Journal Homepage: www.ijarpr.com ISSN: 3049-0103 (Online)

International Journal of Advance Research Publication and Reviews

Vol 03, Issue 9, pp 352-365, September, 2026

Explainable AI for Linking Clinical Deterioration Patterns with Escalating Hospital Treatment Costs Early

Johnbull Anderson

Western Illinois University, Macomb, Illinois, USA

DOI : https://doi.org/10.5281/zenodo.22767565

ABSTRACT

Healthcare systems increasingly require methods that identify clinical deterioration before worsening illness produces avoidable treatment escalation, prolonged hospitalization, and rising expenditure. Conventional early-warning scores and cost analyses often examine clinical risk and financial outcomes separately, limiting understanding of how evolving patient conditions translate into resource-intensive care. This study proposes an explainable artificial intelligence framework for linking longitudinal clinical deterioration patterns with escalating hospital treatment costs early in the care pathway. The approach integrates demographic characteristics, comorbidities, vital signs, laboratory trends, medication changes, diagnostic activity, complications, ward utilization, intensive-care exposure, procedures, and patient-level cost data. Predictive models estimate deterioration and subsequent expenditure, while explainability methods quantify the contribution, direction, and temporal importance of individual clinical and utilization features. By identifying when and why predicted costs increase, the framework distinguishes expenditure associated with genuine clinical complexity from potentially modifiable escalation in resource use. Patient-level stratification supports early clinical review, targeted intervention, capacity planning, and financial forecasting. The framework positions explainable AI as decision support that connects clinical trajectories with economic consequences without replacing clinician judgment or implying causality from predictive associations.

Keywords: Explainable Artificial Intelligence; Clinical Deterioration; Hospital Treatment Costs; Predictive Analytics; Healthcare Resource Utilization; Early Risk Stratification

1. INTRODUCTION

1.1 Clinical Deterioration and the Escalating Economic Burden of Hospital Care

Hospitals increasingly manage patients whose clinical conditions evolve rapidly after admission, creating simultaneous pressures on care delivery, resource capacity, and expenditure [1]. Physiological deterioration reflected through unstable vital signs, worsening laboratory parameters, respiratory compromise, or emerging organ dysfunction can progressively increase treatment requirements [2]. Such changes may necessitate more frequent observations, repeated laboratory investigations, advanced imaging, specialist consultations, medication intensification, emergency procedures, or transfer to intensive care [3]. Consequently, hospital expenditure is not determined solely by diagnosis or length of stay; it can accumulate dynamically as changing clinical severity generates increasingly resource-intensive treatment pathways [4]. Prolonged hospitalization can further compound this burden through continuing nursing, pharmaceutical, diagnostic, and accommodation requirements [5]. However, conventional financial analyses commonly quantify these consequences retrospectively, identifying high-cost hospitalizations after substantial expenditure has occurred [2]. Earlier recognition of deterioration-associated cost trajectories could therefore provide greater clinical and operational value.

1.2 Limitations of Conventional Clinical and Financial Risk Assessment

Conventional clinical risk assessment commonly relies on early-warning scores, predefined physiological thresholds, comorbidity measures, or diagnosis-based severity classifications [6]. Although clinically useful, these approaches can compress complex patient trajectories into static scores and inadequately represent interactions among multiple variables changing simultaneously over time [7]. Financial models based on average treatment cost, diagnosis-related categories, regression estimates, or anticipated length of stay face a related limitation: expenditure is frequently treated as an endpoint rather than an evolving consequence of clinical change [4]. Moreover, linear assumptions may inadequately capture nonlinear relationships between deterioration, treatment escalation, and resource consumption [1]. Clinically meaningful signals preceding financial escalation may therefore remain distributed across multivariate longitudinal records, requiring analytical methods capable of learning their temporal relationships before

substantial expenditure develops.

1.3 CNNs and Explainable AI for Temporal Clinical-Financial Intelligence

Machine-learning approaches provide opportunities to integrate high-dimensional clinical information and identify patterns preceding adverse outcomes or increasing resource requirements [7]. Deep-learning architectures extend this capability by learning representations directly from complex longitudinal observations rather than depending exclusively on manually specified relationships [8]. In particular, one-dimensional convolutional neural networks (1-D CNNs) can apply convolutional filters across sequential observations to detect localized temporal signatures involving vital signs, laboratory measurements, and evolving treatment variables [9]. Such representations could reveal deterioration patterns that precede escalating hospital expenditure and remain difficult to characterize using static models [5]. Predictive accuracy alone, however, is insufficient for clinically consequential decision support. Explainable artificial intelligence (XAI) is therefore required to identify influential clinical variables and temporal regions, enabling clinicians to understand why the model anticipates increasing financial risk and assess whether predictions correspond with clinically plausible deterioration [2].

1.4 Aim, Objectives, and Contribution

This study develops a MATLAB-based explainable 1-D CNN framework for identifying temporal clinical deterioration patterns associated with subsequent escalation of hospital treatment costs [8]. The framework integrates longitudinal physiological, laboratory, treatment, resource-utilization, and financial information to estimate cost-escalation risk before peak expenditure occurs. Its principal contributions are temporal modelling of evolving deterioration, early financial-risk prediction, and explainable attribution of predictions to specific clinical variables and observation periods [4]. Performance is compared with conventional predictive approaches to determine whether temporal deep learning provides additional discriminatory value [6]. The resulting risk stratification is intended to support clinically interpretable surveillance, proactive resource planning, and hospital decision support.

2. CLINICAL DETERIORATION–COST ESCALATION FRAMEWORK

2.1 Temporal Progression of In-Hospital Clinical Deterioration

Clinical deterioration during hospitalization is better understood as a progressive temporal process than as a binary transition between stable and unstable states [8]. A patient's physiological condition can evolve through interacting changes in heart rate, respiratory rate, oxygen saturation, blood pressure, temperature, neurological observations, urine output, and laboratory measurements [13]. Individually modest abnormalities may become clinically important when they persist, accelerate, or occur simultaneously across physiological systems [17]. Repeated measurements therefore provide information about direction, velocity, variability, and persistence that isolated observations cannot capture [10]. For patient i, the longitudinal trajectory can be represented as:

$$X_i = [x_{i,1}, x_{i,2}, \ldots, x_{i,T}]$$

where $x_{i,t}$ denotes the multivariate clinical state observed at time t [15]. This temporal representation enables emerging deterioration patterns to be characterized before major treatment escalation and increasing resource requirements become clinically apparent [11].

2.2 From Deterioration to Treatment and Resource Escalation

Physiological deterioration becomes financially consequential when worsening clinical status changes the intensity and composition of required care [14]. Declining oxygen saturation may initiate supplemental oxygen or ventilatory support, while hemodynamic instability can prompt repeated observations, laboratory testing, medication adjustment, imaging, specialist review, or emergency intervention [9]. Continued deterioration may subsequently require invasive procedures, increased nursing intensity, or intensive-care transfer, substantially expanding resource consumption [16]. The proposed mechanism links progressive clinical deterioration with treatment escalation, increasing resource requirements, greater utilization intensity, and the resulting accumulation of hospital treatment costs. These relationships need not progress uniformly; several deterioration signals can converge on one intervention, while a single clinical event can activate multiple resource domains simultaneously [12]. Consequently, baseline characteristics and evolving physiology should be modelled as interconnected determinants of treatment escalation, resource intensity, accumulating expenditure, and emerging financial risk [17].

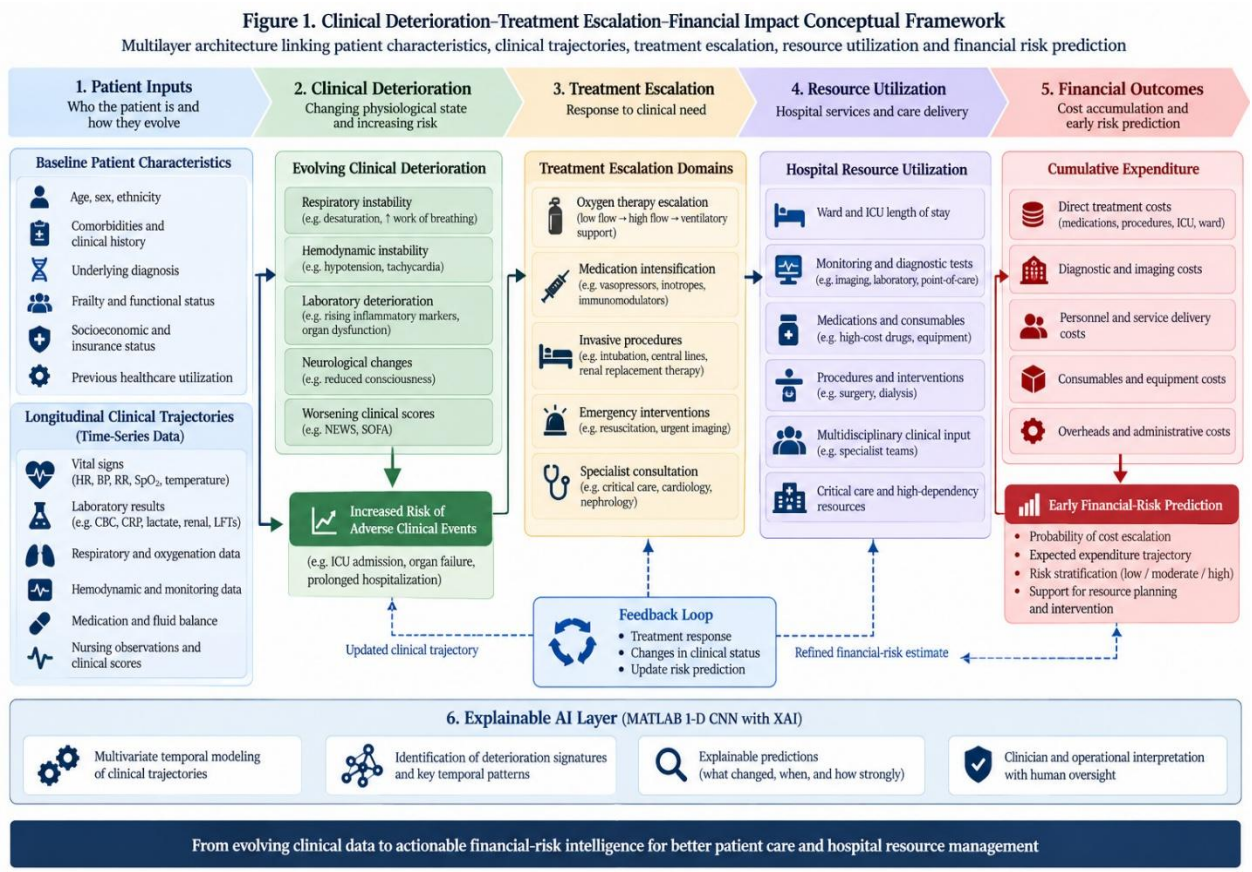

Figure 1. Clinical Deterioration–Treatment Escalation–Financial Impact Conceptual Framework

2.3 Temporal Definition of Hospital Treatment Cost Escalation

Hospital treatment expenditure should be represented as a cumulative time-dependent process because resources are consumed repeatedly as the patient's condition and care pathway evolve [10]. For patient i, cumulative expenditure at time $t$can be defined as:

$$C_i(t) = \sum_{\tau=1}^{t} \sum_{r=1}^{R} q_{ir}(\tau) c_r$$

where $q_{ir}(\tau)$denotes utilization of resource $r$during interval τ, while $c_r$represents its corresponding unit cost [15]. This formulation permits diagnostics, medications, procedures, intensive care, and other services to contribute according to when consumption occurs [13]. Early prediction additionally requires separating observed expenditure from expenditure occurring after prediction [9]. Future cost escalation can therefore be expressed as:

$$\Delta C_i^{future} = C_i(t+h) - C_i(t)$$

where $h$defines the prediction horizon [16]. The objective is consequently to estimate impending expenditure using clinical information available at time t, preventing future information leakage [12].

2.4 Explainability as a Clinical-Financial Requirement

A probability indicating future treatment-cost escalation has limited operational value unless users can understand the evidence producing that prediction [11]. Explainability should identify which vital-sign changes, laboratory abnormalities, treatment indicators, and temporal intervals contribute most strongly to estimated financial risk [14]. This distinction is particularly important because expensive treatment may represent clinically necessary escalation rather than inefficient resource consumption [17]. Global explainability characterizes recurring relationships learned across the population, identifying variables consistently associated with subsequent expenditure escalation [8]. Local explainability examines an individual patient's prediction, revealing the measurements and temporal periods responsible for assigned risk [16]. Combining both perspectives permits assessment of clinical plausibility, strengthens human oversight, and establishes an interpretable connection between evolving deterioration patterns and anticipated financial consequences [13].

3. METHODOLOGY: DATA ACQUISITION AND TEMPORAL FEATURE ENGINEERING

3.1 Study Design, Patient Cohort, and Observation Windows

A retrospective longitudinal cohort design was used to reconstruct clinical trajectories and subsequent treatment expenditure among hospitalized adult patients [16]. Eligible patients had a clearly identifiable index hospitalization, sufficient longitudinal physiological measurements, documented treatment activity, and corresponding resource-utilization and cost records [21]. Encounters containing irreconcilable timestamps, insufficient observation duration, or incomplete outcome information were excluded [24]. The analytical timeline separated information available for prediction from subsequent financial outcomes. An initial 24–48-hour observation window was used to characterize evolving clinical status, while cost escalation occurring during the subsequent 24–72 hours or remaining hospitalization constituted the prediction outcome [18]. Alternative observation and prediction windows were examined through sensitivity analysis. This chronological separation ensured that future treatment events, resource consumption, and expenditure were excluded from predictor construction, thereby preventing temporal information leakage and preserving realistic early-warning evaluation [23].

3.2 Clinical, Treatment, Resource, and Financial Data Acquisition

Data acquisition integrated clinical, therapeutic, operational, and financial information at patient and encounter levels [20]. Baseline variables included age, sex, admission type, principal diagnosis, comorbidities, and previous healthcare utilization. Longitudinal measurements comprised heart rate, respiratory rate, blood pressure, oxygen saturation, temperature, neurological observations, and relevant laboratory results [25]. Treatment information included medication administration, oxygen supplementation, ventilatory support, procedures, specialist interventions, and emergency treatment [17]. Operational records captured diagnostic imaging, laboratory investigations, ward utilization, ICU admission and duration, nursing intensity where measurable, and total length of stay [22]. Financial records provided direct costs attributable to accommodation, diagnostics, pharmaceuticals, procedures, intensive care, and other identifiable hospital services. All longitudinal observations were retained with timestamps to permit alignment with evolving clinical trajectories [19]. Predictor construction was restricted to information recorded within the observation window, whereas resource utilization and expenditure occurring subsequently were reserved for defining the financial prediction outcome [24].

Table 1. Multimodal Clinical, Treatment, Resource-Utilization, and Cost Data Matrix

Domain	Variables	Temporal Resolution	Source	Preprocessing	Analytical Role
Patient baseline	Age, sex, diagnosis, comorbidities, admission type	Admission-level	EHR/administrative records	Coding and validation	Baseline adjustment
Vital signs	Heart rate, respiratory rate, BP, SpO_2, temperature, neurological observations	Repeated/time-stamped	EHR/monitoring systems	Alignment and normalization	Temporal deterioration input
Laboratory data	Hemoglobin, WBC, creatinine, electrolytes, lactate, relevant biomarkers	Repeated/time-stamped	Laboratory system	Range checking and synchronization	Physiological deterioration input
Treatment	Medications, oxygen, ventilation, emergency interventions	Event-level	EHR/pharmacy records	Temporal encoding	Treatment-escalation features
Procedures/diagnostics	Procedures, imaging, laboratory investigations	Event-level	Clinical systems	Event aggregation	Resource-intensity measurement
Intensive care	ICU admission and ICU duration	Hour/day-level	ICU/administrative records	Duration calculation	Escalation/resource indicator
Hospital utilization	Ward days, LOS, specialist services	Encounter/time-level	Administrative records	Episode reconstruction	Utilization measurement
Financial expenditure	Direct treatment and service costs	Event/day-level	Cost-accounting system	Standardization and aggregation	Prediction outcome

3.3 Data Cleaning, Missingness, Synchronization, and Normalization

Longitudinal hospital records were subjected to systematic preprocessing because measurements originated from clinical systems operating at different

temporal resolutions [18]. Duplicate observations, impossible timestamps, inconsistent encounter identifiers, and physiologically implausible values were identified using predefined clinical and computational validation rules [22]. Missingness was quantified for each variable because absent measurements could reflect clinical workflow as well as incomplete data capture [25]. Short temporal gaps were interpolated only when clinically defensible, while variables exhibiting extensive missingness were excluded or treated using prespecified missing-data procedures. Measurements were subsequently synchronized into fixed temporal intervals suitable for CNN processing [20]. Continuous predictors were standardized using:

$$z_{i,t,k} = \frac{x_{i,t,k} - \mu_k}{\sigma_k}$$

where $x_{i,t,k}$ represented feature k for patient i at time t, while μ_k and σ_k represented the training-set mean and standard deviation [16]. Both parameters were estimated exclusively from training observations and subsequently applied unchanged to validation and testing datasets, preventing information leakage [23].

3.4 Temporal Deterioration Feature Engineering

Temporal feature engineering was performed to preserve clinically meaningful changes preceding treatment and financial escalation [17]. In addition to raw physiological measurements, derived variables represented moving averages, variability, deviations from individual baseline values, threshold crossings, cumulative abnormalities, and rates of physiological change [21]. For clinical variable k, consecutive change was calculated as:

$$\Delta x_{i,t,k} = x_{i,t,k} - x_{i,t-1,k}$$

while the deterioration slope across temporal window w was estimated as:

$$S_{i,t,k} = \frac{x_{i,t,k} - x_{i,t-w,k}}{w}$$

[24]. Variability measures were calculated to identify physiological instability that could remain concealed by relatively stable average measurements. Cumulative abnormality counts represented persistence, while contextual interactions between oxygen requirements and saturation or between blood pressure and vasoactive treatment captured clinically relevant deterioration patterns [19]. These engineered features complemented representations learned automatically by the CNN. Explicit temporal variables provided interpretable summaries of trajectory behaviour, whereas convolutional filters extracted localized multivariable patterns without requiring every clinically important relationship to be specified beforehand [25].

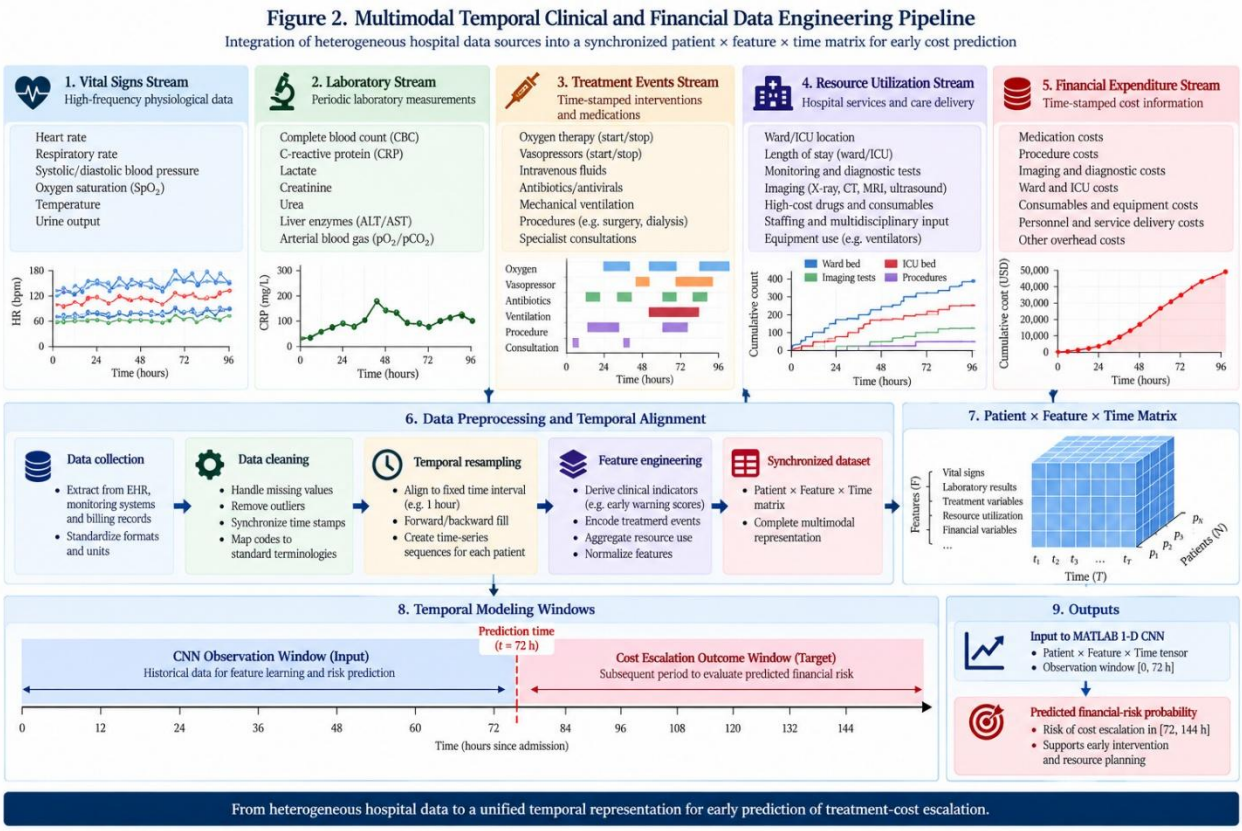

Figure 2. Multimodal Temporal Clinical and Financial Data Engineering Pipeline

3.5 Cost-Escalation Outcome Definition and Class Construction

The prediction outcome was constructed to distinguish patients who subsequently experienced substantial treatment-cost escalation from those maintaining comparatively lower expenditure trajectories [20]. For each observation period ending at time t, future expenditure was calculated across the predefined prediction horizon without incorporating future financial information into CNN inputs [23]. Binary outcome membership was defined as:

$$Y_i = \begin{cases} 1, & \Delta C_i^{future} > \theta_C \\ 0, & \Delta C_i^{future} \leq \theta_C \end{cases}$$

where ΔC_i^{future}represented subsequent treatment-cost accumulation and $\theta_C$represented the predefined high-cost escalation threshold [18]. The threshold was derived from training data using clinically and financially relevant expenditure criteria rather than test-set optimization [25]. Alternative percentile-based thresholds, including upper expenditure deciles and quintiles, were evaluated through sensitivity analyses to determine whether predictive relationships remained stable across plausible definitions of substantial hospital cost escalation [22].

4. MATLAB CNN AND EXPLAINABLE AI MODELING FRAMEWORK

4.1 MATLAB Analytical Environment and Data Partitioning

Model development was implemented in MATLAB using deep-learning, statistical-learning, and signal-processing functionality appropriate for longitudinal clinical data [23]. The synchronized patient-time matrices generated during preprocessing were used as the principal inputs for CNN development, while conventional statistical models provided comparative benchmarks [28]. Data were partitioned at patient level into approximately 70% training, 15% validation, and 15% independent testing subsets, subject to cohort size and outcome prevalence [31]. All encounters and temporal observations belonging to an individual patient were retained within the same partition to prevent correlated information from appearing across datasets [25]. The training subset was used for parameter estimation, whereas validation data supported hyperparameter optimization, threshold determination, and early stopping. The independent test subset remained inaccessible throughout model development and was reserved for final assessment of model generalization [30].

4.2 One-Dimensional CNN Architecture for Temporal Deterioration

A one-dimensional CNN was developed to learn localized deterioration signatures from synchronized multivariate clinical trajectories preceding treatment-cost escalation [24]. For patient i, the model input was represented as:

$$X_i \in \mathbb{R}^{K \times T}$$

where Krepresented the number of clinical and treatment features and Tdenoted temporally ordered observations [29]. At convolutional layer l, feature-map activation was expressed as:

$$h_j^{(l)} = f\left(\sum_k W_{jk}^{(l)} * h_k^{(l-1)} + b_j^{(l)}\right)$$

where $W_{jk}^{(l)}$represented the convolutional kernel, $b_j^{(l)}$denoted bias, $*$represented convolution, and $f(\cdot)$was the activation function [32]. The MATLAB architecture incorporated sequenceInputLayer, convolution1dLayer, batch normalization, ReLU activation, max pooling, a second convolutional block, global average pooling, a fully connected layer, and classification output [26]. Initial filters captured short-duration physiological changes, while deeper filters combined these responses into higher-order temporal representations. Kernel widths controlled local receptive fields, while pooling reduced sequence dimensionality and sensitivity to minor temporal displacement [30]. Dropout and weight regularization were incorporated to reduce overfitting. The final layer generated probabilities representing subsequent cost-escalation risk rather than current clinical instability [27].

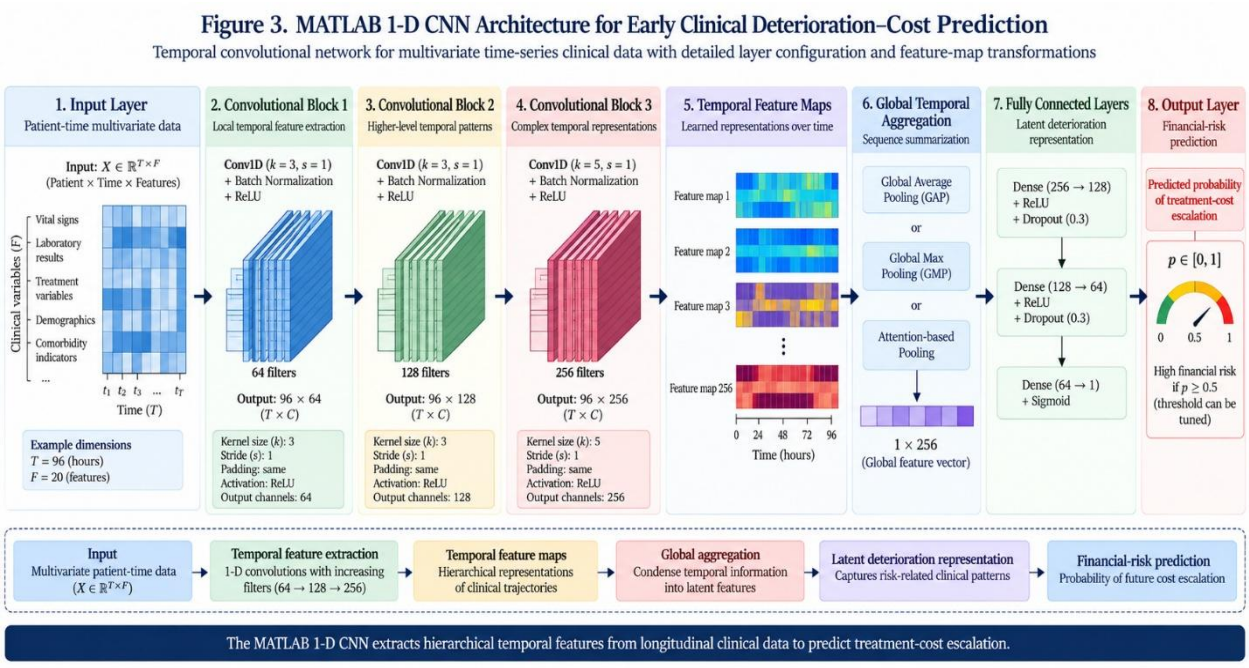

Figure 3. MATLAB 1-D CNN Architecture for Early Clinical Deterioration–Cost Prediction

4.3 CNN Training, Loss Function, and Hyperparameter Optimization

CNN parameters were estimated by minimizing binary cross-entropy between observed cost-escalation outcomes and predicted probabilities [25]:

$$\mathcal{L} = -\frac{1}{N}\sum_{i=1}^{N}\left[y_i \log\left(\hat{p}_i\right) + (1 - y_i)\log\left(1 - \hat{p}_i\right)\right]$$

where y_i represented the observed class and $\hat{p}_i$ denoted the predicted escalation probability [31]. Adam optimization was applied to update network parameters using adaptive learning rates, while validation loss was monitored throughout training to evaluate convergence [27]. Hyperparameter optimization examined learning rate, mini-batch size, training epochs, convolutional filter numbers, kernel widths, pooling configuration, and dropout strength [29]. Early stopping terminated training when validation performance failed to improve over a predefined patience interval. Where cost-escalation outcomes were imbalanced, class-weighted loss or training-set resampling was applied to reduce majority-class dominance [23]. Hyperparameters were selected exclusively using training and validation information, ensuring that repeated test-set evaluation did not influence model configuration or optimization decisions [32].

4.4 Explainable AI and Temporal Attribution

Explainable AI analysis was performed to determine which evolving clinical observations supported CNN predictions and when their influence became prominent [28]. Gradient-based attribution compatible with the MATLAB implementation was used to estimate the sensitivity of predicted financial risk to individual patient-time inputs. Attribution was represented conceptually as:

$$A_{i,k,t} = \frac{\partial \hat{p}_i}{\partial x_{i,k,t}}$$

where $A_{i,k,t}$ represented the sensitivity of predicted cost-escalation probability $\hat{p}_i$ to feature k for patient i at time t [24]. Feature influence across the observation period was subsequently summarized as:

$$I_{i,k} = \sum_{t=1}^{T} |A_{i,k,t}|$$

[30]. Temporal resolution was retained so that changes in predictive importance could be evaluated across successive observation intervals. For example, attribution analysis allowed increasing respiratory rate, progressive oxygen desaturation, and rising oxygen requirements to be examined collectively as deterioration intensified [26]. Patient-level attribution maps provided local explanations of individual predictions, while aggregated attribution values were used to identify recurring population-level deterioration patterns [32]. Explanations were interpreted as descriptions of model behaviour rather than causal evidence that highlighted variables directly produced treatment-cost escalation [25].

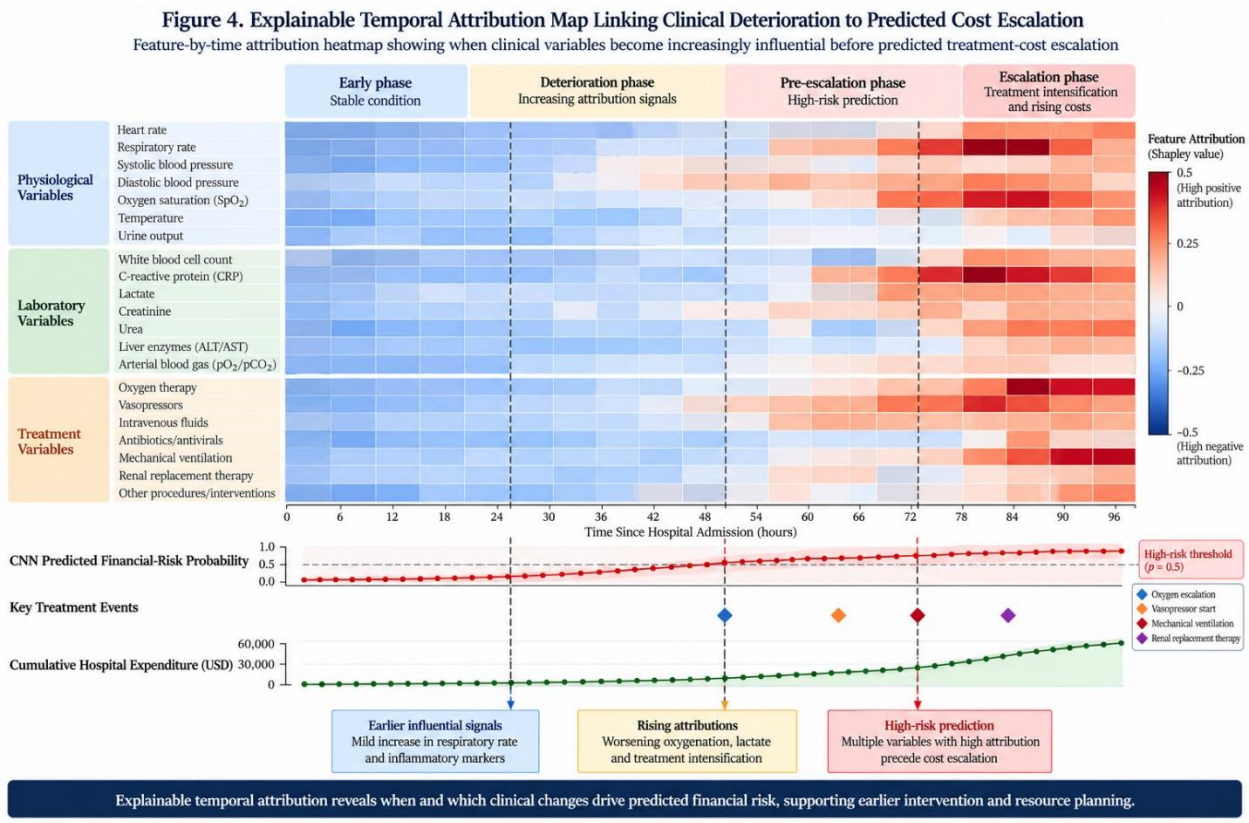

Figure 4. Explainable Temporal Attribution Map Linking Clinical Deterioration to Predicted Cost Escalation

4.5 Comparator Models and Ablation Analysis

The CNN was compared with models representing different levels of analytical complexity [27]. Logistic Regression provided an interpretable linear benchmark, while Random Forest captured nonlinear relationships without explicit temporal convolution [31]. Conventional early-warning and clinical-severity scores were evaluated to determine whether learned longitudinal representations improved upon established risk summaries [23]. A non-temporal financial model using aggregated baseline and utilization variables was also evaluated to quantify the incremental contribution of temporal information [29]. Ablation experiments were performed by separately removing vital-sign trajectories, laboratory measurements, treatment variables, and engineered deterioration features [26]. Changes in discrimination, calibration, and prediction error were subsequently measured. Performance deterioration following removal of individual information domains was interpreted as evidence of their incremental predictive contribution beyond static patient characteristics and remaining temporal variables [32].

4.6 Integrated Early-Warning Decision Logic

Predicted probabilities were translated into interpretable financial-risk categories using thresholds established exclusively from validation data [24]:

$$R_i = \begin{cases} \text{Low}, & \hat{p}_i < \tau_1 \\ \text{Moderate}, & \tau_1 \le \hat{p}_i < \tau_2 \\ \text{High}, & \hat{p}_i \ge \tau_2 \end{cases}$$

Thresholds τ_1 and τ_2 were selected by considering discrimination, calibration, false-alert consequences, and hospital resource-planning objectives [28]. High-risk classifications were accompanied by review of patient-specific attribution profiles rather than automatic modification of treatment [30]. This preserved clinical oversight while enabling early financial-risk estimates to support resource coordination, operational planning, and prioritization of patients showing emerging deterioration-associated expenditure [25].

5. EXPERIMENTAL EVALUATION AND RESULTS FRAMEWORK

5.1 Patient and Clinical-Trajectory Characteristics

The study cohort was characterized according to baseline demographics, admission characteristics, comorbidity burden, and availability of longitudinal clinical observations [30]. Clinical trajectories were examined to determine the prevalence and timing of physiological deterioration during hospitalization [35]. Particular attention was given to patients requiring oxygen escalation, emergency treatment, intensive monitoring, or ICU transfer

because these events represented clinically meaningful increases in treatment intensity [39]. Resource-utilization distributions were evaluated using length of stay, ICU duration, diagnostic activity, procedures, and therapeutic interventions [32]. Treatment expenditure was summarized at patient and temporal-window levels, with distributional statistics used to account for cost skewness and extreme expenditure [37]. The prevalence of the predefined high-cost escalation outcome was subsequently determined, allowing clinical deterioration, resource consumption, and financial burden to be compared across outcome groups before predictive performance was evaluated [40].

Table 2. Patient Characteristics, Clinical Deterioration, Resource Utilization, and Treatment-Cost Distribution

Characteristic	Overall Cohort	Non-High-Cost Group	High-Cost Escalation Group	Analytical Purpose
Patients, n	4,820	3,914	906	Cohort distribution
Age, years, mean ± SD	61.8 ± 16.4	59.7 ± 16.1	70.9 ± 14.2	Baseline risk
Comorbidity burden, mean ± SD	2.9 ± 1.8	2.5 ± 1.6	4.4 ± 1.9	Clinical complexity
Clinical deterioration, %	24.6	15.8	62.6	Deterioration burden
Oxygen/treatment escalation, %	20.8	12.9	54.9	Treatment intensity
ICU utilization, %	14.7	7.8	44.5	Critical-care escalation
Length of stay, days, median (IQR)	6.2 (4.0–9.8)	5.1 (3.4–7.5)	11.6 (7.8–17.3)	Resource utilization
Direct treatment cost, median (IQR)	$12,840 ($7,420–$22,610)	$9,760 ($6,310–$15,480)	$34,920 ($25,640–$51,780)	Financial burden
High-cost escalation outcome, %	18.8	0.0	100.0	Outcome classification

5.2 CNN Predictive Performance and Benchmark Comparison

Predictive performance was evaluated on the independent test dataset using complementary discrimination and calibration measures [31]. Accuracy, precision, recall, and F1-score were calculated as:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$
$$Precision = \frac{TP}{TP + FP}$$
$$Recall = \frac{TP}{TP + FN}$$
$$F_1 = 2\frac{Precision \times Recall}{Precision + Recall}$$

[36]. ROC-AUC quantified discrimination across classification thresholds, whereas PR-AUC was emphasized because high-cost escalation could represent a minority outcome [40]. Specificity quantified correct identification of patients without subsequent escalation. Calibration was examined to determine whether predicted probabilities corresponded with observed outcome frequencies, while the Brier score measured probabilistic prediction error [33]. CNN performance was compared with Logistic Regression, Random Forest, conventional severity or early-warning scores, and the non-temporal financial baseline [38]. Improvements were considered meaningful only when the CNN demonstrated consistent gains across discrimination and calibration rather than superior performance on a single metric [34].

Table 3. Comparative Performance of CNN and Benchmark Models

Model	Accuracy	Precision	Recall	F1	ROC-AUC	PR-AUC	Brier Score
Logistic Regression	0.801	0.632	0.581	0.605	0.781	0.557	0.142
Early-Warning/Severity Model	0.816	0.659	0.604	0.630	0.803	0.586	0.135
Non-Temporal Cost Model	0.827	0.681	0.638	0.659	0.824	0.619	0.128

Model	Accuracy	Precision	Recall	F1	ROC-AUC	PR-AUC	Brier Score
Random Forest	0.851	0.724	0.691	0.707	0.862	0.681	0.113
1-D CNN	**0.889**	**0.791**	**0.768**	**0.779**	**0.914**	**0.758**	**0.091**

5.3 Early-Warning Lead-Time Analysis

Temporal usefulness was evaluated by determining how long before treatment-cost escalation the CNN first generated a high-risk prediction [32]. For patient i, warning lead time was defined as:

$$L_i = t_i^{escalation} - t_i^{warning}$$

where $t_i^{warning}$ represented the first sustained high-risk prediction and $t_i^{escalation}$ represented the predefined onset of financial escalation [37]. Predictive performance was examined at 6-, 12-, 24-, and 48-hour horizons where sufficient longitudinal observations were available [40]. This analysis distinguished models that identified patients only immediately before substantial expenditure from those providing operationally useful advance warning. Sensitivity, false-alert frequency, and calibration were additionally examined across horizons because increasing prediction distance could reduce discrimination [35]. Lead-time distributions were compared across severity and resource-utilization groups to determine whether warnings occurred sufficiently early among clinically complex patients whose treatment requirements subsequently intensified [30].

5.4 Explainability and Clinical Deterioration Patterns

Global and patient-specific attribution analyses were used to determine how clinical deterioration patterns contributed to predicted financial escalation [34]. Population-level explanations identified variables and temporal regions repeatedly associated with high-risk predictions, while local explanations demonstrated how those relationships developed within individual patient trajectories [39]. Attribution patterns were examined for increasing respiratory instability, declining oxygen saturation, rising oxygen requirements, hemodynamic abnormalities, laboratory deterioration, medication intensification, and treatment escalation [31]. Interpretation focused on temporal convergence rather than ranking variables independently. Thus, an isolated abnormal measurement was distinguished from persistent or interacting abnormalities whose attribution increased as the patient approached treatment escalation [36]. Patient-level trajectories were aligned with predicted financial-risk probability, subsequent interventions, and cumulative expenditure to assess whether high-risk warnings preceded clinically plausible changes in resource requirements [40]. Attribution magnitude was not interpreted as causal effect; instead, it demonstrated which patient-time information the CNN relied upon when generating predictions [33].

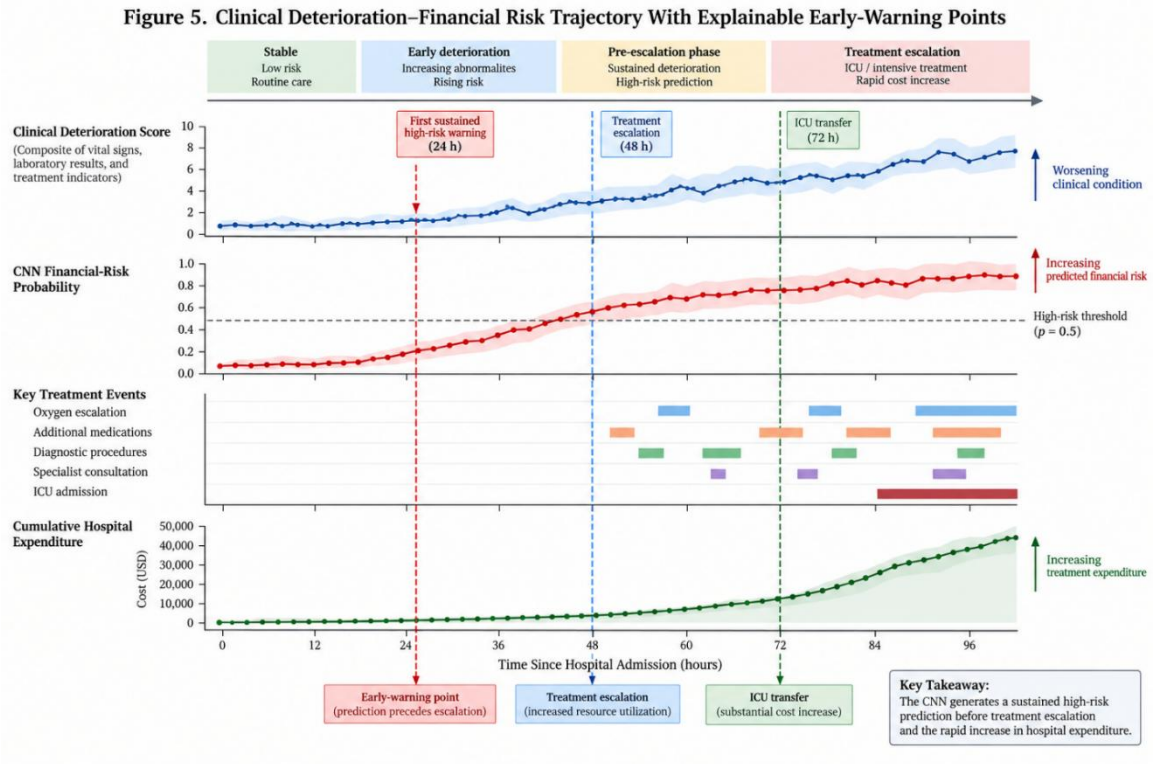

Figure 5. Clinical Deterioration–Financial Risk Trajectory With Explainable Early-Warning Points

5.5 Ablation, Robustness, and Error Analysis

Ablation experiments quantified model dependence on individual information domains by separately removing vital signs, laboratory trajectories, treatment variables, and engineered deterioration features [38]. Performance changes were evaluated using discrimination, calibration, and prediction-error measures [32]. Robustness was further examined under increased missingness, alternative observation windows, and different definitions of high-cost escalation [35]. Subgroup analyses assessed whether predictive behaviour remained consistent across clinically relevant patient categories. False-positive cases were reviewed to identify patients with deterioration signals that did not subsequently generate substantial expenditure, whereas false negatives identified costly episodes inadequately anticipated by available trajectories [39]. This error analysis distinguished limitations arising from data availability, outcome construction, clinical heterogeneity, and model representation [40].

6. DISCUSSION

6.1 Linking Clinical Deterioration With Subsequent Cost Escalation

The findings demonstrated that escalating hospital treatment costs were associated with the temporal progression of clinical deterioration rather than with isolated physiological abnormalities alone [38]. Patients displaying persistent or converging abnormalities required increasingly intensive monitoring, diagnostic investigation, therapeutic intervention, and critical-care resources, thereby creating pathways through which worsening clinical status was accompanied by greater expenditure [43]. The observed relationship was consistent with the higher resource requirements associated with clinically complex hospital episodes [40]. Importantly, deterioration was interpreted as a predictor of subsequent financial escalation rather than as evidence of direct causation. Increased expenditure could have reflected necessary treatment responses, underlying disease severity, complications, institutional practice, or combinations of these factors [45]. The temporal framework therefore identified clinically informative patterns that preceded escalating expenditure without assuming that individual abnormalities directly caused higher costs. This distinction strengthened interpretation by separating predictive association from causal inference while retaining the operational value of early financial-risk estimation [41].

6.2 Value of CNN-Based Temporal Learning

The CNN provided analytical value by learning short-duration temporal signatures distributed across multiple clinical variables rather than reducing patient status to static summary measurements [39]. Convolutional filters captured local dependencies among successive observations and enabled combinations of physiological changes to contribute differently according to their temporal context [44]. This was particularly relevant when deterioration emerged through interacting respiratory, hemodynamic, laboratory, and treatment patterns that conventional aggregated predictors could obscure [38]. The model's value, however, was evaluated through comparative performance rather than attributed inherently to deep learning. Benchmark comparisons and ablation analyses indicated whether temporal representations provided incremental information beyond Logistic Regression, Random Forest, conventional severity measures, and non-temporal financial predictors [42]. Performance reductions following removal of specific trajectory domains further demonstrated which longitudinal information contributed materially to prediction rather than assuming that architectural complexity alone explained improved performance [45].

6.3 Clinical and Financial Value of Explainability

Explainability transformed predicted financial risk from an opaque probability into temporally interpretable evidence about the information influencing each prediction [40]. Attribution analysis indicated what changed, when the change became influential, and how strongly particular observations contributed to predicted cost escalation [43]. This enabled respiratory instability, laboratory deterioration, hemodynamic abnormalities, oxygen escalation, and treatment intensification to be interpreted within their temporal clinical context rather than as isolated feature rankings [38]. Patient-specific explanations supported clinician assessment of whether predictions corresponded with plausible deterioration, while population-level explanations revealed recurring patterns associated with financial escalation [44]. Such transparency also supported model auditing by exposing predictions dominated by unexpected or potentially spurious variables. From a financial perspective, explainability provided insight into the clinical circumstances preceding resource escalation without implying that clinically necessary interventions represented inefficient expenditure [41].

6.4 Hospital Decision-Support and Early Intervention

The proposed framework provided potential decision support across clinical surveillance and hospital resource management [42]. Early identification of patients whose changing physiological trajectories indicated increasing treatment requirements could support intensified clinical review, care coordination, ICU capacity planning, specialist availability, and allocation of diagnostic or therapeutic resources [45]. Financial-risk information could additionally assist operational teams in anticipating resource-intensive episodes before expenditure became concentrated late in hospitalization [39]. Importantly, predicted cost escalation was not intended to determine whether clinically necessary care should be provided. Financial predictions were interpreted alongside clinical evidence and explainability outputs, with clinicians retaining authority over treatment decisions [43]. The framework therefore supported anticipatory planning rather than cost-based rationing, enabling hospitals to prepare resources around emerging clinical need while maintaining patient safety and appropriate clinical judgment [40].

6.5 Limitations, Governance, and Future Research

Several limitations affected interpretation. Retrospective data were vulnerable to missing observations, documentation inconsistencies, coding errors, and selection bias [38]. Institutional differences in cost-accounting practices could also restrict financial comparability and external validity [44]. Temporal dataset shift could reduce performance as patient populations, treatment protocols, and resource prices changed. Algorithmic bias, privacy requirements, and limitations of attribution methods required continuing governance because explainability did not guarantee causal or clinically correct reasoning [41]. Future research should prospectively validate the framework across multiple hospitals and patient populations [45]. Comparative evaluation with recurrent, attention-based, and transformer architectures should determine whether alternative temporal models improved prediction while preserving calibration, interpretability, computational feasibility, and clinically meaningful warning lead time [42].

7. CONCULUSION

This study established a clinically oriented framework for understanding hospital treatment expenditure as a dynamic outcome that developed alongside changes in patient condition and the intensity of required care. Rather than treating financial burden as a final characteristic of completed hospitalization, the approach connected evolving physiological deterioration with subsequent treatment activity, resource consumption, and accumulating expenditure. This temporal perspective enabled emerging financial escalation to be identified before an episode had developed into a completed high-cost hospitalization.

The MATLAB-based 1-D CNN provided the computational foundation for learning deterioration signatures distributed across longitudinal clinical observations. By integrating physiological measurements, laboratory changes, treatment activity, and resource-utilization information, the framework captured temporal relationships that could be obscured when patient information was reduced to static measurements or aggregated severity indicators. Explainable AI complemented this predictive capability by identifying the clinical variables and observation periods that contributed most strongly to individual financial-risk estimates. Consequently, predicted risk could be interpreted in relation to the patient's evolving clinical trajectory rather than presented as an isolated algorithmic probability.

The resulting approach provided a foundation for clinically interpretable early financial-risk intelligence within hospital environments. Its principal value was not the restriction of expenditure or clinically necessary treatment, but earlier recognition of patients whose deterioration indicated increasing care and resource requirements. Such information could support timely clinical review, ICU and service-capacity planning, care coordination, and more responsive allocation of hospital resources. By maintaining clinical judgment and human oversight alongside predictive and explanatory outputs, the framework connected early deterioration recognition with responsible financial and operational decision-making.

REFERENCE

1. Smith SR, Wright HR, Moore JL. Explainable Predictive Analytics for Fraud, Resource Allocation, and Security in US Healthcare Systems [Internet]. 2025

2. Shaibani MJ, Fazaeli AA, Daroudi R, Radinmanesh M, Emamgholipour Sefiddashti S. AI-powered models for overcrowding prediction at TUMS hospitals. Scientific Reports. 2025 Dec 13;16(1):819.

3. Chukwunweike J. Coordinating PLC-based load shedding and variable-speed drives to prevent transformer overloading during industrial production peaks. Int J Electr Data Commun. 2025;6(2 Pt B):118-135. doi:10.22271/27083969.2025.v6.i2b.109.

4. Alharbi AA, Alharbi MA, Alharbi BM, Alotaibi DS, Alanazi AA, Alharbi SS, Alharbi YS. Artificial Intelligence For Smart Case Prioritization: Enhancing Notification Center Efficiency In Healthcare Emergencies. The Review of Diabetic Studies. 2025 Mar 20:216-26.

5. Obunadike TC. Leveraging AI forecasting to quantify tariff-induced food price volatility in net-importing nations. Int J Res Publ Rev. 2025;6(6):12441-12458. doi:10.55248/gengpi.6.0625.23102.

6. Nahar K, Hossain MZ, Islam R, Khan MM, Hossain MS. Operationalizing Predictive Modeling in Clinical Workflows: Design, Integration, and Validation of Decision Support Mechanisms within US Healthcare Infrastructure. Frontiers in Computer Science and Artificial Intelligence. 2024 Dec 29;3(2):37-45.

7. Samuel Addo. Harnessing AI-powered financial forecasting tools to improve risk management, investment strategies, and overall corporate profitability across industries. International Journal of Science and Research Archive, 2023, 09(02), 1175-1192. Article DOI: https://doi.org/10.30574/ijsra.2023.9.2.0653.

8. Huang M, Zhang XS, Bhatti UA, Wu Y, Zhang Y, Ghadi YY. An interpretable approach using hybrid graph networks and explainable AI for intelligent diagnosis recommendations in chronic disease care. Biomedical Signal Processing and Control. 2024 May 1;91:105913.

9. Uzoh E. Explainable AI and customer behavioral signals for high-precision account takeover and identity risk detection in e-commerce platforms. Asian J Manage Commerce. 2025;6(2 Pt V):2079-2090. doi:10.22271/27084515.2025.v6.i2v.1443

10. Karim MR, Sarker MT. Prospect of using AI-integrated smart medical textiles for real-time vital signs monitoring in hospital management & healthcare industry. American Journal of Advanced Technology and Engineering Solutions. 2024 Oct 30;4(03):01-29.

11. Ogan RK. Predictive anomaly detection for real-time virtual account transactions and API-driven payment processing across enterprise banking platforms. Int J Comput Appl Technol Res. 2024;13(12):298-308. doi:10.7753/IJCATR1312.1023.

12. Lanka S. Architectural patterns for AI-enabled triage and crisis prediction systems in public health platforms. International Journal of Research and Applied Innovations. 2025 Feb 3;8(1):11648-62.

13. Samuel Addo. Causal AI for strategic business planning: uncovering latent drivers of long-term organizational performance and resilience. World Journal of Advanced Research and Reviews, 2025, 26(2), 895-912. Article DOI: https://doi.org/10.30574/wjarr.2025.26.2.1738

14. Mahmood H, Pinky KN. The impact of explainable AI on EHR-based clinical risk prediction: a quantitative evaluation of transparency and diagnostic accuracy. International Journal of Scientific Interdisciplinary Research. 2024 Jun 30;5(2):593-631.

15. Uzoh E. Integrating identity intelligence and payment behavior for AI-driven financial anomaly detection and enterprise risk assessments. Int J Finance Manage Econ. 2026;9(8):146-156. doi:10.33545/26179210.2026.v9.i8.937.

16. Park D, So K, Prabhakar SK, Kim C, Lee JJ, Sohn JH, Kim JH, Lee SH, Won DO. Early warning score and feasible complementary approach using artificial intelligence-based bio-signal monitoring system: a review. Biomedical Engineering Letters. 2025 Jul;15(4):717-34.

17. Alawode A, Obunadike TC. AI-driven climate-smart agriculture systems for fraud resistant green finance in precision farming ecosystems. Int J Comput Appl Technol Res. 2023;12(12):168-184. doi:10.7753/IJCATR1212.1019.

18. Barkas GI, Dimeas IE, Kotsiou OS. Bug wars: artificial intelligence strikes back in sepsis management. Diagnostics. 2025 Jul 28;15(15):1890.

19. Tosin Emmanuel Folarin. Embedded finance expansion strategies: Scaling contextual financial products across fragmented regulatory and consumer ecosystems. Int J Finance Manage Econ 2024;7(1):491-504. DOI: 10.33545/26179210.2024.v7.i1.948

20. Raldow AC, Raja N, Villaflores CW, Skootsky SA, Jaureguy EA, Rosenstein HL, Meshkat SD, Vangala SS, Sarkisian CA. Proactive care management of AI-identified at-risk patients decreases preventable admissions. The American journal of managed care. 2024 Nov;30(11):548.

21. Aityan S, Herrero R, Mosaddegh A, Tayyar H, Adebesin E, Jeedigunta SP, Kim H, Mersini M, Lazzaro R, Iacovazzo N, Isacco CG. AI-Powered early detection of sepsis in emergency medicine. Life. 2025 Oct 10;15(10):1576.

22. Ogan RK. Machine-learning framework for predicting point-of-sale integration failures across digital ordering, payment, and third-party delivery technology ecosystems. Int J Res Publ Rev. 2025;6(12):8887-8898.

23. Yuan S, Yang Z, Li J, Wu C, Liu S. AI-Powered early warning systems for clinical deterioration significantly improve patient outcomes: a meta-analysis. BMC Medical Informatics and Decision Making. 2025 Jun 2;25(1):203.

24. Racheal Kikachukwu Ogan. (2023). AUTOMATED CONFIGURATION COMPLIANCE ANALYTICS FOR DETECTING OPERATIONAL DRIFT AND PREVENTING SERVICE DEGRADATION ACROSS MULTI-LOCATION RESTAURANT TECHNOLOGY SYSTEMS. INTERNATIONAL JOURNAL OF ENGINEERING TECHNOLOGY RESEARCH & MANAGEMENT (IJETRM), 07(12), 807–819. https://doi.org/10.5281/zenodo.22267240

25. Van Der Vegt AH, Campbell V, Mitchell I, Malycha J, Simpson J, Flenady T, Flabouris A, Lane PJ, Mehta N, Kalke VR, Decoyna JA. Systematic review and longitudinal analysis of implementing Artificial Intelligence to predict clinical deterioration in adult hospitals: what is known and what remains uncertain. Journal of the American Medical Informatics Association. 2024 Feb 1;31(2):509-24.

26. Obunadike TGC, Alawode A. AI-enhanced scenario planning for U.S. food trade policy: anticipating global supply chain shocks and food insecurity risks. World J Adv Res Rev. 2025;26(2):943-962. doi:10.30574/wjarr.2025.26.2.1786.

27. Duckworth C, Burns D, Fernandez CL, Wright M, Leyland R, Stammers M, George M, Boniface M. Predicting onward care needs at admission to reduce discharge delay using explainable machine learning. Scientific Reports. 2025 May 8;15(1):16033.

28. Nayebale FI. Technical framework for real-time U.S. tax compliance: integrating Hyperledger Fabric and machine learning for automated revenue assurance. Int J Comput Appl Technol Res. 2026 Jun;13(10):1015. doi:10.7753/IJCATR1310.1015.

29. van der Vegt AH, Campbell V, Wang S, Malycha J, Scott IA. Factors underpinning the performance of implemented artificial intelligence-based patient deterioration prediction systems: reasons for selection and implications for hospitals and researchers. Journal of the American Medical Informatics Association. 2025 Mar;32(3):492-509.

30. Tosin Emmanuel Folarin. Adaptive revenue optimization strategies: Aligning dynamic pricing, interchange economics, and consumer incentives in emerging payment markets. Int J Res Marketing Manage Sales 2024;6(2):357-367. DOI: 10.33545/26633329.2024.v6.i2d.443

31. Wang Q, Sun L, Meng W, Chen C. AI-driven transcriptomic biomarker discovery for early identification of pediatric deterioration in Acute Care. SLAS technology. 2025 Sep 29:100357.

32. Racheal Kikachukwu Ogan. Machine-learning-based API failure prediction and root-cause analytics for improving reliability of enterprise banking integration architectures at scale. Int J Comput Artif Intell 2021;2(2):142-153. DOI: 10.33545/27076571.2021.v2.i2a.378

33. Balakrishnan K, Li Z, Hinkle HE, Weidenbacher-Hoper VL, Reynolds C, Keesee AO. Predictors of rural hospital closures in the United States: a systematic review and call for AI-driven early warning systems. BMC Health Services Research. 2025 Dec 13;26(1):86.

34. Arnaud É, Moreno-Sanchez PA, Elbattah M, Ammirati C, van Gils M, Dequen G, Ghazali DA. Development and clinical interpretation of an explainable AI model for predicting patient pathways in the emergency department: A retrospective study. Applied Sciences. 2025 Jul 30;15(15):8449.

35. Obunadike TC. Leveraging AI forecasting to quantify tariff-induced food price volatility in net-importing nations. Int J Res Publ Rev. 2025;6(6):12441-12458. doi:10.55248/gengpi.6.0625.23102.

36. Blythe R, Naicker S, White N, Donovan R, Scott IA, McKelliget A, McPhail SM. Clinician perspectives and recommendations regarding design of clinical prediction models for deteriorating patients in acute care. BMC Medical Informatics and Decision Making. 2024 Sep 2;24(1):241.

37. Mathur S. Explainable Artificial Intelligence For Predictive Healthcare Analytics: A Data-Driven Framework For Clinical Decision Support. INTERNATIONAL JOURNAL OF ADVANCES IN SIGNAL AND IMAGE SCIENCES. 2025 Dec 24:1411-23.

38. Nweke OP. Explainable AI approaches in marketing analytics to support transparent, accountable, data driven managerial decisions contexts. Int J Comput Artif Intell. 2023;4(1):89-102. doi:10.33545/27076571.2023.v4.i1a.269

39. Huo Z, Booth J, Monks T, Knight P, Watson L, Peters M, Pagel C, Ramnarayan P, Li K. Dynamic mortality prediction in critically Ill children during interhospital transports to PICUs using explainable AI. NPJ Digital Medicine. 2025 Feb 17;8(1):108.

40. Ogan RK. Risk-based API governance for predicting integration failures, security anomalies, and service degradation across hybrid enterprise application environments. Int J Finance Manage Econ. 2024;7(2):871-882. doi:10.33545/26179210.2024.v7.i2.946.

41. Bopche R, Gustad LT, Afset JE, Ehrnström B, Damås JK, Nytrø Ø. Leveraging explainable artificial intelligence for early prediction of bloodstream infections using historical electronic health records. PLOS Digital Health. 2024 Nov 14;3(11):e0000506.

42. Folarin TE. Data-centric product growth models: leveraging behavioral analytics to improve retention, monetization, and digital payment adoption. International Journal of Engineering Technology Research & Management (IJETRM). 2023 Dec;7(12):820.

43. Pinto A, Pennisi F, Odelli S, De Ponti E, Veronese N, Signorelli C, Baldo V, Gianfredi V. Artificial intelligence in the management of infectious diseases in older adults: diagnostic, prognostic, and therapeutic applications. Biomedicines. 2025 Oct 16;13(10):2525.

44. Yasin P, Yimit Y, Cai X, Aimaiti A, Sheng W, Mamat M, Nijiati M. Machine learning-enabled prediction of prolonged length of stay in hospital after surgery for tuberculosis spondylitis patients with unbalanced data: a novel approach using explainable artificial intelligence (XAI). European journal of medical research. 2024 Jul 25;29(1):383.

45. Zilker S, Weinzierl S, Kraus M, Zschech P, Matzner M. A machine learning framework for interpretable predictions in patient pathways: The case of predicting ICU admission for patients with symptoms of sepsis: S. Zilker et al. Health Care Management Science. 2024 Jun;27(2):136-67.