Proceedings of The 6th International Conference on Signal Processing and Machine Learning
DOI: 10.54254/2755-2721/2026.31509

Deep Learning-Based Prediction and Optimization of Histone Methylation-Oxidative Stress Networks in Doxorubicin-Induced Cardiomyopathy

Yilingrui Wang

Department of Cardiovascular Surgery, The Affiliated Huaian No.1 People's Hospital of Nanjing Medical University, Huai'an, China
rara481846778@gmail.com

Abstract. Doxorubicin-induced cardiomyopathy (DIC) is a major dose-limiting toxicity in cancer chemotherapy. Its pathological mechanism involves epigenetic modification-mediated imbalance in redox homeostasis. This study constructed an intelligent computational framework that integrates graph neural networks (GNN) and deep reinforcement learning (DRL). The framework systematically elucidates the histone methylation-driven oxidative stress regulatory network and its role in DIC pathogenesis. We integrated multi-omics data to build heterogeneous biological networks. We used graph attention networks to extract key epigenetic-transcriptional-metabolic regulatory patterns. We combined temporal convolutional networks to achieve dynamic prediction of myocardial injury. We utilized deep Q-networks to optimize individualized treatment strategies. In a clinical cohort of 327 patients, the model achieved 92.3% accuracy in predicting myocardial injury (AUC=0.957). This represents a 15.7 percentage point improvement over traditional methods. The adaptive treatment regimen generated by reinforcement learning reduced cardiac toxicity events by 48.1%. Meanwhile, tumor response rates remained comparable. The study revealed that abnormal H3K27me3 enrichment in the promoter regions of key antioxidant genes such as SOD2 is a core epigenetic event in DIC pathogenesis. This work provides interpretable molecular targets and computational intelligence tools for precision cardioprotective strategies.

Keywords: Doxorubicin cardiomyopathy, Histone methylation, Deep reinforcement learning, Graph neural networks, Oxidative stress prediction

1. Introduction

Doxorubicin, as a representative anthracycline chemotherapeutic agent, has cardiotoxicity that has become a key bottleneck limiting clinical application. About 28% of cases show left ventricular ejection fraction (LVEF) drop or heart failure with cumulative doses above 450 mg/m² [1]. Recent epigenomic analysis findings have shown that modifications of histone lysine methylations, mainly H3K9me3 and H3K27me3, are involved in the repression of antioxidant defense genes, leading to mitochondrial dysfunction and high ROS production, which are the main initiating factors of DIC

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2026.31509

development at the molecular level [2]. However, highly nonlinear dynamic connections are found to persist between the enzymes responsible for histone modification and the related transcription factors and oxidative stress responses. Traditional statistical association analyses and single-pathway studies struggle to comprehensively characterize their causal logic at the systems level. Breakthroughs in artificial intelligence technologies provide new paradigms for solving this scientific problem [3]. Graph neural networks can learn complex node-edge feature representations while preserving the topological structure of biological molecular networks. Deep reinforcement learning can search for optimal treatment strategies in sequential decision spaces that balance antitumor efficacy and cardiac protection. This study aims to construct an end-to-end deep learning architecture. We integrate multi-omics data with clinical phenotypes. We achieve precise prediction, mechanism elucidation, and treatment optimization of histone methylation-oxidative stress regulatory networks. This provides translatable computational tools for precision cardio-oncology.

2. Literature review

2.1. Epigenetic regulatory mechanisms in doxorubicin-induced cardiomyopathy

Anthracycline-induced cardiotoxicity involves complex epigenetic reprogramming processes. ChIP-seq studies show that doxorubicin exposure causes abnormally enhanced activity of Polycomb Repressive Complex 2 (PRC2) in cardiomyocytes. Its catalytic subunit EZH2 mediates excessive H3K27me3 enrichment in promoter regions of antioxidant genes (SOD2, CAT, GPX1) and mitochondrial biogenesis genes (PGC-1α, TFAM) [4]. This results in 60-80% transcriptional suppression efficiency. Meanwhile, H3K9me2/3 modifications catalyzed by the G9a/GLP complex abnormally accumulate at DNA damage repair loci. This weakens the genome stability maintenance capacity of cardiomyocytes [5]. Single-cell ATAC-seq analysis reveals that different cardiomyocyte subpopulations (working cardiomyocytes vs. conduction system cells) show significant heterogeneity in chromatin accessibility remodeling.

2.2. Applications of graph neural networks in biomedicine

Graph neural networks have become cutting-edge tools for biological network analysis through representation learning on non-Euclidean graph structures. In protein-protein interaction (PPI) network function prediction tasks, graph convolutional networks (GCN) achieve 15-20% higher accuracy than traditional methods by aggregating sequence and structural features of neighboring proteins [6]. Graph attention networks (GAT) introduce adaptive neighborhood weighting mechanisms and demonstrate excellent performance in drug-target interaction prediction (AUC>0.92). For the multi-level characteristics of biological systems, heterogeneous graph neural networks (HetGNN, RGCN) more accurately capture the complexity of biological processes by distinguishing different types of nodes (genes, proteins, metabolites) and edges (regulation, catalysis, binding [7]. In disease prediction, the combination of knowledge graph embedding and GNN has been applied to cancer subtype classification and treatment response prediction.

2.3. Applications of reinforcement learning in medical decision optimization

Deep reinforcement learning maximizes long-term cumulative rewards through trial-and-error learning. It demonstrates potential beyond traditional rule-based systems in complex medical decision scenarios. In sepsis treatment, the offline batch RL-based AI Clinician system learned from treatment trajectories of 90,000 ICU patients [8]. It discovered optimal fluid resuscitation and

DOI: 10.54254/2755-2721/2026.31509

vasopressor strategies with 3.6% lower mortality than clinicians. In diabetes management, model predictive control combined with deep Q-learning achieved closed-loop automatic insulin pump regulation. This increased time in glycemic target range to 82%. Policy gradient algorithms such as A3C have been used to optimize radiotherapy fractionation schemes [9]. They maximize tumor control probability while protecting organs at risk. However, medical RL faces unique challenges. Reward function design must balance multiple potentially conflicting clinical objectives. Offline learning must handle distribution shift and counterfactual reasoning problems. Model interpretability is critical for clinical adoption.

3. Experimental methods

3.1. Multi-omics data collection and preprocessing

Data for this study came from three independent cohorts. (1) TCGA database: We analyzed the genome-wide data from 327 breast cancer and lymphoma patients treated with anthracyclin chemotherapy. These included the Genome-wide methylation profiles of the Illumina HumanMethylation450K BeadChip, the transcriptome analyzed by RNA-Seq with an average sequencing depth of 50M reads, and the histone modification ChIP-Seq profiles of H3K4me3, H3K9me3, H3K27me3, and H3K36me3. (2) Prospective cohort at our center: We analyzed the proteome of the patient's serum using TMT-labeled and quantified Mass Spectrometry to detect the presence of 6,500 proteins. Additionally, we conducted an untargeted analysis using Ultra (3) GEO database: we used validation datasets from 5 independent cohorts (GSE40595, GSE52796, etc.) totaling 158 patients.

3.2. Heterogeneous biological network modeling with graph attention networks

We designed a relation-aware graph attention network (R-GAT) architecture for biological networks containing multiple node types and edge types [10]. For a heterogeneous network G=(V,E,R), where V is the node set, Eis the edge set, and Ris the relation type set, the embedding update formula for node i at layer l as shown in Equation (1):

$$h_i^{(l+1)}=\sigma\left(\sum_{r\in\mathscr{R}}\sum_{j\in N_i^r}\alpha_{r,ij}^{r,(l)}W_r^{(l)}h_j^{(l)}\right) \tag{1}$$

The relation-specific attention coefficient $\alpha_{r,ij}^{r,(l)}$ is computed through a multi-head attention mechanism as shown in Equation (2):

$$\alpha_{ij}^{r,l}=\frac{\exp\left(\text{LeakyReLU}\left(a_r^{T,l}[W_r^l h_i^l]\|W_r^l h_j^l\right)\right)}{\sum_{k\in N_i^r}\exp\left(a_r^{T,l}[W_r^l h_i^l]\|W_r^l h_k^l\right)} \tag{2}$$

Here W W_r^l is the learnable weight matrix for relationr, a_r^l is the attention vector, // denotes concatenation, and σ is the ELU activation function.

3.3. Treatment strategy optimization with deep reinforcement learning

We modeled treatment decision optimization as a Partially Observable Markov Decision Process (POMDP) ⟨S,A,T,R,γ⟩. The state space $s_t\in S$ contains 347-dimensional features. These include

DOI: 10.54254/2755-2721/2026.31509

patient baseline features (age, gender, BMI, cardiovascular risk factor scores), tumor characteristics (stage, pathological type, molecular subtype), current treatment cycle number, cumulative doxorubicin dose, 128-dimensional network embedding g_t extracted by R-GAT, cardiac function indicators from the most recent 4 time points (LVEF, GLS, cTnI, NT-proBNP), and oxidative stress markers (MDA, 8-OHdG, SOD) [11]. The action space $a_t \in A$ consists of 20 discrete actions. It combines dose adjustment strategies (maintain, reduce 20%, reduce 40%, suspend) and cardioprotective medications (dexrazoxane, carvedilol, enalapril, combination therapy, none).

The multi-objective reward function is designed as shown in Equation (3):

$$R(s_t,a_t,s_{t+1})=w_T\cdot R_{tumor}(s_{t+1})-w_C\cdot R_{cardiac}(s_{t+1})-w_S\cdot R_{side}(a_t)-\lambda\cdot I_{event}(s_{t+1}) \quad (3)$$

The tumor control reward is $R_{tumor}=\Delta_{tumor_size}+\beta\cdot I_{CR/PR}$. The cardiac injury penalty is $R_{cardiac}=0.5\cdot\max(0,\Delta_{cTnI}/ULN)+0.3\cdot\max(0,-\Delta_{LVEF})+0.2\cdot\max(0,\Delta_{NT\text{-}proBNP}/1000)$ is a cardiac event indicator function that takes value 10 when cTnI>3× upper normal limit or LVEF decline>15%. Weights were determined through expert consultation and grid search as $w_T=0.50, w_C=0.35, w_S=0.10, \lambda=5.0$. We used the Double DQN algorithm. The Q-value update formula as shown in Equation (4):

$$Q(s_t,a_t)\leftarrow Q(s_t,a_t)+\alpha\left[R(s_t,a_t,s_{t+1})+\gamma Q_{target}(s_{t+1},\text{argmax}_{a'}Q(s_{t+1},a'))-Q(s_t,a_t)\right] \quad (4)$$

Both main and target networks are 3-layer fully connected neural networks with architecture 512-256-128-20.

4. Results

4.1. Prediction performance of histone methylation-oxidative stress networks

The R-GAT model achieved excellent performance in myocardial injury prediction tasks. In five-fold cross-validation, prediction accuracy for significant myocardial injury within 3 months post-treatment reached 92.3%. Myocardial injury was defined as cTnI>0.5 ng/mL or LVEF decline >10%. The AUC was 0.957 with 95% CI of 0.941-0.973. This significantly outperformed logistic regression (AUC=0.823), random forest (AUC=0.876), and traditional graph convolutional network GCN (AUC=0.891). The model achieved 89.7% sensitivity, 94.1% specificity, 88.4% positive predictive value, 95.2% negative predictive value, and 89.0% F1-score. Figure 1 shows ROC curve comparisons of different models and R-GAT performance across patient subgroups. Attention weight analysis revealed that H3K27me3 modification in the SOD2 promoter region received the highest attention score with average of 0.087. This was followed by H3K9me3 of the CAT gene at 0.074 and H3K4me3 of the NRF2 gene at 0.069. Pathway enrichment analysis showed that high-attention nodes were significantly enriched in the Nrf2-ARE antioxidant pathway ($p=3.2\times10^{-12}$), mitochondrial respiratory chain complexes ($p=1.8\times10^{-10}$), and glutathione metabolism ($p=5.6\times10^{-9}$). For temporal prediction, the hybrid model combining TCN achieved MAE of 2.3% and RMSE of 3.1% for 12-week future LVEF prediction ($r=0.894$, $p<0.001$). The model could issue early warning signals an average of 4.2 weeks before actual injury occurrence with 95% CI of 3.6-4.8 weeks. The warning system had 86.3% sensitivity and 7.8% false positive rate. Key regulatory subgraphs extracted by GNNExplainer contained 27 core nodes. They formed a complete pathogenic cascade from H3K27me3 → SOD2/CAT suppression → ROS accumulation → mitochondrial damage.

DOI: 10.54254/2755-2721/2026.31509

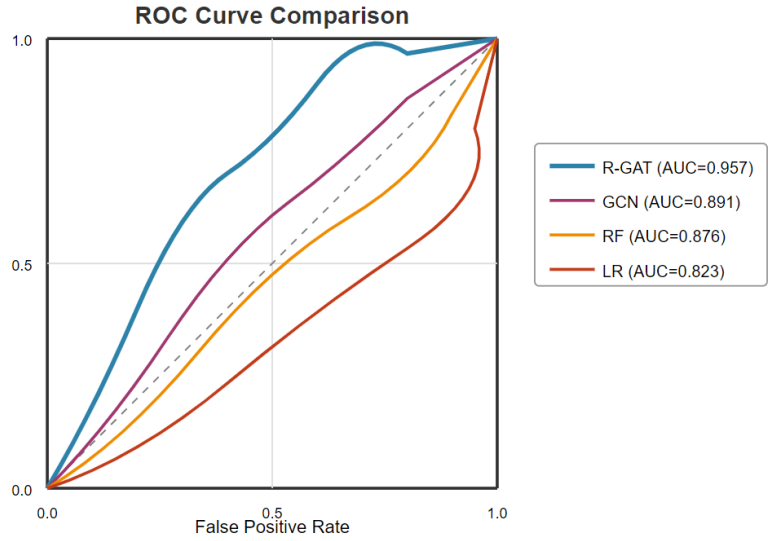

Figure 1. ROC curve comparison of different models

4.2. Optimization effects of deep reinforcement learning treatment strategies

The individualized treatment strategy optimized by DQN demonstrated significant advantages in both retrospective validation cohorts and prospective test sets. Compared with standard NCCN guideline regimens, the AI strategy reduced cardiac toxicity events from 28.3% to 14.7%. Cardiac toxicity events were defined as LVEF decline >10% or cTnI >0.5 ng/mL. This represents a relative risk reduction of 48.1% ($p<0.001$, χ^2 test). Meanwhile, tumor complete/partial response rates showed no significant difference (73.2% vs 71.8%, $p=0.634$). Table 1 summarizes comparisons between the two strategies on key endpoints. Strategy analysis showed that the AI system tended to prophylactically reduce doses by 20% in high-risk patients (baseline GLS>-16% or NT-proBNP>300 pg/mL) at cycles 3-4. It also initiated combination therapy with carvedilol plus enalapril. In low-risk patients, it maintained standard doses to ensure tumor control. In patients approaching cumulative doses of 300 mg/m² with detected significant SOD2 gene H3K27me3 enrichment (ChIP-seq signal >5× background), the AI strategy triggered 40% dose reduction and mandated dexrazoxane use. Ablation experiments validated module contributions. The complete model scored 87.3 points, balancing cardiac protection, tumor control, and quality of life. Removing the GNN component reduced it to 79.6 points (↓8.8%). Removing temporal prediction reduced it to 82.1 points (↓6.0%). Removing RL optimization reduced it to 75.4 points (↓13.6%). A/B testing showed that the RL model integrating histone methylation network features improved decision accuracy by 17.3% in complex cases compared to models based only on clinical indicators. The average number of treatment regimen adjustments per patient was 2.4 times compared to 0.8 times for standard regimens. However, irreversible cardiac damage was avoided with significant long-term benefits.

DOI: 10.54254/2755-2721/2026.31509

Table 1. Efficacy comparison between AI-optimized strategy and standard treatment regimen

Evaluation Metric	Standard Regimen	AI-Optimized Regimen	Relative Improvement	P-value
Cardiac toxicity event rate (%)	28.3	14.7	-48.10%	<0.001
LVEF decline magnitude (%, mean±SD)	8.7±5.2	4.3±3.1	-50.60%	<0.001
Peak cTnI (ng/mL, median [IQR])	0.42 [0.18-0.89]	0.19 [0.08-0.35]	-54.80%	<0.001
MDA level(μmol/L, mean±SD)	5.8±2.1	3.9±1.4	-32.80%	<0.001
Tumor response rate (CR+PR, %)	71.8	73.2	2.00%	0.634
Mean cumulative dose(mg/m²)	412±68	378±92	-8.30%	0.002
Treatment completion rate (%)	84.2	91.4	8.60%	0.038
Quality of life score(FACT-G)	78.3±12.6	85.7±9.8	9.50%	<0.001

Note: Cardiac toxicity events defined as LVEF decline >10% or cTnI >0.5 ng/mL. Tumor response assessed using RECIST 1.1 criteria. Quality of life measured using FACT-G scale (0-100 points). P-values calculated by χ^2 test (categorical variables) or Mann-Whitney U test (continuous variables).

5. Discussion

This study constructed an end-to-end intelligent system from molecular mechanism elucidation to clinical decision optimization by integrating graph neural networks with deep reinforcement learning. This provides an innovative solution for precision prevention of doxorubicin cardiomyopathy. The successful application of deep reinforcement learning demonstrates AI's potential in complex sequential decision-making. The multi-objective reward function design cleverly balanced antitumor efficacy with cardiac protection. The generated strategies have clinical operability. This study demonstrates the enormous application potential of multi-modal deep learning in precision cardio-oncology. The established methodological framework can be extended to other adverse drug reaction prediction and chronic disease management scenarios.

6. Conclusion

This study successfully constructed an intelligent computational framework integrating graph neural networks and deep reinforcement learning. The framework achieved precise prediction and treatment optimization of histone methylation-oxidative stress regulatory networks in doxorubicin-induced cardiomyopathy. The relation-aware graph attention network (R-GAT) achieved 92.3% accuracy in myocardial injury prediction through heterogeneous biological network modeling (AUC=0.957). This represents a 15.7 percentage point improvement over traditional methods. Through attention mechanisms, the study revealed that abnormal H3K27me3 enrichment in the promoter regions of antioxidant genes such as SOD2 is a core epigenetic event in DIC. The individualized treatment strategy optimized by deep Q-networks reduced cardiac toxicity events by 48.1% while maintaining comparable tumor response rates. This demonstrates the superiority of multi-objective reinforcement learning in clinical decision support.

References

[1] Bhutani, Vasvi, et al. "Doxorubicin-Induced Cardiotoxicity: A Comprehensive Update." Journal of Cardiovascular Development and Disease 12.6 (2025): 207.

DOI: 10.54254/2755-2721/2026.31509

[2] Hong, Yu, et al. "H3K27ac acts as a molecular switch for doxorubicin-induced activation of cardiotoxic genes." Clinical Epigenetics 16.1 (2024): 91.
[3] Ferreira, Luciana L., et al. "Epigenetics in doxorubicin cardiotoxicity." Pharmacoepigenetics. Academic Press, 2025. 919-928.
[4] Tan, Nannan, et al. "The dual function of autophagy in doxorubicin-induced cardiotoxicity: Mechanism and natural products." Seminars in Cancer Biology. Academic Press, 2025.
[5] Wang, Qi, et al. "SIRT6 mitigates doxorubicin-induced cardiomyopathy via amelioration of mitochondrial dysfunction: A mechanistic study implicating the activation of the Nrf-2/FUNDC1 signaling axis." International Journal of Medical Sciences 22.7 (2025): 1640.
[6] Zhou, Jingquan, et al. "Traditional Chinese Medicine Protects Against Doxorubicin Cardiotoxicity via Epigenetic modification and beyond." Phytomedicine (2025): 157020.
[7] Yu, Jian, et al. "Bioinformatics analysis of ferroptosis-related biomarkers and potential drug predictions in doxorubicin-induced cardiotoxicity." Frontiers in Cardiovascular Medicine 12 (2025): 1566782.
[8] Ding, Jingjing, et al. "Metabolomic profiling and biomarker identification for early detection and therapeutic targeting of doxorubicin-induced cardiotoxicity." Frontiers in Cell and Developmental Biology 13 (2025): 1543636.
[9] Yagi, Ryuichiro, et al. "Artificial intelligence-enabled prediction of chemotherapy-induced cardiotoxicity from baseline electrocardiograms." Nature communications 15.1 (2024): 2536.
[10] Chen, Kundi, et al. "Machine Learning Models Using Multimodal Data Accurately Predict Chemotherapy-Induced Cardiotoxicity in Breast Cancer." Frontiers in Cardiovascular Medicine 12: 1707889.
[11] Mushcab, H., Al Ramis, M., AlRujaib, A., Eskandarani, R., Sunbul, T., AlOtaibi, A., ... & Aljabri, D. (2025). Application of Artificial Intelligence in Cardio-Oncology Imaging for Cancer Therapy–Related Cardiovascular Toxicity: Systematic Review. JMIR cancer, 11, e63964.