MODELING LONGITUDINAL CLINICAL DYNAMICS WITH COVARIANCE-AWARE NEURAL KALMAN FILTERING FOR MULTIMODAL SURVIVAL PREDICTION

Anonymous authors
Paper under double-blind review

ABSTRACT

Cancer prognosis reflects both baseline molecular characteristics and clinical evolution during follow-up. We propose a covariance-aware neural Kalman framework that integrates baseline RNA-seq and demographic covariates with irregular longitudinal clinical observations. A neural observation encoder maps clinical visits into a latent state space, where explicit covariance propagation and an analytically computed Kalman gain govern sequential state updates. Elapsed-time-dependent process covariance incorporates irregular visit intervals into the correction mechanism. Independently trained molecular and clinical predictors are combined through risk-level fusion selected by inner cross-validation. We evaluate the framework on 556 patients with 95 observed events from the MMRF CoMMpass multiple myeloma cohort using a 365-day landmark design and five repeated nested cross-validation partitions. RNA+CovarianceKalman achieves the highest mean pooled out-of-fold C-index among the evaluated models, 0.716 ± 0.013, compared with 0.662 ± 0.028 for the static molecular model and 0.702 ± 0.012 for RNA+GRU. Its one- and two-year post-landmark AUCs are 0.743 ± 0.019 and 0.740 ± 0.021, respectively, with a two-year Brier score of 0.121 ± 0.004. Ablations support the contributions of covariance-derived correction and elapsed-time-dependent process covariance. These results demonstrate the value of covariance-aware state estimation for integrating molecular risk with longitudinal clinical trajectories in this cohort.

1 INTRODUCTION

Accurate survival prediction is an important component of precision oncology. Modern cancer cohorts contain heterogeneous prognostic information, including molecular profiles, clinical variables, medical imaging, and histopathology. High-dimensional molecular measurements can characterize tumor biology that is not fully reflected by conventional clinical variables, motivating multimodal survival models that integrate complementary patient information (Tran et al., 2025; Hu et al., 2025; Nikolaou et al., 2025). Recent work has extended this paradigm through graph learning, neural representation learning, multimodal uncertainty modeling, and structured state-space architectures (Gogoshin & Rodin, 2023; Guo et al., 2025; Vaithiyanathan & Bala, 2026; Chen et al., 2025b). Neural survival models have also been studied across multiple cancer types, while systematic comparisons continue to emphasize the importance of strong Cox-based references (Yan et al., 2024; Bae et al., 2023; Huang et al., 2025).

Multimodal survival models that integrate measurements acquired near baseline capture complementary molecular, pathological, and imaging information. Extending these models to longitudinal clinical observations enables prognosis to reflect the patient's subsequent clinical trajectory. Baseline transcriptomics and later clinical observations encode complementary prognostic information: molecular profiles characterize tumor biology, whereas repeated laboratory and clinical assessments track time-varying clinical state. Representations of longitudinal clinical trajectories can therefore complement baseline molecular predictors (Rastogi et al., 2025).

Longitudinal clinical data are irregularly sampled, partially observed, and generated by an unobserved clinical state. GRUs and Transformers summarize such histories through implicit temporal states

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

(Nguyen et al., 2024; Kadhim et al., 2026). We introduce an explicit state-estimation inductive bias that tracks both the latent clinical state and its covariance across visits.

Kalman filtering provides a natural framework for sequential state estimation: a latent state is predicted, covariance is propagated, an incoming observation defines an innovation, and a covariance-derived gain determines the correction of the prior state. Classical filtering assumes a specified observation space and dynamical model. Neural-assisted filters learn components that are difficult to specify analytically (Revach et al., 2022; Chen et al., 2025a; Sabbaqi et al., 2026). Other hybrid neural–Kalman constructions further demonstrate the broader utility of combining learned representations with model-based recursions (Milstein et al., 2024).

We combine these ideas in a **covariance-aware neural Kalman state estimator**. A neural observation encoder maps each irregular clinical visit into a low-dimensional latent clinical-state space. Within this learned space, the model retains explicit Kalman covariance propagation: process uncertainty grows with elapsed visit time, predicted and measurement covariance determine the Kalman gain, and both latent state and covariance are updated after each observation. The final posterior state is mapped to a clinical Cox risk.

The longitudinal branch is combined with a baseline RNA-seq branch through prediction-level late fusion. The branches are optimized independently, and the fusion coefficient is selected from inner out-of-fold (OOF) predictions. We use a 365-day landmark design: patients must be observed beyond day 365, clinical inputs are restricted to days 0–365, and the target is residual post-landmark survival.

Across five repeated nested-CV partitions, RNA+CovarianceKalman achieves a pooled OOF C-index of 0.716 ± 0.013, compared with 0.662 ± 0.028 for static molecular prediction and 0.702 ± 0.012 for RNA+GRU. Our main contributions are: (1) a neural clinical observation encoder coupled to explicit latent covariance propagation and analytically computed Kalman correction; (2) a multimodal landmark prediction framework that combines this longitudinal state estimator with baseline molecular risk through independently trained branches; and (3) a repeated nested cross-validation evaluation showing consistent gains over the evaluated baselines and quantifying the contributions of covariance-derived gain and elapsed-time-dependent process covariance.

2 Related Work

Multimodal cancer survival prediction. Cancer survival modeling has expanded from conventional clinical regression toward integration of genomic, transcriptomic, epigenomic, imaging, and histopathological information (Tran et al., 2025). CATfusion integrates histology with genomic and epigenomic modalities through cross-attention (Hu et al., 2025), while SurvMamba uses selective state-space modeling for pathology–transcriptomic interactions (Chen et al., 2025b). Prediction-level fusion has also been explored for heterogeneous cancer data (Nikolaou et al., 2025). Related oncology work includes graph neural networks, deep imaging survival prediction, contractive multi-omics representation learning, and multimodal uncertainty modeling (Gogoshin & Rodin, 2023; Rastogi et al., 2025; Guo et al., 2025; Vaithiyanathan & Bala, 2026). These studies establish multimodal prognosis as a strong modeling paradigm, while repeated clinical follow-up remains largely outside their scope.

Longitudinal clinical modeling. RNNs, GRUs, and Transformers are widely used to learn temporal dependencies from longitudinal medical records. CLIMAT combines imaging and auxiliary clinical information for disease-trajectory forecasting (Nguyen et al., 2024), and state-space Transformers have been applied to clinical deterioration prediction (Kadhim et al., 2026). Landmark survival prediction provides a common prediction origin for incorporating longitudinal histories into censored survival analysis. Within this setting, our method uses explicit state and covariance recursions to encode irregular clinical trajectories.

Model-based and neural state estimation. Classical Kalman filtering propagates both latent state and covariance, with the gain determined by the relative uncertainty of prediction and observation. Neural model-based approaches learn components that are difficult to specify analytically. KalmanNet replaces classical gain computation with a recurrently learned gain while retaining the prediction–correction structure (Revach et al., 2022); later extensions consider meta-learning and graph-temporal dynamics (Chen et al., 2025a; Sabbaqi et al., 2026). Hybrid Kalman/neural designs have also appeared

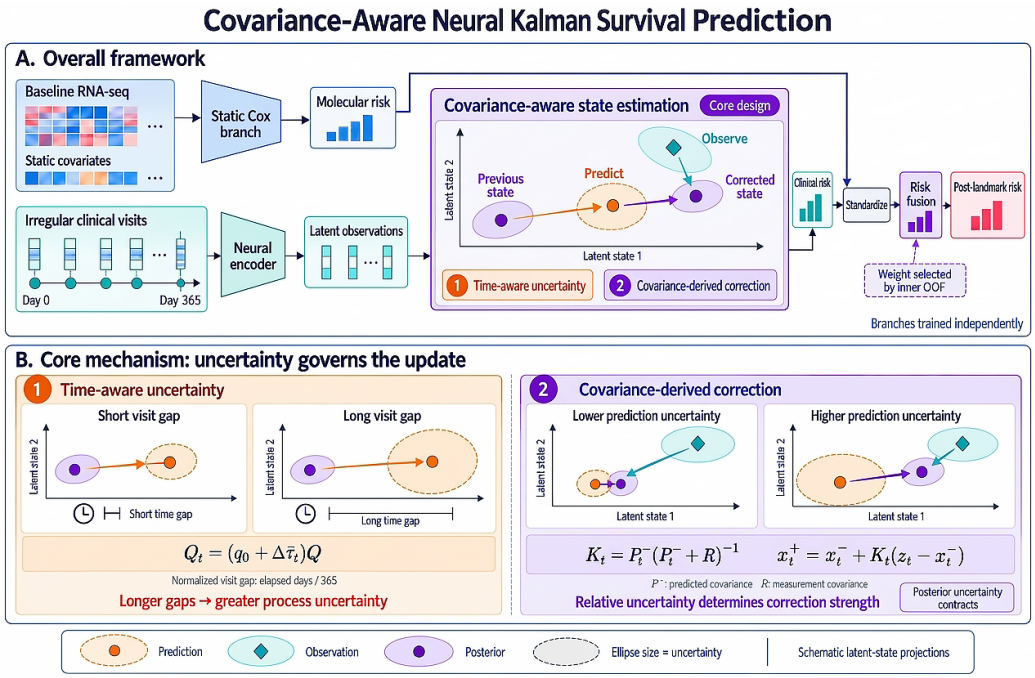


Figure 1: **Overview of the proposed covariance-aware neural Kalman survival framework. (A) Overall framework.** Baseline RNA-seq and static covariates are mapped to a molecular Cox risk, while irregular clinical visits from days 0–365 are encoded as latent observations and processed by a covariance-aware state estimator. The longitudinal branch performs sequential prediction, observation-driven correction, and covariance updating to produce a clinical Cox risk. The molecular and clinical risks are standardized and combined by prediction-level fusion, with the fusion weight selected from inner out-of-fold predictions; the two branches are trained independently. **(B) Core mechanism.** The state estimator uses two coupled covariance-aware operations. First, process covariance increases with the normalized inter-visit interval, adding a larger process-noise contribution to the predicted covariance when visits are farther apart. Second, the predicted covariance and measurement covariance determine the Kalman gain, which controls how strongly the encoded observation corrects the predicted state. Ellipse size schematically represents latent-state covariance, illustrating the process-noise contribution during prediction and covariance contraction during correction.

in other sequential domains (Milstein et al., 2024). Our approach learns the clinical observation representation and latent transition while retaining explicit covariance propagation and a covariance-derived gain. Heterogeneous clinical measurements are thereby embedded into a tractable latent state space with structured prediction–correction dynamics.

3 Method

We combine a static molecular risk branch with a covariance-aware neural Kalman encoder for longitudinal clinical trajectories. The neural component maps irregular observations into a latent state space; within that space, the longitudinal branch performs explicit state prediction, covariance propagation, innovation-based correction, and posterior covariance update. Figure 1 summarizes the framework.

3.1 Landmark Survival Formulation

For patient i, let T_i denote observed overall-survival time and $\delta_i \in \{0, 1\}$ the event indicator. We fix a landmark $L = 365$ days, retain patients satisfying $T_i > L$, and predict residual survival

$$\widetilde{T}_i = T_i - L. \tag{1}$$

Each patient has a static representation $\mathbf{s}_i$ containing baseline RNA-seq and time-invariant covariates and a pre-landmark history

$$\mathcal{X}_i^{(L)} = \{(\tau_{i,t}, \mathbf{y}_{i,t}, \mathbf{m}_{i,t})\}_{t=1}^{n_i}, \qquad 0 \le \tau_{i,t} \le L, \tag{2}$$

where $\mathbf{y}_{i,t} \in \mathbb{R}^D$ is the clinical measurement vector, $\mathbf{m}_{i,t} \in \{0,1\}^D$ is its observation mask, and $\Delta\tau_{i,t} = \tau_{i,t} - \tau_{i,t-1}$. No measurement after day 365 enters the prediction model.

3.2 Static Molecular Risk Encoder

Because RNA-seq is high-dimensional relative to cohort size, all preprocessing is fold-contained. Within each training partition, the 5,000 most variable genes are retained, standardized with training statistics, and reduced to 64 principal components. These components are concatenated with predefined static demographic covariates. Two compact MLP Cox predictors with different regularization strengths produce training-standardized risks $\widehat{r}_i^{\text{strong}}$ and $\widehat{r}_i^{\text{light}}$, combined as

$$\widehat{r}_i^{\text{stat}} = (1-\lambda)\widehat{r}_i^{\text{strong}} + \lambda\widehat{r}_i^{\text{light}}, \qquad \lambda \in \{0, 0.05, \dots, 1\}, \tag{3}$$

with λ selected by inner-OOF concordance. We refer to this as the *static molecular branch*; it contains RNA-seq together with the predefined static covariates.

3.3 Longitudinal Observation Encoding

At visit t, clinical variables are standardized using the corresponding training-fold statistics. Missing entries are numerically filled with training medians while the original mask is retained. We append normalized absolute visit time and inter-visit time,

$$\mathbf{u}_{i,t} = [\widetilde{\mathbf{y}}_{i,t}; \mathbf{m}_{i,t}; \bar{\tau}_{i,t}; \Delta\bar{\tau}_{i,t}], \quad \bar{\tau}_{i,t} = \tau_{i,t}/365, \quad \Delta\bar{\tau}_{i,t} = \Delta\tau_{i,t}/365. \tag{4}$$

A neural encoder maps each visit to

$$\mathbf{z}_{i,t} = E_\phi(\mathbf{u}_{i,t}) \in \mathbb{R}^H, \qquad H = 16. \tag{5}$$

The encoder uses a 24-dimensional hidden layer, layer normalization, ReLU, and dropout with rate 0.05. The first valid observation initializes $\mathbf{x}_{i,1}^+ = \mathbf{z}_{i,1}$ and $P_{i,1}^+ = \mathrm{Diag}(\mathbf{p}_0)$, where $\mathbf{p}_0 = \mathrm{softplus}(\widetilde{\mathbf{p}}_0) + \epsilon$ ensures positive initial variance.

3.4 Time-Aware State and Covariance Prediction

For later visits, the model predicts both latent clinical state and covariance before incorporating the current observation. We parameterize a bounded latent transition generator $A = 0.1\tanh(\widetilde{A})$ and define

$$F_{i,t} = I + \Delta\bar{\tau}_{i,t}A, \qquad \mathbf{x}_{i,t}^- = F_{i,t}\mathbf{x}_{i,t-1}^+. \tag{6}$$

The positive diagonal base process covariance is $Q = \mathrm{Diag}(\mathbf{q})$ with $\mathbf{q} = \mathrm{softplus}(\widetilde{\mathbf{q}}) + \epsilon$. To reflect irregular sampling, process uncertainty increases with elapsed time,

$$Q_{i,t} = (q_0 + \Delta\bar{\tau}_{i,t})Q, \qquad q_0 = 0.05, \tag{7}$$

and prior covariance is propagated as

$$P_{i,t}^- = F_{i,t}P_{i,t-1}^+F_{i,t}^\top + Q_{i,t}. \tag{8}$$

Longer visit intervals increase the additive process-noise contribution to the predicted covariance, allowing elapsed time to influence the subsequent Kalman correction.

3.5 Covariance-Derived Kalman Correction

In the learned latent space we use identity observation matrix $H_{\text{obs}} = I$ and positive diagonal measurement covariance $R = \mathrm{Diag}(\mathbf{r})$, with $\mathbf{r} = \mathrm{softplus}(\widetilde{\mathbf{r}}) + \epsilon$. The innovation, innovation covariance, and Kalman gain are

$$\mathbf{e}_{i,t} = \mathbf{z}_{i,t} - \mathbf{x}_{i,t}^-, \qquad S_{i,t} = P_{i,t}^- + R, \tag{9}$$

$$K_{i,t} = P_{i,t}^-S_{i,t}^{-1} = P_{i,t}^-(P_{i,t}^- + R)^{-1}. \tag{10}$$

The gain is computed analytically from the propagated covariance and the learned measurement covariance. It varies with the visit schedule through the covariance recursion, while the innovation incorporates the encoded clinical observation into the state update. State and covariance are corrected by

$$\mathbf{x}_{i,t}^{+} = \mathbf{x}_{i,t}^{-} + K_{i,t}\mathbf{e}_{i,t}, \tag{11}$$

$$P_{i,t}^{+} = (I - K_{i,t})P_{i,t}^{-}(I - K_{i,t})^{\top} + K_{i,t}RK_{i,t}^{\top}. \tag{12}$$

We use the Joseph form for the covariance update. Padded sequence positions do not update either state or covariance. After the final valid visit n_i, the clinical survival risk is

$$r_i^{\text{clin}} = \mathbf{w}_r^{\top}\mathbf{x}_{i,n_i}^{+} + b_r. \tag{13}$$

The covariance recursion and observation encoder are optimized jointly through the downstream survival objective.

3.6 Censored Survival Learning and Multimodal Fusion

Each modality-specific predictor is trained independently using Cox partial likelihood. For an event patient i, define the post-landmark risk set $\mathcal{R}_i = \{j : \widetilde{T}_j \geq \widetilde{T}_i\}$. The negative partial log-likelihood is

$$\mathcal{L}_{\text{Cox}} = -\sum_{i:\delta_i=1}\left[r_i - \log\sum_{j\in\mathcal{R}_i}\exp(r_j)\right], \tag{14}$$

with Breslow handling of tied event times. Static and clinical risks are standardized using outer-training statistics, $\mathcal{Z}(r) = (r - \mu_{\text{train}})/\sigma_{\text{train}}$, and fused as

$$r_i^{\text{fusion}} = (1-\alpha)\widehat{r}_i^{\text{stat}} + \alpha\widehat{r}_i^{\text{clin}}, \qquad \alpha \in \{0, 0.05, \ldots, 0.70\}. \tag{15}$$

The fusion coefficient is selected exclusively from inner-OOF predictions and fixed before evaluation on the corresponding outer validation fold.

4 Experiments

4.1 Cohort and Preprocessing

We evaluate the framework on the Multiple Myeloma Research Foundation (MMRF) CoMMpass cohort (Skerget et al., 2024). CoMMpass molecular data are part of the broader ecosystem of cancer genomic resources that support large-scale integrative analysis (Heath et al., 2021). We use a 365-day landmark design: patients are included only if observed overall survival exceeds day 365, clinical measurements after the landmark are discarded, and the target is residual post-landmark survival. The resulting cohort contains 556 patients, including 95 observed deaths and 461 censored observations, with median post-landmark observed time 607.5 days (range 7–1619).

All data-dependent preprocessing is performed within the corresponding training partition. For RNA-seq, the 5,000 most variable genes are retained, standardized using training statistics, and reduced to 64 principal components; static demographic covariates use the same training-only statistics. Longitudinal variables are selected within each training partition from prespecified clinical variable categories. Fields related to survival, progression, treatment, dates, or patient identifiers are excluded. Clinical records are restricted to days 0–365, ordered chronologically, and same-day observations are averaged. Missing values are training-median imputed with binary masks retained, and visit time and inter-visit intervals are normalized by 365 days.

4.2 Baselines and Implementation

We compare static, summary-based, recurrent, and covariance-aware models. **Static Molecular** is the inner-OOF-selected ensemble of two regularized MLP Cox predictors operating on fold-contained RNA PCA features and static covariates. **Clinical Summary** represents each trajectory using summary statistics and temporal descriptors. **Clinical GRU** processes the same visit representation with a gated recurrent unit and maps its hidden representation to a Cox risk. **Clinical Covariance**

Kalman is the proposed state estimator. The standardized static risk is combined with each clinical predictor using the same late-fusion procedure, yielding **RNA+Summary**, **RNA+GRU**, and **RNA+CovarianceKalman**.

All neural models are optimized using AdamW and Cox partial likelihood with full-batch optimization. For the covariance-aware longitudinal model, we use a learning rate of 4×10^{-4}, weight decay of 0.02, dropout rate of 0.05 in the neural observation encoder, and gradient clipping at 1.0. Candidate checkpoints are $\{10, 20, 30, 40, 60, 80, 100\}$ epochs and are selected within the inner cross-validation procedure. The matched fusion protocol uses the same longitudinal visit representation and model-selection procedure across the sequence-based multimodal models.

4.3 Repeated Nested-CV Evaluation

Model selection and final evaluation are separated using nested cross-validation. Each repeat contains five outer folds, and each outer-training partition contains three inner folds. Event status and discretized residual survival time are jointly used for stratification. Inner OOF predictions determine the checkpoint, static ensemble coefficient $\lambda \in \{0, 0.05, \ldots, 1.0\}$, and fusion coefficient $\alpha \in \{0, 0.05, \ldots, 0.70\}$. All model-selection decisions are completed within the inner folds. The selected model is then refit on the complete outer-training partition and evaluated on the outer fold.

We repeat the complete nested cross-validation procedure five times with different random seeds. For each repeat, predictions from the five disjoint outer folds are concatenated into a pooled out-of-fold prediction vector covering all 556 patients. We report Harrell's C-index as the mean $\pm$ standard deviation of the five repeat-level pooled estimates.

Secondary evaluation uses cumulative/dynamic time-dependent AUC and IPCW Brier score at 365 and 730 post-landmark days and IBS over days 90–730. Because model outputs are relative Cox risks, probability-based metrics use a one-dimensional Cox recalibration fitted exclusively on the corresponding outer-training partition. Calibration compares predicted survival probability with Kaplan–Meier estimates in prediction-defined groups.

4.4 Main Survival Prediction Results

Figure 2 summarizes survival discrimination across the five repeated nested-CV partitions. RNA+CovarianceKalman achieves the highest mean pooled OOF C-index, 0.716 ± 0.013. Static Molecular reaches 0.662 ± 0.028. Among clinical-only predictors, Clinical Summary achieves 0.636 ± 0.025, Clinical GRU 0.632 ± 0.014, and Clinical Covariance Kalman 0.662 ± 0.007.

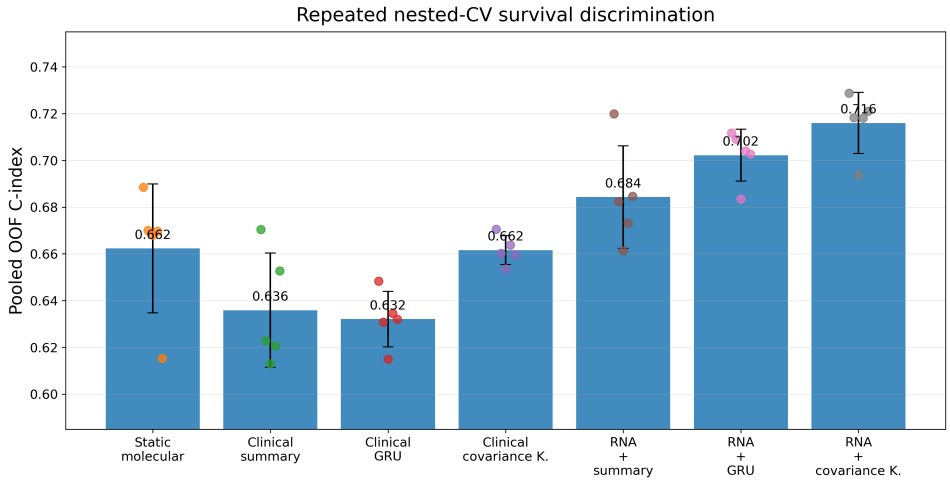


Figure 2: **Survival discrimination across repeated nested cross-validation.** Bars show the mean pooled out-of-fold C-index across five repeats, error bars indicate the standard deviation across repeats, and points show individual repeat-level estimates. RNA+CovarianceKalman achieves the highest mean C-index among the evaluated models (0.716 ± 0.013).

All three longitudinal fusion models improve on Static Molecular: RNA+Summary achieves 0.684 ± 0.022, RNA+GRU achieves 0.702 ± 0.012, and RNA+CovarianceKalman achieves 0.716 ± 0.013. The covariance-aware encoder yields the highest mean discrimination among these alternatives under the shared late-fusion protocol.

Across the five repeated nested-CV runs, RNA+CovarianceKalman improves over RNA+GRU by 0.0137 ± 0.0058 in repeat-level paired C-index, with a positive difference in all five repeats. Relative to Static Molecular, the repeat-level paired improvement is 0.0536 ± 0.0177, again positive in all five repeats. Combining the molecular branch with Clinical Covariance Kalman improves the C-index by 0.0544 ± 0.0191, supporting the complementarity of baseline molecular information and pre-landmark clinical trajectories. In the clinical-only setting, Covariance Kalman achieves 0.662 ± 0.007, compared with 0.632 ± 0.014 for GRU.

Table 1 summarizes the primary and secondary survival prediction metrics across the five repeated nested-CV runs.

The C-index improvement is positive across all five repeated nested-CV runs; paired statistical comparisons across repeats are reported in Appendix A.

Table 1: **Primary and secondary survival prediction performance.** Values are mean ± standard deviation across five repeated nested cross-validation runs. Higher is better for C-index and AUC; lower is better for Brier score and IBS.

Metric	Static Molecular	RNA+GRU	**RNA+CovarianceKalman**
C-index ↑	0.662 ± 0.028	0.702 ± 0.012	$\mathbf{0.716 \pm 0.013}$
AUC_{365} ↑	0.659 ± 0.034	0.720 ± 0.020	$\mathbf{0.743 \pm 0.019}$
AUC_{730} ↑	0.674 ± 0.026	0.725 ± 0.011	$\mathbf{0.740 \pm 0.021}$
Brier_{365} ↓	0.0964 ± 0.0014	0.0943 ± 0.0024	$\mathbf{0.0926 \pm 0.0018}$
Brier_{730} ↓	0.1362 ± 0.0091	0.1246 ± 0.0025	$\mathbf{0.1206 \pm 0.0039}$
$\text{IBS}_{90\text{–}730}$ ↓	0.0940 ± 0.0027	0.0922 ± 0.0019	$\mathbf{0.0907 \pm 0.0015}$

4.5 Time-Dependent Survival Evaluation

Complementary time-dependent evaluation supports the primary concordance results. RNA+CovarianceKalman achieves AUCs of 0.743 ± 0.019 and 0.740 ± 0.021 at one and two years after the landmark, respectively, compared with 0.720 ± 0.020 and 0.725 ± 0.011 for RNA+GRU. Static Molecular reaches 0.659 ± 0.034 and 0.674 ± 0.026 at the same horizons.

Across the five repeated nested-CV runs, the one-year AUC advantage of RNA+CovarianceKalman over RNA+GRU is $+0.0230$ on average and is positive in all five repeats. At two years, the average AUC difference is $+0.0150$ and is positive in four of five repeats. Relative to Static Molecular, the corresponding one- and two-year AUC improvements are $+0.0845$ and $+0.0658$, respectively, with positive differences in all five repeats at both horizons. These results show that the discrimination advantage extends beyond global concordance to clinically relevant post-landmark horizons.

Figure 3 shows time-dependent survival discrimination after the 365-day landmark. Curves show the mean cumulative/dynamic AUC across five repeated nested-CV runs, and shaded regions denote ± 1 standard deviation across repeats.

RNA+CovarianceKalman obtains one- and two-year IPCW Brier scores of 0.0926 ± 0.0018 and 0.1206 ± 0.0039, respectively, compared with 0.0943 ± 0.0024 and 0.1246 ± 0.0025 for RNA+GRU. Its integrated Brier score over post-landmark days 90–730 is 0.0907 ± 0.0015, compared with 0.0922 ± 0.0019 for RNA+GRU and 0.0940 ± 0.0027 for Static Molecular.

After outer-training Cox recalibration, RNA+CovarianceKalman achieves calibration MAE of 0.0251 ± 0.0078 at one year and 0.0253 ± 0.0073 at two years.

Paired t-tests across the five repeated nested-CV runs show that the one-year AUC improvement over RNA+GRU is statistically significant ($p = 0.007$), while the two-year AUC shows a consistent positive trend. Complete paired statistical comparisons are reported in Table 2.

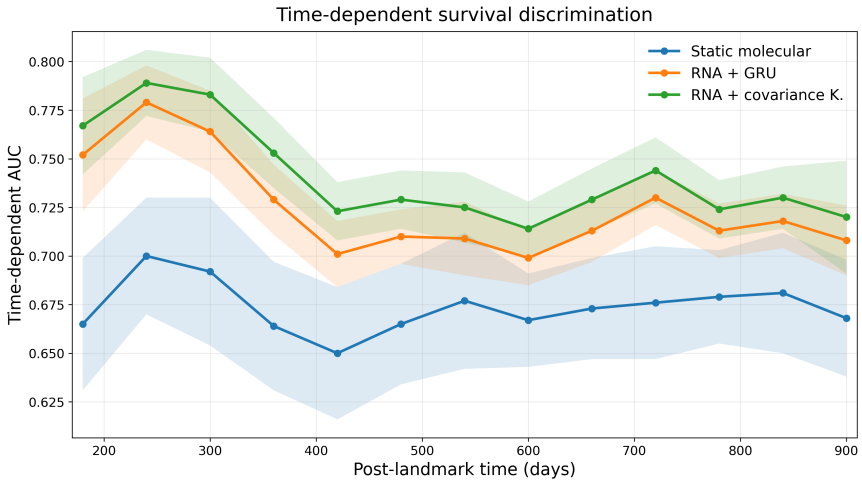


Figure 3: **Time-dependent survival discrimination after the 365-day landmark.** Curves show the mean cumulative/dynamic AUC across five repeated nested-CV partitions, and shaded regions denote variability across repeats. RNA+CovarianceKalman generally maintains higher discrimination than RNA+GRU and the static molecular model across the evaluated post-landmark horizon.

4.6 Ablation of Covariance-Aware Updating

We ablate the two structural components of the proposed state estimator. In the **Fixed Gain** variant, the covariance-derived gain is replaced by $K_t = 0.5I$ while retaining sequential state updating. In the **Constant** Q variant, process covariance no longer increases with elapsed visit time. The complete **Covariance Kalman** model retains both covariance-derived correction and time-aware process uncertainty. Figure 4 summarizes these ablation results across the five repeated nested-CV runs.

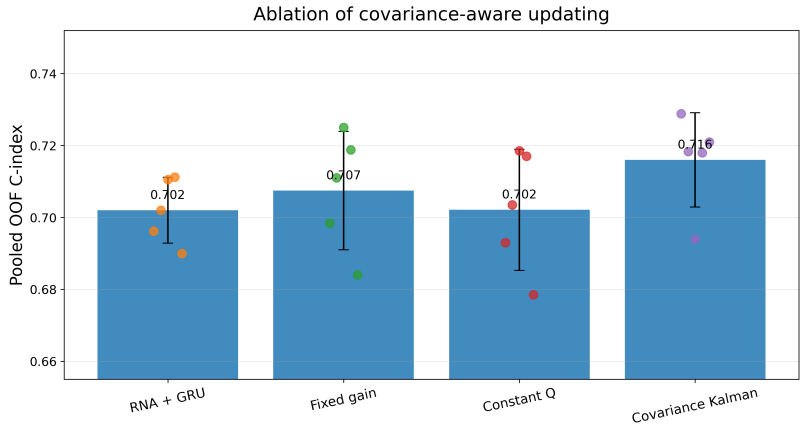


Figure 4: **Ablation of covariance-aware longitudinal updating.** The proposed model is compared with RNA+GRU, a fixed-gain estimator, and a variant with constant process covariance.

Replacing the covariance-derived gain with a fixed gain decreases pooled OOF C-index by 0.0087 ± 0.0069 on average, with the complete model higher in all five repeats. Removing elapsed-time-dependent process uncertainty reduces performance further: the full model improves over Constant-Q by 0.0131 ± 0.0145 and wins all five repeats. These comparisons support two components of the proposed estimator: covariance-derived correction and elapsed-time-dependent process covariance. Together, they connect irregular visit intervals to the correction mechanism through the paths $\Delta\bar{\tau}_{i,t} \rightarrow Q_{i,t} \rightarrow P^{-}_{i,t}$ and $(P^{-}_{i,t}, R) \rightarrow K_{i,t} \rightarrow \mathbf{x}^{+}_{i,t}$.

4.7 Covariance and State-Update Behavior

The covariance recursion exhibits the contraction induced by the Kalman correction. Across evaluated outer folds, the mean diagonal prior and posterior covariances are 0.0688 ± 0.0006 and 0.0396 ± 0.0002, respectively. Posterior variance is lower than prior variance for all valid updates, with mean relative reduction $38.8\% \pm 0.3\%$.

The mean covariance-derived gain is 0.3896 ± 0.0030, with average within-fold SD 0.0912 ± 0.0010, showing substantial variation in correction strength across updates. Mean covariance scales are $Q = 0.0200$, $R = 0.1014$, and $P_0 = 0.0991$, with $Q/R \approx 0.197$.

5 Conclusion

We presented a covariance-aware neural Kalman framework that integrates baseline molecular risk with irregular longitudinal clinical trajectories. The clinical encoder combines learned visit representations with explicit covariance propagation, elapsed-time-dependent process covariance, and analytically computed Kalman correction. Independently trained molecular and clinical branches are combined through risk-level fusion.

On the MMRF CoMMpass 365-day landmark cohort, RNA+CovarianceKalman achieves a pooled out-of-fold C-index of 0.716 ± 0.013 across five repeated nested cross-validation partitions, compared with 0.662 ± 0.028 for Static Molecular and 0.702 ± 0.012 for RNA+GRU. The proposed model also achieves higher one- and two-year AUCs and lower Brier scores than RNA+GRU. Ablations support the contributions of covariance-derived correction and elapsed-time-dependent process covariance, highlighting the value of structured state estimation for multimodal longitudinal survival prediction.

AI Use Statement

Generative AI tools were used to assist with manuscript drafting and editing, code generation and debugging, and research brainstorming. All AI-assisted text, code, analyses, and claims were reviewed and verified by the authors, who take responsibility for the final content of the work.

References

Chul-Young Bae, Bo-Seon Kim, Sun-Ha Jee, Jong-Hoon Lee, and Ngoc-Dung Nguyen. A Study on Survival Analysis Methods Using Neural Network to Prevent Cancers. *Cancers*, 15(19):4757, September 2023. ISSN 2072-6694. doi: 10.3390/cancers15194757. URL https://www.mdpi.com/2072-6694/15/19/4757.

Shanli Chen, Yunfei Zheng, Dongyuan Lin, Peng Cai, Yingying Xiao, and Shiyuan Wang. MAML-KalmanNet: A Neural Network-Assisted Kalman Filter Based on Model-Agnostic Meta-Learning. *IEEE Transactions on Signal Processing*, 73:988–1003, 2025a. ISSN 1053-587X, 1941-0476. doi: 10.1109/TSP.2025.3540018. URL https://ieeexplore.ieee.org/document/10883047/.

Ying Chen, Jiajing Xie, Yuxiang Lin, Yuhang Song, Chen Zhang, Wenxian Yang, and Rongshan Yu. SurvMamba: State Space Model with Multi-Grained Multi-Modal Interaction for Survival Prediction. In *2025 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pp. 3464–3469, Wuhan, China, December 2025b. IEEE. ISBN 979-8-3315-1557-7. doi: 10.1109/BIBM66473.2025.11356727. URL https://ieeexplore.ieee.org/document/11356727/.

Grigoriy Gogoshin and Andrei S. Rodin. Graph Neural Networks in Cancer and Oncology Research: Emerging and Future Trends. *Cancers*, 15(24):5858, December 2023. ISSN 2072-6694. doi: 10.3390/cancers15245858. URL https://www.mdpi.com/2072-6694/15/24/5858.

Mengke Guo, Xiucai Ye, Dong Huang, and Tetsuya Sakurai. Robust feature learning using contractive autoencoders for multi-omics clustering in cancer subtyping. *Methods*, 233:52–60, January 2025. ISSN 10462023. doi: 10.1016/j.ymeth.2024.11.013. URL https://linkinghub.elsevier.com/retrieve/pii/S1046202324002500.

Allison P. Heath, Vincent Ferretti, Stuti Agrawal, et al. The NCI Genomic Data Commons. *Nature Genetics*, 53:257–262, 2021. doi: 10.1038/s41588-021-00791-5.

Yongfei Hu, Xinyu Li, Ying Yi, Yan Huang, Guangyu Wang, and Dong Wang. Deep learning-driven survival prediction in pan-cancer studies by integrating multimodal histology-genomic data. *Briefings in Bioinformatics*, 26(2):bbaf121, March 2025. ISSN 1467-5463, 1477-4054. doi: 10.1093/bib/bbaf121. URL https://academic.oup.com/bib/article/doi/10.1093/bib/bbaf121/8089949.

Yinan Huang, Shadi Bazzazzadehgan, Jieni Li, Arman Arabshomali, Mai Li, Kaustuv Bhattacharya, and John P. Bentley. Comparison of machine learning methods versus traditional Cox regression for survival prediction in cancer using real-world data: a systematic literature review and meta-analysis. *BMC Medical Research Methodology*, 25(1):243, October 2025. ISSN 1471-2288. doi: 10.1186/s12874-025-02694-z. URL https://bmcmedresmethodol.biomedcentral.com/articles/10.1186/s12874-025-02694-z.

Ruqaia J. Kadhim, Hasanain J. Alsaedi, Adel S. Hussain, Mohammad A. Tashtoush, and Rana A. Almuttalibi. An Explainable State-Space Transformer Framework for Predicting Clinical Deterioration in Non-Pulmonary Tuberculosis Patients. *Ingénierie des systèmes d information*, 31(6):2011–2022, June 2026. ISSN 16331311, 21167125. doi: 10.18280/isi.310622. URL https://iieta.org/journals/isi/paper/10.18280/isi.310622.

Amit Milstein, Guy Revach, Haoran Deng, Hai Morgenstern, and Nir Shlezinger. Neural Augmented Kalman Filtering With Bollinger Bands for Pairs Trading. *IEEE Transactions on Signal Processing*, 72:1974–1988, 2024. ISSN 1053-587X, 1941-0476. doi: 10.1109/TSP.2024.3385984. URL https://ieeexplore.ieee.org/document/10494678/.

Huy Hoang Nguyen, Matthew B. Blaschko, Simo Saarakkala, and Aleksei Tiulpin. Clinically-Inspired Multi-Agent Transformers for Disease Trajectory Forecasting From Multimodal Data. *IEEE Transactions on Medical Imaging*, 43(1):529–541, January 2024. ISSN 0278-0062, 1558-254X. doi: 10.1109/TMI.2023.3312524. URL https://ieeexplore.ieee.org/document/10242080/.

Nikolaos Nikolaou, Domingo Salazar, Harish RaviPrakash, Miguel Gonçalves, Rob Mulla, Nikolay Burlutskiy, Natasha Markuzon, and Etai Jacob. A machine learning approach for multimodal data fusion for survival prediction in cancer patients. *npj Precision Oncology*, 9(1):128, May 2025. ISSN 2397-768X. doi: 10.1038/s41698-025-00917-6. URL https://www.nature.com/articles/s41698-025-00917-6.

Deependra Rastogi, Prashant Johri, Massimo Donelli, Seifedine Kadry, Arfat Ahmad Khan, Giuseppe Espa, Paola Feraco, and Jungeun Kim. Deep learning-integrated MRI brain tumor analysis: feature extraction, segmentation, and Survival Prediction using Replicator and volumetric networks. *Scientific Reports*, 15(1):1437, January 2025. ISSN 2045-2322. doi: 10.1038/s41598-024-84386-0. URL https://www.nature.com/articles/s41598-024-84386-0.

Guy Revach, Nir Shlezinger, Xiaoyong Ni, Adria Lopez Escoriza, Ruud J. G. Van Sloun, and Yonina C. Eldar. KalmanNet: Neural Network Aided Kalman Filtering for Partially Known Dynamics. *IEEE Transactions on Signal Processing*, 70:1532–1547, 2022. ISSN 1053-587X, 1941-0476. doi: 10.1109/TSP.2022.3158588. URL https://ieeexplore.ieee.org/document/9733186/.

Mohammad Sabbaqi, Riccardo Taormina, and Elvin Isufi. GKNet: Graph Kalman Filtering and Model Inference via Model-Based Deep Learning. *IEEE Transactions on Signal Processing*, 74:2331–2346, 2026. ISSN 1053-587X, 1941-0476. doi: 10.1109/TSP.2026.3692244. URL https://ieeexplore.ieee.org/document/11514114/.

Sheri Skerget, Daniel Penaherrera, Ajai Chari, et al. Comprehensive molecular profiling of multiple myeloma identifies refined copy number and expression subtypes. *Nature Genetics*, 56(9):1878–1889, 2024. doi: 10.1038/s41588-024-01853-0.

Dao Tran, Ha Nguyen, Van-Dung Pham, Phuong Nguyen, Hung Nguyen Luu, Liem Minh Phan, Christin Blair DeStefano, Sai-Ching Jim Yeung, and Tin Nguyen. A comprehensive review of

cancer survival prediction using multi-omics integration and clinical variables. *Briefings in Bioinformatics*, 26(2):bbaf150, March 2025. ISSN 1467-5463, 1477-4054. doi: 10.1093/bib/bbaf150. URL https://academic.oup.com/bib/article/doi/10.1093/bib/bbaf150/8113022.

M. Vaithiyanathan and K. Bala. A multimodal attention–driven semantic-geometric fusion and dual-mode uncertainty quantification for tumorigenesis prediction. *Biomedical Signal Processing and Control*, 128:111208, December 2026. ISSN 17468094. doi: 10.1016/j.bspc.2026.111208. URL https://linkinghub.elsevier.com/retrieve/pii/S1746809426017623.

Xu Yan, Weimin Wang, Mingxuan Xiao, Yufeng Li, and Min Gao. Survival Prediction Across Diverse Cancer Types Using Neural Networks. In *Proceedings of the 2024 7th International Conference on Machine Vision and Applications*, pp. 134–138, Singapore Singapore, March 2024. ACM. ISBN 979-8-4007-1655-3. doi: 10.1145/3653946.3653966. URL https://dl.acm.org/doi/10.1145/3653946.3653966.

A Paired Statistical Analysis of Performance Metrics

We perform two-sided paired t-tests across the five matched repeated nested-CV runs. For each comparison, performance values are paired by random seed and the test is applied to the five within-repeat differences. For C-index and AUC, positive differences indicate higher discrimination for RNA+CovarianceKalman. For Brier score and IBS, positive differences indicate lower prediction error for RNA+CovarianceKalman. Table 2 summarizes the paired comparisons.

Table 2: **Paired statistical comparison across repeated nested-CV runs.** Values in the Δ column are mean $\pm$ standard deviation of the five paired differences. Two-sided paired t-tests are performed across the five matched repeats.

Comparison	Metric	Δ	p
RNA+CovKalman vs. RNA+GRU	C-index	$+0.0137 \pm 0.0058$	0.006
	AUC_{365}	$+0.0230 \pm 0.0102$	0.007
	AUC_{730}	$+0.0150 \pm 0.0133$	0.065
	Brier_{365}	$+0.0017 \pm 0.0015$	0.070
	Brier_{730}	$+0.0040 \pm 0.0029$	0.038
	IBS_{90-730}	$+0.0015 \pm 0.0010$	0.034
RNA+CovKalman vs. Static Molecular	C-index	$+0.0536 \pm 0.0177$	0.003
	AUC_{365}	$+0.0845 \pm 0.0225$	0.001
	AUC_{730}	$+0.0658 \pm 0.0190$	0.001
	Brier_{365}	$+0.0039 \pm 0.0019$	0.010
	Brier_{730}	$+0.0156 \pm 0.0051$	0.002
	IBS_{90-730}	$+0.0034 \pm 0.0026$	0.046