

---

# Integration of Graph Neural Network and Neural-ODEs for Tumor Dynamic Prediction

---

**Omid Bazgir <sup>\*†1</sup>, Zichen Wang <sup>\*†1</sup>, Marc Hafner <sup>2</sup>, James Lu <sup>1</sup>**

<sup>1</sup> Modeling & Simulation/ Clinical Pharmacology, Genentech

<sup>2</sup> Oncology Bioinformatics/ Computational Sciences, Genentech

{bazgir.omid, wang.zichen, hafner.marc, lu.james}@gene.com

## Abstract

In anti-cancer drug development, a major scientific challenge is disentangling the complex relationships between high-dimensional genomics data from patient tumor samples, the corresponding tumor’s organ of origin, the drug targets associated with given treatments and the resulting treatment response. Furthermore, to realize the aspirations of precision medicine in identifying and adjusting treatments for patients depending on the therapeutic response, there is a need for building tumor dynamic models that can integrate both longitudinal tumor size as well as multimodal, high-content data. In this work, we take a step towards enhancing personalized tumor dynamic predictions by proposing a heterogeneous graph encoder that utilizes a bipartite Graph Convolutional Neural network (GCN) combined with Neural Ordinary Differential Equations (Neural-ODEs). We applied the methodology to a large collection of patient-derived xenograft (PDX) data, spanning a wide variety of treatments (as well as their combinations) on tumors that originated from a number of different organs. We first show that the methodology is able to discover a tumor dynamic model that significantly improves upon an empirical model which is in current use. Additionally, we show that the graph encoder is able to effectively utilize multimodal data to enhance tumor predictions. Our findings indicate that the methodology holds significant promise and offers potential applications in pre-clinical settings.

## 1 Introduction

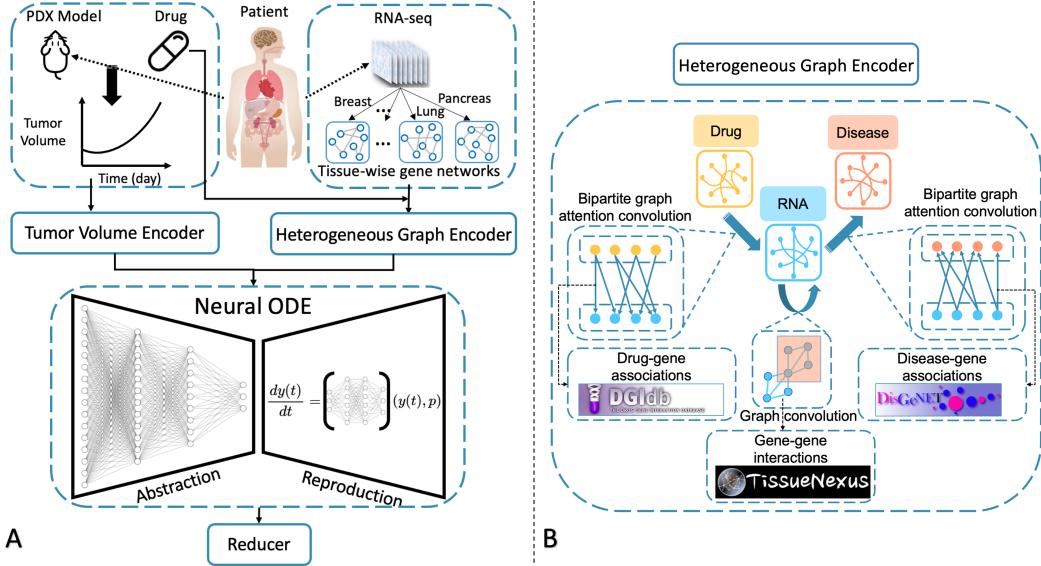
In the development of novel anti-cancer therapies, patient-derived xenografts (PDX) have become an important platform for addressing key questions, such as evaluating the treatment response to therapeutic agents and the combinations thereof, identifying the relevant biomarkers of response and the mechanisms of resistance development Byrne et al. [2017]. Furthermore, as PDX models are obtained by surgically removing patients’ tumor and implanting them in mice, co-clinical avatar trials Byrne et al. [2017] can be performed, whereby treatment of the patients occur simultaneously to the treatment of the corresponding (pre-clinical) PDX models generated from the same patients. Such avatar studies could enable for real-time clinical decision making, to identify the best treatment(s) that the patients could respond to and help deliver some of the promises of precision medicine.

Given the myriad applications of PDX models, an important computational task is the prediction of their dynamic response from the baseline -omics data and/or early tumor size data. This is a particularly challenging task due to the need to meld high-dimensional omics data measured on baseline (e.g., RNA-seq pre-treatment), with the low-dimensional but serially assessed tumor size

---

<sup>\*</sup>Equal contribution

<sup>†</sup>Contributed to this work during internship at Genentech.



**Figure 1: Model architecture overview.** **A)** Integration of GCNs and Neural-ODEs for tumor dynamic prediction. **B)** **Heterogeneous Graph Encoder:** two bipartite graph attention convolution NN to extract disease-gene, and drug-gene association associations, and a graph convolution NN to extract gene-gene interactions, and integrated into emdedding space in baseline level.

measurements under treatment. While empirical Zwep et al. [2021] and spline-based Forrest et al. [2020] tumor dynamic models have been proposed, there has been little progress in melding such dynamic models with high dimensional -omics data. In Zwep et al. [2021], an machine learning (ML) approach has been proposed to use a model based on least absolute shrinkage and selection operator (LASSO) to predict tumor dynamic parameters from copy-number variations (CNVs) of genes from a large PDX data set consisting of various treatments Gao et al. [2015]. In Ma et al. [2021], a few-shot learning (multi-layer perceptron) approach, and in Peng et al. [2022], a heterogeneous GCN approach, have been proposed to learn drug responses from in-vitro (cell-line) data and predict in-vivo (PDX) outcomes within two response categories. While the predictions made by the models in Ma et al. [2021] and Peng et al. [2022] show promising correlations with true responses, they do not necessarily capture the tumor dynamic, which is crucial for clinical decision-making. While the proposed method shows promise, improving the predictivity of the model and including omics data remains an important topic for further research.

Over the past few years, Neural-Ordinary Differential Equations (NODEs) Chen et al. [2018] has emerged as a promising deep learning (DL) methodology for making predictions from irregularly sampled temporal data. Recently, a methodology based on NODE has been developed for tumor dynamic modeling in the clinical trial setting and the generated embeddings from patients' tumor data have been demonstrated to be effective for predicting their Overall Survival (OS) Laurie and Lu [2023]. While the Tumor Dynamic NODE (TDNODE) has set the mathematical foundations for modeling longitudinal tumor data Laurie and Lu [2023], a methodology for the incorporation of high-dimensional, multimodal data into such models has yet to be developed.

In this work, we propose a novel way to combine the previously developed tumor volume encoder with a heterogeneous graph encoder (see Fig. 1). The latter takes as inputs multimodal data consisting of drugs, disease and RNA-seq by incorporating graphs that encapsulate drug-gene associations, disease-gene associations and gene-gene interactions respectively. In application to the PDX data of Gao et al. [2015], we show that by leveraging the multimodal data including RNA-seq in conjunction with early tumor response data, the proposed methodology can significantly improve predictive performance of future tumor response. Finally, we summarize the findings and discuss the potential future applications of the proposed approach in enabling precision medicine in oncology.

## 2 Method

Within our multi-modal framework, we constructed a multi-relational network using three large datasets covering interactions between drugs, genes, and diseases. We used RNA-seq data as node features within tissue-specific knowledge graphs and further integrated it with drug targets of treatments as a heterogeneous graph Zhang et al. [2019].

This integration allowed us to learn an embedding that captures the complex relationships between genes, tumors, and treatments. The following 3 graphs were utilized to represent both biological and pharmacological knowledge:

- **gene-gene graph:** tissue-specific functional gene networks obtained from TissueNexus Lin et al. [2022], which provides gene-gene associations based on gene expression across 49 human tissues.
- **drug-gene graph:** obtained from DGIdb Cotto et al. [2018], which consolidates information on drug target and interacting genes from 30 disparate sources through expert curation and text mining.
- **disease-gene graph:** obtained from DisGeNET Piñero et al. [2016], one of the largest collections of genes and variants associated with human diseases.

For each PDX model, the pre-treatment embeddings using the above multi-modal graphs are used in conjunction with the early, observed tumor measurements to predict the future tumor dynamic profiles. The following subsections cover these two respective aspects.

### 2.1 Heterogeneous Graph Encoder

We formulate the PDX representation learning task as one of graph embedding, by fusing information from a heterogeneous network that incorporates drug, disease, and gene relationships. We apply multi-layer GCN Welling and Kipf [2016] in the RNA domain to model the gene interactions (*gene-gene graph*). Additionally, we use a bipartite graph attention convolutions Wang et al. [2020], Nassar [2018] for message passing from drugs to target genes (*drug-gene graph*), as well as gene embeddings to the disease domain (*disease-gene graph*). The mathematical formulation of our proposed framework is described as follows.

**Bipartite Graphs Attention Convolution.** Conventional GCN assumes that all nodes belong to the same category. However, in our scenario there are heterogeneous attributes across various node types, such as genes, diseases, and drug targets. This limitation becomes evident when node attributes span across different domains. Consequently, we adopt a bipartite graph as a natural representation for modeling inter-domain interactions among distinct node types. We adapt GCN to operate within a bipartite graph, where node feature aggregation exclusively occurs over inter-domain edges. Specifically, let us denote graphs as  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ , where  $\mathcal{V}$  represents the set of vertices, given by  $\{v_i\}_{i=1}^N$ , and  $\mathcal{E}$  is the set of edges. We consider a bipartite graph  $\mathcal{BG}(\mathcal{U}, \mathcal{V}, \mathcal{E})$  defined as a graph  $\mathcal{G}(\mathcal{U} \cup \mathcal{V}, \mathcal{E})$ , where  $\mathcal{U}$  and  $\mathcal{V}$  represent two sets of vertices (nodes) corresponding to the two respective domains. Here,  $u_i$  and  $v_j$  denote the  $i$ -th and  $j$ -th node in  $\mathcal{U}$  and  $\mathcal{V}$ , respectively, where  $i = 1, 2, \dots, M$  and  $j = 1, 2, \dots, N$ . All edges within the bipartite graph exclusively connect nodes from  $\mathcal{U}$  and  $\mathcal{V}$  (i.e.,  $\mathcal{E} = \{(u, v) | u \in \mathcal{U}, v \in \mathcal{V}\}$ ). The features of the two sets of nodes are denoted by  $X_u$  and  $X_v$ , where  $X_u \in \mathbb{R}^{M \times P}$  is a feature matrix with  $\vec{x}_{u_i} \in \mathbb{R}^P$  representing the feature vector of node  $u_i$ , and  $X_v \in \mathbb{R}^{N \times Q}$  is defined similarly. For the message passing  $\text{MP}_{v \rightarrow u}$  from domain  $\mathcal{V}$  to  $\mathcal{U}$ , we define a general bipartite graph convolution (*bg*) as:

$$bg_{\mathcal{E}}(u_i) = \rho \left( \text{agg} \left( \{W_{u_i, v_j} \vec{x}_{v_j} | v_j \in \mathcal{N}_{u_i}^{\mathcal{E}}\} \right) \right), \quad (1)$$

where  $\mathcal{N}_{u_i}^{\mathcal{E}}$  represents the neighborhood of node  $u_i$  connected by  $\mathcal{E}$  in  $\mathcal{BG}(\mathcal{U}, \mathcal{V}, \mathcal{E})$  (i.e.,  $\mathcal{N}_{u_i}^{\mathcal{E}} \subset \mathcal{V}$ ).  $W_{u_i, v_j} \in \mathbb{R}^{M \times N}$  is a feature weighting kernel transforming  $N$ -dimensional features to  $M$ -dimensional features, the  $\text{agg}$  is a permutation-invariant aggregation operation, and the  $\rho$  operator can be a non-linear activation function. In our work, we used element-wise mean-pooling and ReLU Nair and Hinton [2010] for  $\text{agg}$  and  $\rho$  respectively.

Our bipartite graph convolution layers utilize the graph attention network Veličković et al. [2017] as the backbone on the node features, resulting in the bipartite graph attention convolution layer

(*bga*). Since the attention mechanism considers features of two sets of nodes, we specifically define a learnable matrix  $W^u \in \mathbb{R}^{P \times S}$  (resp.  $W^v \in \mathbb{R}^{Q \times S}$ ) for  $X_u$  (resp.  $X_v$ ). The *bga* can be formulated as:

$$bga_{\mathcal{E}}(u_i) = \text{ReLU} \left( \sum_{v_j \in \mathcal{N}_{u_i}^{\mathcal{E}}} \alpha_{u_i, v_j} W^v \vec{x}_{v_j} \right), \quad (2)$$

The attention mechanism is a single-layer feedforward neural network, parameterized by a weight vector  $\vec{a}$  and applying the LeakyReLU non-linearity function. The attention weight coefficients can be expressed as:

$$\alpha_{u_i, v_j} = \frac{\exp(\rho(\vec{a}^T [W^u \vec{x}_{u_i} \| W^v \vec{x}_{v_j}]))}{\sum_{v_k \in \mathcal{N}_{u_i}^{\mathcal{E}}} \exp(\rho(\vec{a}^T [W^u \vec{x}_{u_i} \| W^v \vec{x}_{v_k}]))}, \quad (3)$$

where  $T$  and  $\|$  represent the matrix transposition and concatenation operations respectively.

**Information Fusion.** We represent the heterogeneous network as an undirected graph  $\mathcal{G}(\mathcal{V}, \mathcal{E})$  with the following three sets of nodes: drugs ( $\mathcal{V}^A$ ), diseases ( $\mathcal{V}^B$ ), and genes ( $\mathcal{V}^C$ ). The initial features of these three sets of nodes are denoted as  $X_{\mathcal{V}^A}$ ,  $X_{\mathcal{V}^B}$ , and  $X_{\mathcal{V}^C}$ , respectively. The edges consist of two inter-domain sets representing drug-gene associations ( $\mathcal{E}^{AC}$ ) and disease-gene associations ( $\mathcal{E}^{BC}$ ), as well as an intra-domain set representing gene network connections ( $\mathcal{E}^{CC}$ ).

First, we apply multi-layer GCNs to the RNA domain to model the gene interactions. Importantly, the module parameters and gene node features are initialized by a pre-trained variational graph auto-encoder (VGAE). The VGAE was trained using the RNA-seq data from all the PDX models in this study, whereby each PDX model was represented as a graph with genes as the nodes of the graph. The tumor-specific graphs were obtained from TissueNexus Lin et al. [2022].

In the second step, we apply a single bipartite graph attention convolution layer to propagate the message from drugs to target genes. Conceptually, this step can be viewed as projecting information from the macro level (e.g., from the domain of drugs) to the micro level (e.g., the domain of genes). To formulate the message-passing step, we represent the hidden embeddings of node  $v_i^c$  as  $h_{v_i^c}^{(k)}$ , where  $k$  is the computational step and when  $k = 0$ ,  $h_{v_i^c}^{(0)} = x_{v_i^c}$ . The  $h_{v_i^c}^{(k)}$  is computed as follows:

$$\text{MP}_{\mathcal{V}^A \rightarrow \mathcal{V}^C}^{(k)} : h_{v_i^c}^{(k)} = bga_{\mathcal{E}^{AC}}(v_i^c) + h_{v_i^c}^{(k-1)}, \quad (4)$$

Similarly in the last step, we utilize the non-linear graph information captured by gene nodes to update the hidden embeddings of the disease nodes. In particular, we apply another bipartite graph attention convolution layer to project gene embeddings to the disease domain. Therefore, the third step can be viewed as an attentional pooling of the disease gene subgraph. The updated feature representations of disease nodes are computed as follows:

$$\text{MP}_{\mathcal{V}^C \rightarrow \mathcal{V}^B}^{(k)} : h_{v_i^b}^{(k)} = bga_{\mathcal{E}^{CB}}(v_i^b) + h_{v_i^b}^{(k-1)}. \quad (5)$$

We concatenate the updated drug and disease embeddings into a unified representation for PDX experiments (tumor-treatment combinations) as  $\beta_1$ , which can be used as feature inputs for downstream tasks.

## 2.2 Tumor Volume Prediction

We formulate a tumor volume dynamic model using a Neural-ODE which utilizes both the embedding generated from the heterogeneous graph encoder, as well as the embedding generated from feeding early tumor volume data into an encoder. The model aims to utilize early tumor volume data together with baseline embedding (encapsulating drug, disease and RNA-seq data) in order to make personalized predictions.

**Tumor Volume Encoder.** In a similar manner to Laurie and Lu [2023], we implemented a tumor volume encoder to inform the Neural-ODE in order to make predictions individualized to a specific PDX model. This recurrent neural network (RNN) based encoder maps a short window of early observed tumor volumes (of an arbitrary length) into an embedding that we denote as  $\beta_2$ .

**Neural-ODE.** In the clinical context, an approach to model tumor dynamic data using Neural-ODE has been developed, demonstrating the ability of such a formalism to discover the right dynamical

model from longitudinal patient tumor size data across treatment arms Laurie and Lu [2023]. In this work, we generalized the methodology to meld longitudinal measurements with the multi-modal data into a single deep learning framework and demonstrated in the setting of dynamic modeling of PDX data. We consider a dynamical system of the following form:

$$\frac{dy(t)}{dt} = f_\theta(y(t), \beta), t \in [0, T] \quad (6)$$

where 0 and  $T$  denote the start time of PDX experimentation and the end of the prediction time respectively,  $f_\theta$  is a neural network parameterized by a set of weights  $\theta$  to be learned across all PDX data, and  $\beta = [\beta_1 || \beta_2]$  as the PDX embedding obtained by concatenating the outputs of the heterogeneous graph and tumor volume encoders respectively. Thus, a dynamical law represented by  $f_\theta$  is learned across all PDX data, with the concatenated embedding  $\beta$  serving to provide the initial condition for the Neural-ODE specific to the PDX model of interest. After simulating eqn. 6 to obtain the time evolution of state  $y(t)$ , we then reconstruct the tumor volume data using a two-layer MLP reducer. The Neural-ODE approach has the benefits of handle variable-length sequences, accommodate varying observation intervals, and incorporate positional encoding for capturing temporal context. To enhance model simplicity and generalization, we keep the ODE system dimension no larger than the existing tumor growth inhibition (TGI) model Zwep et al. [2021]. The TGI model captured the longitudinal tumor volume measurements, per PDX, with two empirical ODEs through estimation of three parameters: growth rate, treatment efficacy, and time-dependent resistance development ( $k_r$ ).

**Loss Function.** As a measure of the overall fit of the model prediction to the observed tumor volume values, we incorporated the summation of the mean squared error (MSE) and the mean absolute percentage error (MAPE) De Myttenaere et al. [2016]. The MSE is used to maintain the model accuracy, and MAPE is used to balance in capturing the tumor volume growth or shrinkage trends in the training set.

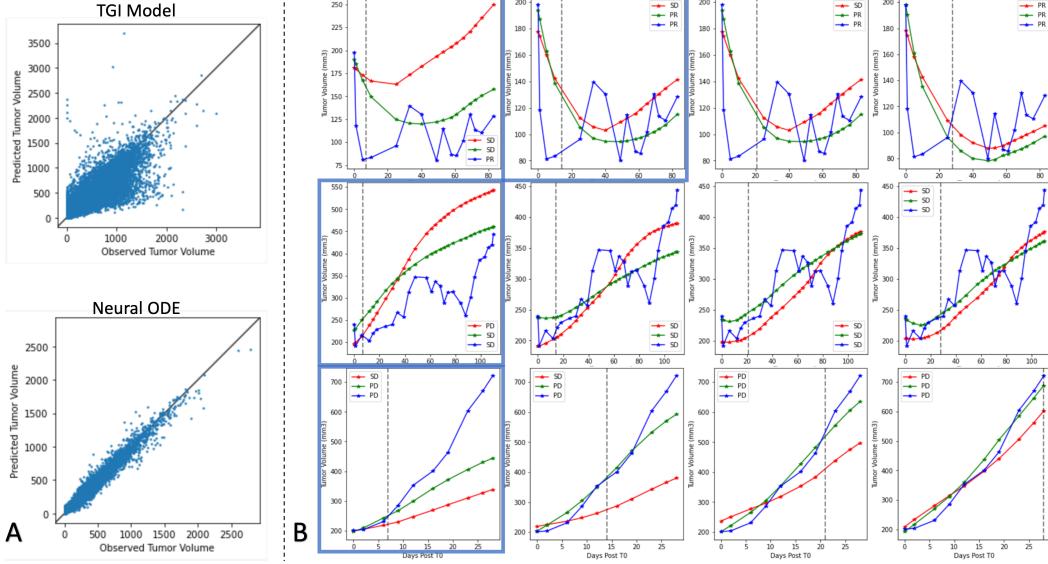
### 3 PDX dataset

The primary dataset utilized in this study was obtained from a large-scale pre-clinical study conducted in PDX mice models, as detailed in Gao et al. [2015]. This dataset encompassed more than 1000 PDX models, each characterized by their baseline mRNA expression levels prior to treatment. In total, the dataset covered 62 distinct treatments across six different diseases, with tumor volume measurements every 2-3 days. For our analysis, we included data from 191 unique tumors and 59 different treatments, resulting in a comprehensive dataset of 3470 PDX experiments (combinations of tumor and treatment) spanning 5 tumor types. This selection was based on the availability of RNA-seq data.

### 4 Results

To assess the effectiveness of the heterogeneous graph encoder in describing changes in tumor volume, we trained the encoder using mRECIST Therasse et al. [2000], Gao et al. [2015] response labels. Specifically, we consolidated the response categories (CR, PR, and SD) into a single response category, and PD as the second category (Table A.1 in the Appendix). To predict these response categories, we used a binary classifier using a three-layer MLP neural network, that takes the heterogeneous graph encoder embedding as the input. The dataset was split by PDX models, and standard 5-fold cross-validation experiments were performed. We evaluated the prediction performance using three metrics: balanced accuracy, area under the receiver operating characteristic (AUROC), and F1-score. Our model’s performance was compared against several competing approaches, including a non-graph-based deep learning method and traditional machine learning methods. To investigate the impact of the pre-training strategy, we also implemented a variant of our model with a randomly initialized gene graph.

**GCNs outperforms baseline methods in modeling RNA-seq for treatment response prediction.** As shown in Table 1, GCN-based methods demonstrated superior performance in the responder classification task across all evaluation metrics. These models outperformed non-graph-based approaches by more than 9% in F1 score and 5% in AUROC. Additionally, we observed that



**Figure 2: Tumor dynamic prediction:** **A)** A comparison on tumor volume trace fitting between tumor growth inhibition (TGI) model proposed by Zwep et al. [2021] and our proposed Neural-ODE system. **B)** Tumor volume trace prediction using the proposed model for 3 different PDXs on 3 different row of curves. The blue curves represent the ground-truth, the green curves represent our proposed model prediction, and the red curves represent the TGI model predictions. Each column shows, one observation window, and the highlighted blocks are the observations windows where our proposed model reaches to correct classification of mRECIST response category (same as the ground-truth).

our model achieved a noticeable improvement of approximately 4% in F1 score and 3% in AUROC when utilizing gene graph pre-training. This result highlights the effectiveness of graph reconstruction pre-training in enhancing gene latent representations.

Table 1: The summary of model performance on treatment response prediction using RNA-seq data. The mean over 5-fold cross validation is reported.

Method	Balanced Accuracy	AUROC	F1
Pretrained GCN	<b>0.681</b>	<b>0.741</b>	<b>0.623</b>
GCN	0.652	0.713	0.582
MLP	0.621	0.688	0.525
SVM/RF	0.618	0.676	0.532

#### 4.1 Tumor Volume Trace Fitting

We aimed to evaluate capability of the Neural-ODE in capturing the longitudinal tumor volume data trend in comparison with the state-of-the-art TGI model for PDX proposed by Zwep et al. [2021]. We utilized all available tumor volume data for each PDX experiment in this experiment. Our approach involved encoding the longitudinal tumor volume data into a latent space using an RNN-based encoder. Subsequently, we used this latent space as an part of the initial condition to solve an ODE system using the Neural-ODE model to reconstruct the dynamic tumor volume data. The population level results are shown in Panel (A) of Figure 2. Notably, the Neural-ODE model outperformed the TGI model with R2 of 0.96 compared to 0.71, and Spearman correlation of 0.96 compared to 0.86.

#### 4.2 Tumor Volume Trace Prediction

We evaluated the ability of our model (presented in the Figure 1) to predict of future tumor volume dynamic based on a limited observed longitudinal tumor data. We selected the observation windows

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 embedding improves the prediction of response categories across all observation windows.

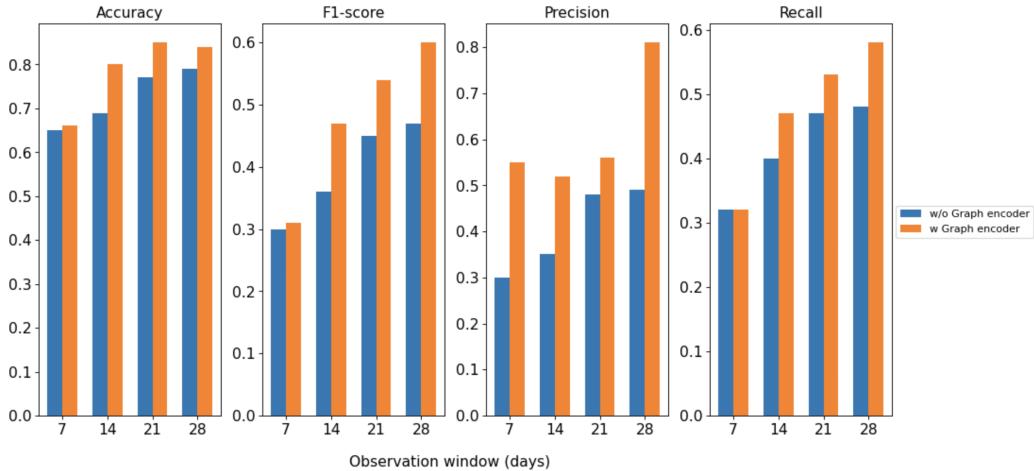


Figure 3: Predictive performance of our proposed model as a classifier for mRECIST categories, with and without the heterogeneous graph encoder and considering different lengths of observation windows.

## 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.

## References

- Annette T Byrne, Denis G Alférez, Frédéric Amant, Daniela Annibali, Joaquín Arribas, Andrew V Biankin, Alejandra Bruna, Eva Budinská, Carlos Caldas, David K Chang, et al. Interrogating open issues in cancer precision medicine with patient-derived xenografts. *Nature Reviews Cancer*, 17(4):254–268, 2017.
- Ricky TQ Chen, Yulia Rubanova, Jesse Bettencourt, and David K Duvenaud. Neural ordinary differential equations. *Advances in neural information processing systems*, 31, 2018.
- Kelsy C Cotto, Alex H Wagner, Yang-Yang Feng, Susanna Kiwala, Adam C Coffman, Gregory Spies, Alex Wollam, Nicholas C Spies, Obi L Griffith, and Malachi Griffith. Dgidb 3.0: a redesign and expansion of the drug–gene interaction database. *Nucleic acids research*, 46(D1):D1068–D1073, 2018.
- Arnaud De Myttenaere, Boris Golden, Bénédicte Le Grand, and Fabrice Rossi. Mean absolute percentage error for regression models. *Neurocomputing*, 192:38–48, 2016.
- William F Forrest, Bruno Aliche, Oleg Mayba, Magdalena Osinska, Michal Jakubczak, Paweł Piatkowski, Lech Choniawko, Alice Starr, and Stephen E Gould. Generalized additive mixed modeling of longitudinal tumor growth reduces bias and improves decision making in translational oncology. *Cancer Research*, 80(22):5089–5097, 2020.
- Hui Gao, Joshua M Korn, Stéphane Ferretti, John E Monahan, Youzhen Wang, Mallika Singh, Chao Zhang, Christian Schnell, Guizhi Yang, Yun Zhang, et al. High-throughput screening using patient-derived tumor xenografts to predict clinical trial drug response. *Nature medicine*, 21(11):1318–1325, 2015.
- Mark Laurie and James Lu. Explainable deep learning for tumor dynamic modeling and overall survival prediction using neural-ode. *arXiv preprint arXiv:2308.01362*, 2023.
- Cui-Xiang Lin, Hong-Dong Li, Chao Deng, Yuanfang Guan, and Jianxin Wang. Tissuenexus: a database of human tissue functional gene networks built with a large compendium of curated rna-seq data. *Nucleic acids research*, 50(D1):D710–D718, 2022.
- Jianzhu Ma, Samson H Fong, Yunan Luo, Christopher J Bakkenist, John Paul Shen, Soufiane Mourragui, Lodewyk FA Wessels, Marc Hafner, Roded Sharan, Jian Peng, et al. Few-shot learning creates predictive models of drug response that translate from high-throughput screens to individual patients. *Nature Cancer*, 2(2):233–244, 2021.
- Vinod Nair and Geoffrey E Hinton. Rectified linear units improve restricted boltzmann machines. In *Proceedings of the 27th international conference on machine learning (ICML-10)*, pages 807–814, 2010.
- Marcel Nassar. Hierarchical bipartite graph convolution networks. *arXiv preprint arXiv:1812.03813*, 2018.
- Wei Peng, Hancheng Liu, Wei Dai, Ning Yu, and Jianxin Wang. Predicting cancer drug response using parallel heterogeneous graph convolutional networks with neighborhood interactions. *Bioinformatics*, 38(19):4546–4553, 2022.
- Janet Piñero, Àlex Bravo, Núria Queralt-Rosinach, Alba Gutiérrez-Sacristán, Jordi Deu-Pons, Emilio Centeno, Javier García-García, Ferran Sanz, and Laura I Furlong. Disgenet: a comprehensive platform integrating information on human disease-associated genes and variants. *Nucleic acids research*, page gkw943, 2016.
- Patrick Therasse, Susan G Arbuck, Elizabeth A Eisenhauer, Jantien Wanders, Richard S Kaplan, Larry Rubinstein, Jaap Verweij, Martine Van Glabbeke, Allan T van Oosterom, Michael C Christian, et al. New guidelines to evaluate the response to treatment in solid tumors. *Journal of the National Cancer Institute*, 92(3):205–216, 2000.
- Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Lio, and Yoshua Bengio. Graph attention networks. *arXiv preprint arXiv:1710.10903*, 2017.
- Zichen Wang, Mu Zhou, and Corey Arnold. Toward heterogeneous information fusion: bipartite graph convolutional networks for in silico drug repurposing. *Bioinformatics*, 36(Supplement\_1):i525–i533, 2020.
- Max Welling and Thomas N Kipf. Semi-supervised classification with graph convolutional networks. In *J. International Conference on Learning Representations (ICLR 2017)*, 2016.
- Chuxu Zhang, Dongjin Song, Chao Huang, Ananthram Swami, and Nitesh V Chawla. Heterogeneous graph neural network. In *Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining*, pages 793–803, 2019.
- Laura B Zwep, Kevin LW Duisters, Martijn Jansen, Tingjie Guo, Jacqueline J Meulman, Parth J Upadhyay, and JG Coen van Hasselt. Identification of high-dimensional omics-derived predictors for tumor growth dynamics using machine learning and pharmacometric modeling. *CPT: pharmacometrics & systems pharmacology*, 10(4):350–361, 2021.

## Appendix

The time-dependent tumor response is determined by comparing tumor volume change at time  $t$  to its baseline (i.e., initial) value:

$$\Delta V(t) = 100\% \times \frac{V(t) - V_{initial}}{V_{initial}} \quad (\text{A.1})$$

The best response is defined as the minimum value of  $\Delta V(t)$  for  $t \geq 10$  days.

mRECIST category	Description
Complete Response (CR)	best response < -95%
Partial Response (PR)	-95% < best response < -50%
Stable Disease (SD)	-50% < best response < -35%
Progressive Disease (PD)	otherwise

Table A.1: **mRECIST categories:** The modified Response Evaluation Criteria in Solid Tumors (mRECIST) Therasse et al. [2000], based progression indication in less than 64 days. The best response is computed using the percentage of changes in the tumor volume using the equation A.1. For further details please refer to Gao et al. [2015]