EMORY UNIVERSITY | Department of Data and Decision Sciences

EMORY | Undergraduate Research Programs Pathways Center

# Causal Representation Learning under Informative Missingness for Clinical Multimodal Prediction and Offline Decision-Making

**Zihan Liang**
Emory College of Arts and Sciences
Department of Data and Decision Sciences
**Advisor: Prof. Ruoxuan Xiong**

## Clinical Missingness Is MNAR

- Clinical EHR data are **systematically incomplete** rather than randomly incomplete.
- The probability that a modality or measurement is observed often depends on **latent patient state and clinician decision-making**.
- This is a **missing-not-at-random (MNAR) problem**, not a nuisance that can be removed by imputation alone.
- We study two scales of informative missingness:
  - **Patient-level MMNAR**: which modalities are available for a patient.
  - **Step-level monitoring intensity MNAR**: how often variables are measured over time.

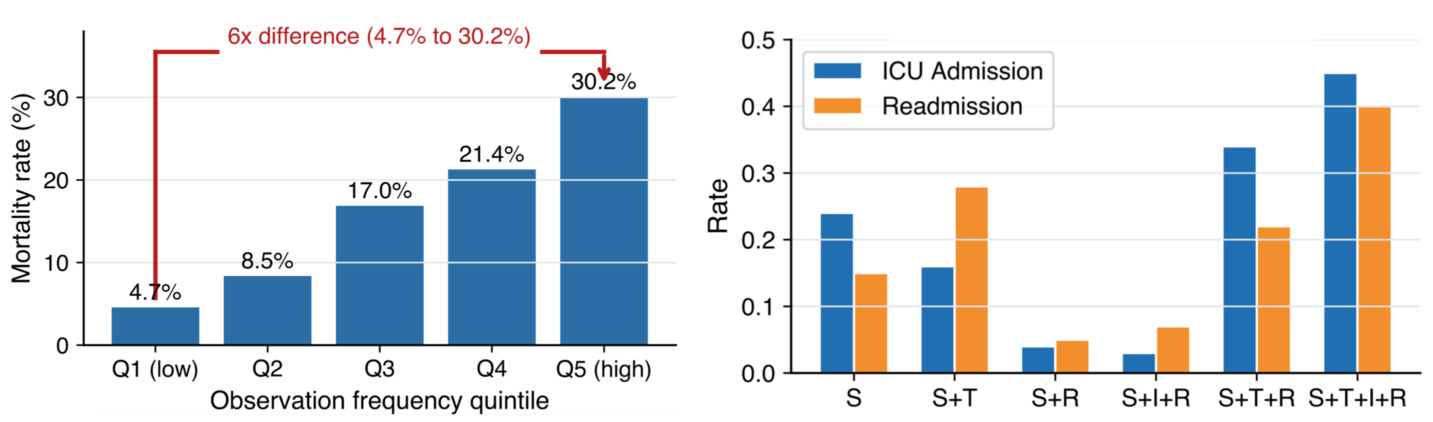


**Figure 1 and 2.** Informative missingness in clinical data.

## The Thesis Question and Dataset

***How should observation patterns in multimodal clinical data — at both patient and step levels — be modeled to learn representations that are sufficient for prediction and control, and what methodological interventions correct the distinct failure modes that arise when such patterns are ignored?***

**INSTANTIATION I: STATIC EPISODE-LEVEL PREDICTION.**

- MIMIC-IV cohort: 20,000 adult ICU patients.
- Modalities: structured EHR, chest X-ray, discharge summary, radiology report.
- Tasks: 30-day readmission, post-discharge ICU admission, 180-day mortality.

**INSTANTIATION II: SEQUENTIAL DECISION-MAKING.**

- MIMIC-III, MIMIC-IV and eICU cohort.
- 72-hour window → H = 18 decision steps at 4-hour intervals
- Action space: IV fluids and vasopressors.

## Unified Causal View of MNAR in Clinical Multimodal Learning

**STATIC SETTING (PATIENT-LEVEL MODALITY MISSINGNESS)**

- Latent health $h_i^*$ → multimodal contents $x_i$
- Latent health $h_i^*$ → observation pattern $\delta_i$
- Latent health $h_i^*$ → clinical outcome $y_i$
- Critical extra edge: $\delta_i \to y_i$ directly *(↑tests → earlier interv. → chg. outcome)*

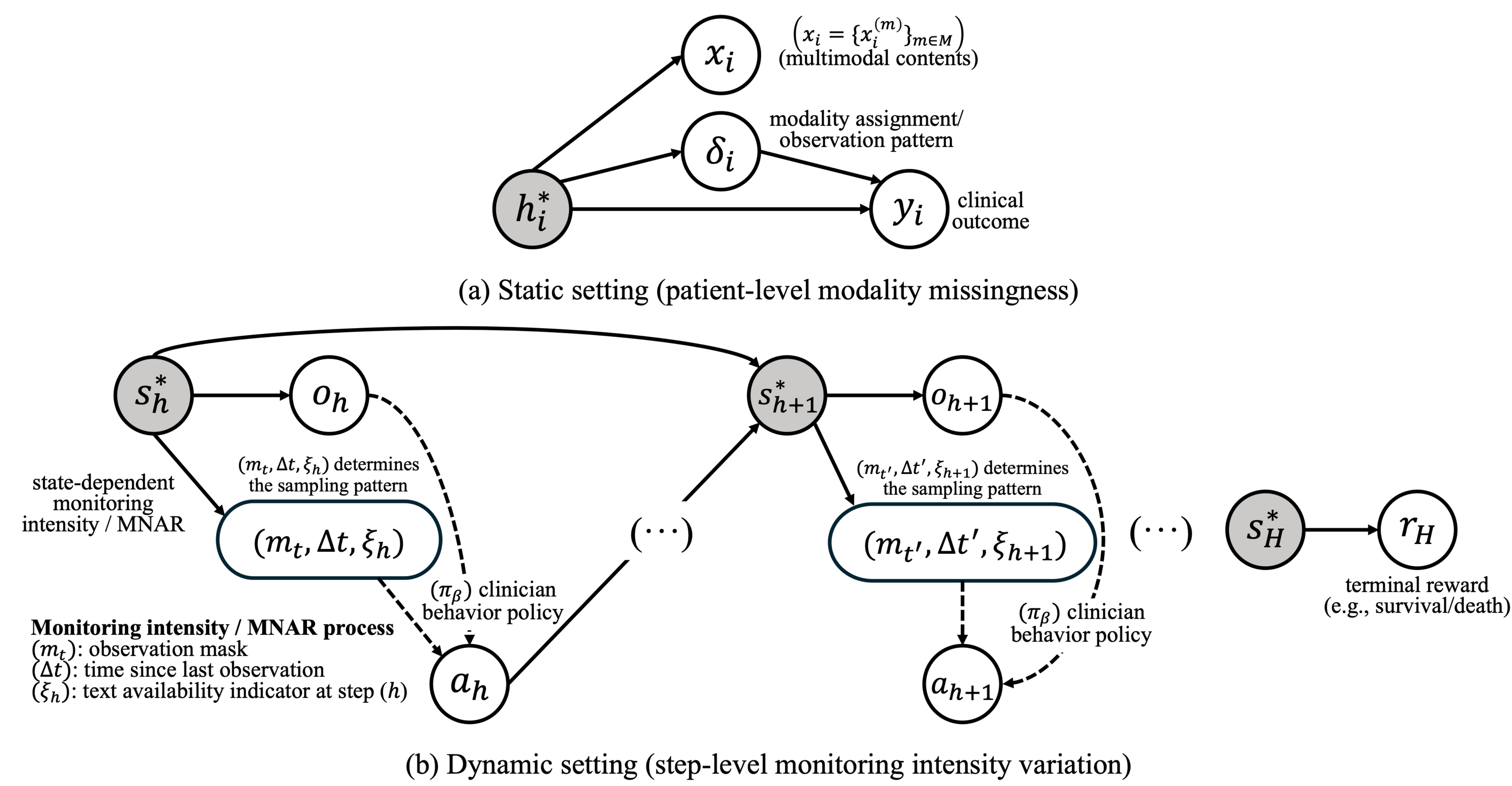


**Figure 3.** Unified causal structure for clinical multimodal learning under MNAR.

**DYNAMIC SETTING (STEP-LEVEL MONITORING INTENSITY)**

- Latent state $s_h^*$ → observations $o_h$ and monitoring intensity $(m_t, \Delta t, \xi_h)$
- $(s_h^*, a_h) \to s_{h+1}^*$ *(treatment changes the patient's state)*
- Terminal state $s_H^*$ → reward $r_H$ *(survival or death)*

## Methodology

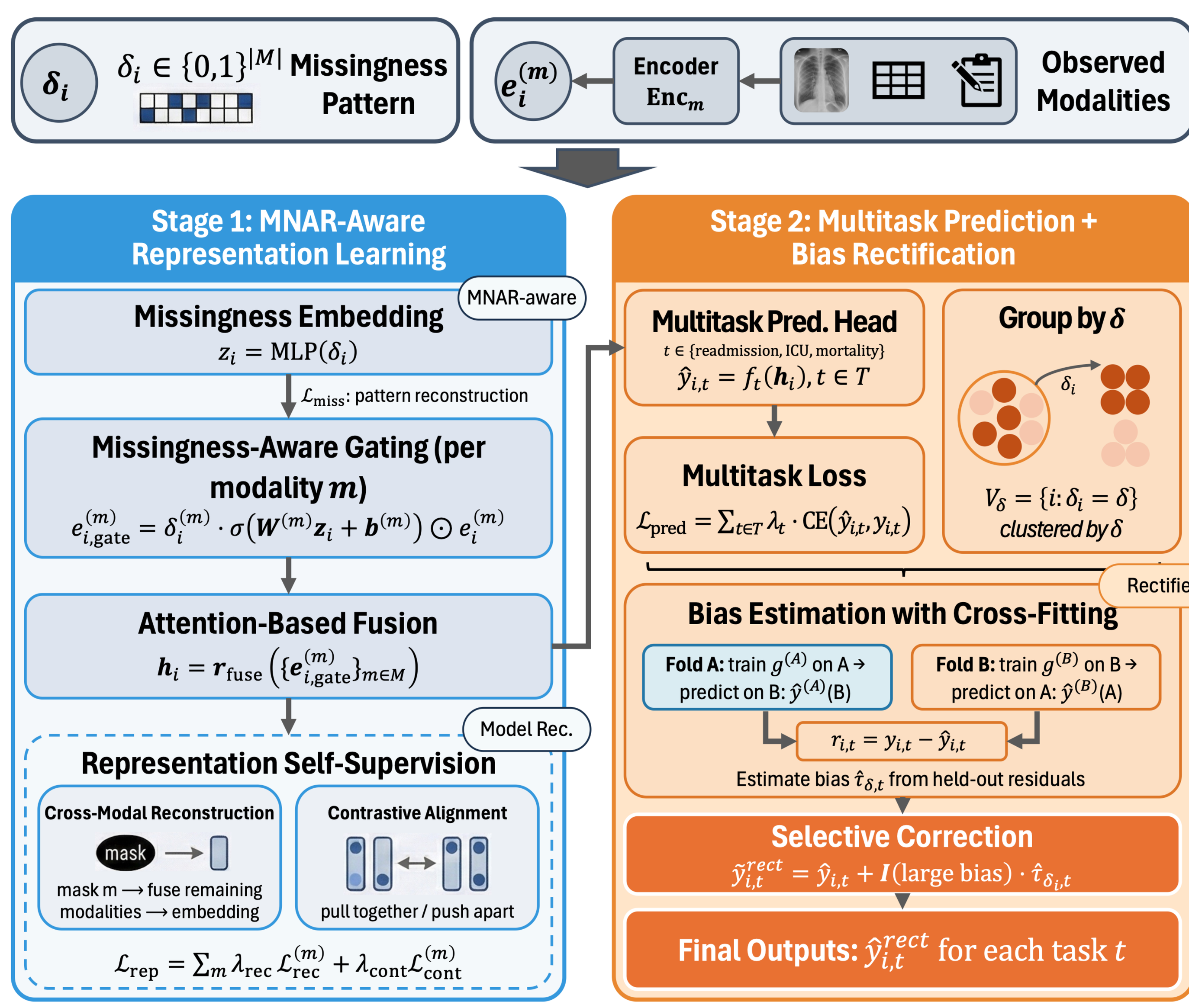


**Figure 4.** Instantiation I (static MMNAR prediction): two-stage pipeline for patient-level multimodal learning under informative missingness.

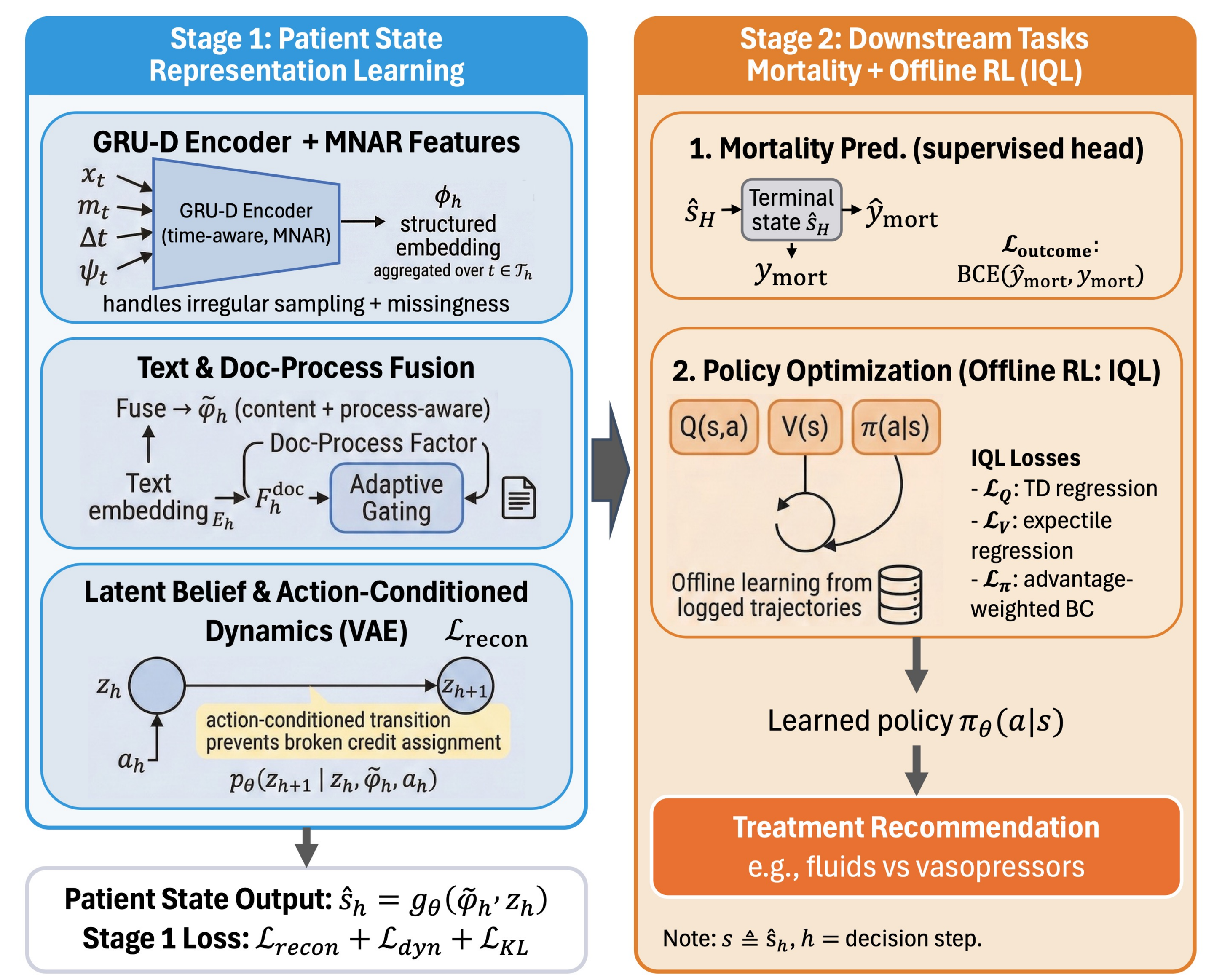


**Figure 5.** Instantiation II (dynamic MNAR offline control): MNAR-aware state learning and downstream policy optimization.

## Results

### Results I: Static Prediction under MMNAR

| Task | Best Baseline | Ours | Gain |
|---|---|---|---|
| ICU Admission (MIMIC-IV) | 0.869 (DrFuse) | **0.982** | **+13.0%** |
| 30-Day Readmission (MIMIC-IV) | 0.799 (MUSE+) | **0.866** | **+8.4%** |
| In-Hospital Mortality (MIMIC-IV) | 0.905 (MUSE+) | **0.947** | **+4.6%** |

**Our method outperforms all 13 baselines on both MIMIC-IV and eICU.**

### Results II: Offline Decision-Making under Dynamic MNAR

| Method | MIMIC-IV FQE | vs. Clinician | eICU FQE | vs. Clinician |
|---|---|---|---|---|
| Clinician (BC) | 0.521 | — | 0.534 | — |
| CQL | 0.418 | −19.8% | 0.408 | — |
| Tabular Q (AI Clinician) | 0.491 | −5.8% | 0.478 | — |
| **Ours (IQL)** | **0.634** | **+21.7%** | **0.604** | **+13.1%** |

**Our policy outperforms clinician behavior and all RL baselines on both datasets.**

### Robustness and Cross-Institutional Generalization

| Shift Severity | Ours | MUSE+ | Gap |
|---|---|---|---|
| None (original) | 0.866 | 0.799 | +6.7% |
| Mild (KL = 0.2) | 0.854 | 0.771 | +8.3% |
| Moderate (KL = 0.5) | 0.838 | 0.734 | +10.4% |
| **Severe (KL = 1.0)** | **0.812** | **0.687** | **+12.5%** |

| Protocol | Train → Test | FQE | vs. Clinician |
|---|---|---|---|
| In-distribution | MIMIC-IV → MIMIC-IV | 0.634 | +21.7% |
| **Zero-shot transfer** | **MIMIC-IV → eICU** | **0.568** | **+6.4%** |
| Fine-tuned transfer | MIMIC-IV + eICU FT → eICU | 0.594 | +11.2% |

**Related work accepted at EMNLP 2025 (Main) and ACL 2026 (Findings).** **github.com/CausalMLResearch/**