Mr.Dec: Daily-Scale Longitudinal Multimodal Modeling for 30-Day Readmission Prediction

Minjun Kim*[0009−0009−8429−5240] and Jong Hak Moon*[0000−0002−6708−3918]

Yeji X, Seoul, South Korea
{mj.kim, jh.moon}@yejix.ai

Abstract. Predicting 30-day hospital readmission is essential for assessing patient stability and optimizing healthcare resources. As clinical risk evolves with the accumulation of evidence during hospitalization, capturing these dynamic trajectories is essential. However, many existing approaches compress the complex longitudinal history into fixed representations, often losing the granular, day-level clinical signals that reflect a patient's evolving physiological state. To address this, we propose Mr.Dec (**M**ultimodal **R**eadmission-risk prediction **De**coder), which models each admission as a natural chronological sequence of daily multimodal events. By leveraging a Transformer Decoder, Mr.Dec integrates daily Electronic Health Record(EHR) updates and intermittent Chest X-ray(CXR) findings in a time-aligned stream, reflecting the actual clinical workflow. To ensure robustness, we utilize Disease-Specific Supervised Contrastive Learning as an auxiliary regularization to induce a diagnosis-aware structure in the latent space. Evaluations on the MIMIC-IV and MIMIC-CXR datasets show that Mr.Dec achieves state-of-the-art performance by preserving the integrity of the clinical sequence. Furthermore, our model identifies "Critical Days" within an admission, providing actionable and clinically grounded interpretations for real-time risk stratification. Code is available at: https://github.com/yejix-ai/MR.DEC

Keywords: Causal Clinical Trajectory Modeling · Multimodal Fusion · 30-Day Hospital Readmission Prediction

1 Introduction

Predicting 30-day all-cause hospital readmission is a widely used benchmark of healthcare quality and patient stability. Unplanned readmissions worsen outcomes and impose substantial economic burden, costing billions annually. While post-discharge environmental factors also shape risk [7, 10, 16], a central engineering objective is to maximize discriminative signal from in-hospital data to identify high-risk cohorts before discharge [1, 3, 8, 9, 13, 17, 18]. The operational goal is a robust screening mechanism that interprets a patient's evolving trajectory to support triage and resource allocation. Accurate longitudinal modeling is essential to capture subtle signs of instability that may warrant intervention.

* Equal contribution

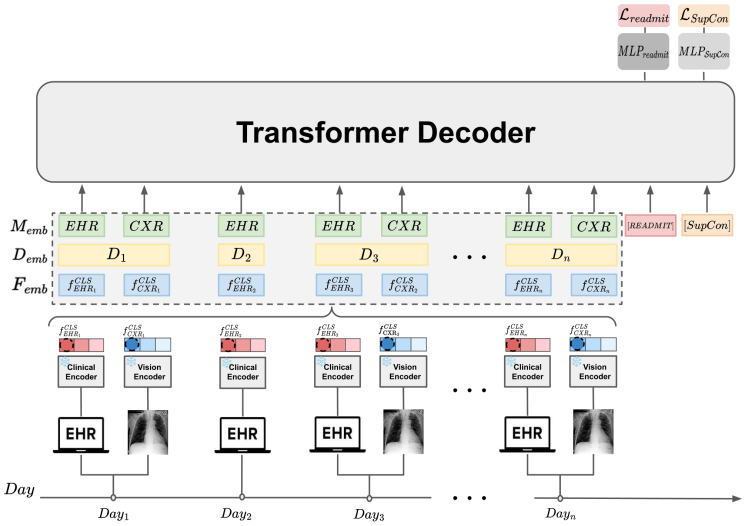

Fig. 1. Mr.Dec (Multimodal Readmission-risk Prediction Decoder). For each hospital day n, textualized EHR (e.g., ICD codes mapped to natural-language descriptors) is encoded by a clinical encoder and same-day CXR is encoded by a vision encoder, producing modality-specific `[CLS]` representations $\{f^{\text{CLS}}_{\text{EHR}_n}, f^{\text{CLS}}_{\text{CXR}_n}\}$ (with $f^{\text{CLS}}_{\text{CXR}_n}$ omitted when no CXR is acquired). Day-wise features are serialized in chronological order and projected with modality-specific layers; modality and day embeddings are added to obtain the embedded token sequence $F_{\text{emb}} = (h_1, \ldots, h_L)$. We then append two learnable special tokens, yielding the decoder input $\tilde{F}_{\text{emb}} = [F_{\text{emb}}, \texttt{[READMIT]}, \texttt{[SupCon]}]$. The final hidden states of `[READMIT]` and `[SupCon]` are fed to separate MLP heads for readmission risk estimation optimized with $\mathcal{L}_{\text{readmit}}$ and admission-level representation learning optimized with $\mathcal{L}_{\text{SupCon}}$, respectively.

Toward this end, readmission modeling has evolved toward richer patient representations. Early approaches relied on administrative variables such as ICD codes as static categorical features [11, 13], but these lack the granularity needed to reflect complex clinical progression. More recent work adopted multimodal learning, recognizing that structured codes alone are insufficient. This includes incorporating clinical notes with LLMs [3, 9] and integrating CXR using deep neural networks [8, 9, 13, 17, 18]. Such fusion improves prediction by combining complementary signals, including semantic context from text and visual evidence from imaging.

Within this multimodal landscape, fusing time-series EHR and routine CXR provides a strong foundation for day-level trajectory tracking. Although clinical notes can add semantic detail, they are often irregular, subjective, and retrospective [3, 9], which can introduce reporting delays and observer bias. In contrast, time-series EHR (e.g., labs) and CXR serve as more direct proxies of physiological status: numerical signals provide immediate quantitative indicators, while CXR offers an observer-independent view of internal pathology such as pulmonary edema

progression that may not be captured by codes [18]. High-fidelity tracking of daily clinical evolution therefore benefits from emphasizing these objective snapshots. However, existing SOTA (state-of-the-art) systems such as MM-STGNN [18] and MuST [9] often aggregate variable-length admissions into static representations to exploit population-level structure, which can dilute intra-patient temporal granularity and weaken causal dependencies. This motivates an alternative design that prioritizes sequence fidelity over global graph topology, especially when the timing of physiological change is clinically decisive.

To bridge this gap, we propose MR.DEC (**Multimodal Readmission-Risk Prediction Decoder**), a daily-scale causal Transformer that preserves intra-admission temporal structure by modeling each hospitalization as a chronological token stream of daily EHR records interleaved with sporadic same-day CXR observations. **Our contributions are threefold:** (i) we introduce a new multimodal trajectory architecture that jointly models longitudinal EHR and intermittent CXR and achieves SOTA performance on 30-day readmission prediction, (ii) we propose an efficient daily-recording tokenization that supports longer sequences for fine-grained risk tracking across the full hospital stay without retraining separate unimodal models, and (iii) we develop clinically grounded training and analysis components, including disease-specific supervised contrastive regularization for robustness under label noise and imbalance, and gradient-based "Critical Days" identification that links predicted risk changes to day-level EHR and CXR.

2 Methods

Problem Formulation We formulate 30-day all-cause readmission prediction as a *causal in-hospital trajectory modeling* task, motivated by evidence that readmission risk evolves as clinical evidence accumulates during hospitalization [2, 4]. A single admission is represented as a chronological sequence of hospital days $\mathcal{S} = (d_1, d_2, \ldots, d_N)$ over N days, where each day is a multimodal record $d_n := \left(x_n^{\mathrm{EHR}}, x_n^{\mathrm{CXR}}\right)$. Here, x_n^{EHR} denotes the structured EHR observed on day n, and x_n^{CXR} denotes the chest radiograph acquired on day n (if available; otherwise $x_n^{\mathrm{CXR}} = \varnothing$). Given the full in-hospital history $\mathcal{H}_N = (d_1, \ldots, d_N)$, our model predicts the admission-level outcome

$$\hat{y} = f_\theta(\mathcal{H}_N),$$

while enforcing causality inside the decoder via masked self-attention so that each token aggregates evidence only from preceding tokens in the interleaved stream.

2.1 MR.DEC (Multimodal Readmission-risk prediction Decoder)

Hierarchical Multimodal Tokenization We encode the full admission history $\mathcal{H}_N$ as a day-aligned, variable-length token stream that reflects the fact that CXR is not acquired every day (Fig. 1). For each hospital day $n \in \{1, \ldots, N\}$,

textualized EHR is encoded by a frozen Clinical Encoder (e.g., BioClinical ModernBERT [15]) to obtain $f^{\mathrm{CLS}}_{\mathrm{EHR}_n}$, and same-day CXR (if present) is encoded by a frozen Vision Encoder (e.g., EVA-X-Base [20]) to obtain $f^{\mathrm{CLS}}_{\mathrm{CXR}_n}$.

We serialize day-wise tokens in chronological order by emitting one EHR token per day and inserting the CXR token immediately after it only on imaging days. We denote the resulting interleaved token features for the admission as

$$F = (f_1, \ldots, f_L),$$

where

$$F = \big[\underbrace{f^{\mathrm{CLS}}_{\mathrm{EHR}_1}, f^{\mathrm{CLS}}_{\mathrm{CXR}_1}}_{\text{Day 1}}, \underbrace{f^{\mathrm{CLS}}_{\mathrm{EHR}_2}}_{\text{Day 2}}, \underbrace{f^{\mathrm{CLS}}_{\mathrm{EHR}_3}, f^{\mathrm{CLS}}_{\mathrm{CXR}_3}}_{\text{Day 3}}, \ldots, \underbrace{f^{\mathrm{CLS}}_{\mathrm{EHR}_N}, f^{\mathrm{CLS}}_{\mathrm{CXR}_N}}_{\text{Day } N}\big].$$

Thus, the token length is

$$L = N + \sum_{n=1}^{N} \mathbb{I}\big[x_n^{\mathrm{CXR}} \neq \varnothing\big] .$$

Each token f_i has modality $m_i \in \{\mathrm{ehr}, \mathrm{cxr}\}$ and hospital-day index $n_i \in \{1, \ldots, N\}$. We project f_i into the model dimension d_{model} with modality-specific projections and add modality and day embeddings:

$$h_i = \mathrm{GELU}\big(W^{m_i}_{\mathrm{proj}} f_i + b^{m_i}_{\mathrm{proj}}\big) + M_{\mathrm{emb}}(m_i) + D_{\mathrm{emb}}(n_i), \tag{1}$$

where $M_{\mathrm{emb}}(\cdot)$ and $D_{\mathrm{emb}}(\cdot)$ denote learnable modality and day embedding lookups, respectively. We denote the resulting decoder input embedding sequence as $F_{\mathrm{emb}} := (h_1, \ldots, h_L)$.

Multimodal Readmission-Risk Prediction Decoder We model the admission trajectory by feeding the decoder input $\tilde{F}_{\mathrm{emb}} = [F_{\mathrm{emb}}, \texttt{[READMIT]}, \texttt{[SupCon]}]$ into a causal Transformer decoder (Fig. 1). Masked self-attention enforces chronological information flow: each position aggregates information only from preceding tokens, so the appended summary tokens can attend to the entire admission stream while respecting its temporal order.

Training objective. We jointly optimize two objectives using the appended summary tokens in $\tilde{F}_{\mathrm{emb}}$. Let $\mathbf{z}_{\mathrm{readmit}}$ and $\mathbf{z}_{\mathrm{supcon}}$ denote the final hidden states of `[READMIT]` and `[SupCon]`, respectively. We map $\mathbf{z}_{\mathrm{readmit}}$ to the admission-level probability $\hat{y}$ using an MLP classification head, and map $\mathbf{z}_{\mathrm{supcon}}$ to a projected representation $\tilde{\mathbf{z}}_{\mathrm{supcon}}$ using a separate MLP projection head. This decouples the predictive and contrastive signals, and the model is trained end-to-end with

$$\mathcal{L}_{\mathrm{total}} = \mathcal{L}_{\mathrm{readmit}} + \lambda\, \mathcal{L}_{\mathrm{SupCon}},$$

where $\lambda = 0.3$ scales the contrastive contribution.

Readmission prediction. We supervise the admission-level prediction with binary cross-entropy:

$$\mathcal{L}_{\text{readmit}} = -\big[y \log \hat{y} + (1-y)\log(1-\hat{y})\big]. \tag{2}$$

Disease-specific supervised contrastive learning. We apply supervised contrastive loss to the projected representation $\tilde{\mathbf{z}}_{\text{supcon}}$ to impose diagnosis-supergroup-aware structure in the latent space and improve robustness to noisy readmission labels. We define positives using a coarse diagnostic label `SUPER_GROUP`[1]; each admission is assigned a single `SUPER_GROUP` as its most frequent coarse ICD-10 category (14 total). For an anchor admission i in a batch $\mathcal{B}$, positives $\mathcal{P}(i)$ are admissions that share both the same readmission label y and the same `SUPER_GROUP`. To ensure positives exist for all anchors, each mini-batch includes at least two samples for every (`SUPER_GROUP`, y) combination, while combinations are sampled proportionally to their prevalence. Here, $\text{sim}(\cdot,\cdot)$ denotes cosine similarity and $\tau = 0.1$ is a temperature hyperparameter. Let $\tilde{\mathbf{z}}_i$ denote the projection-head output of the `[SupCon]` token for admission i. The loss is

$$\mathcal{L}_{\text{SupCon}} = \sum_{i\in\mathcal{B}} \frac{-1}{|\mathcal{P}(i)|} \sum_{p\in\mathcal{P}(i)} \log \frac{\exp(\text{sim}(\tilde{\mathbf{z}}_i, \tilde{\mathbf{z}}_p)/\tau)}{\sum_{a\in\mathcal{B}\setminus\{i\}} \exp(\text{sim}(\tilde{\mathbf{z}}_i, \tilde{\mathbf{z}}_a)/\tau)}. \tag{3}$$

3 Experiments

3.1 Experimental Setup

Dataset. We conducted experiments on MIMIC-IV [5] linked with MIMIC-CXR [6]. Following prior studies [9, 18], we include adult admissions (age $\geq$ 18) with a length of stay $\geq$ 48 hours and at least two CXRs during the stay. This yields 13,821 admissions from 11,972 patients. We apply a stratified patient-level split (90%/10%), yielding 12,438 training and 1,383 test admissions with no patient overlap. Daily EHR summaries are constructed by concatenating demographics, ICD-10 diagnosis subgroups, abnormal laboratory results, and medications grouped by therapeutic class, with duplicates removed. For imaging, we restrict to frontal-view radiographs (AP/PA).

Baselines. We compare MR.DEC against two categories of baselines. For a comprehensive evaluation, we include the reported metrics of MuST [9] as an established literature reference. To ensure a strictly fair and identical-cohort comparison, we prioritize a direct baseline evaluation by re-implementing MM-STGNN [18] on our unified MIMIC-IV cohort.

[1] We use diagnostic super-groups with the following categories: Circulatory (19.4%), Endocrine & Metabolic (13.1%), Symptoms & Signs (8.4%), Respiratory (7.9%), Trauma & Poisoning (7.6%), Digestive (7.2%), Genitourinary (5.9%), Factors & Services (5.2%), Blood & Immune (5.2%), Mental Disorders (4.9%), Nervous System (4.5%), Musculoskeletal & Skin (4.5%), Infectious (3.7%), and Neoplasms (2.5%).

LVLMs (Qwen3-VL [12], MedGemma [14], Lingshu [19]) serve as an upper-bound probe for general-purpose multimodal reasoning. Given their scale (4B–32B parameters) and broad medical pretraining, they represent a strong zero-shot alternative to task-specialized trajectory modeling, assessing whether purpose-built longitudinal architectures remain necessary in the foundation model era.[2]

3.2 Performance Analysis and Clinical Robustness

Table 1 compares MR.DEC against all baselines across modality settings via modality masking without retraining. In the multimodal (CXR–EHR) setting, MR.DEC achieves the best overall performance (AUC 0.814, F1 0.752), outperforming the primary task-matched baseline MM-STGNN by +0.014 in AUC and +0.184 in F1. These gains are especially meaningful under class imbalance, where ACC can be inflated by majority-class predictions: MM-STGNN attains high ACC (0.857) but much lower F1 (0.568), whereas MR.DEC improves positive-case detection while maintaining competitive ACC (0.744). Compared to its Binary Cross-Entropy(BCE)-only ablation, MR.DEC shows comparable AUC (0.814 vs. 0.809) but substantially higher Rec(N-C) (0.548 vs. 0.329), confirming that $\mathcal{L}_{\text{SupCon}}$ improves sensitivity to non-critical readmission cases under label noise and class imbalance—precisely where clinical value is highest.

Table 1. Performance comparison for 30-day readmission prediction. All models predict the same 30-day post-discharge outcome; the day limit refers to the length of the *input* in-hospital sequence, not the prediction horizon. Unless otherwise noted, models are trained on sequences of up to 10 hospital days; MR.DEC (Max 30-days) extends the input window to 30 days. For the CXR+EHR setting, we report outcome-stratified recall (y=1): *Rec(C)* for critical cases and *Rec(N-C)* for non critical cases. Bold and underlined values indicate the best and second-best performance, respectively. 'Params' and 'Dom.' denote trainable parameters and training domain (Gen.: General, Med.: Medical), respectively.

			CXR			EHR			CXR+EHR				
Model	**Params**	**Dom.**	**AUC**	**ACC**	**F1**	**AUC**	**ACC**	**F1**	**AUC**	**ACC**	**F1**	**Rec(C)**	**Rec(N-C)**
Readmission-specialized models													
MuST[1]	—	Med.	0.721	0.744	—	**0.784**	**0.850**	—	0.799	**0.859**	—	—	—
MM-STGNN	3M	Med.	0.769	**0.757**	0.477	0.781	0.848	0.556	0.800	0.857	0.568	0.714	0.214
Large Vision-Language Models (LVLMs)													
Qwen3-VL	32B	Gen.	0.525	0.550	0.595	0.609	0.500	0.667	0.605	0.505	0.669	**1.000**	0.007
MedGemma-1.5v	4B	Med.	0.546	0.554	0.625	0.517	0.496	0.659	0.461	0.462	0.600	0.753	0.030
MedGemma	27B	Med.	0.557	0.518	**0.670**	0.541	0.514	0.673	0.541	0.500	0.666	**1.000**	0.007
Lingshu	32B	Med.	0.517	0.507	0.635	0.561	0.527	**0.676**	0.489	0.518	0.672	**1.000**	0.038
Ours													
MR.DEC (BCE only)	19M	Med.	0.717	0.633	0.483	0.673	0.676	0.633	0.809	0.722	0.697	0.821	0.329
MR.DEC	19M	Med.	**0.787**	0.711	0.656	0.750	0.707	0.665	**0.814**	0.744	**0.752**	0.911	**0.548**
MR.DEC (Max 30-days)	19M	Med.	0.778	0.689	0.629	0.697	0.661	0.672	0.793	0.726	0.736	0.917	0.506

[2] Full prompting details for all LVLM baselines are available at https://github.com/yejix-ai/MR.DEC.

Table 2. Category-stratified performance for 30-day readmission prediction in the paired CXR–EHR setting. Categories are grouped by the primary source of clinical evidence (radiographic findings vs. structured EHR).

Model	Categories with Radiographic Manifestations									EHR-Predominant Categories					
	Circulatory			Respiratory			Infectious			Endocrine & Metabolic			Genitourinary		
	AUC	ACC	F1	AUC	ACC	F1	AUC	ACC	F1	AUC	ACC	F1	AUC	ACC	F1
Readmission-specialized models															
MM-STGNN	0.798	**0.766**	0.499	0.794	**0.745**	0.514	**0.799**	**0.733**	0.543	0.791	**0.763**	0.499	**0.812**	0.737	0.511
Large Vision-Language Models (LVLMs)															
Qwen3-VL	0.587	0.525	0.687	0.591	0.588	0.739	0.585	0.558	0.714	0.596	0.523	0.687	0.571	0.567	0.724
MedGemma-1.5v	0.472	0.483	0.616	0.458	0.500	0.654	0.424	0.470	0.613	0.473	0.480	0.615	0.456	0.500	0.639
MedGemma	0.518	0.520	0.684	0.513	0.584	0.737	0.561	0.553	0.710	0.512	0.521	0.685	0.485	0.571	0.725
Lingshu	0.467	0.530	0.686	0.432	0.584	0.734	0.481	0.567	0.717	0.478	0.528	0.686	0.468	0.574	0.725
Ours															
MR.DEC (BCE only)	0.803	0.716	0.699	0.791	0.701	0.714	0.794	0.692	0.720	**0.800**	0.712	0.701	0.787	0.688	0.710
MR.DEC	**0.815**	0.747	**0.761**	**0.798**	0.732	**0.771**	0.781	0.724	**0.770**	**0.800**	0.735	**0.756**	0.800	**0.743**	**0.782**
MR.DEC (Max 30-days)	0.795	0.733	0.750	0.761	0.716	0.760	0.780	0.722	0.769	0.789	0.720	0.739	0.780	0.725	0.768

To separate critical deterioration-driven sensitivity from non-critical readmission detection, we report outcome-stratified recall for readmission-positive admissions ($y = 1$). Specifically, we stratify the positive cohort into two subgroups based on clinical severity at discharge: *Critical* cases (i.e., end-of-life) and *Non-critical* readmission cases, denoted as Rec(C) and Rec(N-C), respectively. LVLM baselines show high Rec(C) but near-zero Rec(N-C) (0.007–0.038), indicating that their sensitivity is heavily concentrated on critical cases. In contrast, MR.DEC preserves strong Rec(C) (0.911) while substantially improving Rec(N-C) to 0.548. MR.DEC (Max 30-days) extends the temporal window to 30 days for longer-stay admissions that baselines cannot natively handle. Despite the longer and noisier sequences, it remains highly competitive (AUC 0.793, F1 0.736) with a markedly higher F1 than MM-STGNN (+0.168), demonstrating graceful scalability without architectural modification.

Consistent Performance Across Clinical Categories Table 2 reports category-stratified results for 30-day readmission prediction in the paired CXR–EHR setting. The five diagnostic groups span conditions with strong radiographic correlates (Circulatory, Respiratory, Infectious) as well as EHR-predominant categories where risk-relevant evidence is primarily reflected in structured daily records (Endocrine & Metabolic, Genitourinary). Across all groups, MR.DEC maintains consistently strong performance (F1: 0.756–0.782), indicating that it leverages multimodal evidence in a category-appropriate manner. In contrast, general-purpose LVLM baselines exhibit only moderate discrimination, with substantially lower AUC (roughly 0.42–0.60), suggesting limited stability for specialized clinical risk stratification. Notably, MR.DEC retains robust category-wise performance when extending the temporal window to 30 days: MR.DEC (Max

[1] Due to the absence of publicly available source code, performance metrics for MuST are cited from the original publication. F1-scores were not available and are denoted by '–'.

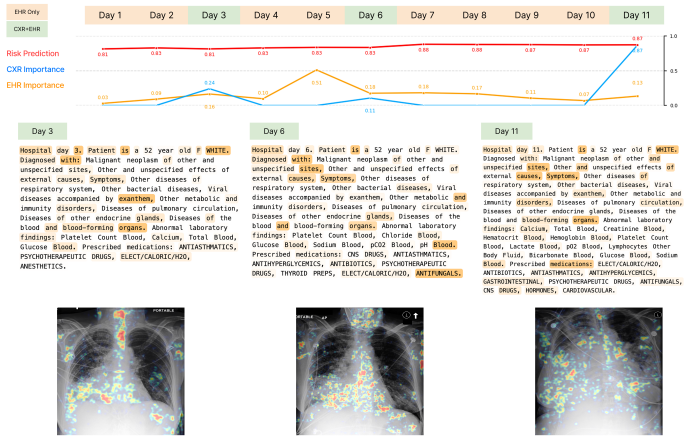

Fig. 2. Readmission prediction in a deteriorating oncology admission. CXRs were obtained only on Hospital Days 3, 6, and 11 (shown with modality-specific heatmaps). Risk rises from 0.81, peaks at 0.88 on Day 7, and remains at 0.87 through Day 11.

30-days) achieves F1 of 0.739–0.769 across groups, supporting its scalability to longer admission trajectories without sacrificing consistency.

Interpretability and Clinical Trajectory Analysis Figure 2 illustrates MR.DEC's trajectory-level interpretability on a representative deteriorating oncology admission. We define *Critical Days* as hospital days where token-level attributions peak and predicted risk changes most markedly, providing temporally grounded signals for real-time risk stratification. Heatmaps are derived by back-projecting gradient-based importance scores from the `[READMIT]` token through the decoder sequence to the original EHR tokens and CXR patch embeddings. CXRs were acquired only on Days 3, 6, and 11; modality-specific heatmaps are shown at these time points. The predicted risk rises from 0.81 to 0.87 by discharge, with EHR heatmaps emphasizing diagnosis and comorbidity tokens throughout, where Day 6 highlights anti-infective medications and Day 11 shifts toward abnormal-lab expressions. In parallel, CXR heatmaps show lung-field activations expanding from Day 3 and peaking by Day 11, together yielding temporally coherent cross-modal explanations.

4 Conclusion

We presented MR.DEC (**M**ultimodal **r**eadmission-risk prediction **Dec**oder), a framework that redefines readmission prediction as a daily-scale causal trajectory modeling task. By jointly modeling longitudinal EHR and intermittent CXR in a single causal decoder, MR.DEC outperforms both graph-based and large

foundation model baselines, particularly in identifying critical deterioration trajectories. Disease-Specific Supervised Contrastive Learning further captures inter-patient phenotypic structure without explicit graph construction, and token-level interpretability provides temporally grounded signals for real-time risk stratification. Future work will explore extending the multimodal token stream to additional clinical modalities such as ECG and clinical notes, broadening the framework's applicability across diverse inpatient settings.

Acknowledgements This work was supported by the TIPS (Tech Incubator Program for Startup) Program of the Ministry of SMEs and Startups and Korea Startup Foundation under Grant No. RS-2025-25467010.

Disclosure of Interests The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Alarcón, I.D.S., Santana-Velásquez, A., Duitama, M.J.F., Salazar-Sánchez, M.B., Hernández-Arango, A.: Prediction of readmission risk in hospital patients using artificial intelligence techniques. In: 2024 3rd International Congress of Biomedical Engineering and Bioengineering (CIIBBI). pp. 1–6. IEEE (2024)
2. Davis, S., Greiner, R.: Survival models and longitudinal medical events for hospital readmission forecasting. BMC Health Services Research **24**(1), 1394 (2024)
3. Jiang, L.Y., Liu, X.C., Nejatian, N.P., Nasir-Moin, M., Wang, D., Abidin, A., Eaton, K., Riina, H.A., Laufer, I., Punjabi, P., et al.: Health system-scale language models are all-purpose prediction engines. Nature **619**(7969), 357–362 (2023)
4. Jiang, W., Siddiqui, S., Barnes, S., Barouch, L.A., Korley, F., Martinez, D.A., Toerper, M., Cabral, S., Hamrock, E., Levin, S.: Readmission risk trajectories for patients with heart failure using a dynamic prediction approach: Retrospective study. JMIR Medical Informatics **7**(4), e14756 (2019)
5. Johnson, A., Bulgarelli, L., Pollard, T., Gow, B., Moody, B., Horng, S., Celi, L.A., Mark, R.: MIMIC-IV. PhysioNet (2024), version 3.1, DOI: 10.13026/kpb9-mt58
6. Johnson, A., Pollard, T., Mark, R., Berkowitz, S., Horng, S.: MIMIC-CXR database. PhysioNet (2024), version 2.1.0, DOI: 10.13026/4jqj-jw95
7. Lax, Y., Martinez, M., Brown, N.M.: Social determinants of health and hospital readmission. Pediatrics **140**(5), e20171427 (2017)
8. Lv, J., Zhang, M., Fu, Y., Chen, M., Chen, B., Xu, Z., Yan, X., Hu, S., Zhao, N.: An interpretable machine learning approach for predicting 30-day readmission after stroke. International Journal of Medical Informatics **174**, 105050 (2023)

9. Miao, Y., Yu, L.: MuST: Multimodal spatiotemporal graph-transformer for hospital readmission prediction. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 276–285. Springer (2023)
10. Navathe, A.S., Zhong, F., Lei, V.J., Chang, F.Y., Sordo, M., Topaz, M., Navathe, S.B., Rocha, R.A., Zhou, L.: Hospital readmission and social risk factors identified from physician notes. Health Services Research **53**(2), 1110–1136 (2018)
11. Okolie, A., Bello, A., Ikhifa, M.O., Ibiyeye, A.O., Agbeso, D.O., Alumona, P.: Machine learning approaches for predicting 30-day hospital readmissions: Evidence from Massachusetts healthcare data. World Journal of Advanced Research and Reviews **28**(1), 1–12 (2025)
12. Qwen Team: Qwen3 technical report (2025)
13. Reddy, N.K., Anirudh, T., Varshitha, T., Kumar, B.C., Kumar, N.G., Jabbar, M.: Hospital readmission risk predictor for diabetic patients using XGBoost. In: 2025 3rd World Conference on Communication & Computing (WCONF). pp. 1–6. IEEE (2025)
14. Sellergren, A., Kazemzadeh, S., Jaroensri, T., Kiraly, A., Traverse, M., Kohlberger, T., Xu, S., Jamil, F., Hughes, C., Lau, C., et al.: MedGemma technical report (2025)
15. Sounack, T., Davis, J., Durieux, B., Chaffin, A., Pollard, T.J., Lehman, E., Johnson, A.E., McDermott, M., Naumann, T., Lindvall, C.: BioClinical ModernBERT: A state-of-the-art long-context encoder for biomedical and clinical NLP (2025)
16. Sterling, M.R., Ringel, J.B., Pinheiro, L.C., Safford, M.M., Levitan, E.B., Phillips, E., Brown, T.M., Nguyen, O.K., Goyal, P.: Social determinants of health and 30-day readmissions among adults hospitalized for heart failure in the REGARDS study. Circulation: Heart Failure **15**(1), e008409 (2022)
17. Subasi, A.: Hospital readmission forecasting using artificial intelligence. In: Applications of Artificial Intelligence in Healthcare and Biomedicine, pp. 455–520. Elsevier (2024)
18. Tang, S., Tariq, A., Dunnmon, J.A., Sharma, U., Elugunti, P., Rubin, D.L., Patel, B.N., Banerjee, I.: Predicting 30-day all-cause hospital readmission using multimodal spatiotemporal graph neural networks. IEEE Journal of Biomedical and Health Informatics **27**(4), 2071–2082 (2023)
19. Xu, W., Chan, H.P., Li, L., Aljunied, M., Yuan, R., Wang, J., Xiao, C., Chen, G., Liu, C., Li, Z., et al.: Lingshu: A generalist foundation model for unified multimodal medical understanding and reasoning (2025)
20. Yao, J., Wang, X., Song, Y., Zhao, H., Ma, J., Chen, Y., Liu, W., Wang, B.: EVA-X: A foundation model for general chest X-ray analysis with self-supervised learning. npj Digital Medicine **8**(1), 678 (2025)