Supplementary Materials

Improving Severity Grading of Chemotherapy-Induced Myelosuppression in AML via Data-Driven and Model-Based Deep Learning

Yu Zhou[1], Yuyang Xiao[1], Qian Wang[1], Chenfeng Mo, Rui Cao, Hongli Liu, Xin Ma, Li Fu, Huimin Gao, Li Xu, Suoqin Jin, Fuling Zhou[*], Xiufen Zou[*]

Contents

1 Supplementary Methods **1**
1.1 Data preprocessing 1
1.2 The Deficiency Period Calculation 2
1.3 Fitting Model Parameters 2
1.4 The Gaussian mixture model 4
1.5 Data Generator Based on Radial Basis Function Networks . . 4
1.5.1 Dynamic Decay Adjustment in RBF Networks (RBF-DDA) 4
1.5.2 Generating Data from RBF Networks 5

2 Supplementary Results **6**
2.1 Model Fitting Results for All Patients and Four Blood Cell Types 6
2.2 Correlation of Model Parameters with Duration of Grade IV myelosuppression 7
2.3 Correlation Between Time-Series Features and Clinical Indicators 8
2.4 Comparison of Density Distributions between Virtual and Real Patient Data 8

3 Supplementary Figures **9**

4 Supplementary Tables **15**

1. Supplementary Methods

1.1. Data preprocessing

Laboratory data were restructured by transposing, standardizing column names, and removing irrelevant columns. Features with more than 50% miss-

ing data were removed[1]. Discrete variables with missing values were imputed using the mode, and continuous variables were imputed using the median. For continuous variables, abnormal values were identified and adjusted using the 1.5*IQR rule, following the method described in Pezoulas et al.[2, 3].

We extracted time-series features from the routine blood test data of patients. These features include two key elements: the initial value and the average value prior to treatment.Similarly, for virtual patients, we generated these time-series features.

1.2. The Deficiency Period Calculation

In the treatment of AML patients, myelosuppression grading post-chemotherapy is based on the World Health Organization's criteria for acute and subacute toxic responses to anticancer drugs [4]. To facilitate subsequent analyses, we quantified the severity of post-chemotherapy myelosuppression by measuring the duration during which four types of blood cell counts remained at grade IV myelosuppression, in line with WHO criteria. The criteria for IV myelosuppression are as follows: WBC $< 1 \times 10^9/L$, neutrophils $< 0.5 \times 10^9/L$, PLT $< 25 \times 10^9/L$, and HGB $< 65g/L$. For example, the neutrophils deficiency period is the interval during which neutrophils count stays below $0.5 \times 10^9/L$.

Using post-chemotherapy routine blood count data, we calculated the duration of grade IV myelosuppression for these blood cells to assess the severity of the patient's myelosuppression. As depicted in Figure S1, if a patient enters grade IV myelosuppression multiple times, the cumulative duration from the first start to the last end of this grade is considered as the duration of grade IV myelosuppression. A longer duration of grade IV myelosuppression for any of these blood cells indicates more severe myelosuppression, suggesting a lower suitability for the patient to chemotherapy.

1.3. Fitting Model Parameters

Considering the significance of the parameters and the dynamic characteristics of the model, we selected five variable parameters to capture patient heterogeneity, namely B, $slope_A$, $slope_D$, k_{tr}, and γ. The first four parameters reflect the patient's steady-state blood cell count, sensitivity to the drug, and the time required for the drug to exert its effect. The parameter γ is a bifurcation parameter, which influences the system's steady state[5]. We kept

other parameters constant and varied only these heterogeneity parameters to fit the experimental patient data.

To identify the parameters in our model, we converted the parameter identification into the problem of optimization by minimizing the following objective function, which is defined mathematically as the error between the simulation results and the time series experimental data:

$$\min_{\mathbf{K}} J(\mathbf{K}) = \sqrt{\frac{1}{m}\sum_{j=1}^{m}(x_{ma}(t_j,\mathbf{K}) - x_{ma}^{D}(t_j))^2 + \phi\|\mathbf{K}\|_2^2} \tag{1}$$

Here, $x_{ma}^{D}(t_j)$ represents the clinical experimental data of the patient's mature circulating cell count at time-point t_j, and $x_{ma}(t_j, \mathbf{K})$ denotes the mature circulating cell count at time t_j obtained through the numerical solution of the ODE model with parameter set $\mathbf{K}$. ϕ is set as 0.0001. The particle swarm optimization (PSO) algorithm[6] is used to optimize this objective function and the pseudocode for the PSO algorithm is provided in Algorithm 1. w, c_1, and c_2 are parameters that control the influence of the respective components on the velocity update. We set the number of particles to 100 and the maximum number of iterations to 150.

Algorithm 1 Particle Swarm Optimization (PSO)

1: Initialize a population of particles with random positions(x) and velocities(v) in the problem space(parameter space).
2: Evaluate the fitness($J(\mathbf{K})$) of each particle.
3: Update the particle's best known position ($pBest$) if the current position is better.
4: Update the global best known position ($gBest$) if any particle's position is better than $gBest$.
5: **repeat**
6: **for** each particle i **do**
7: Calculate new velocity: $v_i = w{\cdot}v_i + c_1{\cdot}(pBest_i - x_i) + c_2{\cdot}(gBest - x_i)$
8: Update position: $x_i = x_i + v_i$
9: Evaluate the fitness ($J(\mathbf{K})$) of the particle.
10: Update the $pBest_i$ and $gBest$ if necessary.
11: **end for**
12: **until** a termination criterion is met.

1.4. The Gaussian mixture model

The Gaussian mixture model[7] is defined as a weighted sum of M component Gaussian densities given by the equation:

$$p(x|\lambda) = \sum_{i=1}^{M} w_i \cdot g(x|\mu_i, \Sigma_i) \tag{2}$$

where x represents the data points (in this case, the B_{sum}), λ represents the parameters of the mixture model, w_i are the mixture weights, and $g(x|\mu_i, \Sigma_i)$ are the component Gaussian densities. Each component density is a multivariate Gaussian function of the form:

$$g(x|\mu_i, \Sigma_i) = \frac{1}{(2\pi)^{D/2}|\Sigma_i|^{1/2}} \exp\left(-\frac{1}{2}(x-\mu_i)'\Sigma_i^{-1}(x-\mu_i)\right) \tag{3}$$

where D is the number of dimensions, μ_i is the mean vector, and Σ_i is the covariance matrix for the i-th component. The parameters $\lambda = \{w_i, \mu_i, \Sigma_i\}$ are estimated using the Expectation-Maximization (EM) algorithm[8].

1.5. Data Generator Based on Radial Basis Function Networks

1.5.1. Dynamic Decay Adjustment in RBF Networks (RBF-DDA)

The RBF-DDA algorithm enhances the capability of Radial Basis Function networks by introducing dynamic network adjustments during the training phase[9]. This method facilitates the network's automatic learning of the optimal number of kernels (hidden units), their centers, radii, and associated weights for each category of data. The algorithm proceeds as follows.

Given a set of training data $\{(\mathbf{x}_i, c)\}$, where $\mathbf{x}_i$ is the input vector and c is the associated class label, the RBF-DDA network adjusts its kernels using the following steps:

1. **Initialization:** Set the activation weight A_k of all kernel p_k to zero.
2. **Training Epoch:** Present each training pattern $(\tilde{\mathbf{x}}, c)$ to the network:
   - If there exists a kernel p_k of the correct class c such that its response $R_k(\tilde{\mathbf{x}})$ is above the positive threshold θ^+, increment its weight $A_k+ = 1$.
   - Otherwise, introduce a new kernel p_{m_c+1} for class c with its center $\mathbf{r}_{m_c+1} = \tilde{\mathbf{x}}$ and set its weight $A_{m_c+1} = 1$. Determine its radius σ_{m_c+1} to be the maximum σ such that for all kernel p_j not in class

c, $R_{m_c+1}(\mathbf{r}_j) < \theta^-$, ensuring minimal activation for non-target classes.

- **Radius Adjustment:** Adjust the radius σ of conflicting kernels(all kernels not in class c) to minimize activation for non-target patterns by setting $\sigma = \max\{\sigma : R(\tilde{\mathbf{x}}) < \theta^-\}$.

The response $R_k(\mathbf{x})$ of a kernel p_k to an input $\mathbf{x}$ is computed as $R_k(\mathbf{x}) = \exp\left(-\frac{||\mathbf{x}-\mathbf{r}_k||^2}{(\sigma_k)^2}\right)$, where $\mathbf{r}_k$ is the center and σ_k is the radius of kernel p_k. The algorithm iterates over all training data until the conditions specified by equations (4) and (5) are met, indicating that each pattern $\tilde{\mathbf{x}}$ of class c activates at least one kernel above θ^+ and all kernels of other classes below θ^-.

$$\forall c, \exists p_i : R_i(\tilde{\mathbf{x}}) \geq \theta^+ \tag{4}$$

$$\forall k \neq c, 1 \leq j \leq m_k : R_j(\tilde{\mathbf{x}}) < \theta^- \tag{5}$$

The thresholds θ^+ and θ^- are set to values such as $\theta^+ = 0.4$ and $\theta^- = 0.2$. They control the sensitivity of the network to the correct class and the adjustment of the kernels. After training, the RBF-DDA network can generate new data points by sampling from the adjusted Gaussian kernels, thereby creating virtual patient instances that retain the statistical properties of the training data.

1.5.2. Generating Data from RBF Networks

Upon the completion of the RBF-DDA training, the network comprises a set of Gaussian kernels with centers $\mathbf{r}_k$, weights A_k, and standard deviations σ_k. The Semiartificial package in R was utilized to construct the RBF-DDA network and to generate virtual patient data[10]. To generate new data, we perform the following steps:

1. **Kernel Selection:** After training the RBF-DDA network, kernels with weights below a certain threshold ($minW$) are discarded to prevent overfitting. The remaining kernels retain their centers ($\mathbf{r}_k$), weights (A_k), and class labels (c_k).
2. **Instance Assignment to Kernels and Standard Deviation Adjustment:** Each instance in the training set is assigned to the kernel that yields the maximum activation value for that instance. This assignment is determined by the activation function of the kernels as

follows:

$$z_i = \arg\max_{k \in \mathcal{M}} \exp\left(-\frac{||\mathbf{x}_i - \mathbf{r}_k||^2}{\sigma_k^2}\right)$$

where $\mathbf{x}_i$ is the i-th instance, and $\mathcal{M}$ is the kernel set in the hidden layer. The standard deviation for each kernel (σ_k) is recalculated to expand the kernel's coverage and include a wider range of instances. For each kernel in the network, the new standard deviation is derived from the spread of the associated instances.

3. **Virtual Patient Data Generation:** Virtual patient instances are generated from the adjusted kernels. The number of instances n_k generated by each kernel is proportional to its weight relative to the total weight of all kernels within the same class:

$$n_k = \frac{A_k}{\sum_{k' \in C_k} A_{k'}} \cdot N_{class}$$

where C_k denotes the set of kernels belonging to the same class as kernel k, and N_{class} is the total number of new instances to be generated for that class. New data points ($\mathbf{x}_{new}$) are sampled from the multivariate normal distribution centered at each kernel's center with its adjusted standard deviation:

$$\mathbf{x}_k^{new} \sim \mathcal{N}(\mathbf{r}_k, (\sigma_k^{new})^2), \quad \text{for } k = 1, \ldots, K$$

Where $\mathcal{N}$ denotes the Gaussian distribution and K is the total number of kernels. The generated instances $\mathbf{x}_k^{\text{new}}$ are then combined with the original dataset to enhance the diversity and representativeness of the data used for further analysis.

After detailing the steps of the RBF-DDA algorithm and how virtual patient data is generated, it is important to note that the entire process can be efficiently implemented using the Semiartificial package in R[10]. This package simplifies the generation of data by encapsulating the aforementioned steps into function calls, thereby streamlining the creation of virtual patient instances.

2. Supplementary Results

2.1. Model Fitting Results for All Patients and Four Blood Cell Types

We demonstrated the fitting results of WBC for eight representative patients, showing that our model can accurately capture the dynamic processes

of white blood cells with different trends, as shown in Figure S2. To assess the fitting performance of our ODE models for all patients and four types of blood cells (PLT, WBC, NEUT, and HGB), we calculated the Normalized Root Mean Square Error (NRMSE) for each patient's time-series data. The NRMSE provides a measure of the goodness of fit, with lower values indicating better fitting performance. The NRMSE is calculated using the following formula:

$$NRMSE = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}}{\max(y) - \min(y)}, \tag{6}$$

where y_i are the observed values, $\hat{y}_i$ are the predicted values from the ODE model, and n is the number of observations. The denominator, $\max(y) - \min(y)$, normalizes the RMSE by the range of observed values, making it a dimensionless measure.

Due to the large number of patients (479), we did not present the individual fitting results for each patient. Instead, we summarized the fitting performance by plotting the distribution of NRMSE values for each blood cell type. Figure S3 shows the distribution of NRMSE values for the four blood cell types across all patients. The distributions illustrate that the majority of patients have low NRMSE values, indicating that the models generally provide a good fit to the time-series CBC data.

2.2. Correlation of Model Parameters with Duration of Grade IV myelosuppression

We calculated the correlation between the patient heterogeneity parameters of the ODE models, each established for different types of cells (PLT, WBC, NEUT, and HGB), and the duration of grade IV myelosuppression specific to each cell type using Pearson's correlation coefficient. For instance, for the WBC, we assessed the correlation between the parameters in the ODE model of WBC and the duration of grade IV myelosuppression of WBC. Significant correlations with a p-value less than 0.05 were identified, as shown in Table S2. We plotted scatter plots of these parameters against the corresponding cells' duration of grade IV myelosuppression and observed that four parameters exhibited a clear negative correlation trend, while others did not, as shown in Figure S4. Based on these results, we selected the parameters B_{NEUT}, B_{WBC}, B_{PLT}, and B_{HGB} to construct our quantitative index.

We plotted scatter plots of these parameters against the corresponding cell type's duration of grade IV myelosuppression and observed that four parameters exhibited a clear negative correlation trend, as shown in Figure S4. Based on these results, we selected the parameters B_{NEUT}, B_{WBC}, B_{PLT}, and B_{HGB} to construct our quantitative index.

2.3. Correlation Between Time-Series Features and Clinical Indicators

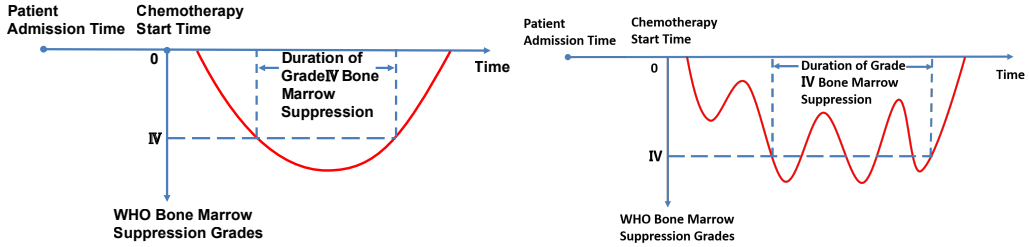
We calculated the correlation between the patient's clinical indicators and time-series features, as shown in Figure S5. Due to the large number of clinical indicators, only the most critical features are displayed in the figure. It can be observed that these features exhibit correlations with each other. Therefore, instead of generating virtual patient data for each feature separately, we treated the patient's time-series features and clinical characteristics as a large set of covariates.

2.4. Comparison of Density Distributions between Virtual and Real Patient Data

We conducted a Kolmogorov-Smirnov (KS) test to compare the density distributions of the generated virtual patient data and the real patient data for various continuous features. The KS statistic and p-value for each feature are presented in Table S3. A p-value greater than 0.01 indicates that the distributions of the virtual and real patient data for that feature are not significantly different, demonstrating that our generated virtual patient data closely match the distribution of the real patient data without being identical. This ensures that the virtual data effectively mimic the real patient data while maintaining some variability.

In addition, we evaluated the preservation of multivariate structure using Hotelling's T-squared tests and visualized the result with t-SNE in Figure S6.

3. Supplementary Figures



(a) Normal Recovery Pattern (b) Fluctuating Recovery Pattern

Figure S1: Patterns of grade IV myelosuppression duration. (a) Normal Recovery Pattern: Illustrates a patient experiencing a single, continuous episode of grade IV myelosuppression. The duration is measured from the onset to the end of this episode. (b) Fluctuating Recovery Pattern: Illustrates a patient experiencing multiple episodes of grade IV myelosuppression. The duration is the cumulative time from the first onset to the last end of these episodes.

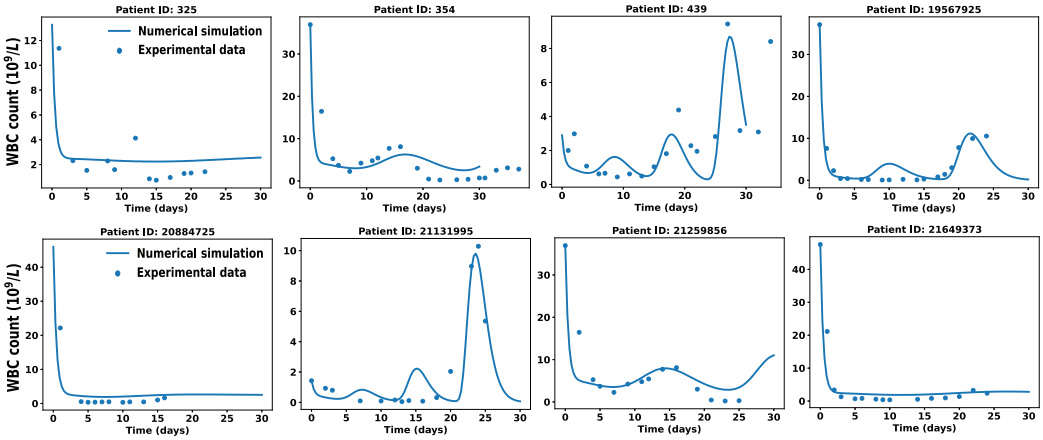


Figure S2: Fitting results of WBC count for 8 representative patients. Each subplot represents data from a different patient, with the dots indicating experimental data and the solid lines representing numerical simulation results. The patient IDs are indicated above each subplot.

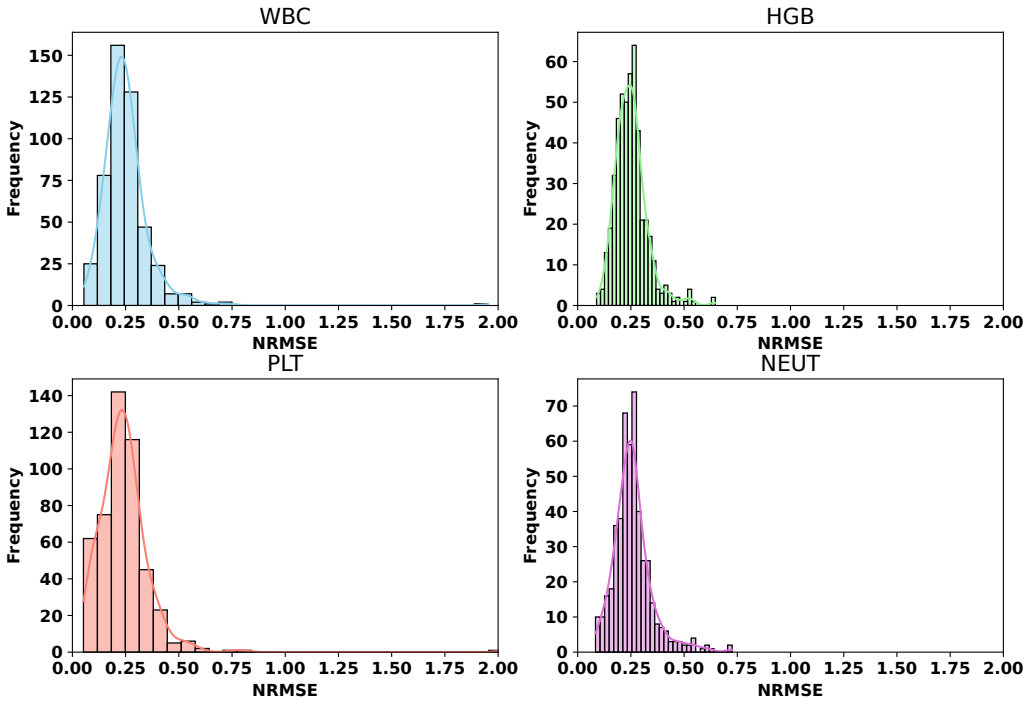


Figure S3: The histogram shows the distribution of NRMSE across all patients, indicating the overall fitting performance of the ODE models for four blood cell types.

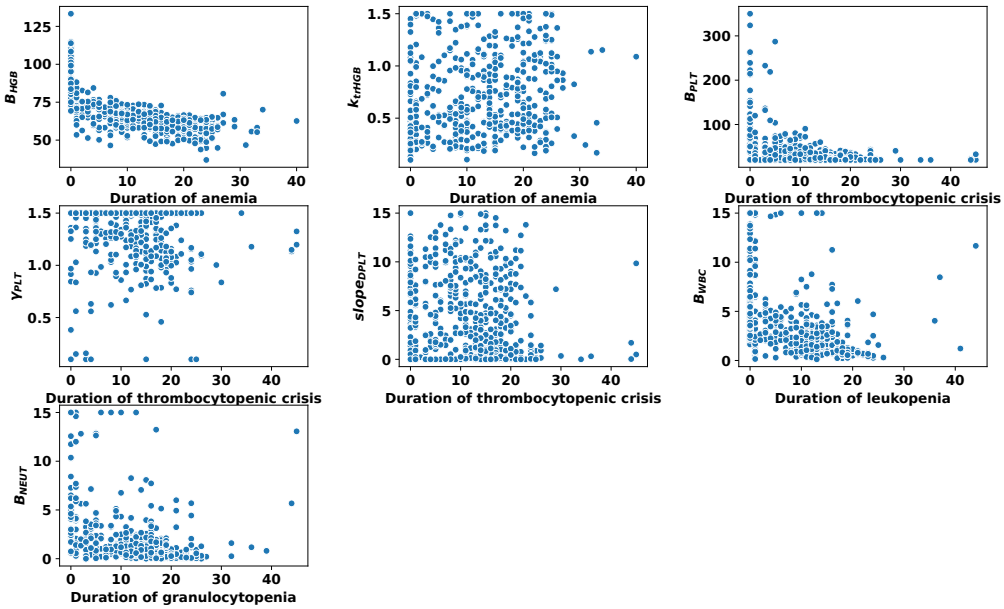


Figure S4: Scatter plots showing the correlation of selected parameters with the duration of grade IV myelosuppression for corresponding cell types.

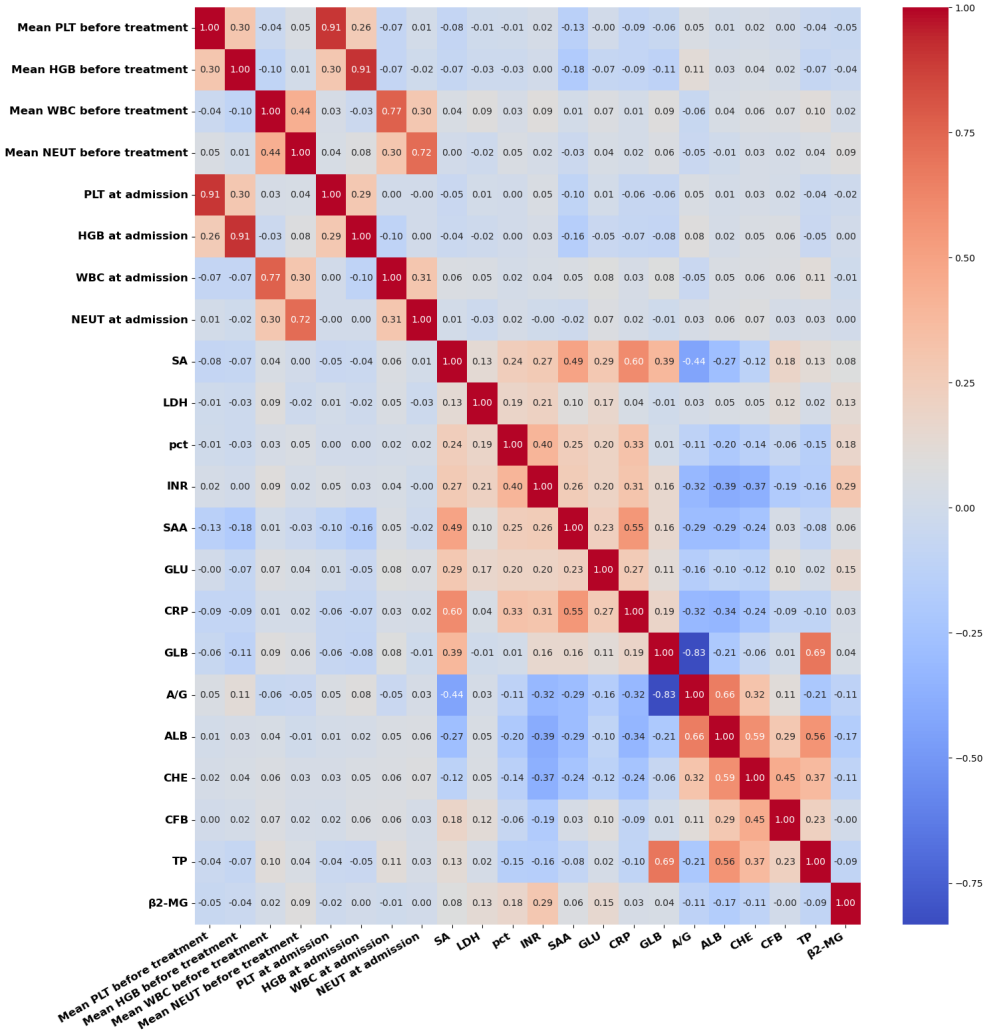


Figure S5: This heatmap illustrates the correlation between various clinical indicators and time-series features in AML patients.

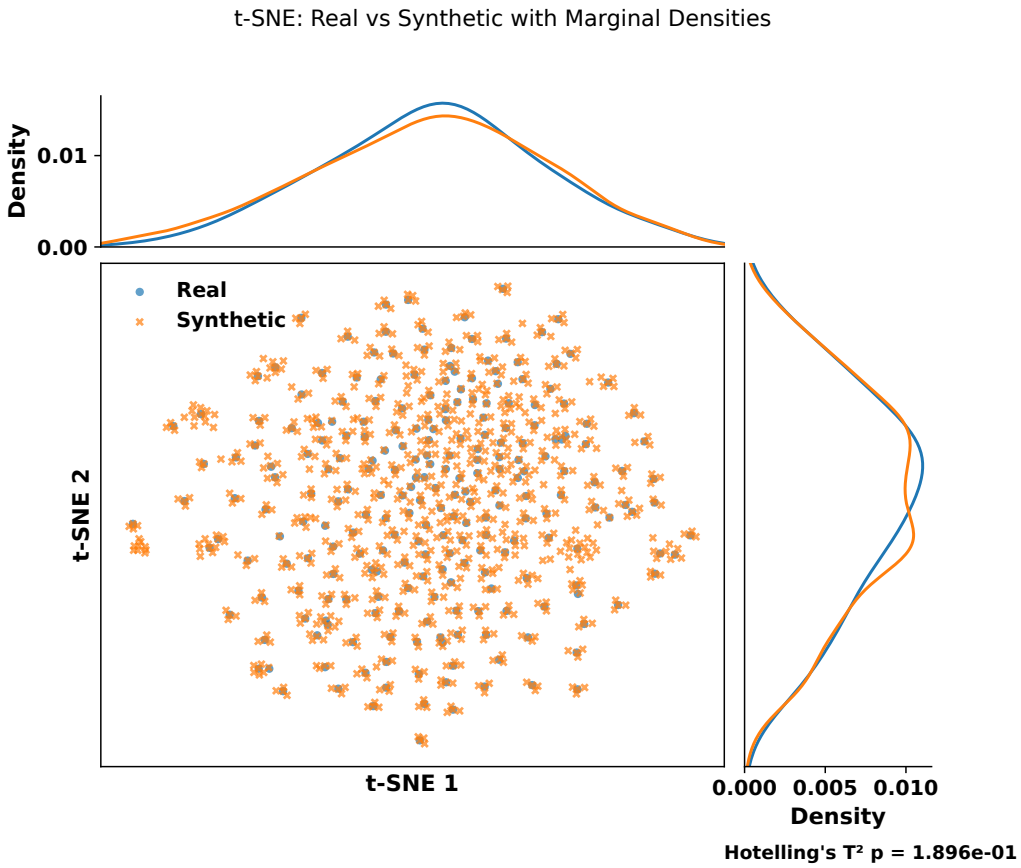


Figure S6: This is a t-SNE visualization of the real data and the virtual data.

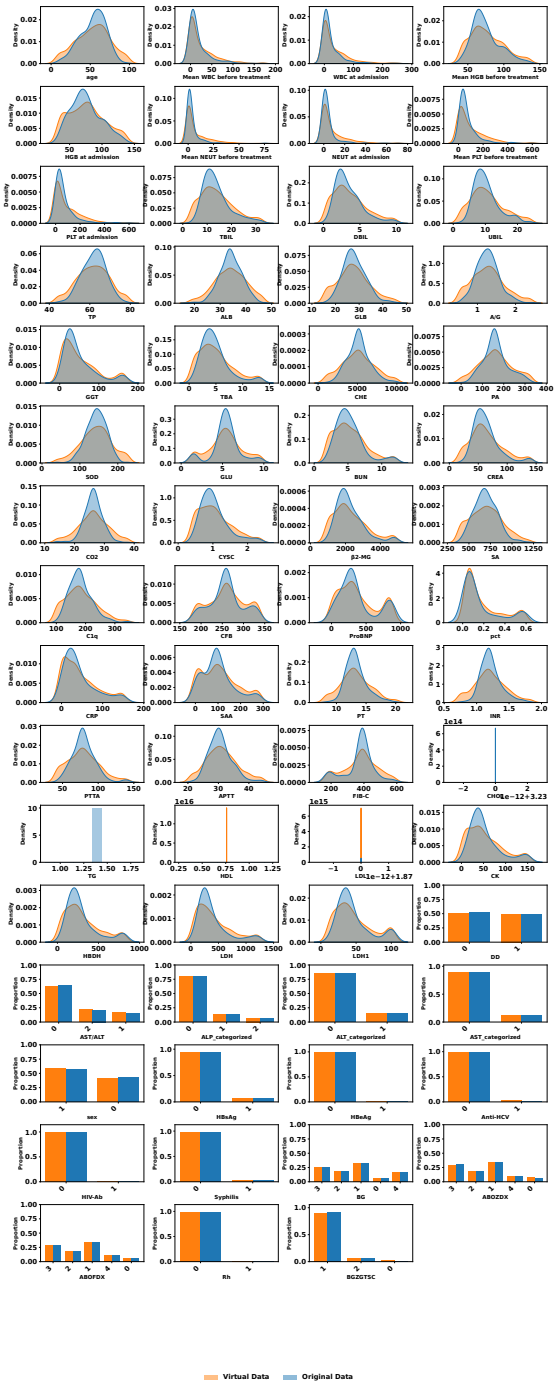


Figure S7: The comparison of each feature's distribution.

4. Supplementary Tables

Table S1: Table of Parameters

Parameters	Biological meaning	Units
k_{pr}	the proliferation rate	1/day
k_{tr}	Transition rate	1/day
γ	Feedback exponent	–
$slope_A$	drug sensitivity of Ara-C	L/μmol
$slope_D$	drug sensitivity of Daunorubicin	L/μmol
B	Baseline WBC count	$10^9/L$
MM_A	Molecular Mass of Ara-C	g/mol
MM_D	Molecular Mass of Daunorubicin	g/mol
V_c	Volume of central compartment	L
k_{ma}	Apoptosis rate	1/day
BSA	Body Surface Area	m^2

Table S2: Significant Correlations between Model Parameters and Duration of Grade IV myelosuppression

Parameter	**Cell Type**	**Correlation Coefficient**	**P-value**
B_{HGB}	HGB	-0.647054	3.659753e-58
k_{trHGB}	HGB	0.176466	1.033606e-04
B_{PLT}	PLT	-0.448335	4.614572e-25
γ_{PLT}	PLT	-0.097804	3.234745e-02
$slope_{DPLT}$	PLT	-0.170414	1.788683e-04
B_{WBC}	WBC	-0.400448	7.100693e-20
B_{NEUT}	NEUT	-0.360299	3.955852e-16

Table S3: KS Test Results for Virtual and Real Patient Data

Feature	KS Statistic	p-value	Feature	KS Statistic	p-value
CFB	0.16	4.46e-03	SOD	0.16	4.46e-03
LDH	0.16	6.13e-03	INR	0.16	6.13e-03
TBIL	0.16	6.13e-03	DBIL	0.15	1.13e-02
ProBNP	0.14	1.51e-02	PA	0.14	1.51e-02
β2-MG	0.14	1.51e-02	SA	0.14	2.01e-02
HBDH	0.14	2.65e-02	A/G	0.13	3.46e-02
LDH1	0.13	4.48e-02	APTT	0.13	4.48e-02
GLU	0.13	4.48e-02	SAA	0.13	4.48e-02
GLB	0.13	4.48e-02	TBA	0.12	5.75e-02
CYSC	0.12	5.75e-02	TP	0.12	7.31e-02
PTTA	0.12	7.31e-02	C1q	0.11	1.15e-01
CRP	0.11	1.15e-01	BUN	0.11	1.15e-01
ALB	0.11	1.15e-01	age	0.10	2.15e-01
CREA	0.09	3.10e-01	ABOFDX	0.03	1.00e+00
HDL	0.00	1.00e+00	TG	0.00	1.00e+00
CHOL	0.00	1.00e+00	LDL	0.00	1.00e+00
Mean NEUT before treatment	0.32	3.56e-11	CO2	0.20	1.72e-04
Mean PLT before treatment	0.27	2.99e-08	FIB-C	0.19	2.55e-04
Mean WBC before treatment	0.25	4.43e-07	PT	0.17	1.63e-03
Mean HGB before treatment	0.12	7.31e-02	UBIL	0.16	3.21e-03
NEUT at admission	0.29	5.32e-09	GGT	0.20	1.72e-04
WBC at admission	0.26	2.63e-07	CK	0.18	7.97e-04
PLT at admission	0.25	4.43e-07	pct	0.17	2.30e-03
HGB at admission	0.12	7.31e-02	CHE	0.16	3.21e-03

Table S4: The features of the dataset, including derived indicators.

White Blood Cell count (WBC)	Hemoglobin (HGB)
Neutrophils (NEUT)	Platelet count (PLT)
Sex	Age
Total Bilirubin (TBIL)	Direct Bilirubin (DBIL)
Unconjugated Bilirubin (UBIL)	Total Protein (TP)
Albumin (ALB)	Globulin (GLB)
Albumin/Globulin ratio (A/G)	Gamma-Glutamyl Transferase (GGT)
Total Bile Acid (TBA)	Cholinesterase (CHE)
Prealbumin (PA)	Sodium (SOD)
Glucose (GLU)	Blood Urea Nitrogen (BUN)
Creatinine (CREA)	Carbon Dioxide (CO2)
Cystatin C (CYSC)	Beta-2 Microglobulin (β2-MG)
Serum Amyloid A / SA (SA)	Complement Component 1q (C1q)
Factor B (B)	Pro-Brain Natriuretic Peptide (ProBNP)
Procalcitonin (pct)	C-Reactive Protein (CRP)
Serum Amyloid A (SAA)	Prothrombin Time (PT)
International Normalized Ratio (INR)	Prothrombin Time Activity (PTTA)
Activated Partial Thromboplastin Time (APTT)	Fibrinogen Concentration (FIB-C)
D-dimer (DD)	Creatine Kinase (CK)
Hydroxybutyrate Dehydrogenase (HBDH)	Lactate Dehydrogenase (LDH)
Lactate Dehydrogenase Isoenzyme 1 (LDH1)	Hepatitis B Surface Antigen (HBsAg)
Hepatitis B e Antigen (HBeAg)	Hepatitis C Virus Antibody (Anti-HCV)
Human Immunodeficiency Virus Antibody (HIV-Ab)	Syphilis test (Syphilis)
Blood Group (BG)	ABO Blood Type (forward typing, ABOZDX)
ABO Blood Type (reverse typing, ABOFDX)	Rhesus factor (Rh)
Blood Group Crossmatch Test (BGZGTSC)	

References

[1] Pezoulas V. C, Grigoriadis G. I, Tachos N. S, Barlocco F, Olivotto I, Fotiadis D. I, Generation of virtual patient data for in-silico cardiomyopathies drug development using tree ensembles: a comparative study, in: 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), IEEE, 2020, pp. 5343–5346.

[2] Pezoulas V. C, Kourou K. D, Kalatzis F, et al., Medical data quality assessment: On the development of an automated framework for medical data curation, *Computers in biology and medicine* 107 (2019) 270–283.

[3] Pezoulas V. C, Kourou K. D, Kalatzis F, et al., Enhancing medical data quality through data curation: a case study in primary sjögren's syndrome, *Clin Exp Rheumatol* 118 (3) (2019) 90–96.

[4] Grading standard for acute and subacute toxicity of anticancer drugs(who standard), *Cancer* (3) (1992) 254.

[5] Fornari C, Pin C, Yates J. W, Mettetal J. T, Collins T. A, Importance of stability analysis when using nonlinear semimechanistic models to describe drug-induced hematotoxicity, *CPT: Pharmacometrics & Systems Pharmacology* 9 (9) (2020) 498–508.

[6] Marini F, Walczak B, Particle swarm optimization (pso). a tutorial, *Chemometrics and Intelligent Laboratory Systems* 149 (2015) 153–165.

[7] Reynolds D. A, others, Gaussian mixture models., *Encyclopedia of biometrics* 741 (659-663) (2009).

[8] Moon T. K, The expectation-maximization algorithm, *IEEE Signal processing magazine* 13 (6) (1996) 47–60.

[9] Berthold M, Diamond J, Boosting the performance of rbf networks with dynamic decay adjustment, *Advances in neural information processing systems* 7 (1994).

[10] Robnik-Sikonja M, semiArtificial: Generator of Semi-Artificial Data, r package version 2.4.1 (2021).
URL `https://CRAN.R-project.org/package=semiArtificial`