

ORIGINAL RESEARCH

Open access

Self-Supervised Representation Learning for Healthcare Operations Events Using Timestamped Orders, Patient Transfers, Staff Actions, Queue Transitions, System Interaction Logs, and Unit-Level Workflow Signals

Yuki Yamamoto^{1*}, Kenji Ito¹

Abstract

Hospitals generate dense streams of timestamped operational events, including orders, transfers, staff actions, queue changes, and system interactions. These events describe how care actually unfolds, yet much of their value remains unused because they are rarely labeled for prediction tasks. Existing operational predictive models often depend on task-specific labels, handcrafted features, and local workflow assumptions. This limits their ability to scale across hospitals, departments, and evolving operational conditions. This manuscript designs a self-supervised representation learning model that pre-trains on diverse healthcare operations event streams. The goal is to learn a generalizable embedding of hospital operational state that can be adapted to multiple downstream prediction tasks. The proposed model uses a transformer-based architecture trained with masked event modeling and temporal contrastive learning. Timestamped orders, transfers, staff actions, queue transitions, system interaction logs, and unit-level workflow signals are represented as time-aware event sequences, and the pre-trained backbone is later fine-tuned for specific operational tasks. Conceptually, the model could learn semantic and temporal regularities of hospital workflow, such as common discharge sequences, clustered STAT order activity, and operational precursors to bottlenecks. These representations would be expected to support downstream tasks such as delay forecasting, anomaly detection, and resource demand estimation when labeled data are limited. Self-supervised learning could unlock the latent value of healthcare operations logs by creating reusable representations of hospital workflow. Such a model could become a foundation for operational analytics, enabling faster and more adaptable development of predictive tools.

Keywords Transformer models, Hospital operations, Self-supervised learning, Operational prediction, Event sequences, Healthcare workflow

*Correspondence:

Yuki Yamamoto
yuki.yamamoto@gmail.com

¹ Department of Healthcare Analytics and Informatics, Faculty of Medicine, Osaka University, Osaka, Japan

Introduction

Hospitals increasingly operate through digital systems that record orders, transfers, documentation, medication administration, queue movement, and other workflow events as timestamped data. Although these streams are rich descriptions of operational activity, their use in predictive analytics remains constrained by fragmentation, local coding practices, and the cost of creating labeled outcomes for each operational question. Clinical representation learning has

shown that large-scale electronic health records can support prediction when raw longitudinal events are transformed into learned patient embeddings, but operational event streams remain less systematically modeled [1-3]. This creates an opportunity to treat hospital operations data not merely as audit trails, but as sequential evidence of how care delivery systems behave over time.

Self-supervised learning offers a conceptual path for learning from unlabeled event sequences before a specific downstream task is defined. The success of bidirectional transformer pre-training in language modeling demonstrated that masked token prediction can produce reusable contextual representations, and related time-series models suggest that temporal structure can be learned from sequence context rather than manual feature engineering [4-6]. Hospital operations events have an analogous grammar: orders precede task queues, transfers alter unit state, and staff actions mediate the relationship between demand and throughput. A model that learns these dependencies could represent the evolving operational state of a patient episode, unit shift, or service line.

Pre-trained healthcare models such as BEHRT, Med-BERT, CEHR-BERT, and related transformer frameworks show that structured clinical events can be encoded into representations that support downstream prediction. These models mainly emphasize diagnoses, medications, procedures, or broader clinical timelines, but their design principles can be adapted to operational signals such as bed status, queue transitions, patient movement, and staff workflow events [1, 2, 7-9]. Foundation-oriented approaches in electronic health records further suggest that one learned backbone could be adapted to multiple prediction settings rather than rebuilt for each outcome [10-13]. In hospital operations, such a backbone could support delay prediction, bottleneck detection, capacity forecasting, and anomaly triage without starting from task-specific feature construction each time.

The central thesis of this manuscript is that a self-supervised model can learn robust representations from raw hospital operations events and transfer them to diverse operational prediction tasks. Rather than treating each task as a separate modeling exercise, the proposed MDL framework pre-trains a time-aware event encoder on heterogeneous operational streams and fine-tunes it through lightweight task heads. Prior work on temporal health event prediction, contrastive learning, and health trajectory modeling provides the methodological basis for this design, while the operational adaptation remains a conceptual model-development contribution. The proposed architecture is therefore positioned as a reusable representation layer for hospital analytics rather than a single-purpose supervised model.

Background

Timestamped operational data in hospitals

Timestamped operational data in hospitals include medication, laboratory, and imaging orders; admission-discharge-transfer movements; nurse documentation; medication administration; queue entry and exit events; bed-management status updates; consult requests; and system interaction logs. These events are heterogeneous because they arise from different information systems, but they share a temporal structure that makes them suitable for sequence modeling. Longitudinal EHR representation learning has already shown that structured event histories can be transformed into contextual embeddings, and this same principle can be extended from clinical history to operational workflow [1, 3, 7]. The key modeling challenge is to preserve event type, time interval, metadata, and institutional context while allowing the model to learn cross-stream dependencies.

Supervised deep learning for specific operations tasks

Supervised deep learning in healthcare has often focused on specific prediction problems such as risk prediction, disease progression, adverse events, no-show risk, or clinical deterioration. These models can be powerful, but they typically require labels, target-specific feature pipelines, and local validation procedures that must be rebuilt when the outcome changes [14-16]. Even scalable EHR models can depend on task-specific supervision when adapted to a defined endpoint, which limits flexibility for operational questions that change with hospital priorities [3, 17, 18]. For hospital

operations, this creates a mismatch between the abundance of unlabeled workflow traces and the scarcity of curated labels for every delay, bottleneck, or resource utilization target.

Self-supervised learning on event sequences

Self-supervised learning on event sequences uses the structure of the sequence itself as a training signal, allowing models to learn representations without manually assigned labels. Masked event modeling asks the model to recover hidden events from context, autoregressive pre-training asks it to predict future events from the past, and contrastive objectives encourage representations of related temporal segments to be closer than unrelated or corrupted segments [4-6, 19]. In irregularly sampled healthcare data, temporal neighborhood coding and multivariate time-series representation learning are especially relevant because the timing between events carries operational meaning [5, 19]. For hospital workflow, these objectives can be adapted so that the model learns not only which event occurs, but when it occurs and how it relates to surrounding unit and patient states.

Pre-training and transfer learning in healthcare AI

Pre-training and transfer learning have become central to healthcare AI because they allow models to learn from broad longitudinal data before adaptation to specific tasks. BEHRT, Med-BERT, CEHR-BERT, and TransformEHR illustrate how transformer models can encode medical event sequences and transfer contextual knowledge to prediction tasks involving diagnoses, outcomes, and patient trajectories [1, 2, 7, 9]. More recent generative and foundation-model approaches suggest that patient timelines can be modeled as sequences with reusable structure, supporting adaptation across tasks and settings [10-12, 20, 21]. However, these developments have primarily centered on clinical events, leaving operational and administrative event streams underdeveloped as a target for self-supervised pre-training.

The promise of a foundation model for hospital operations

A foundation model for hospital operations would learn the grammar of workflow: how orders trigger queues, how transfers reshape capacity, how staff actions respond to changing demand, and how unit-level states condition future events. Foundation-model thinking in healthcare emphasizes broad pre-training, flexible adaptation, and robustness across temporal and institutional shifts, which are also essential properties for hospital operations analytics [10, 13, 22]. Such a model would not replace task-specific validation, but it could provide a reusable representation layer that reduces repeated feature engineering and supports rapid fine-tuning. In this sense, the proposed operational foundation model is analogous to EHR timeline models, but its semantic units are workflow events rather than only clinical concepts [11, 23, 24].

Model Development Overview

High-level pre-training and fine-tuning framework

The proposed framework begins with self-supervised pre-training on unlabeled operational event sequences and then adapts the learned backbone to specific operational prediction tasks. During pre-training, the model could learn contextual embeddings for patient episodes, unit shifts, and institutional workflow states by reconstructing masked events and contrasting temporally coherent sequences against corrupted alternatives. During fine-tuning, a lightweight prediction head would use the learned representation for tasks such as delay risk, bottleneck detection, or resource demand estimation. This design follows the general logic of healthcare pre-training models, where broad event-sequence learning supports later task-specific adaptation [1, 2, 9, 12].

Figure 1 illustrates the proposed self-supervised transformer architecture, from raw timestamped hospital operations events through tokenization, time-aware sequence construction, pre-training, reusable representation learning, task-specific fine-tuning, and governed operational deployment.

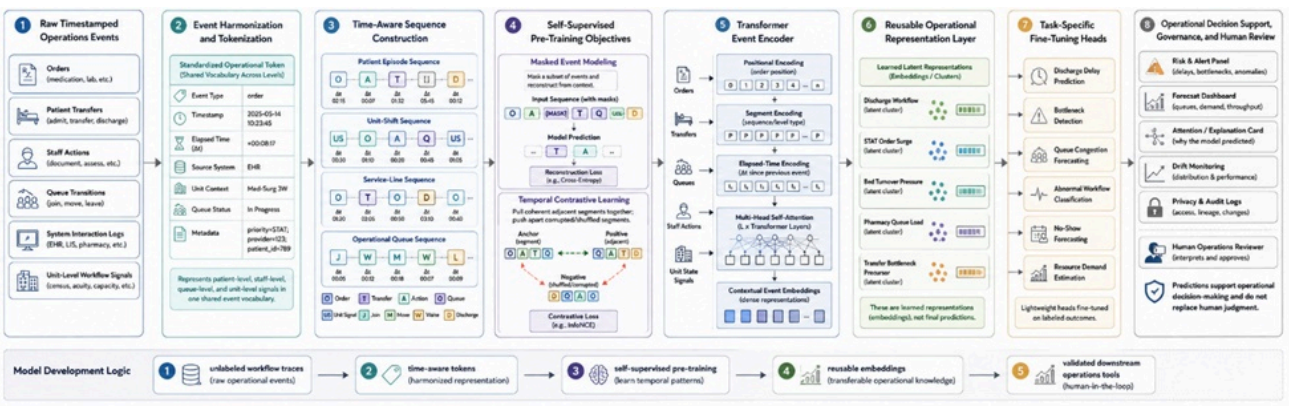


Figure 1. Self-Supervised Transformer Architecture for Learning Reusable Hospital Operations Event Representations

Core event streams used for pre-training

The pre-training corpus would combine timestamped orders, patient transfers, staff actions, queue transitions, system interaction logs, and unit-level workflow signals into a unified event stream. Orders could include medication, laboratory, imaging, diet, and consult requests, while transfers would encode admission, discharge, bed assignment, and inter-unit movement. Staff actions and system logs would represent the operational work required to move patients through care pathways, extending the event vocabulary beyond traditional diagnosis and medication histories [7, 25, 26]. By integrating these streams, the model could learn dependencies between patient-level actions and unit-level constraints that are usually separated across analytics pipelines.

Design principles

The model is guided by five design principles: modality-agnostic tokenization, time-aware encoding, broad institutional pre-training, efficient fine-tuning, and interpretable representations. Modality-agnostic tokenization allows orders, transfers, queues, and staffing signals to share a common representation space, while time-aware encoding preserves irregular event spacing and operational tempo. Efficient fine-tuning is important because many operational labels are limited, and parameter-efficient adaptation would be consistent with transfer-learning approaches used in clinical event modeling [2, 10, 27]. Interpretability is also essential because operational users need to understand whether embeddings and attention patterns reflect meaningful workflow structures rather than opaque correlations [15, 22].

Data Sources and Event Stream Construction

Extraction and tokenization of heterogeneous events

Each operational event is converted into a structured token that includes event type, timestamp, time interval, source system, and relevant metadata. A medication order token, for example, could encode priority and route, while a transfer token could encode origin unit, destination unit, bed status, and transfer reason. This mirrors the way clinical transformer models represent diagnosis, medication, or procedure concepts as sequence elements, while extending the vocabulary to include operational workflow events [1, 2, 7, 8]. Tokenization should preserve enough semantic detail for reconstruction and prediction while avoiding unnecessary identifiers that could impair privacy or generalization.

Constructing event sequences

Event streams can be assembled at different units of analysis, including patient episodes, unit shifts, service lines, or operational queues. Within each sequence, events are ordered by timestamp, and elapsed-time encodings are inserted so the model can distinguish rapid clusters from long gaps. This construction follows temporal EHR models that treat

healthcare histories as ordered event sequences, while also drawing on time-series representation learning methods that emphasize irregular sampling and local temporal neighborhoods [5, 11, 19, 20]. Patient-centered sequences would capture care trajectories, whereas unit-centered sequences would capture workload, congestion, and capacity dynamics.

Table 1 defines the proposed operational event representation schema, showing how heterogeneous timestamped workflow signals can be transformed into reusable time-aware tokens for self-supervised hospital operations modeling.

Table 1. Operational Event Representation Schema for Self-Supervised Hospital Workflow Modeling

Representation layer	Operational elements encoded	Token-level design logic	Temporal information preserved	Why this strengthens transfer learning	Example downstream relevance
Order-event tokens	Medication, laboratory, imaging, diet, consult, procedure, and discharge-related orders	Encodes order type, urgency, priority, service, source system, and relevant order metadata	Order time, interval since prior event, clustering of repeated orders, STAT-order bursts	Allows the model to learn how clinical demand enters operational queues and how order patterns precede delays or resource pressure	Pharmacy backlog prediction, radiology queue forecasting, discharge delay prediction
Patient-transfer tokens	Admission, discharge, bed assignment, inter-unit transfer, transport request, bed-cleaning status	Encodes origin unit, destination unit, transfer type, bed state, movement status, and transfer reason	Transfer sequence, dwell time, time from order to movement, repeated transfer attempts	Captures patient-flow grammar and bed-management dependencies across units	Boarding risk, placement delay, bed-turnover prediction, unit congestion detection
Staff-action tokens	Documentation, medication administration, task completion, sign-off, escalation, handoff, queue acknowledgement	Encodes role category, action type, task status, escalation status, and operational context without exposing unnecessary identifiers	Response time, action latency, shift-period timing, task completion intervals	Represents the labor-mediated pathway between demand and throughput	Workload forecasting, staffing pressure detection, missed-task anomaly identification
Queue-transition tokens	Entry into queue, reassignment, prioritization, hold status, completion, cancellation, queue exit	Encodes queue type, priority status, queue stage, queue load, and service-line context	Queue residence time, transition latency, queue accumulation rate, sequence of handoffs	Teaches the model bottleneck dynamics rather than isolated event counts	Transport delay, pharmacy verification delay, imaging queue congestion, consult backlog detection

System-interaction tokens	EHR clicks, message routing, dashboard updates, order modifications, alert acknowledgements, scheduling interactions	Encodes interaction type, source application, task category, and workflow role	Interaction density, repeated modification timing, time from alert to acknowledgement	Captures hidden administrative and coordination work that is often absent from supervised feature sets	Administrative burden prediction, alert fatigue monitoring, documentation workload detection
Unit-level workflow tokens	Census, occupancy, staffing ratios, pending discharges, queue load, open beds, service-line demand	Encodes unit state as periodic or state-change tokens linked to patient and queue events	State snapshots, shift-level tempo, demand surges, occupancy transitions	Contextualizes patient-level events within operational capacity and local workflow pressure	Capacity forecasting, command-center dashboards, cross-unit bottleneck detection
Institutional-context tokens	Hospital site, unit type, service line, shift category, calendar period, policy environment	Encodes site and workflow context in privacy-aware categories rather than identifiable labels	Temporal drift, seasonal effects, weekend/weekday patterns, policy-change periods	Supports adaptation across hospitals and evolving operational conditions while allowing local calibration	Generalizability testing, site-specific fine-tuning, workflow drift monitoring
Derived sequence embeddings	Patient episode embeddings, unit-shift embeddings, queue embeddings, service-line embeddings	Aggregates contextual token representations into reusable operational state vectors	Preserves sequential dependencies and learned temporal proximity	Provides the reusable backbone for multiple downstream tasks without rebuilding handcrafted features	Delay prediction, anomaly detection, resource forecasting, operational risk triage

Unit-level workflow signals as aggregate features

Unit-level workflow signals such as census, staffing ratios, bed occupancy, pending discharges, and queue load can be represented as periodic or state-change tokens within the same event sequence. These aggregate signals are important because the meaning of a patient-level event often depends on the operational state in which it occurs. For example, an imaging order may carry different implications when queue load is high, staffing is constrained, or bed turnover is delayed. Multimodal and graph-aware healthcare representation learning supports this idea by showing that richer context can improve the structure of learned embeddings, and the same principle can be applied to operational state modeling [13, 16, 25].

Self-Supervised Pre-Training Architecture

Transformer encoder for event sequences

The proposed model uses a transformer encoder to process heterogeneous operational tokens with positional, segment, and elapsed-time encodings. The encoder would represent each event in relation to surrounding orders, transfers, staff actions, queues, and unit-level state, allowing contextual embeddings to reflect both event identity and workflow position. Transformer-based EHR models demonstrate that attention mechanisms can encode long-range dependencies in healthcare sequences, and time-series transformer frameworks show how similar architectures can be adapted to multivariate temporal data [1, 2, 4, 9]. For hospital operations, the architecture should be efficient enough for healthcare IT environments while still expressive enough to model cross-stream dependencies and irregular timing.

Masked event modeling objective

Masked event modeling randomly hides selected operational tokens or metadata fields and trains the model to infer them from surrounding context. The masked token could be an order type, transfer destination, queue transition, staff action, or time-interval category, depending on which aspect of workflow is being reconstructed. This objective is conceptually aligned with bidirectional language-model pre-training and with healthcare models that adapt masked prediction to structured EHR sequences [1, 2, 6, 7]. In hospital operations, successful reconstruction would indicate that the model has learned contextual regularities such as typical discharge workflows, order sequencing, and dependencies between unit state and subsequent actions.

Temporal contrastive learning objective

Temporal contrastive learning adds a second self-supervised signal by encouraging representations of coherent adjacent workflow segments to be closer than representations of corrupted, shuffled, or unrelated segments. This would force the model to learn operational continuity rather than simply memorizing event co-occurrence, because negative samples could be drawn from different patient episodes, unrelated unit shifts, or temporally implausible event orderings. Contrastive approaches in healthcare and temporal representation learning suggest that such objectives can help models capture meaningful sequence structure under limited labels [5, 16, 19, 27, 28]. Applied to operations, the objective could help distinguish genuine workflow progression from anomalous temporal patterns that may later support bottleneck detection or delay explanation.

Fine-Tuning for Downstream Operational Tasks

Downstream task examples

After pre-training, the operational backbone could be fine-tuned for discharge delay prediction, bottleneck detection, no-show forecasting, abnormal workflow classification, queue congestion prediction, and resource demand estimation. Each downstream task would attach a lightweight prediction head to a shared representation of patient-level, staff-level, queue-level, and unit-level events. This approach follows the transfer logic of pre-trained EHR models, where one contextual backbone supports multiple prediction tasks after task-specific adaptation [2, 9, 10, 12]. For hospital operations, the same structure would allow a single representation layer to support many operational questions without rebuilding the entire model for each use case.

Fine-tuning with limited labeled data

Fine-tuning with limited labeled data is central to the proposed MDL framework because many operational labels are difficult to define, validate, and maintain over time. A pre-trained event representation could reduce the burden of handcrafted feature engineering by transferring workflow knowledge learned from unlabeled sequences into task-specific models. Prior work on healthcare transfer learning and contrastive pre-training suggests that contextual representations can support adaptation when labels are constrained, although each operational task would still require local validation [1, 14, 27]. The expected value is not a guaranteed performance gain, but a more reusable modeling pipeline for operational analytics.

Multi-task and continual learning

The pre-trained model could also support multi-task fine-tuning, where related operational outcomes are learned jointly through task-specific heads connected to a shared event encoder. For example, discharge delay, bed turnaround risk, pharmacy verification delay, and radiology queue congestion may share workflow precursors even if their labels differ. Continual learning would allow periodic updates as new operational patterns emerge, while safeguards would be needed to avoid overwriting previously learned workflow structure [10, 13, 22]. This is particularly important because hospital operations change through staffing models, new EHR configurations, service-line redesigns, and evolving care pathways.

Interpretability of Learned Representations

Understanding the pre-trained embedding space

The embedding space produced by the pre-trained model should be examined to determine whether operationally meaningful events occupy coherent regions. Discharge-related orders, bed-management events, transport requests, and post-discharge documentation could cluster if the model has learned a latent discharge workflow, while medication-verification, administration, and nurse documentation events could form another operational neighborhood. Prior work on patient embeddings and graph-based clinical representations supports the idea that learned healthcare representations can reveal clinically meaningful structure, but operational embeddings would need review by domain experts [18, 24, 25]. Such visualization would be exploratory rather than proof of model validity.

Attention analysis for root cause investigation

Attention analysis could help operational teams inspect which prior events contributed most strongly to a downstream prediction such as a potential discharge delay or queue bottleneck. A delay prediction might attend to missing bed-cleaning completion, delayed consult sign-off, repeated order modifications, or prolonged queue residence, depending on the learned event context. Transformer-based healthcare studies have used attention and contextual representations to support interpretability, but attention patterns should be treated as explanatory aids rather than definitive causal evidence [1, 4, 15]. In the proposed model, attention review would be paired with workflow expertise to support root-cause investigation and model governance.

Integration Into Hospital Analytics Infrastructure

Deploying the pre-trained model as a service

The pre-trained operational model could be deployed as a service that receives updated event streams and returns embeddings or task-specific predictions to dashboards, command centers, and EHR-integrated applications. This service-oriented design would separate representation learning from downstream application development, allowing multiple analytics tools to use the same contextual event backbone. Scalable EHR modeling and foundation-model frameworks highlight the importance of building reusable infrastructure rather than isolated models for each prediction task [3, 12, 13]. In hospital operations, deployment would also require monitoring for workflow drift, interface changes, and shifts in event-coding practices.

Enabling rapid prototyping of new operational models

With a pre-trained operational backbone, analysts could prototype new models for emerging operational questions by defining a label, selecting relevant prediction windows, and fine-tuning a lightweight head. A pharmacy delay model, transport bottleneck model, or admission boarding risk model could use the same underlying event embeddings rather than starting with a new feature design process. This mirrors the promise of foundation models for healthcare, where

broad pre-training supports flexible adaptation across tasks and changing data contexts [11, 13, 21, 23]. The model would still require prospective evaluation and workflow integration before use in operational decision support.

Evaluation Strategy

Pre-training quality

Pre-training quality should be evaluated through representation-focused tests rather than claims of operational effectiveness. Linear probing could assess whether frozen embeddings contain information relevant to operational tasks, while masked event reconstruction could test whether the model has learned contextual regularities in orders, transfers, queues, and staff actions. Time-series and masked-sequence representation learning provide established conceptual tools for these assessments, although operational interpretation would depend on the event vocabulary and workflow setting [4-6, 19]. Evaluation should also inspect whether learned representations reflect plausible temporal structure rather than shortcut correlations.

Fine-tuning performance vs. Task-specific models

For each downstream task, the fine-tuned pre-trained model should be compared conceptually against task-specific models trained from scratch on the same labeled data. Relevant outcomes could include discrimination, calibration, error, robustness, interpretability, and operational usability, but this manuscript does not report empirical results or performance numbers. Prior EHR modeling studies demonstrate the importance of benchmarking pre-trained representations against supervised baselines, especially when claims concern transferability or generalization [2, 10, 20, 26]. In hospital operations, comparisons should also consider implementation burden, label quality, and whether the model's predictions align with actionable workflow interventions.

Robustness and generalizability

Robustness should be evaluated across time periods, hospital units, and, where possible, separate hospitals within a health system. Temporal shift is especially important because staffing models, patient demand, EHR build, and operational policies can change, altering the relationship between event sequences and downstream outcomes. Recent work on EHR foundation models and temporal distribution shift emphasizes that generalization cannot be assumed simply because a model is pre-trained on large event corpora [10, 12, 22]. The operational model should therefore be evaluated for transportability, drift sensitivity, and the need for periodic re-calibration or re-training.

Table 2 presents an evaluation and governance framework that separates representation quality, transfer performance, operational actionability, interpretability, privacy protection, and lifecycle monitoring.

Table 2. Evaluation and Governance Framework for a Self-Supervised Hospital Operations Foundation Model

Evaluation domain	Core question	Recommended assessment approach	Evidence produced	Operational interpretation	Governance implication

Pre-training signal quality	Has the model learned meaningful workflow structure from unlabeled event streams?	Masked event reconstruction, temporal-order corruption tests, contrastive segment discrimination	Reconstruction accuracy, temporal plausibility scores, contrastive separation metrics	Demonstrates whether the model recognizes regularities in orders, transfers, queues, staff actions, and unit state	Poor reconstruction may indicate weak token design, inconsistent source-system mapping, or insufficient event harmonization
Embedding utility	Do frozen representations contain information useful for downstream operations tasks?	Linear probing and few-shot prediction using frozen embeddings	Probe performance across delay, congestion, anomaly, and demand tasks	Tests whether the backbone captures transferable operational state before full fine-tuning	Supports decisions about whether the representation layer is reusable or task-specific retraining is needed
Fine-tuning performance	Does the pre-trained backbone improve task adaptation compared with models trained from scratch?	Head-to-head comparison with supervised baselines using the same labeled data	Discrimination, calibration, prediction error, decision-curve utility, label-efficiency curves	Determines whether pre-training reduces label requirements or improves operational prediction	Claims of foundation-model value should be limited unless transfer gains are demonstrated across tasks
Temporal robustness	Does performance remain stable when workflow patterns change over time?	Train-test splits across months, seasons, policy changes, EHR build changes, and staffing model changes	Temporal performance drift, calibration drift, representation-shift metrics	Identifies whether embeddings remain useful under evolving operational conditions	Requires drift monitoring, periodic recalibration, and governance review after workflow redesign
Cross-unit and cross-site transportability	Can the representation generalize across departments, service lines, or hospitals?	External validation by unit, hospital, service line, and operational context	Transportability metrics, subgroup calibration, site-stratified error profiles	Distinguishes reusable workflow structure from locally memorized event patterns	Determines whether local fine-tuning, site adapters, or institution-specific calibration are required
Interpretability of learned workflow states	Are embeddings and attention patterns operationally meaningful to domain experts?	Embedding visualization, nearest-neighbor event review, attention inspection, expert adjudication	Clusters of workflow states, attention explanations, expert plausibility ratings	Shows whether the model's internal structure aligns with recognizable operational processes	Interpretability review should be mandatory before high-stakes command-center use

Actionability of downstream outputs	Do predictions correspond to feasible operational interventions?	Workflow simulation, user review, decision-path mapping, intervention alignment analysis	Mapping from predicted risk to possible actions, escalation pathways, resource allocation options	Prevents technically accurate predictions from becoming operationally unusable alerts	Deployment should require predefined response protocols and accountable human review
Privacy and re-identification risk	Could timestamped operational sequences expose sensitive patient, staff, or institutional patterns?	De-identification audit, rare-sequence review, access-control testing, metadata minimization review	Privacy-risk register, retained-field justification, audit-log evidence	Recognizes that timestamps, units, and rare workflow combinations may remain sensitive	Requires role-based access, audit trails, privacy review, and careful metadata governance
Equity and workload impact	Could model outputs worsen disparities or shift burden unfairly across units or staff groups?	Subgroup analysis, unit-level burden review, alert-distribution monitoring, fairness diagnostics	Error profiles by patient group, unit, shift, and service line; alert burden metrics	Evaluates whether the model improves operations without creating hidden inequities	Governance should include fairness review, escalation oversight, and monitoring of unintended workload redistribution
Lifecycle monitoring	Can the model remain safe and useful after deployment?	Prospective monitoring of calibration, drift, alert acceptance, override patterns, and operational outcomes	Model monitoring dashboard, update logs, retraining triggers, user feedback patterns	Treats the model as infrastructure requiring continuous stewardship rather than a one-time prediction tool	Establishes accountability for maintenance, recalibration, retirement, and post-deployment evaluation

Limitations

Computational and data infrastructure requirements

Pre-training a transformer on heterogeneous operational event streams would require substantial data engineering, governance, storage, and computing infrastructure. Smaller facilities may lack the resources to harmonize orders, transfers, staffing events, queue logs, and system interaction traces into a consistent time-aware representation. Foundation-model discussions in healthcare emphasize that scale can create value but also introduces practical barriers related to computation, reproducibility, and implementation capacity [13, 22]. The proposed MDL framework should therefore be considered a conceptual design that may be most feasible initially in larger health systems or multi-institutional collaborations.

Privacy implications of event sequences

Operational event sequences may contain re-identifiable patterns because combinations of timestamps, unit locations, staff actions, and rare workflow events can reveal sensitive information even when obvious identifiers are removed. Privacy protection would require de-identification, access control, audit logging, role-based governance, and careful review of which metadata are retained for modeling. Large-scale EHR modeling studies show the promise of learning from longitudinal healthcare data, but they also imply a need for strong governance when broad patient timelines are used for pre-training [3, 11, 22]. For operational foundation models, privacy auditing should be treated as part of the model-development process rather than an afterthought.

Conclusion

A self-supervised representation learning model for healthcare operations events would treat hospital workflow as a rich temporal sequence rather than a set of isolated operational indicators. By pre-training on orders, transfers, staff actions, queue transitions, system interaction logs, and unit-level workflow signals, the model could learn contextual representations of the evolving operational state. These representations could then be adapted to downstream tasks through lightweight fine-tuning rather than rebuilt for every new prediction problem.

The major strength of the proposed framework is its ability to unlock value from vast unlabeled operational data. It could provide transferable representations for delay prediction, resource forecasting, bottleneck detection, and workflow anomaly identification. It could also support faster prototyping of operational models while giving analysts tools to inspect embeddings and attention patterns. This combination of reuse, adaptability, and interpretability makes self-supervised learning especially well suited to hospital operations analytics.

Important challenges remain before such a model could be responsibly deployed. Data engineering requirements are substantial, and operational event streams may vary widely across hospitals and EHR configurations. Privacy risks are also significant because detailed timestamped workflow traces can encode sensitive patterns. Real-world validation would be necessary to determine whether the learned representations are robust, useful, and safe in operational decision-making.

A logical next step is the creation of a multi-institutional consortium to pre-train an open operational foundation model on diverse hospital event streams. Such a consortium could define shared event schemas, governance standards, evaluation protocols, and adaptation methods for hospital operations tasks. This would make the model more analogous to large language models, but grounded in the temporal grammar of healthcare delivery. If developed carefully, an operational foundation model could become a common infrastructure layer for safer, faster, and more adaptive hospital analytics.

Acknowledgements

None

Conflict of interest

None

Financial support

None

Ethics statement

None

Received: 04 Mar 2026

Revised: 25 Mar 2026

Accepted: 01 May 2026

Published online: 20 July 2026

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

Li Y, Rao S, Soares JR, Hassaine A, Ramakrishnan R, Canoy D, et al. BEHRT: transformer for electronic health records. *Sci Rep.* 2020;10(1):7155.

Rasmy L, Xiang Y, Xie Z, Tao C, Zhi D. Med-BERT: pretrained contextualized embeddings on large-scale structured electronic health records for disease prediction. *NPJ Digit Med.* 2021;4(1):86.

Rajkomar A, Oren E, Chen K, Dai AM, Hajaj N, Hardt M, et al. Scalable and accurate deep learning with electronic health records. *NPJ Digit Med.* 2018;1(1):18.

Zerveas G, Jayaraman S, Patel D, Bhamidipaty A, Eickhoff C. A transformer-based framework for multivariate time series representation learning. In: *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*; 2021. p. 2114-24.

Franceschi JY, Dieuleveut A, Jaggi M. Unsupervised scalable representation learning for multivariate time series. *Adv Neural Inf Process Syst.* 2019;32:4650-61.

Devlin J, Chang MW, Lee K, Toutanova K. BERT: pre-training of deep bidirectional transformers for language understanding. In: *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*; 2019. p. 4171-86.

Pang C, Jiang X, Kalluri KS, Spotnitz M, Chen R, Perotte A, et al. CEHR-BERT: incorporating temporal information from structured EHR data to improve prediction tasks. In: *Proceedings of Machine Learning for Health*. PMLR; 2021. p. 239-60.

Zeng X, Linwood SL, Liu C. Pretrained transformer framework on pediatric claims data for population specific tasks. *Sci Rep.* 2022;12(1):3651.

Yang Z, Mitra A, Liu W, Berlowitz D, Yu H. TransformEHR: transformer-based encoder-decoder generative model to enhance prediction of disease outcomes using electronic health records. *Nat Commun.* 2023;14(1):7857.

Guo LL, Steinberg E, Fleming SL, Posada J, Lemmon J, Pfohl SR, et al. EHR foundation models improve robustness in the presence of temporal distribution shift. *Sci Rep.* 2023;13(1):3767.

Kraljevic Z, Bean D, Shek A, Bendayan R, Hemingway H, Yeung JA, et al. Foresight—a generative pretrained transformer for modelling of patient timelines using electronic health records: a retrospective modelling study. *Lancet Digit Health.* 2024;6(4):e281-

90.

Guo LL, Fries J, Steinberg E, Fleming SL, Morse K, Aftandilian C, et al. A multi-center study on the adaptability of a shared foundation model for electronic health records. *NPJ Digit Med.* 2024;7(1):171.

Moor M, Banerjee O, Abad ZS, Krumholz HM, Leskovec J, Topol EJ, et al. Foundation models for generalist medical artificial intelligence. *Nature.* 2023;616(7956):259-65.

Chen H, Lundberg SM, Erion G, Kim JH, Lee SI. Forecasting adverse surgical events using self-supervised transfer learning for physiological signals. *NPJ Digit Med.* 2021;4(1):167.

Rao S, Li Y, Ramakrishnan R, Hassaine A, Canoy D, Cleland J, et al. An explainable transformer-based deep learning model for the prediction of incident heart failure. *IEEE J Biomed Health Inform.* 2022;26(7):3362-72.

Zang C, Wang F. SceHR: supervised contrastive learning for clinical risk prediction using electronic health records. In: *Proceedings of the IEEE International Conference on Data Mining (ICDM)*; 2021. p. 857-62.

Ma F, Chitta R, Zhou J, You Q, Sun T, Gao J. Dipole: diagnosis prediction in healthcare via attention-based bidirectional recurrent neural networks. In: *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*; 2017. p. 1903-11.

Choi E, Bahadori MT, Song L, Stewart WF, Sun J. GRAM: graph-based attention model for healthcare representation learning. In: *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*; 2017. p. 787-95.

Tonekaboni S, Eytan D, Goldenberg A. Unsupervised representation learning for time series with temporal neighborhood coding. *arXiv.* 2021;arXiv:2106.00750.

Renc P, Jia Y, Samir AE, Was J, Li Q, Bates DW, et al. Zero-shot health trajectory prediction using transformer. *NPJ Digit Med.* 2024;7(1):256.

Makarov N, Bordukova M, Quengdaeng P, Garger D, Rodriguez-Esteban R, Schmich F, et al. Large language models forecast patient health trajectories enabling digital twins. *NPJ Digit Med.* 2025;8(1):588.

Wornow M, Xu Y, Thapa R, Patel B, Steinberg E, Fleming S, et al. The shaky foundations of large language models and foundation models for electronic health records. *NPJ Digit Med.* 2023;6(1):135.

Redekop E, Wang Z, Kulkarni R, Pleasure M, Chin A, Hassanzadeh HR, et al. Zero-shot medical event prediction using a generative pretrained transformer on electronic health records. *J Am Med Inform Assoc.* 2025;32(12):1833-42.

Xian S, Grabowska ME, Kullo IJ, Luo Y, Smoller JW, Walunas TL, et al. Transformer patient embedding using electronic health records enables patient stratification and progression analysis. *NPJ Digit Med.* 2025;8(1):521.

Liu Y, Zhang Z, Mi J, Pan S, Chen T, Guo Y, et al. GatorCLR: personalized predictions of patient outcomes on electronic health records using self-supervised contrastive graph representation. *J Biomed Inform.* 2025;168:104851.

Liu T, Liang L, Che C, Liu Y, Jin B. A transformer-based framework for temporal health event prediction with graph-enhanced representations. *J Biomed Inform.* 2025;166:104826.

Feng W, Wu H, Ma H, Tao Z, Xu M, Zhang X, et al. Applying contrastive pre-training for depression and anxiety risk prediction in type 2 diabetes patients based on heterogeneous electronic health records: a primary healthcare case study. *J Am Med Inform Assoc.* 2024;31(2):445-55.

Wanyan T, Honarvar H, Jaladanki SK, Zang C, Naik N, Somani S, et al. Contrastive learning improves critical event prediction in COVID-19 patients. *Patterns.* 2021;2(12):100389.