical integration as a constrained decision process. A dynamic perioperative risk estimate is passed through a thresholding policy that must respect alert budgets and clinician review, after which graded actions feed back into the monitoring loop that generates future evidence.

projected evidence vectors

$$\begin{aligned} z_t^m &= P_m e_t^m, \\ \Omega_t^m &= P_m u_t^m P_m^\top, \\ z_t^m &\in \mathbb{R}^d. \end{aligned} \tag{3.3}$$

The shared dimension d is not interpreted as a generic feature count but as the dimension of a latent perioperative state. In this state, components can be designed or learned to correspond approximately to perfusion adequacy, inflammatory activation, systemic stress, microbial vulnerability, technical complexity, or surveillance intensity. Full interpretability is not assumed, but semantically organized axes make auditing easier.

Fusion should be reliability-aware. A perfusion video with poor exposure, a waveform contaminated by motion artifact, or a molecular assay with low sequencing depth should not update the state as aggressively as a high-quality measurement. Let $\rho_t^m \in 0, 1$ be a learned reliability weight:

$$\begin{aligned} \rho_t^m &= \sigma\big(\alpha_m - \beta_m \delta_t^m - \chi_m \nu_t^m \;\; \zeta_m s_t^m\big)\,, \\ \delta_t^m &= \text{time since last valid observation}, \\ \nu_t^m &= \text{missingness intensity}, \\ s_t^m &= \text{quality score}, \end{aligned} \tag{3.4}$$

where σ is the logistic map. Reliability is thus a dynamic quantity that decays with staleness and poor quality but may increase with signal richness or context. Importantly, the model need not ignore low-reliability modalities entirely; it can attenuate them continuously.

A principled shared-state estimate can be obtained by solving a regularized consensus problem at each update time:

$$\begin{aligned} \hat{x}_t = \arg\min_{x \in \mathbb{R}^d} & \sum_{m \in \mathcal{M}_t} \rho_t^m \, \|x - z_t^m\|^2_{\Omega_t^{m,-1}} \\ & \lambda_t x^\top L_t x \;\; \mu_t \, \|x - \bar{x}_{t^-}\|_2^2\,, \end{aligned} \tag{3.5}$$

where $\mathcal{M}_t$ is the set of modalities available at time t, L_t is a graph Laplacian encoding temporary compatibility constraints among latent coordinates, and $\bar{x}_{t^-}$ is the state immediately before assimilation. The first term pulls the state toward available modality projections, the second preserves latent structural relationships, and the third prevents unstable jumps inconsistent with the recent trajectory unless the evidence is strong. This estimator can be interpreted as a Bayesian precision-weighted fusion step under quadratic priors.

The graph term deserves emphasis because multimodal fusion often fails not by ignoring a modality but by forcing incompatible measurements into a single overly rigid representation. Suppose one latent coordinate tracks perfusion integrity and another tracks systemic inflammatory burden. During intraoperative acquisition, imaging may constrain the former far more than the latter. During early postoperative surveillance, the reverse may happen (9). A time-varying Laplacian permits smooth coordination among latent dimensions without assuming a fixed covariance pattern throughout the episode. The graph can be learned from data or partly specified from physiology and workflow.

Temporal context should also enter the representation. Rather than a single embedding per modality, the encoder can emit local trajectory summaries such as slope, curvature, burstiness, or recurrence statistics. For waveforms, this captures not just level shifts but oscillatory instability and loss of autoregulatory structure. For perfusion video, it captures dye arrival delay, wash-out asymmetry, or texture homogenization. For structured records, it captures order sequences and laboratory kinetics. These summaries can be stacked into augmented embeddings so that the shared latent state reflects both current value and recent motion.

A subtle but important issue arises from modality-specific time scales. Molecular profiles may change slowly but encode long-range predisposition. Waveforms change rapidly and encode acute instability. Video-derived perfusion may provide

Table 1. Modalities and Acquisition Characteristics

Modality	Sampling Type	Timing Advantage	Noise Source
Waveforms	Continuous	Pre-chart	Motion artifacts
Imaging	Episodic	Intraoperative	Exposure variation
Molecular	Sparse	Preoperative	Sequencing depth
EHR Data	Discrete	Delayed	Abstraction loss

Table 2. State-Space Model Components

Component	Symbol	Role	Property
Latent State	x_t	System state	Continuous-time
Drift	Ax_t	Baseline evolution	Linear/nonlinear
Jump Process	J_t	Async updates	Sparse events
Noise	W_t	Uncertainty	Gaussian

punctate but highly informative snapshots. Aligning them on a single clock can create either oversmoothing or overreaction. The consensus estimator above is therefore only one layer of the system. It provides local assimilation, but the evolution of the shared state is governed separately by a continuous-time dynamical model that can preserve each modality's effective persistence. That dynamical component is introduced next.

4. Continuous-Time Latent Hazard Dynamics

The shared latent representation is useful only if it evolves in a way that can connect sparse predisposition, punctate intraoperative observations, and dense physiologic fluctuations into a coherent trajectory of risk (10). A natural approach is a continuous-time state-space model in which the hidden state encodes the evolving perioperative condition and a hazard process maps that condition into event risk. Continuous time is essential because charting and intervention schedules are irregular, modality arrivals are asynchronous, and clinically meaningful lead can be measured in minutes or hours rather than in arbitrary discretization bins.

Let $x_t \in \mathbb{R}^d$ denote the latent perioperative state. We model its evolution as a stochastic controlled differential system:

$$dx_t = \big(Ax_t \;\; Bu_t \;\; bt\big)\,dt \\ \sum_{m\in\mathcal{M}} K_m\, dJ_t^m \;\; \Sigma^{12}x_t, t\, dW_t, \tag{4.1}$$

where A is a baseline drift matrix, u_t is a vector of observed or inferred interventions and physiologic controls, bt is a deterministic time-varying drift relative to the surgical phase, J_t^m is a jump process carrying modality-specific evidence arrivals, K_m maps jumps into state perturbations, and W_t is standard Brownian motion. The continuous drift captures gradual evolution, the jump terms absorb sudden informational updates from arriving modalities, and the diffusion term represents unresolved variability and unmodeled perturbations.

The inclusion of jump processes is particularly important for earlier-than-EHR settings. A perfusion video provides a burst of information during a narrow intraoperative interval. A laboratory result appears episodically. A molecular assay may be available only once. These are not well approximated by constant-rate observations on a grid. By allowing modality arrivals to induce state jumps, the model preserves the fact that some observations suddenly collapse uncertainty in clinically relevant directions. The latent state can thus change sharply after a high-information acquisition even when the underlying physiology changes smoothly, reflecting improved knowledge rather than abrupt biologic transformation.

Observation models relate raw embeddings to the latent state. For projected modality vectors z_t^m obtained from the previous section, write

$$z_t^m = C_m x_t \;\; D_m x_t \;\; \epsilon_t^m, \\ \epsilon_t^m \sim \mathcal{N}\big(0, \Omega_t^m \rho_t^m\big), \\ m \in \mathcal{M}_t, \tag{4.2}$$

where C_m maps the latent state to modality space, D_m allows sensitivity to local state velocity, and the effective observation covariance is scaled by reliability. The derivative term is useful because many modalities reflect not only state level but change rate. For example, a rising inflammatory slope may be more informative than a mildly elevated but stable value.

The event process is specified by a hazard function driven by the latent state. Let N_t denote the counting process for anastomotic failure recognition or onset proxy. Conditional on the latent path, define (11)

$$\lambda t \mid x_t, x_t, \xi_t = \lambda_0 t \exp\big(a^\top x_t \;\; b^\top x_t \;\; c^\top \xi_t \\ x_t^\top H x_t\big), \tag{4.3}$$

where $\lambda_0 t$ is a baseline hazard, ξ_t contains external contextual variables such as procedure stage or surveillance intensity, and H is a low-rank interaction matrix. The quadratic term permits synergistic risk amplification, for instance when impaired perfusion and systemic inflammatory activation jointly generate greater danger than the sum of their marginal effects. The derivative contribution allows acceleration toward instability to influence hazard even before absolute state values become extreme.

Table 3. Risk and Lead-Time Definitions

Quantity	Symbol	Meaning	Dependency
Risk	Rt, Δ	Failure probability	History $\mathcal{H}t$
EHR Risk	R^{ehr}	Delayed estimate	Chart-time
Lead Gain	Λ	Time advantage	Threshold θ
Alarm Time	τ^*	Trigger point	Policy-based

Table 4. Encoder Outputs by Modality

Modality	Embedding e_t	Uncertainty u_t	Feature Type
Waveform	$\mathbb{R}^{d_m}$	PSD matrix	Temporal patterns
Imaging	$\mathbb{R}^{d_m}$	PSD matrix	Spatial dynamics
Molecular	$\mathbb{R}^{d_m}$	PSD matrix	Compositional
Structured	$\mathbb{R}^{d_m}$	PSD matrix	Sparse events

Conditional survival then follows from standard time-to-event calculus:

$$\begin{aligned} St \mid \mathcal{H}_t &= \exp\big(-\int_0^t \lambda s \mid x_s, x_s, \xi_s \, ds\big), \\ Rt, \Delta &= 1 - \frac{St \;\; \Delta \mid \mathcal{H}_{t\Delta^-}}{St \mid \mathcal{H}_t}. \end{aligned} \tag{4.4}$$

Although this expression resembles familiar survival modeling, the informational content is markedly different because the filtration is acquisition-time based and the latent state is updated by raw multimodal evidence rather than only by charted summaries.

A notable benefit of the state-space construction is the ability to encode phase-specific perioperative behavior. The period before incision, the interval of anastomosis creation, immediate postoperative recovery, and subsequent ward or intensive-care monitoring do not share the same dynamics. One can represent this by switching drift regimes or by augmenting the state with a phase indicator π_t. In the simplest form,

$$\begin{aligned} dx_t = \big(A_{\pi_t} x_t \;\; B_{\pi_t} u_t \;\; b_{\pi_t} t\big)\, dt \\ \sum_m K_{m,\pi_t} \, dJ_t^m \;\; \Sigma_{\pi_t}^{12} \, dW_t, \end{aligned} \tag{4.5}$$

where regime-dependent parameters reflect the different interpretations of signals across phases. A transient drop in blood pressure during induction has a different meaning from a sustained perfusion abnormality during bowel reconstruction or from vasopressor-dependent instability after transfer. Phase-aware dynamics are therefore not optional decoration but a requirement for coherent interpretation.

Because the observed event time may lag latent onset, the hazard can be split into biologic onset and recognition subhazards. Let T_i^{bio} be latent onset and T_i^{rec} the first recognition proxy. Then

$$\begin{aligned} \lambda_{\text{bio}} t &= \lambda_{0,\text{bio}} t \exp\big(\eta_{\text{bio}} x_t, x_t, \xi_t\big), \\ \lambda_{\text{rec}} t &= \lambda_{0,\text{rec}} t \exp\big(\eta_{\text{rec}} x_t, \zeta_t\big), \\ T_i^{\text{obs}} &= \min\big(T_i^{\text{bio}} \;\; D_i, \;\; T_i^{\text{rec}}\big), \end{aligned} \tag{4.6}$$

where D_i is a stochastic delay and ζ_t includes observation intensity and clinician response variables. This decomposition separates pathophysiologic risk from the institution's tendency to detect and document it. A model evaluated only on observed recognition time may appear late simply because charting is late; the latent decomposition reduces that bias when enough downstream structure exists to estimate it.

Continuous-time state models are often criticized for poor interpretability, but in this application interpretability can be structured around state transitions and attributable updates. Each assimilation event can be accompanied by a decomposition of state change into modality contributions: (12)

$$\begin{aligned} \Delta x_t &= x_t - x_{t^-} \\ &= \sum_{m \in \mathcal{M}_t} \Gamma_t^m z_t^m \;\; \Gamma_t^0 x_{t^-}, \end{aligned} \tag{4.7}$$

where Γ_t^m are update operators derived from the local fusion rule. Clinicians can then inspect whether a risk increase was driven primarily by perfusion evidence, molecular predisposition, waveform instability, or delayed chart confirmation. This type of attribution is especially useful when structured EHR variables disagree with raw modalities, because it clarifies whether the model is acting on a pre-charting signal or merely echoing a later charted abnormality.

The hazard formulation also supports monotonic clinical constraints. Certain latent directions should, beyond a threshold, not reduce near-term failure risk. If a coordinate encodes severe hypoperfusion or escalating inflammatory activation, then all else equal the hazard should be nondecreasing in that coordinate over relevant ranges. Such constraints can be imposed through sign restrictions on rows of a or through partial order penalties on the hazard map. These restrictions help prevent implausible inversions caused by local data sparsity, particularly in small or heterogeneous surgical cohorts.

Finally, the continuous-time hazard view changes the meaning of data fusion. The task is not to build a single omniscient representation but to maintain the best possible estimate of a latent trajectory whose risk-relevant properties unfold before, during, and after charting. This point matters because it shifts optimization away from static classification and toward sequential state estimation under delayed observation. That, in turn, determines the learning problem.

Table 5. Reliability Weight Factors

Factor	Symbol	Effect	Interpretation
Staleness	δ_t	Negative	Older data less trusted
Missingness	ν_t	Negative	Sparse signals penalized
Quality	s_t	Positive	High fidelity boosts weight
Bias Term	α	Baseline	Modality prior

Table 6. Training Objective Components

Loss Term	Symbol	Purpose	Challenge Addressed
Survival	$\mathcal{L}_{surv}$	Event likelihood	Censoring
Delay	$\mathcal{L}_{delay}$	Label shift	Recognition lag
Lead	$\mathcal{L}_{lead}$	Early detection	Timing utility
Alignment	$\mathcal{L}_{align}$	Fusion consistency	Modality mismatch

5. Learning Under Delayed Labels, Missingness, and Temporal Utility

Training an earlier-than-EHR predictor requires more than maximizing likelihood under a survival model. The optimization must reconcile censored outcomes, delayed or noisy event labels, asynchronous evidence arrivals, missing-not-at-random modalities, domain heterogeneity, and the practical requirement that early alarms matter more than late ones. A single scalar loss cannot capture all of this unless it is explicitly composite. The proposed learning objective therefore combines several terms, each corresponding to a distinct failure mode of naive training.

The base term is a censored time-to-event likelihood built on the hazard process. For episode i with observed event indicator $\delta_i \in \{0, 1\}$ and observed time $\tilde{T}_i$, the ordinary contribution is

$$\mathcal{L}^i_{\text{surv}} = -\delta_i \log \lambda \tilde{T}_i \; {}_0^{\tilde{T}_i} \lambda s \, ds. \tag{5.1}$$

This is necessary but insufficient because $\tilde{T}_i$ may be a delayed recognition proxy rather than true onset. To handle label delay, one can replace the point event time with a latent interval or recognition kernel. If $T_i \in L_i$, U_i is plausible given recognition data, the event term becomes an integral over onset uncertainty: (13)

$$\mathcal{L}^i_{\text{delay}} = -\log {}_{L_i}^{U_i} \lambda t \, \exp\!\left(-{}_0^t \lambda s \, ds\right) \kappa_i t \, dt, \tag{5.2}$$

where $\kappa_i t$ is a recognition-consistency density. This converts hard recognition times into soft onset windows and reduces the risk of punishing the model for predicting before documentation.

Because the target application values early usable prediction rather than merely eventual ranking, a lead-sensitive term is added. Let $g\tau$ be a nonnegative gain function over lead time $\tau = T_i - t$, increasing over a clinically useful interval and decaying when alarms become too early to be actionable or too close to the event to change management. Define

$$\mathcal{L}^i_{\text{lead}} = -{}_0^{\tilde{T}_i} w_i t \left[Y_i t \log R_i t, \Delta \right. \\ \left. \left(1 - Y_i t\right) \log \left(1 - R_i t, \Delta\right) \right] dt, \tag{5.3}$$

where $Y_i t$ is a horizon label indicating whether failure occurs within Δ after t and $w_i t = g\tilde{T}_i - t$ for cases, with an analogous burden-aware weight for noncases. This term upweights correct predictions occurring in the window where earlier intervention could reasonably alter surveillance, imaging, diversion decisions, or escalation thresholds.

A central difficulty in multimodal learning is that absent data do not imply absent risk. Missingness often reflects workflow, technology availability, or clinical suspicion. Simply imputing missing modalities with global means treats absence as ignorance rather than information. A better approach is to parameterize observation intensity and incorporate it into the state update and loss. Let M^m_t indicate availability of modality m at time t, and let $\pi^m_t = \Pr M^m_t = 1 \mid x_t, \zeta_t$ be an observation model. One may regularize the latent state so that it explains both observed values and observation patterns:

$$\mathcal{L}^i_{\text{obs}} = -{}_m {}_0^{\tau_i} \left[M^m_t \log \pi^m_t \right. \\ \left. \left(1 - M^m_t\right) \log \left(1 - \pi^m_t\right) \right] dt. \tag{5.4}$$

This is not intended to estimate clinician intent perfectly. Rather, it discourages the model from confusing systematic acquisition patterns with pure noise and allows the state to represent surveillance intensity as a distinct factor.

Alignment regularization is needed because modalities can agree on underlying risk while disagreeing locally due to timing offsets, scale distortions, or differential sensitivity. The goal is not to force embeddings to coincide pointwise, which would destroy modality specificity, but to align them in a manner consistent with the shared state. A suitable penalty compares each projected modality embedding to the current latent estimate after accounting for reliability:

$$\mathcal{L}^i_{\text{align}} = {}_m {}_0^{\tau_i} \rho^m_t \, \| z^m_t - C_m x_t \|_2^2 \, dt \\ \gamma \, {}_{m \neq m'} {}_0^{\tau_i} \rho^m_t \rho^{m'}_t \, d_{\text{geo}}\left(z^m_t, z^{m'}_t\right) dt, \tag{5.5}$$

where d_{geo} is a geometry-aware divergence. For compositional modalities this may operate in log-ratio coordinates; for imaging it may operate on spatial summary statistics rather than raw pixels.

Table 7. Evaluation Metrics

Metric	Symbol	Focus	Type
Dynamic AUC	$C_\Delta t$	Discrimination	Time-varying
AUPRC	AUPRC	Precision-recall	Imbalanced data
Calibration Error	CalErr	Probability accuracy	Risk-based
Lead Gain	Δ_{lead}	Timing benefit	Clinical utility

Table 8. Decision Policy Levels

Risk Range	Action Level	Response	Priority
Low	a^0	Monitor	Minimal
Moderate	a^1	Review	Medium
High	a^2	Investigate	Elevated
Critical	a^3	Intervene	Urgent

Calibration should be treated as a primary objective, not a post hoc fix (14). A model that correctly ranks episodes but overstates absolute risk can generate clinically unacceptable alarm burden. Dynamic calibration can be enforced by matching predicted and empirical event frequencies across time and risk strata:

$$\mathcal{L}_{\text{cal}} = \tau \int_0^1 \left(\mathbb{E}\Big[Y_t \mid \hat{R}_t = r \Big] - r \right)^2 d\mu r \, dt, \tag{5.6}$$

where μ is a weighting measure over risk levels. In practice this is approximated by differentiable binning or isotonic-like smooth constraints on the risk path.

To handle domain variation across institutions, procedure types, and surveillance cultures, one can add a transport-robustness term. Let d index domains. The latent representation should preserve outcome-relevant structure without overfitting institution-specific acquisition habits. A combined invariance and discrepancy penalty can be written as

$$\mathcal{L}_{\text{shift}} = \sum_d \|\nabla_{\theta_h} R_d \theta_h, \theta_o\|_2^2 \\ \eta \sum_{d<d'} \text{MMD}\big(\mathcal{Z}_d, \mathcal{Z}_{d'}\big), \tag{5.7}$$

where θ_h are shared parameters, θ_o are output parameters, R_d is domain-specific risk, and $\mathcal{Z}_d$ denotes the latent representation distribution in domain d. The first term discourages reliance on brittle shortcuts that reverse across domains; the second reduces representation drift while preserving predictive structure.

The full training criterion is then

$$\mathcal{L} = \sum_i \big(\mathcal{L}^i_{\text{surv}} \;\; \beta_1 \mathcal{L}^i_{\text{delay}} \;\; \beta_2 \mathcal{L}^i_{\text{lead}} \\ \beta_3 \mathcal{L}^i_{\text{obs}} \;\; \beta_4 \mathcal{L}^i_{\text{align}} \big) \;\; \beta_5 \mathcal{L}_{\text{cal}} \;\; \beta_6 \mathcal{L}_{\text{shift}}. \tag{5.8}$$

The weights β_k encode the trade-off between statistical fit, temporal utility, and robustness. In settings where lead time is the main research question, β_2 should not be treated as a minor regularizer; it should be large enough that the optimizer cannot achieve good loss merely by predicting slightly earlier than chart confirmation.

Optimization itself is nontrivial because the state evolves continuously and observations arrive irregularly. A practical strategy is to parameterize the drift, hazard, and reliability components with differentiable modules while integrating trajectories through adaptive solvers between event times. The jump updates at acquisition events permit exact assimilation points, while continuous intervals are solved numerically. Mini-batching can proceed by grouping episodes with similar event counts rather than similar lengths, reducing integration variance. Stochastic adjoint methods are possible but should be audited for numerical instability when hazards become sharp.

An additional issue is class imbalance (15). Anastomotic failure is uncommon relative to uneventful recovery, and the rare-event structure worsens for fine-grained lead labels because only a small temporal fraction of an episode lies close to failure. Rather than global class weights alone, the lead-sensitive loss already reallocates emphasis toward clinically relevant intervals. However, extreme imbalance may still require focal reweighting around false negatives at useful lead times. The challenge is to do this without destroying calibration. One solution is to separate the hazard-ranking component from the calibration component so that focal emphasis sharpens discrimination while the calibration term restrains probability inflation.

The final learning complication is intervention confounding. High-risk patients receive more imaging, tighter monitoring, broader antibiotics, or lower thresholds for reoperation. A model trained naively on observed outcomes may learn to associate aggressive surveillance with worse outcomes even when surveillance mitigates harm. Some of this can be absorbed by the latent observation and recognition submodels, but a stronger approach is to include treatment processes explicitly in u_t and to penalize risk estimates that become pathologically dependent on post-treatment proxies rather than pre-treatment state. This does not solve causal identification, but it reduces the chance that the model becomes a detector of clinician concern rather than patient deterioration.

Table 9. Hazard Function Inputs

Input	Symbol	Role	Effect
State	x_t	Core signal	Linear impact
Velocity	x_t	Trend	Early warning
Context	ξ_t	External factors	Modulation
Interaction	$x_t^T H x_t$	Synergy	Nonlinear boost

Table 10. Earlier-than-EHR Advantages

Aspect	EHR-Based	Proposed	Benefit
Timing	Delayed	Immediate	Earlier alerts
Resolution	Aggregated	Raw signals	More detail
Adaptivity	Static	Dynamic	Continuous updates
Utility	Limited	Lead-aware	Actionable insight

6. Evaluation, Lead-Time Utility, and Identifiability

Evaluation of earlier-than-EHR systems cannot stop at discrimination metrics computed on final labels. A model can achieve high area under the curve while providing negligible additional lead over chart-based monitoring, or it can produce early alarms that are useless because they are poorly calibrated or operationally overwhelming. The evaluation framework must therefore reflect three linked questions: whether the model predicts the event, whether it predicts it early enough to change care, and whether it does so at acceptable burden. This section formulates metrics consistent with those requirements.

The most direct comparison is between acquisition-time and chart-time risk trajectories. For each episode, define alarm times under a shared decision rule as in the problem formulation. Aggregate lead gain can then be measured by

$$\begin{aligned} \Delta_{\text{lead}}\theta &= \mathbb{E}\big[\Lambda_i\theta \mid \tau_i^\star\theta < \infty\big], \\ \Pi_{\text{useful}}\theta, \tau_0 &= \Pr\big(\Lambda_i\theta \geq \tau_0\big), \end{aligned} \tag{6.1}$$

where τ_0 is a minimum clinically meaningful lead threshold (16). Reporting only mean lead can conceal the fact that gains are concentrated in a few episodes; the probability of achieving useful lead is therefore equally important.

Discrimination should be evaluated dynamically. At each time t, one can assess whether episodes failing within horizon Δ receive higher risk than those remaining stable beyond $t\ \Delta$. Dynamic concordance and horizon-specific precision-recall are suitable:

$$\begin{aligned} \mathrm{C}_\Delta t &= \Pr\Big(\hat{R}_i t, \Delta > \hat{R}_j t, \Delta \mid T_i \in t, t\ \Delta,\ T_j > t\ \Delta\Big), \\ \mathrm{AUPRC}_\Delta t &= \int_0^1 \mathrm{Prec}_\Delta t, r\, d\,\mathrm{Rec}_\Delta t, r. \end{aligned} \tag{6.2}$$

These quantities vary over the perioperative timeline. A model that performs well during intraoperative or immediate postoperative windows may be more useful than one whose average performance is driven by later, easier intervals.

Calibration must also be dynamic. The relevant question is whether a predicted risk of, say, 0.30 over the next six hours corresponds to an actual six-hour event frequency near 0.30 in episodes with similar risk and phase. Time-local calibration error can be defined as

$$\mathrm{CalErr}_\Delta t = \int_0^1 \Big|\Pr\Big(T \leq t\ \Delta \mid T > t, \hat{R}t, \Delta = r\Big) - r\Big|\ d\mu r. \tag{6.3}$$

Because earlier-than-EHR systems may operate before conventional confirmation signals appear, miscalibration can be especially damaging: clinicians will tolerate uncertainty less readily when the model acts before the chart looks abnormal.

Operational burden enters through alarm frequency, dwell time above threshold, and intervention-adjusted positive predictive value. Let $A_i t = \mathbf{1}\{\hat{R}_i t, \Delta \geq \theta\}$ and define

$$\begin{aligned} B\theta &= \mathbb{E}\big[\textstyle\int_0^{\tau_i} A_i t\, dt\big], \\ N_{\text{alerts}}\theta &= \mathbb{E}\Big[\textstyle\sum_k \mathbf{1}\{\text{upcrossing}_k\theta\}\Big], \\ \mathrm{PPV}_{\text{act}}\theta &= \Pr\big(T \leq \tau^\star\theta\ \Delta \mid \tau^\star\theta < \infty\big). \end{aligned} \tag{6.4}$$

Dwell time matters because a model that stays above threshold continuously may be impossible to operationalize even if it has good ranking properties. Upcrossing counts matter because alarm fatigue is driven more by repeated alerts than by continuous displays on a dashboard.

To combine earliness and burden into a single evaluative object, define a lead-time utility. Let $\tau_i^\star$ be the alarm time and let U_i be

$$\begin{aligned} U_i = {} & g T_i - \tau_i^\star\, \mathbf{1}\{\tau_i^\star < T_i\} \\ & - c_{\text{fp}}\, \mathbf{1}\{\tau_i^\star < \infty,\ \tau_i^\star \geq T_i\} \\ & - c_{\text{none}}\, \mathbf{1}\{\tau_i^\star = \infty,\ \delta_i = 1\}, \end{aligned} \tag{6.5}$$

where g is the gain from actionable lead, c_{fp} is false-positive cost, and c_{none} is the cost of failing to alert. Expected utility can then be traced over thresholds. This makes explicit that two models with similar AUC may differ markedly in clinical value because one shifts alarms into a region where management options still diverge.

A defining feature of earlier-than-EHR research is the need to compare models under different information clocks (17). A

fair comparison must evaluate not only the multimodal model against a chart-time baseline, but also the same architecture with acquisition times artificially delayed to chart times. This ablation isolates the value of timing from the value of representation. If the model improves only when raw observations are allowed at acquisition time, then the gain is genuinely earlier-than-EHR rather than merely due to a richer encoder. The clock-ablation design is therefore fundamental.

Identifiability deserves careful treatment. One cannot fully identify latent biologic onset from recognition labels without auxiliary assumptions. Similarly, one cannot infer that earlier alarms cause improved outcomes from retrospective data alone. What can be identified under moderate assumptions is narrower: the statistical relation between acquisition-time observations and future recognized events; the incremental lead relative to chart-time constraints; the calibration of predicted horizons; and the stability of these quantities across domains. The state-space hazard model should thus be evaluated as a predictive device with explicit timing semantics, not as a causal oracle.

This caution is especially important when multimodal signals partially encode clinician behavior. Device ordering patterns, surgical video acquisition, and laboratory intensity can all reflect preexisting concern. Such variables may still be useful predictors, but they blur the line between patient-state inference and workflow-state inference. A transparent evaluation protocol should therefore report performance with and without workflow proxies, and it should test robustness under counterfactual perturbations of observation intensity. One operational approach is to randomize or mask selected workflow variables during evaluation and quantify risk drift:

$$\mathrm{Sens}_{\mathrm{wf}} = \mathbb{E}\left[\left|\hat{R}t, \Delta; \zeta_t - \hat{R}t, \Delta; \zeta_t^{\mathrm{masked}}\right|\right]. \tag{6.6}$$

A large value indicates that the model may be relying excessively on surveillance practice rather than physiology.

Fairness in this context is less about forcing identical risk scores across subgroups than about ensuring comparable calibration, lead gain, and error asymmetry. Because earlier-than-EHR modalities may be differentially available across patient groups or institutions, subgroup evaluation should include modality availability profiles. A model may appear equally accurate overall while systematically delivering later or less certain alarms for groups with reduced access to high-information modalities (18). The evaluation set should therefore stratify by both patient attributes and modality regimes.

In summary, evaluation for earlier-than-EHR multimodal prediction must treat time as first-class. The question is not simply whether the model predicts failure, but whether acquisition-time evidence shifts risk estimation forward in a way that remains calibrated, robust, and operationally tolerable. Any framework that omits the clock comparison, lead-time utility, and burden metrics will underestimate the central contribution of such systems.

7. Clinical Integration as a Constrained Decision Process

Prediction alone does not define clinical usefulness. A perioperative risk score becomes relevant only when embedded in a decision process with limited attention, finite staffing, heterogeneous intervention costs, and uncertain treatment effects. Earlier-than-EHR systems create both opportunity and obligation: because they may fire before the chart appears clearly abnormal, they require more explicit reasoning about what action follows an alert and how much uncertainty clinicians are expected to tolerate. The natural formalism is constrained sequential decision-making rather than passive risk display.

Let $a_t \in \mathcal{A}$ denote an action at time t, where actions can range from no escalation to repeat perfusion assessment, additional imaging, broadened surveillance, hemodynamic optimization, diversion consideration, or early consultation. Let x_t be the latent state and let $\hat{R}_t$ be the risk estimate. The value of an action depends on the hidden trajectory and on the time remaining before failure becomes clinically obvious. One can define an action policy

$$\pi a_t \mid \hat{R}_t, x_t, \chi_t, \tag{7.1}$$

where χ_t contains context such as unit workload, postoperative phase, and current treatment status. The policy should maximize expected utility while respecting alert and intervention budgets:

$$\pi^\star = \arg\max_\pi \mathbb{E}_\pi\left[\int_0^\tau u a_t, x_t, T, \chi_t \, dt\right]$$
$$\text{subject to } \mathbb{E}_\pi\left[N_{\mathrm{alerts}}\right] \le B_1, \, \mathbb{E}_\pi\left[C_{\mathrm{int}}\right] \le B_2, \tag{7.2}$$

where u is a context-sensitive utility, N_{alerts} is the number of alerts, C_{int} is intervention cost, and B_1, B_2 are operational budgets. This formulation prevents the common but unrealistic assumption that every risk threshold crossing should trigger the same response.

The state estimate enables graded rather than binary intervention. A model may indicate that risk is driven primarily by perfusion uncertainty, by systemic physiologic drift, or by a longstanding predisposition with little acute change. Each pattern suggests a different response. The decision layer should therefore use state attribution, not just scalar risk. For example, a perfusion-dominant alarm intraoperatively might favor repeat imaging or revision of the anastomosis (19). A postoperative waveform-dominant alarm might favor earlier imaging, laboratory reassessment, or intensive monitoring. A molecular predisposition signal without acute instability might alter prophylactic planning rather than trigger urgent action.

Operationally, progressive escalation is preferable to a single hard alarm. Define several action bands based on risk and uncertainty:

$$a_t = \begin{cases} a^0, & \hat{R}_t < \theta_1, \\ a^1, & \theta_1 \le \hat{R}_t < \theta_2, \\ a^2, & \theta_2 \le \hat{R}_t < \theta_3, \\ a^3, & \hat{R}_t \ge \theta_3, \end{cases} \tag{7.3}$$

with thresholds chosen under a utility model rather than pure F1 optimization. Lower bands may surface contextual explanations without interruptive alerts, while higher bands may require acknowledgment and action logging. This layered design is especially important for earlier-than-EHR systems because the evidentiary standard for interruptive alerts should be higher when the chart does not yet corroborate the concern.

Human oversight benefits from counterfactual explanations tied to modality timing. A useful display is not merely "risk is high" but "risk is elevated because acquisition-time perfusion heterogeneity and waveform instability jointly exceed the expected postoperative trajectory, while charted values remain within nominal bounds." Such explanations highlight that the alert is earlier-than-EHR by design rather than apparently contradictory to the chart. They may also reduce the understandable tendency to dismiss pre-charting alerts as algorithmic noise.

Latency constraints matter technically. If a modality is earlier-than-EHR only in principle but the inference pipeline takes longer than charting to process it, the advantage disappears. The system therefore needs bounded end-to-end latency. Continuous waveforms and device logs are naturally streamable. Video-derived features require near-real-time preprocessing with quality control. Molecular assays are more variable; some will remain preoperative rather than intraoperative inputs. The architectural implication is that the representation and state updates should be incremental. Full retrospective re-encoding of long histories is incompatible with the intended use (20).

Auditability also becomes more demanding. Since some inputs precede the formal chart, institutions will require a precise account of what data were observed, when they were observed, and how they influenced the state. The jump-update decomposition from the dynamics section provides a basis for such audit trails. Every alert can be accompanied by a modality-time map showing acquisition time, quality score, state shift, and incremental hazard contribution. This is not only a governance requirement. It also supports error analysis when clinicians disagree with the model.

Clinical integration further requires explicit handling of counterfactual data absence. Many centers will not have all modalities. The model should degrade gracefully to subsets rather than assuming full multimodal availability. One can operationalize this through modality-contingent policies and by maintaining uncertainty intervals that widen when high-information channels are absent. Rather than pretending to the same confidence under sparse data, the system should indicate that a risk estimate is driven predominantly by lower-fidelity evidence and therefore should influence action differently.

A final deployment issue concerns feedback loops. If the model triggers earlier imaging or reoperation, future data distributions will change, and apparent incidence may shift because detection becomes earlier or more frequent. This is not merely a postdeployment nuisance. The decision layer should already anticipate that observed labels after deployment reflect a policy-dependent world. An adaptive monitoring program will therefore need ongoing recalibration and possibly off-policy evaluation methods. Earlier-than-EHR systems will likely create stronger feedback loops than standard chart-based scores because they intervene earlier in the clinical course.

Framed as a constrained decision process, multimodal earlier-than-EHR prediction becomes less mysterious (21). The model estimates a state and a hazard under acquisition-time information; the policy maps those estimates into graded actions under workload constraints. The quality of the whole system depends not only on prediction accuracy but on whether the timing advantage is converted into actions that are proportionate, interpretable, and operationally sustainable.

8. Discussion

The framework developed here rests on a simple but consequential premise: the structured electronic health record is an incomplete temporal interface to the patient state. In perioperative anastomotic failure, some of the most informative signals arise before the chart meaningfully reflects them, either because they are raw measurements awaiting abstraction or because they come from modalities that conventional EHR pipelines ignore. Treating the problem as one of earlier-than-EHR inference changes both the data model and the evaluation target. It shifts emphasis from static tables to asynchronous acquisition streams, from feature concatenation to latent state estimation, and from overall discrimination to clinically useful lead under burden constraints.

One implication is that multimodality is valuable primarily because different modalities occupy different positions in the timing and causality landscape. A modality is not helpful merely because it adds dimensions. It is helpful when it contributes evidence that is temporally proximal to the biologic process, informative under current uncertainty, and available before structured documentation catches up. This principle explains why raw waveforms, intraoperative imaging, and selected molecular profiles can matter even when their late charted surrogates already exist. The charted surrogate is both delayed and compressed. The raw modality preserves trajectory shape, transient instability, and quality context that may be essential for early risk estimation.

Another implication is that missingness can no longer be treated as an inconvenience. In multimodal perioperative data, what is observed depends on technology, workflow, suspicion, and institutional habit. Earlier-than-EHR modeling therefore requires explicit observation models. Without them, the system may conflate surveillance intensity with deterioration, or it may overtrust modalities available only in select cases. The proposed framework handles this by modeling reliability, staleness, and observation intensity as part of the latent state dynamics. That design does not eliminate bias, but it makes the bias legible and therefore auditable (22).

The continuous-time hazard formulation is also a conceptual advance over static classification, but it should not be romanticized. Continuous time is useful because the underlying clinical process is irregular, not because continuous models are intrinsically superior. They introduce their own burdens, including numerical instability, sensitivity to solver choices, and more complicated training. In small datasets, a discretized approximation may be more robust. The point is that if the scien-

tific claim concerns lead measured in minutes or hours relative to raw acquisition and chart delays, then the temporal model must be able to represent those intervals faithfully. Otherwise the evaluation of earliness becomes an artifact of arbitrary binning.

Several limitations of the framework are clear. First, latent onset of anastomotic failure is not directly observed. Any decomposition into biologic onset and recognition is assumption-laden and will depend on proxy events, intervention logs, and surveillance intensity. The model therefore estimates a decision-relevant hidden process rather than an incontrovertible mechanistic truth. Second, not all earlier-than-EHR signals will transport well across institutions. Video protocols, waveform fidelity, assay platforms, and charting habits differ. This is why transport-robust objectives and clock-based ablations are necessary. Third, high-information modalities may be unequally distributed across patients and hospitals. A system that performs well only where sophisticated imaging and dense monitoring are available may widen disparities if deployed uncritically.

Interpretability remains a delicate issue. The proposed state attribution and modality-jump decomposition offer a path toward local explanation, but they do not fully solve the challenge of communicating complex temporal evidence to clinicians under pressure. Interpretability in this setting is not just a mathematical property; it is a workflow property (23). A risk explanation is useful only if it helps a clinician decide whether to trust the signal and what to do next. This suggests that future methodological work should measure not only calibration and lead gain but also explanation stability under small perturbations and the extent to which state attributions agree with clinician reasoning patterns.

A related concern is intervention confounding. Much perioperative data is generated in response to concern. An earlier-than-EHR model may detect that concern before it appears in the structured chart, which can be useful operationally, but it raises the question of whether the model is identifying patient deterioration or clinician suspicion. The right response is not to exclude all workflow variables automatically. Clinician concern is often a legitimate summary of subtle cues. The better response is to separate workflow-mediated evidence from direct physiologic evidence, quantify reliance on each, and test sensitivity to their removal. This preserves predictive value while guarding against circular deployment in which the model simply amplifies preexisting behavior.

The paper has emphasized anastomotic failure, but the framework is more general. Any perioperative complication in which raw or semi-structured signals precede structured chart abstraction could be modeled similarly. The general lesson is that risk prediction should be indexed to the earliest reliable observation time, not to the first convenient tabular representation. Once that principle is adopted, the technical agenda expands naturally to include acquisition-time filtering, reliability-aware multimodal fusion, lead-sensitive training, and decision-theoretic deployment.

What remains uncertain is how much of the theoretical timing advantage will survive in full clinical practice. Some modalities may yield modest incremental lead once preprocessing latency, quality failures, and human review are accounted for. Others may prove institution-specific or too costly to maintain. The value of the framework is not that it guarantees large gains, but that it provides a rigorous way to test for them. By distinguishing acquisition time from chart time, biologic onset from recognition, prediction from action, and discrimination from lead utility, it converts an intuitive idea into a falsifiable modeling program (24).

References

[1] V. Kane and M. C. Talpallikar, "Study on risk factors for mortality and morbidity in patients undergoing surgery for colorectal cancer," *International Surgery Journal*, vol. 6, pp. 2766–, 7 2019.

[2] B. Sadanandan, "Data-driven risk stratification for anastomotic leak: A proof-of-concept study using electronic health records," *Journal of Data Science, Predictive Analytics, and Big Data Applications*, vol. 9, no. 6, pp. 1–27, 2024.

[3] W. J. Kim, C. M. Lee, L. An, J. H. Kim, and S. Park, "Effect of biologic material reinforcement on surgical anastomosis after gastrectomy-a pilot study.," *Frontiers in oncology*, vol. 9, pp. 1184–, 11 2019.

[4] C. K. McCarthy, P. K. McGaha, N. S. Rozich, N. A. Yokell, J. S. Lees, and W. L. Berry, "Creation of colonic anastomosis in mice," *Journal of Visualized Experiments*, 1 2019.

[5] E. Martínez-López, A. Martínez-Pérez, S. Navarro-Martínez, J. C. Sebastián-Tomás, N. de'Angelis, and E. García-Granero, "Real-time fluorescence image-guided gastrointestinal oncologic surgery: Towards a new era.," *World journal of gastrointestinal oncology*, vol. 13, pp. 1029–1042, 9 2021.

[6] D. A. Iozsa, R.-I. Spătaru, L. F. Tomescu, O. Stiru, F. Gherghiceanu, F. Furtunescu, D. Radavoi, N. Bacalbasa, I. Balescu, and A. Tulin, "Pediatric esophageal stenoses: Challenges and new surgical device promoting tension-free esophageal anastomosis.," *Experimental and therapeutic medicine*, vol. 23, pp. 220–220, 1 2022.

[7] V. N. Rohith, S. V. Arya, A. Rani, R. K. Chejara, A. Sharma, J. K. Arora, D. S. Kalwaniya, A. Tolat, P. G, and A. Singh, "Preoperative serum albumin level as a predictor of abdominal wound-related complications after emergency exploratory laparotomy.," *Cureus*, vol. 14, pp. e31980–e31980, 11 2022.

[8] C. R. Baker, M. E. Bailey, Y. Soon, S. Singh, and S. R. Preston, "Two-phase laparoscopic-assisted oesophago-gastrectomy: a single-unit experience of 111 consecutive cases and outcomes," *Surgical endoscopy*, vol. 25, pp. 3658–3667, 6 2011.

[9] Y. Yang, Q. Wang, J. Luo, Y. Jiang, R. Zhou, S. Tong, Z. Wang, and Q. Tong, "Superoxide dismutase mimic, mnte-2-pyp enhances rectal anastomotic strength in rats after preoperative chemoradiotherapy.," *Oxidative medicine and cellular longevity*, vol. 2020, pp. 3509859–11, 4 2020.

[10] U. Jogiat, V. Mocanu, D. W. Birch, N. J. Switzer, S. R. Turner, and S. Karmali, "Pulmonary complications are a strong independent predictor of 30-day mortality following elective bariatric surgery.," *Obesity surgery*, vol. 32, pp. 696–703, 1 2022.

[11] A.-G. Chen, R. Zhang, W. Xia, N. Kang, W. Ge, K. Zhu, and Z. cheng Yu, "Combined laparoscopic and thoracoscopic esophagectomy for esophageal carcinoma and gastro-esophageal anastomosis in right thoracic cavity: analysis of 38 cases," *Chinese Journal of Thoracic and Cardiovaescular Surgery*, vol. 29, pp. 525–527, 9 2013.

[12] A. T. Cristaudo, K. Hitos, R. Gunnarsson, A. de Costa, and A. Yuen, "Re: Development and validation of a multivariable prediction model in open abdomen patients for entero-atmospheric fistula.," 12 2023.

[13] A. Bhangu, P. Singh, J. E. Fitzgerald, A. Slesser, and P. Tekkis, "Postoperative nonsteroidal anti-inflammatory drugs and risk of anastomotic leak: Meta-analysis of clinical and experimental studies," *World journal of surgery*, vol. 38, pp. 2247–2257, 3 2014.

[14] M. Sultana, V. Butnari, A. Mansuri, C. Rao, S. Kaul, R. Boulton, J. Huang, and N. Rajendran, "Comparison of early surgical outcomes

of robotic and laparoscopic colorectal cancer resection in a single centre: A cohort study," 5 2023.

[15] D. K. Manatakis, N. Gouvas, J. Souglakos, and E. Xynos, "Neo-adjuvant chemotherapy alone for the locally advanced rectal cancer: a systematic review," *International journal of clinical oncology*, vol. 25, pp. 1570–1580, 7 2020.

[16] R. R, "A clinical study – resection and anastomosis of bowel in our surgical practice," *Academia Journal of Surgery*, vol. 3, pp. 1–7, 5 2020.

[17] B.-U. Nguyen, M. T. Vu, Q.-T. Pham, V.-H. N. Phung, Q.-N. N. Dang, T. Thai, D.-T. L. Vo, T. T. Tran, M. M. Huynh, Q. K. Huynh, T. K. Ngo, and H. T. Dao, "Adopting the swenson-like technique for patients with hirschsprung disease in vietnam.," *Pediatric surgery international*, vol. 39, pp. 124–124, 2 2023.

[18] M. S. Abdelhamid, A. Rashad, M. A. Negida, A. Garib, S. Soliman, and T. Elgaabary, "Emergency laparoscopic left sided colonic resection with primary anastomosis: Feasibility and safety," *Archives of Surgery and Clinical Research*, vol. 2, pp. 031–038, 11 2018.

[19] J. L. Goldwag, L. R. Wilson, S. J. Ivatury, E. M. Pauli, M. J. Tsapakos, and M. Z. Wilson, "Stoma closure and reinforcement (scar): A study protocol for a pilot trial.," *Contemporary clinical trials communications*, vol. 19, pp. 100582–, 6 2020.

[20] A. Otero-Piñeiro, F. B. de Lacy, J. J. V. Laarhoven, B. Martin-Perez, S. Valverde, R. Bravo, and A. M. Lacy, "The impact of fluorescence angiography on anastomotic leak rate following transanal total mesorectal excision for rectal cancer: a comparative study.," *Surgical endoscopy*, vol. 35, pp. 754–762, 2 2020.

[21] J. L. Grosfeld, V. N. Ballantine, and J. F. Csicsko, "A critical evaluation of the duhamel operation for hirschsprung's disease.," *Archives of surgery (Chicago, Ill. : 1960)*, vol. 113, pp. 454–460, 4 1978.

[22] A. H. Sohail, A. Beneat, A. H. Sohail, A. Morel, H. Ali, H. Hakmi, J. Whittington, M. K. Gangwani, S. Islam, S. Maan, and T. B. S. Pacheco, "Non-gastrointestinal complications and short-term outcomes in patients with anastomotic leak after colon resection: A propensity matched analysis of american college of surgeons national surgical quality improvement program database," *International Journal of Surgery*, pp. 1–7, 8 2023.

[23] M. F. Cunha, B. Mendes, P. Mendanha, I. Miguel, J. Rachadell, E. Amorim, and P. Castelo-Branco, "Albumin and crp, which is the best biomarker for postoperative complications on colorectal surgery? prospective cohort study," 11 2021.

[24] J. Ahmed, M. Nasir, K. Flashman, J. Khan, and A. Parvaiz, "Totally robotic rectal resection: an experience of the first 100 consecutive cases.," *International journal of colorectal disease*, vol. 31, pp. 869–876, 2 2016.