

ORIGINAL RESEARCH

Open access

Interpretable Machine Learning Model for Detecting Clinical Pathway Deviations in Hospitalized Patients Using Order Sequences, Vital Sign Trends, Laboratory Monitoring Frequency, and Provider Decision Patterns

Siti Rahman^{1*}, Ahmad Zaki¹, Nurul Huda², Amir Faisal¹, Lim Wei²

Abstract

Clinical pathways are designed to standardize inpatient care for common conditions while allowing clinically justified individualization. Deviations from these pathways are frequent and may reflect either appropriate adaptation to patient complexity or potentially harmful departure from evidence-informed practice. Current deviation detection often depends on retrospective audit, static compliance rules, or aggregate dashboards. These approaches can miss subtle temporal drift in care delivery and rarely explain why a specific patient trajectory diverged from the expected pathway. This article proposes an interpretable machine learning model for detecting clinical pathway deviations in hospitalized patients. The model focuses on order sequences, vital sign trends, laboratory monitoring frequency, and provider decision patterns as dynamic indicators of care-process variation. Conceptually, the model would compare each patient's evolving care trajectory with learned expected pathways using sequence-comparison and outlier-detection logic. SHAP-based or attention-informed explanations would identify the specific features responsible for a deviation flag, such as delayed monitoring, omitted follow-up testing, or unusual ordering behavior. The proposed model could detect when a patient's care trajectory diverges from an expected pathway and provide a transparent rationale for review. For example, it could flag a missing repeat troponin, a delayed antibiotic escalation, or a laboratory monitoring pattern inconsistent with the patient's clinical state. An interpretable pathway-deviation model could shift quality monitoring from manual, sample-based review toward continuous and transparent pathway surveillance. Such a system would support real-time clinical awareness, structured audit, and organizational learning.

Keywords Electronic health records, Process mining, Explainable artificial intelligence, SHAP, Clinical pathways, Pathway deviation

*Correspondence:

Siti Rahman

siti.rahman@outlook.com

¹ Department of Intelligent Healthcare Informatics, Faculty of Medicine, University of Malaya, Kuala Lumpur, Malaysia

² Department of Clinical Data Systems Engineering, Faculty of Engineering, Universiti Teknologi Malaysia, Johor Bahru, Malaysia

Introduction

Clinical pathways are intended to reduce unwarranted variation by translating evidence-informed recommendations into structured sequences of assessment, monitoring, treatment, and reassessment. In hospital settings, however, the pathway specified in a guideline or order set often differs from the care actually delivered at the bedside, because clinical complexity, workload, provider preference, and local routines shape real-world practice. Process-mining studies have

shown that care pathways reconstructed from electronic health records frequently contain variants that depart from nominal workflows, creating an opportunity to distinguish meaningful personalization from avoidable non-adherence [1-6]. For hospitalized patients, the challenge is therefore not merely whether a pathway exists, but whether the evolving sequence of care remains clinically coherent over time [7, 8].

Traditional deviation detection relies heavily on manual chart review, retrospective quality audit, and compliance dashboards that report whether selected actions occurred. These tools are useful for accountability, but they often treat pathway adherence as a static checklist rather than a temporal process involving order timing, monitoring cadence, response to abnormal findings, and provider decision patterns. Studies of clinical workflow and electronic order behavior suggest that deviations may emerge from misalignment between order sets and clinician workflow, from local practice variants, or from incomplete visibility into process context [9-11]. Because many current dashboards do not explain the source of deviation, they can identify that a care element was missed without clarifying why the patient's trajectory became atypical [12, 13].

Timestamped electronic health record data create a richer basis for modeling expected inpatient care trajectories. Computerized provider order entry logs, medication orders, laboratory testing events, vital-sign measurements, and flowsheet documentation can be represented as temporal event streams rather than isolated variables. Sequential modeling, temporal embeddings, and anomaly-detection methods have been applied to clinical event sequences and electronic health record data, suggesting that deviations can be detected as departures from learned patterns of care rather than from simple rule violations [14-20]. This makes it possible to conceptualize pathway adherence as a dynamic alignment problem between the observed patient-day and an expected trajectory for comparable patients [21, 22].

The central thesis of this article is that an interpretable machine learning model could identify patient-level clinical pathway deviations while also explaining the specific care-process features that produced the flag. Rather than acting as a black-box alert, the model would provide a structured rationale, such as delayed laboratory reassessment, unexpected medication sequencing, or reduced monitoring despite physiologic instability. Explainable artificial intelligence methods, including SHAP, local explanation frameworks, and clinician-centered explanation design, are particularly relevant because deviation alerts must be clinically contestable and actionable. In this form, the model would support not only real-time corrective action but also organizational learning about how pathways succeed, drift, or require redesign.

Background

Clinical pathways and their impact on hospital outcomes

Clinical pathways aim to standardize evidence-informed care across diagnoses, services, and care settings while preserving the flexibility needed for patient-specific adaptation. Inpatient pathway adherence can influence quality through timely diagnostics, consistent monitoring, coordinated medication use, and escalation when physiologic deterioration is observed. However, pathway implementation studies and process-mining reviews show that real-world care commonly includes multiple variants, some reflecting legitimate clinical complexity and others reflecting avoidable process drift [2, 5, 6]. Therefore, a pathway-deviation model should not treat every departure as error, but should instead support structured review of whether the deviation is clinically justified [8, 22].

Order sequences and temporal care patterns

Order sequences provide a direct record of how inpatient care unfolds, including diagnostic testing, medication initiation, monitoring orders, imaging, procedures, and consults. Sequence-mining and process-mining methods can reconstruct these event logs to reveal frequent care paths, atypical variants, and misalignment between expected and observed workflows [1-4, 16]. Recent work on medical-order sequence variants and clinical event prediction shows that order timing and event context can be modeled as temporal patterns rather than isolated actions [16-19]. For pathway-deviation

detection, this means that the model should evaluate whether the order occurred, when it occurred, and whether its position in the care sequence was clinically consistent with the expected trajectory [15, 20].

Vital sign and laboratory monitoring as pathway markers

Vital sign trends and laboratory monitoring frequency are important markers of whether care intensity matches patient risk. A patient with worsening physiology may require more frequent monitoring, repeat laboratory assessment, or escalation of therapy, while a stable patient may appropriately transition to lower-intensity observation. Because EHR-based event streams include flowsheet measurements and laboratory orders, monitoring frequency can be represented as part of the pathway rather than as background documentation [14, 15, 17]. In an interpretable deviation model, omitted scheduled tests, delayed repeat measurements, and failure to intensify monitoring after abnormal results would be treated as care-process signals requiring clinical explanation [1, 21].

Provider decision patterns and variation

Provider decision patterns influence pathway adherence because clinicians differ in ordering habits, response to alerts, tolerance for uncertainty, documentation practices, and willingness to follow standardized order sets. Research on pharmacy order interventions, medication ordering errors, and workflow-order set alignment shows that provider actions embedded in the EHR can reveal recurring patterns of variation, some of which may signal opportunities for safer system design [9-11, 20]. These patterns should be interpreted cautiously because variation may reflect expertise, patient complexity, specialty norms, or local constraints rather than poor care. A pathway-deviation model should therefore compare providers with expected pathway logic, peer practice, and their own historical baseline before presenting a deviation as clinically meaningful [9, 10].

Interpretable machine learning for healthcare quality

Interpretable machine learning is essential for clinical quality applications because clinicians and quality teams must understand why a system identifies a process as atypical. SHAP-based explanations, global feature-importance summaries, local explanation displays, and clinician-centered explanation frameworks can translate model output into audit-relevant statements about order timing, monitoring gaps, or unusual event combinations [23-27]. In pathway analytics, interpretability also helps distinguish plausible personalization from unjustified drift by exposing the specific features driving the flag. This supports a governance model in which artificial intelligence augments clinical judgment rather than replacing it [24, 26, 27].

Model Development Overview

High-level deviation detection pipeline

The proposed pipeline would construct a longitudinal care profile for each admission or patient-day using timestamped orders, vital signs, laboratory events, medications, and provider actions. This profile would be compared with a learned expected pathway for clinically similar patients, using sequence-comparison or outlier-detection logic to identify departures in timing, ordering, monitoring, and escalation. Prior work on clinical pathway discovery, electronic health record anomaly detection, and tensor-based sequence modeling supports the idea that atypical care patterns can be learned from high-dimensional event streams without reducing the pathway to a static checklist [7, 14, 15]. The model output would be an interpretable deviation assessment accompanied by a rationale that identifies which care-process elements most strongly contributed to the flag [23, 25].

Figure 1 illustrates the proposed interpretable machine learning architecture for transforming timestamped inpatient care events into clinically reviewable pathway-deviation signals and quality-improvement outputs.

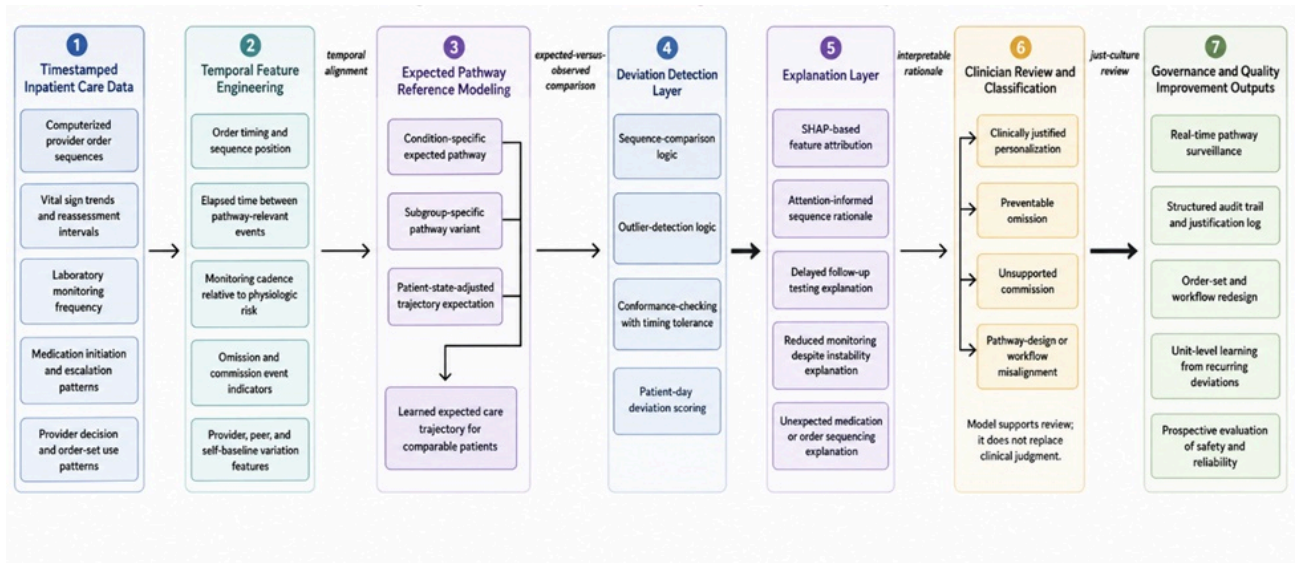


Figure 1. Interpretable Machine Learning Framework for Detecting Clinical Pathway Deviations in Hospitalized Patients.

Core input features

Core input features would include embedded order sequences, timing between key orders, medication administration patterns, vital sign trend summaries, laboratory monitoring cadence, provider-level decision features, and patient context. The order-sequence component would represent the type, timing, and relative position of events, while the monitoring component would capture whether vital signs and laboratory tests were repeated at a cadence consistent with the patient's condition. Personalized event-prediction approaches and medication-order prediction studies suggest that such features can encode both general pathway structure and patient-specific trajectory information [17–20]. Provider decision features would add information about ordering behavior, response to abnormal findings, and local practice variation, while remaining subject to clinical review rather than automatic judgment [9–11].

Design principles

The model should be explainable by default, temporally aware, patient-specific, and sensitive to both omissions and commissions in care delivery. It should not simply mark every nonstandard pathway as incorrect, because clinically appropriate variation is common in hospitalized patients and may be necessary for safety. Process-mining and conformance-checking studies emphasize the importance of representing real clinical pathways as flexible, variant-rich processes rather than rigid templates [1, 5, 6, 21]. Accordingly, the model should integrate interpretable machine learning with quality-improvement workflows so that deviations become prompts for review, learning, and pathway refinement [13, 27, 28].

Data Sources and Feature Engineering

Order sequence extraction and encoding

Order sequence extraction would begin with timestamped events from computerized provider order entry systems, medication administration records, laboratory information systems, imaging orders, consult orders, and nursing documentation. Similar orders could be grouped into clinically meaningful categories to reduce sparsity while preserving distinctions relevant to pathway adherence, such as diagnostic reassessment, therapeutic escalation, and monitoring intensity. Published work on medical-order sequence variants, inpatient medication-order prediction, and clinical event-sequence modeling supports the conceptual use of embeddings or sequence representations to characterize care

trajectories [16–20]. These representations would allow the model to evaluate not only the presence of an order but also its timing and relationship to preceding clinical events [15, 17].

Vital sign and laboratory monitoring features

Vital sign and laboratory features would be engineered to capture trend direction, time since last measurement, monitoring frequency, abnormality persistence, and concordance between physiologic risk and clinical response. For example, the model could represent whether a worsening vital sign was followed by repeat assessment, laboratory reassessment, escalation orders, or documented stabilization. Electronic health record anomaly-detection methods and pathway-preprocessing studies suggest that such temporally structured features can help identify unexpected care trajectories while accounting for noisy hospital data [14, 15, 29]. Feature construction should preserve interpretability by using clinically meaningful concepts such as “delayed repeat measurement” or “monitoring frequency lower than expected for risk state,” rather than opaque latent variables alone [23, 25].

Provider decision and variation features

Provider decision features would summarize ordering preferences, response time to abnormal results, use of order sets, medication changes, consult patterns, and deviation from peer or self-baseline behavior. These features are relevant because pathway adherence is partly shaped by how clinicians interpret information and act within the EHR workflow. Studies of provider action data, medication ordering errors, and order-set workflow alignment show that EHR action traces can reveal both safety-relevant variation and system-design issues [9–11]. In the proposed model, provider variation would be used to contextualize a deviation rather than assign blame, ensuring that explanations support just-culture review and pathway improvement [26, 27].

Interpretable Model Architecture

Model choice and interpretability strategy

A practical architecture could use gradient-boosted decision trees with SHAP explanations for tabular pathway features, while sequence-derived features could be generated from temporal embeddings or process-mining representations. This approach would balance model flexibility with transparency, allowing global explanations for common deviation drivers and local explanations for individual patient-days. SHAP-based methods have been widely discussed as a way to move from local feature attribution toward broader understanding of model behavior, while clinician-centered explanation frameworks emphasize that interpretability must be aligned with clinical decision context [23, 26, 27]. Alternative sequence models with attention visualization could also be considered, but their explanations should be evaluated carefully to ensure that they correspond to clinically meaningful pathway elements [24, 25].

Input feature vector and preprocessing

The input feature vector would convert temporal care data into interpretable components such as counts of order categories, elapsed time between events, monitoring cadence, trend direction, and deviation from expected sequence position. Missingness would be handled as a potentially informative signal when clinically appropriate, because absence of laboratory reassessment or vital sign documentation can itself reflect a pathway-relevant omission. Work on preprocessing patient pathways and modeling uncertain clinical records highlights the importance of handling noisy, incomplete, and variably ordered hospital data before applying conformance or anomaly-detection logic [21, 29]. Feature engineering should therefore preserve the clinical meaning of missing, delayed, repeated, or unusually sequenced events instead of treating them only as statistical artifacts [7, 12].

Output: deviation assessment and explanation

The model output would be a patient-level or patient-day-level deviation assessment that indicates whether the current trajectory appears inconsistent with the expected pathway for comparable clinical context. The explanation layer would identify the leading contributors to that assessment, such as delayed follow-up testing, unexpected medication sequencing, reduced monitoring despite physiologic instability, or provider-ordering patterns that differ from the learned pathway. Visualization tools for clinical pathways and explanation-display research suggest that such output should be presented in a format that supports rapid review, drill-down, and auditability rather than overwhelming clinicians with raw event logs [12, 13, 27]. In this way, the model would transform pathway-deviation detection from a retrospective compliance exercise into an interpretable, real-time quality-improvement signal [6, 23].

Identifying and Characterizing Deviations

Learning expected pathways from historical data

Expected pathways would be learned from historical care trajectories that represent clinically coherent management for defined inpatient conditions. The model could use clustering, sequence mining, or process-discovery methods to identify common order patterns, monitoring rhythms, medication sequences, and escalation points across comparable patient groups [2, 4, 7, 22]. These learned trajectories should be interpreted as empirical representations of usual care rather than as definitive standards, because observed historical practice may contain embedded variation, omissions, or institutional bias [5, 29]. Therefore, expected pathways should be reviewed against clinical guidelines, quality standards, and expert judgment before being used as reference trajectories for deviation detection [6, 28].

Detecting omission and commission deviations

The model should distinguish omission deviations, where an expected action is absent or delayed, from commission deviations, where an unexpected order, medication, or monitoring pattern appears in the trajectory. Omission may include missing repeat laboratory assessment, lack of follow-up vital sign monitoring, or failure to escalate after abnormal results, while commission may include unnecessary testing, unsupported medication changes, or care steps that do not align with the patient's pathway stage. Sequence-based anomaly detection and medical-order prediction studies support the conceptual detection of both absent expected events and unexpected inserted events in clinical timelines [14-16, 20]. This distinction matters because omission and commission deviations imply different quality-improvement responses, such as reminder design, pathway clarification, order-set revision, or clinician education [9-11].

Table 1 classifies pathway deviations into clinically interpretable categories that distinguish potentially preventable care-process problems from justified personalization, documentation artifacts, and pathway-design limitations.

Table 1. Analytic Taxonomy of Clinical Pathway Deviations Detectable by the Proposed Model

Deviation category	Operational definition	Core data signals	Example pathway-deviation pattern	Interpretive question for clinical review	Likely quality-improvement response
Omission deviation	Expected care action is absent or delayed relative to pathway stage	Missing repeat orders, delayed labs, absent monitoring, no escalation order	No repeat troponin after abnormal initial result	Was the omitted step clinically unnecessary, forgotten, delayed, or undocumented?	Reminder redesign, pathway clarification, escalation protocol review
Commission deviation	Unexpected care action occurs outside the	Extra testing, unsupported medication	Broad-spectrum antibiotic escalation	Was the additional action justified by undocumented	Order-set refinement, stewardship

	expected sequence or clinical state	change, atypical consult or imaging order	despite improving clinical status	concern or local practice variation?	review, decision-support tuning
Timing deviation	Correct action occurs but at a clinically atypical time	Prolonged interval between abnormal finding and response	Delayed laboratory reassessment after worsening vital signs	Did workflow, staffing, handoff, or clinical uncertainty explain the delay?	Handoff redesign, alert timing adjustment, staffing review
Monitoring-intensity deviation	Monitoring cadence does not match patient risk state	Vital sign frequency, lab frequency, observation interval	Reduced monitoring despite physiologic instability	Was the patient actually stable, or was deterioration under-recognized?	Nursing workflow support, monitoring protocol adjustment
Sequence-position deviation	Care action appears in an unusual order relative to expected pathway	Event-order embeddings, order sequence comparison	Medication change before confirmatory diagnostic reassessment	Was the sequence clinically intentional or inconsistent with pathway logic?	Pathway education, order-sequence guidance
Provider-pattern deviation	Ordering or response pattern differs from peer, service, or self-baseline	Provider ordering history, order-set use, response latency	Repeated bypassing of standard order set	Does variation reflect expertise, specialty norms, workaround behavior, or unsafe drift?	Just-culture review, order-set usability improvement
Documentation-sensitive deviation	Apparent deviation may reflect incomplete or delayed documentation	Missing notes, delayed event timestamps, sparse flowsheet data	No documented reassessment before therapy change	Is the deviation real or a documentation artifact?	Documentation workflow review, data-quality improvement
Pathway-design deviation	Recurrent deviations suggest the pathway itself may not fit real-world care	Repeated flags across patients, units, or subgroups	Frequent justified deviations in complex comorbidity subgroup	Does the pathway require subgroup-specific adaptation?	Pathway redesign, subgroup-specific protocol development

Temporal and contextual awareness

A pathway-deviation model should account for timing, clinical context, and patient state, because the same care action may be appropriate at one stage of hospitalization and inappropriate at another. For example, reduced monitoring may be expected during recovery but concerning during active deterioration, while delayed medication escalation may be appropriate if laboratory results or vital signs show improvement. Temporal event models and personalized clinical prediction approaches suggest that care trajectories should be interpreted in relation to prior events and evolving patient condition rather than as isolated snapshots [17-19]. Conformance-checking methods that tolerate uncertainty and reordering are particularly relevant because real clinical records often contain delayed documentation, overlapping workflows, and ambiguous event timing [1, 21].

Subgroup-specific pathways

Subgroup-specific pathways are necessary because expected care may differ by comorbidity, age, admission type, service line, disease severity, and baseline physiologic status. A model that uses a single generic pathway may falsely flag appropriate individualization as deviation, especially for complex hospitalized patients whose care requires modified monitoring or treatment sequences. Healthcare pathway discovery and interpretable prediction frameworks support the use of patient-specific or subgroup-aware models that compare each trajectory with clinically relevant peers rather than with an undifferentiated average [8, 22]. Such subgroup logic should remain transparent so that clinicians can see whether the model judged the patient against a medical, surgical, intensive care, or condition-specific reference pattern [12, 13].

Explainability for Clinical Audit and Action

Global explanations for quality improvement teams

Global explanations would summarize the most common drivers of pathway deviation across a unit, service, diagnosis group, or time period. SHAP summary displays or similar interpretable outputs could show whether deviation flags are most often driven by delayed laboratory reassessment, inconsistent vital sign monitoring, unusual medication sequencing, or provider-specific ordering variation [23–25]. These explanations would help quality-improvement teams identify whether variation reflects knowledge gaps, workflow friction, order-set misalignment, staffing constraints, or pathway design problems [11, 28]. The purpose would be to guide targeted improvement rather than produce punitive rankings of clinicians or teams [26, 27].

Local explanations for the attending clinician

Local explanations would translate an individual patient's deviation flag into a clinically interpretable statement that can be reviewed during rounds or handoff. For example, the model could indicate that the patient's pathway appears atypical because follow-up testing has not occurred after an abnormal result, or because monitoring frequency is lower than expected for the current risk state. Local explanation methods are especially important in healthcare because clinicians must be able to contest, contextualize, or override a model-generated flag using bedside knowledge [23, 26, 27]. In this setting, the explanation should identify the relevant order, vital sign, laboratory, or provider-action feature rather than simply displaying a generic risk label [12, 13].

Counterfactual suggestions for pathway realignment

Counterfactual explanations could describe which care-process action would most plausibly move the patient's trajectory closer to the expected pathway. Rather than only indicating that a trajectory is atypical, the model might suggest that repeat laboratory monitoring, medication reassessment, escalation documentation, or renewed vital sign review would reduce the deviation signal if clinically appropriate. Such feedback should be framed as a prompt for review rather than a directive, because not all deviations require correction and some represent intentional individualized care [6, 8]. Explanation-display research and clinician-centered XAI work emphasize that actionability, clarity, and alignment with clinical workflow are central to whether explanations support real decisions [26, 27].

Audit trail and justification logs

Every deviation flag and its explanation should be stored in an audit trail that records the patient context, triggering features, clinician response, and any documented justification for accepting or dismissing the alert. This log would allow retrospective peer review, pathway refinement, and analysis of whether repeated deviations reflect unsafe drift, appropriate personalization, or flaws in the pathway definition. Process-mining studies show that care delivery can be understood as a set of variants rather than a single linear pathway, making audit trails useful for learning how real

workflows diverge from designed processes [1-6]. A just-culture approach is essential so that logged deviations become material for system learning rather than blame assignment [26, 28].

Integration into Clinical Governance and Improvement

Real-time pathway adherence dashboard

A real-time pathway adherence dashboard could display active deviation flags at the patient, unit, and service levels, with each flag paired with a concise explanation. The dashboard should allow users to distinguish high-priority care-process concerns from lower-priority variation, while preserving the ability to inspect the underlying event timeline. Visualization research for clinical pathways and EHR-based displays supports the need for interfaces that make complex temporal care patterns understandable without forcing clinicians to inspect raw logs [12, 14]. For governance, the dashboard should be available to appropriate clinical leaders, charge nurses, quality teams, and attending clinicians so that urgent deviations can be reviewed in context [6, 28].

Feeding deviation data into continuous quality improvement

Aggregated deviation data could inform morbidity and mortality conferences, pathway update committees, resident education, nursing workflow redesign, and order-set maintenance. Recurrent deviation patterns may indicate that the pathway is poorly aligned with real workflow, that clinicians require clearer decision support, or that certain patient subgroups need a modified pathway. Studies of provider ordering behavior, order-set misalignment, and process mining in healthcare suggest that EHR event data can reveal system-level opportunities for redesign when interpreted carefully [4, 9-11]. In this way, the model would function not only as a detection tool but also as part of a learning health system that continuously refines pathway expectations [5, 6, 28].

Table 2 presents a governance framework for ensuring that interpretable pathway-deviation alerts function as reviewable quality-improvement signals rather than autonomous judgments of clinical correctness.

Table 2. Governance Framework for Translating Deviation Alerts into Clinical Action and Organizational Learning

Governance component	Primary purpose	Required model output	Human reviewer role	Safety safeguard	Evidence of successful implementation
Real-time patient-level review	Identify active pathway deviations requiring attention	Patient-day deviation score with local explanation	Determine whether bedside action, documentation, or dismissal is appropriate	Alert must be contestable and reviewable against the chart	Timely review of high-priority deviations without alert fatigue
Local explanation display	Make each deviation clinically understandable	Top contributing features and pathway-stage context	Interpret whether the driver is clinically meaningful	Explanation must use care-process language, not opaque model terms	Clinicians can explain why a flag appeared
Justification log	Preserve clinician reasoning and audit trail	Flag reason, timestamp, response, and documented rationale	Record whether deviation was justified, corrected, or deferred	Logs used for learning, not blame	Consistent documentation of accepted and dismissed alerts

Global deviation summary	Identify recurring system-level patterns	Aggregated SHAP drivers or grouped deviation categories	Quality team interprets unit, service, or diagnosis-level patterns	Avoid punitive provider ranking	Recurrent causes linked to workflow or pathway redesign
Pathway refinement committee	Update pathway definitions using observed variation	Recurrent justified and unjustified deviation patterns	Decide whether pathway logic requires revision	Historical practice must not automatically define best practice	Updated pathways reflect both evidence and real-world workflow
Bias and subgroup review	Prevent unfair or misleading pathway expectations	Subgroup-specific deviation rates and explanation patterns	Assess whether certain groups are over-flagged	Compare model outputs across age, comorbidity, service, and severity strata	Reduced inappropriate flagging of complex or underrepresented subgroups
Prospective validation	Test clinical usefulness before full deployment	Detection accuracy, explanation quality, actionability, and outcome measures	Expert panel adjudicates flagged and unflagged cases	No autonomous clinical judgment or punitive use	Demonstrated improvement in review quality and preventable omission detection
Continuous quality-improvement integration	Convert deviation data into organizational learning	Trends across time, units, order sets, and pathway stages	Translate findings into training, workflow, or order-set changes	Maintain just-culture framing	Deviation insights lead to measurable process improvements

Evaluation Strategy

Deviation detection accuracy

Evaluation should compare model-identified deviations with expert-adjudicated pathway reviews, while recognizing that the gold standard is partly interpretive. Reviewers could assess whether each flagged trajectory reflects an unjustified departure, an appropriate individualized decision, or an artifact of incomplete documentation. Comparisons with rule-based compliance checks would help determine whether temporal sequence modeling identifies deviations that static rules miss, especially where order timing and monitoring cadence matter [1, 15, 16, 21]. Inter-rater review would be important because disagreement among clinicians may reveal ambiguity in the pathway rather than failure of the model [6, 26].

Explanation quality and clinical relevance

Explanation quality should be evaluated by clinicians who review whether the stated drivers of deviation are understandable, clinically relevant, and useful for action. This evaluation should include both local explanations for individual patient-days and global explanations that summarize recurring patterns across services or units. Prior XAI research in healthcare emphasizes that explanations must be designed around clinical use, not merely mathematical transparency, because clinicians need context, actionability, and the ability to judge whether the model's reasoning is

appropriate [23–27]. Evaluation should therefore ask whether the explanation helps clinicians determine whether to realign care, document justification, or revise the pathway itself [12, 13].

Prospective impact on pathway adherence and outcomes

Prospective evaluation should examine whether model-informed feedback changes pathway adherence, improves care-process reliability, and supports earlier recognition of clinically meaningful drift. Such assessment should avoid treating every reduction in deviation flags as improvement, because appropriate individualized care may still require pathway departure. Instead, the evaluation should focus on whether explanations improve review quality, reduce preventable omissions, support timely escalation, and reveal system-level workflow barriers [6, 11, 28]. Outcome evaluation should be conducted cautiously and transparently, because the conceptual model would require real-world validation before claims about clinical benefit could be justified [7, 8, 27].

Limitations

Defining the “gold standard” of appropriate deviation

A major limitation is that appropriate deviation is difficult to define because pathways are not absolute rules and hospitalized patients often require individualized care. A model may flag a trajectory as atypical even when the clinician intentionally modified care because of comorbidity, patient preference, diagnostic uncertainty, or evolving clinical status. Process-mining and pathway-discovery studies show that real-world care contains many variants, so distinguishing harmful drift from justified adaptation requires expert interpretation rather than algorithmic judgment alone [2, 5, 6, 22]. The model should therefore be positioned as a decision-support and audit tool, not as a definitive judge of clinical correctness [26, 27].

Data completeness and sequence granularity

EHR event data may incompletely represent clinical reasoning, verbal communication, bedside reassessment, patient refusal, nursing workload, and informal care coordination. Order sequences and monitoring frequency are valuable process signals, but they may also reflect documentation practices, system constraints, or delayed data entry rather than true clinical intent. Studies of uncertain clinical records, preprocessing bias, and EHR-based pathway analytics highlight the need to address noisy timing, incomplete capture, and variable event granularity before drawing conclusions from sequence models [12, 21, 29]. As a result, deviation explanations should always be reviewable against the chart and clinical context before being used for quality decisions [13, 27].

Conclusion

An interpretable machine learning model for detecting clinical pathway deviations could provide a structured way to monitor whether inpatient care remains aligned with expected trajectories. By combining order sequences, vital sign trends, laboratory monitoring frequency, and provider decision patterns, the model would treat care delivery as a dynamic process rather than a static checklist.

The main strength of this approach is its integration of temporal trajectory analysis with transparent deviation rationale. A model that identifies both the existence of a deviation and the specific care-process feature driving it could support bedside review, audit, and quality-improvement learning. Its explanations could help clinicians determine whether a deviation reflects appropriate personalization, preventable omission, unsupported commission, or pathway-design failure.

Important challenges remain before such a model could be responsibly implemented. These include defining appropriate deviation, preventing bias from historical practice patterns, ensuring that explanations are clinically meaningful, and

validating the system prospectively in real hospital workflows. Trust would depend not only on model transparency but also on careful governance, user-centered design, and a just-culture approach to audit.

Future work should focus on collaborative pathway modeling across clinical teams, informatics specialists, quality leaders, and patients. Implementation studies across diverse hospital units and conditions would be needed to determine how interpretable deviation detection can support safer, more reliable, and more adaptive inpatient care.

Acknowledgements

None

Conflict of interest

None

Financial support

None

Ethics statement

None

Received: 20 Jan 2024

Revised: 11 Mar 2024

Accepted: 09 Apr 2024

Published online: 20 July 2024

Rights and permissions

Open Access The author(s) retain copyright. This article is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. It may be shared and adapted for non-commercial purposes with appropriate attribution, an indication of changes, and distribution of adaptations under the same license. Third-party material may be subject to separate terms identified in its credit line. View the license at <https://creativecommons.org/licenses/by-nc-sa/4.0/>.

References

- Yan H, Van Gorp P, Kaymak U, Lu X, Ji L, Chiau CC, et al. Aligning event logs to task-time matrix clinical pathways in BPMN for variance analysis. *IEEE J Biomed Health Inform.* 2018;22(2):311-7.
- Baker K, Dunwoodie E, Jones RG, Newsham A, Johnson O, Price CP, et al. Process mining routinely collected electronic health records to define real-life clinical pathways during chemotherapy. *Int J Med Inform.* 2017;103:32-41.
- Alvarez C, Rojas E, Arias M, Munoz-Gama J, Sepúlveda M, Herskovic V, et al. Discovering role interaction models in the Emergency Room using process mining. *J Biomed Inform.* 2018;78:60-77.
- Rojas E, Sepúlveda M, Munoz-Gama J, Capurro D, Traver V, Fernandez-Llatas C. Question-driven methodology for analyzing emergency room processes using process mining. *Appl Sci.* 2017;7(3):302.

De Roock E, Martin N. Process mining in healthcare: an updated perspective on the state of the art. *J Biomed Inform.* 2022;127:103995.

Munoz-Gama J, Martin N, Fernandez-Llatas C, Johnson OA, Sepúlveda M, Helm E, et al. Process mining for healthcare: characteristics and challenges. *J Biomed Inform.* 2022;127:103994.

Muyama L, Neuraz A, Coulet A. Machine learning approaches for the discovery of clinical pathways from patient data: a systematic review. *J Biomed Inform.* 2024;160:104746.

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. *Health Care Manag Sci.* 2024;27(2):136-52.

Balestra M, Chen J, Iturrate E, Aphinyanaphongs Y, Nov O. Predicting inpatient pharmacy order interventions using provider action data. *JAMIA Open.* 2021;4(3):ooab083.

King CR, Abraham J, Fritz BA, Cui Z, Galanter W, Chen Y, et al. Predicting self-intercepted medication ordering errors using machine learning. *PLoS One.* 2021;16(7):e0254358.

Li RC, Wang JK, Sharp C, Chen JH. When order sets do not align with clinician workflow: assessing practice patterns in the electronic health record. *BMJ Qual Saf.* 2019;28(12):987-96.

Rostamzadeh N, Abdullah SS, Sedig K. Data-driven activities involving electronic health records: an activity and task analysis framework for interactive visualization tools. *Multimodal Technol Interact.* 2020;4(1):7.

Linhares CD, Lima DM, Ponciano JR, Olivatto MM, Gutierrez MA, Poco J, et al. ClinicalPath: a visualization tool to improve the evaluation of electronic health records in clinical decision-making. *IEEE Trans Vis Comput Graph.* 2023;29(10):4031-46.

Niu H, Omitaomu OA, Langston MA, Olama M, Ozmen O, Klasky HB, et al. EHR-BERT: a BERT-based model for effective anomaly detection in electronic health records. *J Biomed Inform.* 2024;150:104605.

Niu H, Omitaomu OA, Langston MA, Olama M, Ozmen O, Klasky HB, et al. Detecting anomalous sequences in electronic health records using higher-order tensor networks. *J Biomed Inform.* 2022;135:104219.

Le HH, Yamada T, Honda Y, Sakamoto T, Matsuo R, Yamazaki T, et al. Methods for analyzing medical-order sequence variants in sequential pattern mining for electronic medical record systems. *ACM Trans Comput Healthc.* 2023;4(1):1-28.

Lee JM, Hauskrecht M. Personalized event prediction for electronic health records. *Artif Intell Med.* 2023;143:102620.

Lee JM, Hauskrecht M. Modeling multivariate clinical event time-series with recurrent temporal mechanisms. *Artif Intell Med.* 2021;112:102021.

Lee JM, Hauskrecht M. Neural clinical event sequence prediction through personalized online adaptive learning. In: *International Conference on Artificial Intelligence in Medicine*. Cham: Springer; 2021. p. 175-86.

Rough K, Dai AM, Zhang K, Xue Y, Vardoulakis LM, Cui C, et al. Predicting inpatient medication orders from electronic health record data. *Clin Pharmacol Ther.* 2020;108(1):145-54.

Zhang S, Genga L, Dekker L, Nie H, Lu X, Duan H, et al. Re-ordered fuzzy conformance checking for uncertain clinical records. *J Biomed Inform.* 2024;149:104566.

Kempa-Liehr AW, Lin CY, Britten R, Armstrong D, Wallace J, Mordaunt D, et al. Healthcare pathway discovery and probabilistic machine learning. *Int J Med Inform.* 2020;137:104087.

Lundberg SM, Erion G, Chen H, DeGrave A, Prutkin JM, Nair B, et al. From local explanations to global understanding with explainable AI for trees. *Nat Mach Intell.* 2020;2(1):56-67.

Ahmad MA, Eckert C, Teredesai A. Interpretable machine learning in healthcare. In: *Proceedings of the 2018 ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics*. New York: ACM; 2018. p. 559-60.

Vellido A. The importance of interpretability and visualization in machine learning for applications in medicine and health care. *Neural Comput Appl.* 2020;32(24):18069-83.

Tonekaboni S, Joshi S, McCradden MD, Goldenberg A. What clinicians want: contextualizing explainable machine learning for clinical end use. In: *Machine Learning for Healthcare Conference. Proceedings of Machine Learning Research*; 2019. p. 359-80.

Barda AJ, Horvat CM, Hochheiser H. A qualitative research framework for the design of user-centered displays of explanations for machine learning model predictions in healthcare. *BMC Med Inform Decis Mak.* 2020;20(1):257.

Erdogan TG, Tarhan AK. Multi-perspective process mining for emergency process. *Health Informatics J.* 2022;28(1):14604582221077195.

Uhl L, Augusto V, Dalmas B, Alexandre Y, Bercelli P, Jardinaud F, et al. Evaluating the bias in hospital data: automatic preprocessing of patient pathways algorithm development and validation study. *JMIR Med Inform.* 2024;12:e58978.