

Review

Not peer-reviewed version

Dissecting Multi-Omics Data for Cancer Survival Risk Prediction Using Variational Autoencoders: Current Trends, Challenges, and Opportunities

Sarmistha Das^{*}, [Manish Kohli](#), Shukurat Rahmon, Robert A. Franklin, Davendra S. Sohal, [Marepalli B. Rao](#), [Shesh N. Raj](#)

Posted Date: 3 June 2026

doi: 10.20944/preprints202606.0280.v1

Keywords: survival modeling; cox proportional hazards; deep learning; neural network; multi-omics



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC, OpenAlex.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Review

Dissecting Multi-Omics Data for Cancer Survival Risk Prediction Using Variational Autoencoders: Current Trends, Challenges, and Opportunities

Sarmistha Das ^{1,2,*}, Manish Kohli ³, Shukurat Rahmon ¹, Robert A. Franklin ^{1,4},
Davendra S. Sohal ^{1,4}, Marepalli B. Rao ¹ and Shesh N. Rai ^{1,2,3}

¹ Department of Biostatistics, Health Informatics and Data Sciences, University of Cincinnati College of Medicine, Cincinnati, OH

² Biostatistics and Informatics Shared Resources, Cancer Data Science Center, University of Cincinnati College of Medicine, Cincinnati, OH

³ Department of Internal Medicine, University of Utah, Salt Lake City, Utah

⁴ Department of Internal Medicine, University of Cincinnati College of Medicine, Cincinnati, OH

* Correspondence: sarmisthadascu@gmail.com

Abstract

Survival modeling is a crucial area in cancer research and precision oncology, enabling prediction of time-to-event outcomes such as overall, progression-free, and disease-free survival. The Cox proportional hazards model has long been the foundation of prognostic analysis due to its ease of interpretability, but the assumptions of linearity and proportional hazards limit its ability to capture complex, high-dimensional relationships in multi-omics data. Deep learning (DL)-based survival models address these limitations by providing flexible, nonlinear modeling and advanced representation learning. This review provides an overview of advances in survival modeling, tracing the evolution from traditional Cox regression to neural network-based approaches, including feed-forward survival models and modern DL architectures. To predict survival based on molecular and clinical information, two major strategies have emerged: (1) applying neural networks directly to multi-omics and clinical data within a Cox regression framework, (2) using variational autoencoders (VAEs) to learn compact latent representations of multi-omics data that are combined with clinical variables. Here we discuss in detail some recently developed VAE-based methods that improve prognostic performance, focusing on advanced training strategies and architectural designs that integrate unsupervised representation learning with Cox PH or non-linear extension of Cox models. Further, we highlight the opportunities to answer core biological questions and key advances in the DL paradigm such as optimization, regularization, and model interpretability, while noting that challenges remain in reproducibility, benchmarking, and clinical translation. In this review, we underscore the need for robust, interpretable, and standardized approaches to improve risk stratification by uncovering biologically meaningful patterns in multi-omics and clinical data, thereby advancing precision oncology.

Keywords: survival modeling; cox proportional hazards; deep learning; neural network; multi-omics

1. Introduction

Understanding the effect of clinical and molecular signatures on treatment response is critical in determining survival probability of patients. Traditional approaches for associating survival outcomes with clinical and genetic features have largely relied on Cox proportional hazards (PH) models, which assume a linear relationship between predictors and survival. While these models have been widely used, the assumption of linearity may not adequately capture the complex and dynamic interactions underlying disease progression and treatment response. With the increasing

availability of patient-specific multi-omics data, there is a noticeable shift toward approaches that learn patterns directly from the data rather than relying on predefined assumptions. Earlier studies, limited by data scarcity, primarily focused on single-omics analyses, linking DNA alterations or gene expression changes to survival and treatment outcomes. Although these studies provided important insights, they do not fully reflect the multi-layered nature of biological systems. However, it is intuitive that identifying how clinical and multi-omics features interact can help explain variability in patient response and may lead to improved strategies for prognosis and personalized treatment. Thus, integrating multi-omics data to better understand treatment response and survival has therefore become an active area of research, although, significant challenges remain. Predicting treatment response is challenging, as it requires detailed molecular characterization to uncover the factors driving differences in survival outcomes. This difficulty is further amplified in advanced-stage cancers such as metastatic castration-resistant prostate cancer (mCRPC), where patients are highly heterogeneous making it hard to identify consistent patterns associated with treatment response and survival. This review focuses on emerging efforts in modeling survival using clinical and multi-omics data, that aim to better capture these complex relationships.

To provide a structured perspective, we begin by discussing advances in survival modeling, then explore latent representation learning of multi-omics data for survival prediction, followed by methods aimed at improving these models. Finally, we summarize some current deep learning (DL)-based models in this area and discuss advances and limitations in model training and explainability techniques for improving predictive performance.

1.1. Advances in Survival Modeling: From Cox Regression to Deep Survival Models

Patient classification and survival risk prediction are fundamental to contemporary cancer research and precision oncology. Time-to-event endpoints—such as overall survival (OS), progression-free survival (PFS), and disease-free survival (DFS)—capture both the biological aggressiveness and evolutionary dynamics of tumors. Integrating multi-omics data with survival outcomes enables the identification of molecular drivers of disease progression and facilitates the translation of these insights into clinical decision-making. The Cox PH model is a semi-parametric model widely used to estimate the effect of prognostic variables on the hazard of time-to-event outcomes, such as death, relapse, or recurrence of symptoms. The model assumes that the logarithm of the patient's risk of failure is a linear combination of the covariates. While univariate Cox models are useful for preliminary screening, they fail to account for confounding effects; multivariate Cox models address this limitation by estimating adjusted hazard ratios (HRs) and identifying independent predictors, though they retain assumptions of proportional hazards and linear covariate effects. To overcome linearity constraints, neural network (NN) extensions have been proposed. One early neural network for survival analysis was the feed-forward (FF) model proposed by Faraggi and Simon in 1995 [1], which allowed modeling of complex interactions among covariates. However, prior to the recent technical revolution in the DL paradigm over the last decade, such networks generally could not outperform traditional survival models, such as the Cox PH model. Recent advancements in modern DL techniques to optimize the training of a network have paved the way for success of NN models compared to standard Cox PH models to demonstrate improved predictive performance. Risk neural networks can capture complex, nonlinear relationships between prognostic features and an individual's risk of failure, enabling more accurate prediction of patient outcomes. Building on the Faraggi-Simon approach, DeepSurv [2] first successfully extended this framework to predict risk based on both linear and nonlinear covariate effects for time-to-event outcomes. Other FF NN models such as Cox-nnet [3] based on advanced regularization and optimization training strategies have also improved predictive performance in high-dimensional settings such as genomic survival data. Neural networks such as DeepHit [4] unlike DeepSurv, do not assume PH and can capture time-varying and non-proportional effects. Specifically, DeepHit estimates the joint distribution of survival time and event type, explicitly models censored observations, and is trained using a composite likelihood function followed by ranking of loss. Despite their flexibility, FF NNs

are associated with several drawbacks, including low interpretability, substantial computational requirements, potential overfitting, and reliance on large sample sizes. In addition, several related NN-based survival models, including Cox-Time [5], Dynamic-DeepHit [6], Nnet-survival [7], have been proposed to further relax model assumptions and improve flexibility in capturing complex time-to-event dynamics.

1.2. Latent Representation Learning of Multi-omics data with VAEs for Predicting Survival Outcomes

Recent advances in DL have made NN-based Cox PH models a popular and promising framework for capturing the influence of complex multi-omics variables on survival outcomes. In the literature, two primary strategies are often employed: (1) applying a neural network directly to all features, with network outputs serving as the Cox PH model [3,8], and (2) using variational autoencoders (VAEs) to learn compact latent representations of high-dimensional multi-omics data, which are then combined with clinical covariates in Cox PH models [9–13]. VAEs belong to the class of autoencoders (AEs), a family of unsupervised DL models designed to learn efficient, compressed representations (latent features) of high-dimensional data. Typically, they consist of two components: an encoder that maps input data into a lower-dimensional latent space, and a decoder that reconstructs the original input from this compressed representation. By minimizing reconstruction error, AEs learn salient patterns that capture the underlying structure of the data without requiring labeled outcomes. VAEs extend the concept of standard AEs by introducing a probabilistic framework. Instead of mapping inputs to fixed points in latent space, the encoder predicts the parameters of a probability distribution (typically Gaussian) for each latent feature. During training, latent vectors are sampled from these distributions, and the decoder reconstructs the input from these samples. VAEs are trained to minimize both the reconstruction error and the divergence between the learned latent distribution and a prior distribution typically using the Kullback–Leibler (KL) divergence. This formulation of continuous and structured latent space makes VAEs particularly suitable for generating new samples, integrating heterogeneous data, and capturing complex variability in high-dimensional biological datasets. VAE-based survival prediction models typically comprise two phases: (1) unsupervised learning and (2) supervised learning. In the first phase, a VAE is used to extract latent features that are lower-dimensional and capture the most salient structure of the high-dimensional input data. In the second phase, these latent features along with demographic and clinical data are used as inputs to a survival model, commonly a Cox PH or NN model, to predict time-to-event outcomes. The variations among the methods come from construction of the loss function of the network.

VAEs have been previously explored in the multi-omics context without incorporating survival information, primarily for pan-cancer classification tasks using a combination of unsupervised and supervised learning frameworks [14,15]. OmiVAE predicts multi-class classification by training together VAE loss and classification loss function[14]. The idea of regularization of the classifier encourages the network to learn latent representations that do not increase reconstruction accuracy, but classification error is reduced too. Besides, extension of OmiVAE [13] to include survival data, prior studies have explored survival prediction in cancer subgroups by combining multi-omics data using unsupervised learning approaches, typically traditional or sparse autoencoders followed by classifiers [16,17].

1.3. From Black Box to Insight: Interpreting VAE-Based Survival Predictions

In contrast to inherently interpretable Cox PH models where coefficients directly reflect HRs, DNNs capture nonlinear relationships but are often viewed as ‘black boxes’ despite their strong predictive performance. This has constrained clinical trust and biological interpretability; however, recent advances in explainability approaches such as Shapley values from cooperative game theory, enable quantitative attribution of feature contributions and substantially improves model transparency. For example, XOmiVAE [18] introduces DeepSHAP [19] to provide contribution score for each input molecular feature and omics latent dimension for each prediction. DeepSHAP builds

on the key principle from DeepLIFT [20] based on ‘summation-to-delta’ property. This property ensures the sum of all input feature contributions exactly equals the difference for the given input and the output for a chosen reference (baseline) input. It also enables the approximation of Shapley values, which quantify contribution of each feature to the prediction relative to a reference baseline. Larger Shapley values indicate greater importance, highlighting genes more influential in predicting sub-class. Other methods such as AutoSurv [9], MyeVAE [12] also uses DeepSHAP for quantifying the contribution of input features in risk predictions thus aiding identification of potential biomarkers and promoting greater clinical interpretability.

1.4. Advances and Limitations in Modern Deep Learning Era

While breakthroughs in DL architectures and representation learning techniques have facilitated the adaptation of machine learning and DL methods for multi-omics analysis and cancer classification, their application in this domain remains associated with several challenges. Recent advances in optimization and regularization techniques have significantly improved the training of NN models. Modern optimization methods have evolved from stochastic gradient descent (SGD) to adaptive and momentum-based strategies, with algorithms such as RMSProp, Adam, and AdamW improving convergence speed, stability, and generalization through adaptive learning rates and enhanced gradient handling. To address overfitting, a range of complementary strategies are employed. These include model simplification; regularization techniques such as L2 regularization, which constrains model complexity by penalizing large weights; dropout, which reduces feature co-adaptation through stochastic neuron suppression; and early stopping, which prevents overfitting by halting training based on validation performance; transfer learning, that leverages pretrained models to improve performance in data-scarce settings. In addition, activation functions such as ReLU and its variants improve training efficiency by enabling effective gradient propagation, mitigating vanishing gradient issues, and accelerating convergence through computationally efficient nonlinear transformations.

Despite these advances, several limitations remain. DL models are often sensitive to initialization and hyperparameter settings, leading to reproducibility challenges. Their performance might also degrade in small-sample, high-dimensional settings such as genomics, where overfitting remains a major concern. Furthermore, training and tuning these models are computationally intensive, requiring substantial hardware resources and time. Finally, DL models are inherently data-hungry, relying on large, well-annotated datasets that are often scarce in biomedical and clinical domains.

In this review, we explore some of the cutting-edge methods developed for biomedical research to identify prognostic biomarkers and improve risk prediction models, particularly for complex diseases such as cancer, while highlighting the opportunities and challenges entailed in this paradigm. We first discuss in detail some of the most recently published VAE-based NN models for prognosis prediction in cancer, followed by an overview of technical strategies for effectively integrating multi-omics and clinical information using NNs. Finally, we outline the key opportunities and challenges associated with these approaches.

2. VAE-Based Survival Models – Trending Methods

Emerging trends in risk prediction are moving toward **hybrid architectures, leveraging multiple layers of omics data, and translating these insights into clinically actionable predictions**, particularly for complex diseases like cancer. Recently, a number of VAE-based survival prediction models have been published that combines multi-omics, clinical/demographic data for cancer prognosis prediction. Typically, all of them comprises of an unsupervised phase and a supervised phase followed by an activation-based interpretation approach. For interpretability some methods [9,10,12] use SHAP (SHapley Additive exPlanations) values to quantify feature contributions. To demonstrate the current progress, we describe in detail some representative disease-agnostic models

employing VAE models and summarize them, along with other relevant models, in Table 1. We illustrate the overall architecture of these VAE architectures for survival prediction in Figure 1.

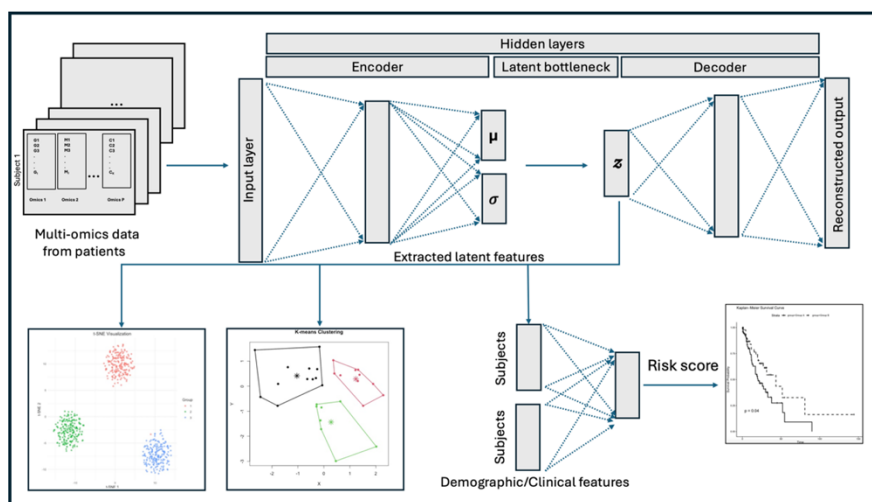


Figure 1. Schematic overview of VAE-based survival prediction framework. The architecture integrates multi-omics data along with clinical and demographic features for cancer survival prediction. It typically consists of an initial unsupervised phase, where VAE learn latent representations from high-dimensional inputs, followed by a supervised phase that combines these representations with clinical and demographic features for survival prediction. To evaluate robustness of latent features derived from unsupervised phase, clustering and other visualization techniques are often used to identify patient subgroups.

Table 1. Disease agnostic deep learning models using VAE architecture.

Reference	Model Architecture	Multi-omics Modalities	Survival and Demographic or Clinical Information	Interpretation	Framework Evaluated On
AutoSurv[9]	KLPMVAE + shallower version of DeepSurv	Transcriptomics (mRNA, miRNA) + pathway	Yes	DeepSHAP	Training/testing: TCGA-BRCA, TCGA-OV, Validation: ICGC-OVAU, Caldas-BC
MyeVAE[12]	VAE + Non-linear version of Cox regression using neural network	Transcriptomics (mRNA, miRNA, lncRNA), Genomics (copy number variants, mutational signature, structural variants)	Yes	DeepSHAP	Training/testing: Multiple Myeloma Research Foundation (CoMMpass) Validation: GSE24080, GSE9782, E-MTAB-4032, GSE19784
VAE-Surv[10]	VAE + DeepSurv	Genomics (Genetic, cytogenetic features)	Yes	Shapley values	Training/testing: Genomed4all MDS cohort

					Validation: IWG-PM
VAE-Cox[11]	VAE + Cox PH (Transfer learning)	Transcriptomics	Yes	Literature survey + functional annotation of the genes in hidden nodes by pathway enrichment analysis	Training/testing: 20 TCGA Pan-cancer datasets Validation: 10 of the above datasets
OmiVAE[14]	VAE + multi-class classifier using neural network	Transcriptomics, Epigenomics	No	Visualization	Training/testing: 33 TCGA Pan-cancer datasets + Normal Samples
OmiEmbed[13]	VAE + multi-class classifier using neural network	Transcriptomics, Epigenomics	Yes	Visualization	Training/testing: GDC Pan-cancer multi-omics datasets + BTM DNA methylation dataset (GSE109381)
XOmiVAE[18]	VAE + multi-class classifier using neural network	Transcriptomics, Epigenomics	No	DeepSHAP	Training/testing: 33 TCGA Pan-cancer datasets + Normal Samples

Abbreviations: Myelodysplastic Syndrome (MDS); International working group for the study of prognosis in MDS cohort (IWG-PM); Genomic Data Commons (GDC); Brain tumor methylation (BTM).

2a. AutoSurv

AutoSurv[9] is a DL framework that first uses an unsupervised method to extract low-dimensional latent features from high-dimensional omics data, then employs a multi-layer perceptron to combine these features with clinical and demographic variables to compute a prognostic index (PI) for each patient, which is subsequently used for supervised classification. At its core, AutoSurv employs a VAE with a KL-annealing strategy (KL-PMVAE) to integrate high-dimensional multi-omics data (e.g., gene expression, miRNA expression) along with pathway-level information, extracting biologically meaningful latent features. These latent representations are then combined with demographic and clinical variables in a survival prediction network (LFSurv) to model patient prognosis. To enhance interpretability, the trained AutoSurv model is analyzed using DeepSHAP, which assigns importance scores to input features, enabling identification of the variables that most strongly distinguish high- and low-risk patients.

In the unsupervised step, the VAE is trained using a composite loss consisting of (i) a **reconstruction loss**, which preserves the essential structure of omics features, and (ii) a **KL divergence** term, which regularizes the latent distribution towards a prior (typically standard normal). This probabilistic regularization encourages a smooth, continuous latent space and improves generalization. To enhance multi-omics integration and stabilize training, AutoSurv adopts a **KL-annealing learning strategy**, in which the KL divergence term is gradually increased during training. This prevents premature over-regularization of the latent space, mitigates posterior collapse, and allows the model to first capture informative biological signals before enforcing distributional constraints. As a result, more robust and biologically meaningful latent features are extracted from heterogeneous omics layers. In the supervised phase, the learned latent features along with **demographic and clinical variables** (e.g., age, race, disease stage) are fed into a fully connected

network which serves as **shallower version of DeepSurv[2]**. The model is optimized to obtain PI for each patient using the **negative log partial likelihood of the Cox PH model**, enabling direct modeling of time-to-event outcomes while preserving compatibility with established survival analysis principles. The overall loss function thus combines VAE reconstruction and KL regularization with Cox-based survival loss, balancing representation learning and prognostic performance. To address interpretability—an

essential requirement in biomedical applications—AutoSurv incorporates **DeepSHAP**, that quantifies the contribution of each input feature (genes, miRNAs, pathways) to individual risk predictions, facilitating identification of potential prognostic biomarkers and enhancing clinical transparency.

2b. MyeVAE

MyeVAE[12] (Myeloma VAE) is a multi-modal VAE designed for personalized risk profiling of newly diagnosed multiple myeloma using multi-omics and clinical features. It extends the standard VAE to jointly model multiple data modalities within a shared latent space and incorporates a **survival sub-network** to generate risk predictions from both learned embeddings and directly observed clinical variables such as age, sex, and disease stage like AutoSurv[9]. The model is trained via **semi-supervised learning**, combining the VAE's evidence lower bound (reconstruction loss plus KL divergence regularization) with a **generalized Cox partial likelihood survival loss**, encouraging the latent space to meaningfully correlate with patient risk. For **interpretability**, SHAP values were computed using DeepSHAP to quantify the contribution of input features across modalities.

2e. VAE-Surv

VAE-Surv is a DL framework developed for genetic-based clustering and survival prediction. The model architecture consists of a VAE for representation learning followed by a DeepSurv based survival prediction module. This two-stage design enables simultaneous dimensionality reduction of high-dimensional genomic features and prognostic risk estimation within a unified pipeline.

In the unsupervised stage, the VAE network learns compact latent representations of genetic data. The model is trained using a composite loss function comprising a reconstruction term and a KL divergence regularization term, which constrains the latent space toward a prior Gaussian distribution and promotes smooth, structured embeddings. These latent variables capture underlying genomic patterns and serve as inputs for downstream analyses. In the supervised stage, the latent representations learned by the VAE are concatenated with clinical variables and provided as input to a DeepSurv network. The survival component is optimized using the Cox PH enabling appropriate modeling of right-censored time-to-event. This architecture supports both molecular clustering of patients in the latent space and individualized risk prediction based on time-to-event data.

Beyond prediction, VAE-Surv incorporates a structured post-training analysis. After model training, latent embeddings are examined to identify patient subgroups and risk-associated genomic patterns. Feature contributions and clustering structures within the latent space are analyzed to uncover genetic alterations associated with survival differences, facilitating biological interpretation and potential biomarker discovery. Overall, VAE-Surv demonstrates how a VAE-based architecture combined with Cox-based survival modeling and systematic post-training analysis can jointly enable genomic stratification and prognostic assessment within a coherent DL framework.

2d. VAECox

VAECox is a DL-based survival analysis framework that integrates VAEs with Cox PH model to improve prognostic prediction from high-dimensional cancer transcriptomic data. It is specifically designed to mitigate overfitting in survival models trained on limited sample sizes by leveraging shared genomic information across cancers through parameter-based transfer learning.

The framework begins with unsupervised training of a VAE on pooled RNA-seq data from 20 cancer types in The Cancer Genome Atlas (TCGA). This step learns compact latent representations of gene expression profiles, capturing shared transcriptomic structures while substantially reducing dimensionality. The VAE is optimized using a composite loss consisting of a reconstruction error

term and a KL divergence term, which regularizes the latent distribution toward a Gaussian prior thus promoting a smooth, continuous, and generalizable latent space suitable for downstream survival modeling. In the supervised phase, tumor-specific gene expression data are passed through the pretrained encoder to generate latent features, which are then input into a Cox PH hazard layer. The survival component is trained by minimizing the negative partial log-likelihood, enabling direct modeling of time-to-event outcomes while appropriately handling right-censored data and estimating relative log hazard ratios in accordance with classical survival analysis principles.

A key innovation of VAECox lies in its transfer learning strategy. For each individual cancer type, the VAECox encoder layer is initialized with weights pretrained on pan-cancer RNA-seq data, and these parameters are subsequently fine-tuned using tumor-specific survival data. By exploiting shared transcriptomic patterns across cancers, this approach improves generalization and reduces overfitting in smaller cohorts. The final network produces individualized log-risk scores directly associated with patient survival outcomes. Although VAECox primarily emphasizes predictive performance, the learned latent features can be interrogated post hoc. Correlating latent dimensions with gene expression levels enables identification of genes associated with survival differences, supporting biological interpretation and potential biomarker discovery. Across multiple TCGA cancer types, VAECox demonstrated superior predictive performance compared with traditional penalized Cox models and other NN survival approaches, as measured by the concordance index (C-index). These results highlight the importance of pan-cancer VAE pretraining and transfer learning for survival prediction in high-dimensional transcriptomic settings.

3. Strategies, Opportunities and Challenges

3.1. Advancement in Technological Strategies

3.1.1. Loss Function Minimization via Optimization

For training an NN model, the idea is to minimize the loss function through parameter optimization. In this context, optimization refers to the process of identifying and updating the network's parameters (weights and biases) so as to minimize the loss over the given dataset. Typically, it involves the iterative refinement of these parameters using gradient-based optimization methods. During training, the network performs a forward pass to generate predictions, evaluates the difference between predictions and ground truth via a loss function, and then applies backpropagation to compute gradients. The gradients are subsequently used by an optimization algorithm to update the parameters in a direction that progressively reduces the loss until convergence, or a stopping criterion is met. Common optimization methods include: (1) Gradient Descent (GD), which updates parameters in the direction of the negative gradient of the loss function to iteratively reduce error; (2) SGD, a practical variant of GD that performs updates using mini-batches of data, improving computational efficiency and enabling faster convergence on large datasets; (3) Momentum-based methods, extends SGD by incorporating an accumulated velocity term from past gradients, and helps accelerate convergence and reduce oscillations in relevant directions. In addition to these, adaptive optimization techniques further improve performance by dynamically adjusting learning rates during training. These include but not limited to: (4) AdaGrad [21], which adapts the learning rate for each parameter based on the historical magnitude of gradients, making it suitable for sparse data but potentially causing overly small learning rates over time; (5) RMSProp, which uses an exponentially decaying average of squared gradients to maintain more stable learning rates; and (6) Adam[22], which combines the benefits of momentum and RMSProp by maintaining both first-order (momentum) and second-order (adaptive scaling) estimates, making it one of the most widely used optimizers in DL due to its robustness and fast convergence.

3.1.2. Overfitting via Regularization

In order for an NN model to effectively reconstruct input data, it is important that it does not simply memorize the training samples. Memorization leads to overfitting, a phenomenon in which the model achieves very low error on the training data but performs poorly on unseen data. Regularization refers to a set of techniques that introduce constraints or penalties during the training of a neural network to mitigate overfitting and improve generalization. Some commonly used regularization methods are: (1) L1 Regularization (Lasso): This technique adds a penalty proportional to the absolute values of the model weights to the loss function. As a result, it encourages sparsity by driving some weights exactly to zero, effectively performing feature selection and simplifying the model, (2) L2 Regularization

(Ridge / Weight Decay): L2 regularization adds a penalty proportional to the square of the weights. Unlike L1, it does not enforce sparsity but instead discourages large weight values, leading to smoother and more stable models that are less sensitive to noise in the training data, (3) **Early Stopping**: In this approach, model training is halted before convergence, based on performance on a validation set. As training progresses, validation error typically decreases initially and then begins to increase when overfitting starts. Early stopping prevents the model from learning noise in the training data by terminating training at the optimal point, (4) Dropout: Dropout randomly deactivates a subset of neurons during each training iteration. This prevents the network from relying too heavily on specific neurons and encourages redundancy and robustness in feature representations. But for inference all neurons are used, typically with scaled weights. These regularization techniques are often used in combination to achieve better generalization performance in neural networks. The KL divergence term in VAE regularizes the latent space by forcing the learned latent distribution to be close to a Gaussian prior distribution. This constraint encourages a smooth, continuous latent space where nearby points produce similar outputs, allowing for meaningful sampling and generation from the prior.

3.1.3. Interpretation of Selected Feature via Explainability

When training NN models across multiple hidden layers for clinical prediction tasks, the non-linear representations learns the important features of the data. These features improve predictive performance but obscure how input clinical and other variables (viz. omics) influence outcomes therefore it poses a challenge to translate the identified features into clinical actionable items. Explainability methods like LIME (Local Interpretable Model-agnostic Explanations), DeepLIFT (Deep Learning Important Features), Layer-wise relevance estimation based on classic equations from co-operative game theory such as Shapley regression values, Shapley sampling values and Quantitative input influence help to interpret predictions. Later on, SHAP[19] values introduced unified the measure of feature importance and modified the ideas of LIME and DeepLIFT into Kernel SHAP and DeepSHAP respectively. These are some of the widely used techniques offering essential post-hoc interpretation that can further refine our understanding, debugging and validating NN training dynamics.

3.1.4. Computational Efficiency via Activation Functions

Activation functions efficiently handle several key roles: they capture non-linear relationships between inputs and outputs, enable the stacking of multiple layers to learn hierarchical features, and improve optimization and convergence. Although being in use for over 70 years, the introduction of RELU (**Rectified Linear Unit**) [23] marked a major breakthrough. It enabled efficient training of deep networks and became the default choice in many architectures. Back in the time when backpropagation became popular smooth and differentiable activation functions like **sigmoid** and **tanh** became dominant. Their differentiability made gradient-based learning feasible in multi-layer networks. As networks became deeper, issues like vanishing gradients highlighted limitations of early choices such as sigmoid and tanh functions. More recent functions like **Leaky ReLU**, **Swish**,

and GELU have been proposed to further improve performance. RELU and its variants produce sparse activations that can reduce computational cost and help prevent overfitting by limiting unnecessary neuron activations.

3.2. Enhanced Opportunities for Risk Prediction from Multi-Omics Data

Often times in biomedical research investigators ask, (1) Which genes, pathways, or molecular signatures predict disease outcome? (2) What molecular programs distinguish patients with favorable versus poor survival outcomes? (3) How can molecular risk stratification inform therapeutic decisions? The answer to the questions lies in appropriately decoding the data to extract biologically meaningful information. In the past, feature selection and other statistical and computational [24–27] methods have been used to extract information from multi-omics data. Although selected features predicated phenotype with better accuracy, challenges occurred when they lacked biological interpretability. Advancements made in the recent past in training DL models have revolutionized their use in biomedical research. In particular, AE architectures are being increasingly recognized as powerful tool for addressing complex biological questions through integration of high dimensional data. In context of cancer and survival analysis, the VAE-based methods mentioned in the article provide novel opportunities to link multi-omics profiles with clinically relevant endpoints such as OS, PFS or DFS by capturing nonlinear relationships and hierarchical structures within the data. DL frameworks extend beyond the capabilities of traditional methods to facilitate biologically meaningful discoveries. Interpretation of selected feature using the concept of Shapley values adds to the advantage of these DL models.

A key advantage of AE-based models is their ability to characterize disease heterogeneity. By integrating and compressing multi-omics data into latent representations, these models capture underlying biological variation and identify molecular subgroups associated with distinct survival outcomes. This stratification provides insights into tumor evolution, metastatic potential, immune context, and recurrence, supporting a more precise understanding of cancer subtypes. DL approaches improve the identification of prognostic biomarkers by capturing nonlinear interactions among genes and pathways that are often missed by linear models. Latent representations from DL networks when combined with Cox proportional hazards models, support estimation of hazard ratios and risk scores, preserving interpretability while enhancing predictive accuracy.

Another important application lies in **guiding personalized therapeutic strategies**. Survival models informed by DL enable robust patient stratification into risk categories, supporting clinical decision-making. High-risk patients may be directed toward more aggressive or combinatorial treatment regimens, whereas low-risk individuals may avoid unnecessary therapeutic burden. Furthermore, survival risk associated molecular subtypes identified through DL-based integration might inform screening programs, enable early intervention strategies or preventive measures, and reveal potential therapeutic targets facilitating biomarker-driven treatment selection and advancing precision oncology. Thus, the impact of these approaches extends to **public health and population-level applications**.

From a methodological standpoint, DL addresses key limitations of classical survival models such as the Cox PH model. While Cox-based approaches remain widely used due to their interpretability and strong statistical foundation, they assume linear relationships between covariates and log-risk, as well as proportional hazards over time – assumptions that are often violated in complex biological systems. Cancer progression is governed by nonlinear regulatory networks, dynamic tumor-immune interactions, and evolving clonal architectures, which are not readily captured by linear models. DNNs overcome these challenges by learning nonlinear, hierarchical representations and enabling the integration of multi-omics data without restrictive assumptions.

Importantly, emerging strategies emphasize **hybrid modeling frameworks** that combine DL-based feature extraction with Cox regression for survival estimation. Although some of the drawbacks of Cox PH model remains, these integrative approach preserves interpretability while leveraging the representational capacity of NNs. By aligning multi-omics features with time-to-event

outcomes through advanced representation learning, DL-based survival modeling provides a robust framework for addressing fundamental questions in cancer biology, including tumor heterogeneity, disease progression, therapeutic response, and patient prognosis. Collectively, these advances highlight the transformative potential of DL in biomedical research, strengthening the link between molecular insights and clinical application, and contributing to the ongoing evolution of precision medicine.

3.3. Challenges and Limitations

Despite the availability of hybrid DL frameworks integrating multi-omics architectures with patient survival outcomes, it remains difficult for users to select the most appropriate method from the existing repertoire. This underscores the need for systematic benchmarking of existing methods, which is still lacking. Benchmarking these approaches has therefore become essential for advancing multi-omics survival modeling toward clinically actionable biomarker discovery. Although a range of architectures including standard[16], variational[9–12], sparse[15] and other AEs—have been developed for dimensionality reduction and data integration, their predictive performance, robustness, and interpretability might differ substantially on users datasets. Without systematic benchmarking, it remains unclear which models yield biologically meaningful and reproducible latent representations.

A key role of benchmarking is to evaluate generalization across independent cohorts. Given the high risk of overfitting in high-dimensional multi-omics data, performance must be validated beyond internal datasets, accounting for cohort heterogeneity, batch effects, and platform variability. Such evaluation is essential for assessing robustness and, ultimately, the potential for clinically actionable insights. Comparative benchmarking can further reveal whether certain architectures more consistently capture druggable pathways, immune-related signals, or DNA repair vulnerabilities, thereby informing downstream treatment strategies such as immunotherapy or targeted therapies, including PARP inhibition. Cross-cohort evaluation across cancer types and sequencing technologies is also critical for establishing robustness and clinical utility. Beyond predictive accuracy, benchmarking must incorporate interpretability and biological coherence. Latent representations should be assessed not only using metrics such as concordance index or time-dependent AUC, but also through pathway enrichment consistency, stability of feature attributions, and reproducibility of biomarker signatures. Alignment with established oncogenic processes—such as immune evasion, cell cycle dysregulation, angiogenesis, and DNA repair deficiency—provides evidence that learned features capture biologically meaningful structure rather than technical noise.

Another critical dimension is the evaluation of integration strategies. Early, intermediate, and late fusion approaches differ in how they balance modality-specific information and shared representation learning. Benchmarking these strategies helps determine which designs yield the most stable and interpretable patient-level embeddings, thereby influencing downstream biomarker discovery and attribution. In addition, interpretability methods applied to DL models require rigorous benchmarking. Post-hoc explanation techniques, including SHAP-based approaches, may yield inconsistent attribution patterns depending on model architecture and data variability. Therefore, evaluating the stability of feature importance across resampling procedures and independent datasets is essential to ensure the robustness of inferred biomarkers. Thus, reproducibility and transparency are central to effective benchmarking. Standardized analytical pipelines, shared datasets, consistent validation frameworks, and open-source implementations collectively enhance comparability across studies and reduce methodological variability. Together, these practices strengthen the reliability of benchmarking outcomes. Overall, benchmarking serves as a critical bridge between methodological innovation and clinical translation, enabling AE based models to evolve from predictive tools into reliable systems for discovering interpretable and clinically actionable multi-omics biomarkers in precision oncology.

4. Discussion

The evolution of survival modeling from classical statistical frameworks to DL-based approaches reflects a broader shift in cancer research toward uncovering biologically meaningful patterns. Integrating high-dimensional multi-omics data with clinical covariates enables the extraction of such information, providing deeper insights into the molecular underpinnings of survival risk. In this context, VAEs have emerged as a particularly compelling strategy, offering a principled approach for learning structured latent representations that capture complex biological variability while enabling robust survival prediction. This review highlights how VAE-based models bridge the gap between data-driven representation learning and clinically interpretable prognostic modeling, positioning them as a key component of next-generation precision oncology frameworks.

VAE-based approaches compress heterogeneous data into lower-dimensional embeddings that preserve essential biological structure. These latent representations not only enhance predictive performance but also facilitate downstream analyses such as patient stratification, subtype discovery, and biomarker identification. Architectures such as AutoSurv, MyeVAE, VAE-Surv, and VAECox exemplify this paradigm, combining unsupervised feature learning with supervised survival objectives in hybrid pipelines that balance learning accuracy and interpretability.

An important advantage of VAE-based survival models lies in their ability to capture disease heterogeneity. Cancer is inherently complex, driven by nonlinear interactions across various omics layers. By learning continuous and structured latent spaces, VAEs enable the identification of molecular subgroups associated with distinct survival outcomes. This stratification has direct clinical relevance, as it can inform risk-adapted treatment strategies, guide therapeutic selection, and support early intervention efforts. Moreover, integration of multi-modal data including clinical and demographic variables further enhances the translational value of these models, aligning computational predictions with real-world clinical decision-making.

Despite these strengths, the “black-box” nature of deep models remains a critical barrier to clinical adoption. The incorporation of explainability techniques, particularly SHAP-based methods, represents an important step toward addressing this limitation. By quantifying feature contributions at both input and latent levels, these approaches enable the identification of biologically relevant drivers of risk and improve transparency in model predictions. However, interpretability remains an evolving area, and the consistency and robustness of explanation methods across datasets and architectures require further investigation.

From a methodological perspective, advances in optimization, regularization, and training strategies have been instrumental in enabling the success of VAE-based survival models. Techniques such as adaptive optimization algorithms, dropout, KL-annealing, transfer learning etc. have improved model stability and generalization, particularly in high-dimensional, low-sample-size settings that is typical in biomedical data. Nonetheless, these models remain sensitive to hyperparameter choices and require substantial computational resources, highlighting the need for standardized training protocols and more efficient architectures.

A major challenge identified in this review is the lack of systematic benchmarking across models and datasets. The growing diversity of architectures including standard autoencoders, VAEs, and hybrid frameworks makes it difficult to determine which approaches are most robust, interpretable, and clinically relevant. Rigorous benchmarking across independent cohorts, cancer types, and data modalities is essential to evaluate generalizability and reproducibility. Importantly, such evaluations should extend beyond predictive accuracy to include biological coherence, stability of latent representations, and consistency of identified biomarkers.

Looking forward, several opportunities can further advance the field. First, the integration of longitudinal and temporal data may enable modeling of dynamic disease progression and treatment response. Second, incorporation of prior biological knowledge, such as gene interaction networks, could enhance interpretability and guide representation learning. Third, the development of standardized, open-source pipelines and shared benchmark datasets will be critical for improving reproducibility and accelerating translation. Finally, closer collaboration between computational

scientists and clinicians will be necessary to ensure that model outputs are actionable, interpretable, and aligned with clinical needs. In summary, VAE-based survival models represent a powerful and evolving framework for integrating multi-omics data with time-to-event outcomes. By combining the strengths of deep representation learning with established survival analysis principles, these approaches offer new opportunities for understanding cancer biology and improving patient risk stratification. However, realizing their full clinical potential will require continued advances in interpretability, robustness, and benchmarking, as well as a sustained focus on translational relevance in precision oncology.

References

1. Faraggi D, Simon R: A NEURAL-NETWORK MODEL FOR SURVIVAL-DATA. *Statistics in Medicine* 1995, 14(1):73-82.
2. Katzman JL, Shaham U, Cloninger A, Bates J, Jiang TT, Kluger Y: DeepSurv: personalized treatment recommender system using a Cox proportional hazards deep neural network. *Bmc Medical Research Methodology* 2018, 18.
3. Ching T, Zhu X, Garmire LX: Cox-nnet: An artificial neural network method for prognosis prediction of high-throughput omics data. *PLoS Comput Biol* 2018, 14(4):e1006076.
4. Ryu JY, Lee MY, Lee JH, Lee BH, Oh KS: DeepHIT: a deep learning framework for prediction of hERG-induced cardiotoxicity. *Bioinformatics* 2020, 36(10):3049-3055.
5. Kvamme H, Borgan O, Scheel I: Time-to-Event Prediction with Neural Networks and Cox Regression. *J Mach Learn Res* 2019, 20.
6. Lee C, Yoon J, Schaar MV: Dynamic-DeepHit: A Deep Learning Approach for Dynamic Survival Analysis With Competing Risks Based on Longitudinal Data. *IEEE Trans Biomed Eng* 2020, 67(1):122-133.
7. Gensheimer MF, Narasimhan B: A scalable discrete-time survival model for neural networks. *PeerJ* 2019, 7:e6257.
8. Wang D, Jing Z, He K, Garmire LX: Cox-nnet v2.0: improved neural-network-based survival prediction extended to large-scale EMR data. *Bioinformatics* 2021, 37(17):2772-2774.
9. Jiang LD, Xu C, Bai YT, Liu AQ, Gong Y, Wang YP, Deng HW: Autosurv: interpretable deep learning framework for cancer survival analysis incorporating clinical and multi-omics data. *Npj Precision Oncology* 2024, 8(1).
10. Rollo C, Pancotti C, Sartori F, Caranzano I, D'Amico S, Carota L, Casadei F, Birolo G, Lanino L, Sauta E et al: VAE-Surv: A novel approach for genetic-based clustering and prognosis prediction in myelodysplastic syndromes. *Computer Methods and Programs in Biomedicine* 2025, 261.
11. Kim S, Kim K, Choe J, Lee I, Kang J: Improved survival analysis by learning shared genomic information from pan-cancer data. *Bioinformatics* 2020, 36:389-398.
12. Chang JG, Chen J, Chew G-L, Chng WJ: MyeVAE: a multi-modal variational autoencoder for risk profiling of newly diagnosed multiple myeloma. *BMC Artificial Intelligence* 2025, 1(1):8.
13. Zhang XY, Xing YT, Sun K, Guo YK: OmiEmbed: A Unified Multi-Task Deep Learning Framework for Multi-Omics Data. *Cancers* 2021, 13(12).
14. Zhang XY, Zhang JQ, Sun K, Yang X, Dai CL, Guo YK: Integrated Multi-omics Analysis Using Variational Autoencoders: Application to Pan-cancer Classification. *Ieee Int C Bioinform* 2019:765-769.
15. Fakoor R, Ladhak F, Nazi A, Huber M: Using deep learning to enhance cancer diagnosis and classification. In: *Proceedings of the international conference on machine learning: 2013*. ACM New York, NY, USA: 3937-3949.
16. Chaudhary K, Poirion OB, Lu L, Garmire LX: Deep learning-based multi-omics integration robustly predicts survival in liver cancer. *Clinical cancer research* 2018, 24(6):1248-1259.
17. Hao J, Kim Y, Mallavarapu T, Oh JH, Kang M: Interpretable deep neural network for cancer survival analysis by integrating genomic and clinical data. *Bmc Med Genomics* 2019, 12.
18. Withnell E, Zhang XY, Sun K, Guo YK: XOmiVAE: an interpretable deep learning model for cancer classification using high-dimensional omics data. *Brief Bioinform* 2021, 22(6).
19. Lundberg SM, Lee SI: A Unified Approach to Interpreting Model Predictions. *Advances in neural information processing systems* 2017, 30.

20. Shrikumar A, Greenside P, Kundaje A: Learning Important Features Through Propagating Activation Differences. *Pr Mach Learn Res* 2017, 70.
21. Duchi J, Hazan E, Singer Y: Adaptive Subgradient Methods for Online Learning and Stochastic Optimization. *J Mach Learn Res* 2011, 12:2121-2159.
22. Kingma DP, Ba J: Adam: A Method for Stochastic Optimization. *CoRR* 2014, abs/1412.6980.
23. Nair V, Hinton GE: Rectified linear units improve restricted boltzmann machines. In: *Proceedings of the 27th international conference on machine learning (ICML-10): 2010*. 807-814.
24. Das S, Majumder PP, Chatterjee R, Chatterjee A, Mukhopadhyay I: A powerful method to integrate genotype and gene expression data for dissecting the genetic architecture of a disease. *Genomics* 2019, 111(6):1387-1394.
25. Das S, Mukhopadhyay I: TiMEG: an integrative statistical method for partially missing multi-omics data. *Scientific Reports* 2021, 11(1).
26. Das S, Srivastava DK: ioSearch: An approach for identifying interacting multiomics biomarkers using a novel algorithm with application on breast cancer data sets. *Genetic Epidemiology* 2023, 47(8):600-616.
27. Miller DM, Yadanapudi K, Rai V, Rai SN, Chen JS, Frieboes HB, Masters A, Mccallum A, Williams BJ: Untangling the web of glioblastoma treatment resistance using a multi-omic and multidisciplinary approach. *Am J Med Sci* 2023, 366(3):185-198.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.