

Table 2: Predictive performance of our proposed model using different observation windows quantified with R2. The mean over 5-fold cross validation is reported.

Observation window	w/o graph encoder	w graph encoder
7 days	0.233	<b>0.302</b>
14 days	0.456	<b>0.479</b>
21 days	0.586	<b>0.608</b>
28 days	0.652	<b>0.659</b>

of 7, 14, 21, and 28 days, to simulate real-world scenarios where early observations are used to forecast the (future unseen) tumor volume trajectory.

We assessed the predictive performance of our model in two ways. Firstly, we employed R2 to quantify the accuracy of our model in predicting unseen tumor volumes. The results in Table 2 indicate the following: 1) The embedding learned from the heterogeneous graph encoder enhances the predictive performance of our proposed model, and 2) as the observation window size increases, our proposed model captures the unseen tumor dynamic more accurately. Additionally, as it is demonstrated in Panel (B) of Figure 2, the model effectively captures the tumor dynamic trend. However the due to noise inherent in the tumor volume measurements and the clinical significance of mRECIST response category prediction, we also assessed our proposed model’s predictive performance as a classifier. The mRECIST categories are derived from the predicted tumor volume time series by applying response criteria. This evaluation measures the model’s performance in correctly classifying the treatment responses based on the predicted tumor volume dynamics. Figure 3 summarizes the classification results, revealing an observable trend in which incorporating the heterogeneous graph encoder improves the prediction of response categories across all observation windows.

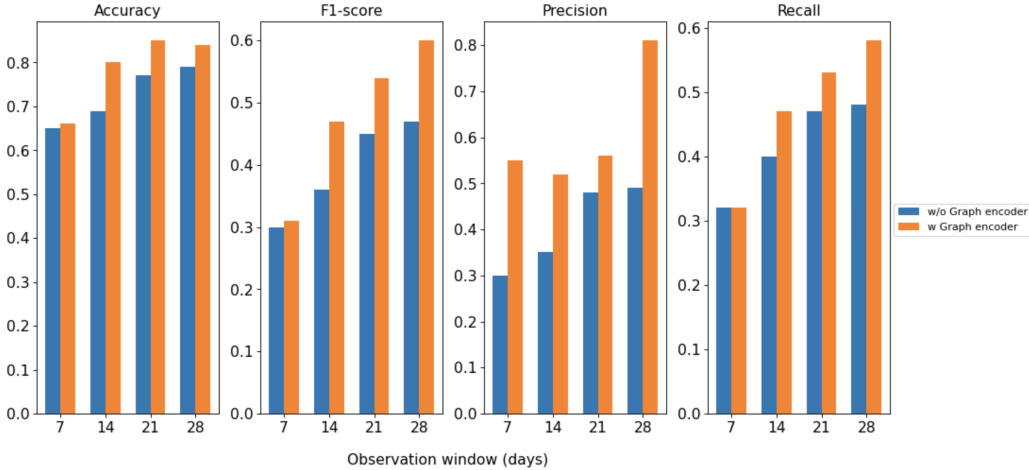


Figure 3: Predictive performance of our proposed model as a classifier for mRECIST categories, with and without the heterogeneous graph encoder and considering different lengths of observation windows.

## 5 Conclusion

In summary, we proposed a novel approach for tumor dynamic prediction that integrates RNA-seq, treatment, disease and longitudinal tumor volume data in an Neural-ODE system in a pre-clinical, PDX setting. We demonstrated that the use of Neural-ODE vastly improved the ability of the model to capture PDX tumor data than a previously proposed TGI model, as well as the benefit of adding the graph encoder to enrich the longitudinal data. As an area for further work, disentangling how the model predictions arise from the multimodal data using explainability techniques and/or attention weights is an important topic to advance our scientific understanding of the complex interplay between gene expression profiles, tumor location and drug targets. This methodology holds significant promise and warrants further evaluations, including in the clinical setting.