Hemoglobin Response Modeling under Erythropoietin Treatment: Physiological Model-Informed Machine Learning Method

Authors:

Zhongyu Zhang, Zukui Li

Date Submitted: 2022-10-21

Keywords: Renal Anemia, Erythropoietin Therapy, Parameter Identification, Physics-Informed Neural Networks

Abstract:

Patients with renal anemia (RA) are usually treated with recombinant human erythropoietin (EPO) because of insufficient renal EPO secretion. The establishment of a good hemoglobin (Hgb) response model is a necessary condition for dose optimization design. The purpose of this paper is to apply physics-informed neural networks (PINN) to build the Hgb response model under EPO treatment. Neural network training is guided by physiological model to avoid overfitting problem. During the training process, the parameters of the physiological model can be estimated simultaneously. To handle differential equations with impulse inputs and time delays, we propose approximate analytical expressions for the pharmacokinetic (PK) model and weighted formulations for the pharmacology (PD) model, respectively. The improved PK/PD model was incorporated into PINN for training. Tests on simulated data show that the proposed method has good performance.

Record Type: Published Article

Submitted To: LAPSE (Living Archive for Process Systems Engineering)

Citation (overall record, always the latest version):	LAPSE:2022.0094
Citation (this specific file, latest version):	LAPSE:2022.0094-1
Citation (this specific file, this version):	LAPSE:2022.0094-1v1

CSChE Systems & Control Transactions

Open Access | Original Research

Conference Proceedings
Canadian Chemical Engineering Conference 2022
October 23-26, Vancouver, BC, Canada

PSEcommunity.org/LAPSE:2022.0094

PSE PRESS

Hemoglobin Response Modeling under Erythropoietin Treatment: Physiological Model-Informed Machine Learning Method

Zhongyu Zhang[a] and Zukui Li[a]*

[a] University of Alberta, Department of Chemical and Materials Engineering, Edmonton, Alberta, Canada
*Corresponding Author: zukui@ualberta.ca

ABSTRACT

Patients with renal anemia (RA) are usually treated with recombinant human erythropoietin (EPO) because of insufficient renal EPO secretion. The establishment of a good hemoglobin (Hgb) response model is a necessary condition for dose optimization design. The purpose of this paper is to apply physics-informed neural networks (PINN) to build the Hgb response model under EPO treatment. Neural network training is guided by physiological model to avoid overfitting problem. During the training process, the parameters of the physiological model can be estimated simultaneously. To handle differential equations with impulse inputs and time delays, we propose approximate analytical expressions for the pharmacokinetic (PK) model and weighted formulations for the pharmacology (PD) model, respectively. The improved PK/PD model was incorporated into PINN for training. Tests on simulated data show that the proposed method has good performance.

Keywords: Renal Anemia, Erythropoietin Therapy, Parameter Identification, Physics-Informed Neural Networks

Date Record: Original manuscript received October 15, 2022. Published October 21, 2022.

INTRODUCTION

Renal anemia (RA) is a disease caused by the deficiency of erythropoietin (EPO) secretion by kidney. This is due to impaired renal function or some toxic substances in the plasma of uremic patients interfering with the production and metabolism of red blood cells [1]. It is a common complication of chronic kidney disease and a risk factor for cardiovascular complications. Now the efficacy of EPO in the treatment of RA has been well documented [2]. EPO is a glycoprotein with 165 amino acids. When oxygen delivery to specific cells within the kidney is reduced, secretion of EPO increases while circulating in the plasma and stimulating bone marrow progenitor cells, thereby increasing erythrocyte production [3]. If the increase in erythrocyte numbers relieves the hypoxic signal, EPO expression is downregulated. Despite its clinical effectiveness, there are potential drug-induced risks in patients treated with EPO. In practice, clinicians usually adjust the frequency and dose of EPO based on current hemoglobin (Hgb) measurements and previous dosing rules. It requires rich clinical experience. While low Hgb level leads to anemia, high Hgb levels can increase the risk of Hgb variation patterns and even mortality for the patient [4]. Therefore, it is important to develop decision support tools that can help the medical staff determine the appropriate dose and frequency of EPO to maintain the target Hgb level and reduce the cost of treatment.

To help physicians make patient-specific decisions on the optimal dosage of EPO treatment, a model that describes the Hgb response to the EPO dosing is necessary. Existing methods of erythropoiesis modeling can be divided into two main categories. One is physiologically driven modeling, which usually uses a combination of pharmacokinetic (PK) and pharmacodynamics (PD) models to describe the dynamics of Hgb concentration following the administration of intravenous EPO [5], [6], [7]. The other is data-driven modeling. It sets EPO dose data and Hgb measurements as input and output, respectively. Then data-driven models like neural network or autoregressive model can be trained to represent the erythropoiesis process [8], [9]. Both approaches have advantages and disadvantages. By building a physiologically driven model, we can get the details of the system states. Moreover, if the theoretical model is correct, the physiologically driven model can work well. But in practice, conventional method often has difficulties in estimating physiological parameters for ill-posed inverse problems [10]. On the other hand, although a data-driven model can approximate complex functions, it is sensitive to data noise and may not perform in prediction.

Given the above problems, this work aims to develop a more efficient method to build an erythropoiesis model. The proposed method uses physics-informed neural networks (PINN) to identify the physiological model parameters. Just like the framework of PINN proposed in paper [11], the front part of the neural network is similar to the ordinary fully connected neural network. With the network output and the associated gradients calculated from auto-differentiation, the physiological model equation residuals are incorporated in the loss function to enforce the physiological model information.

In this way, PINN achieves good estimation and robustness to noise and disturbances.

PHYSIOLOGICAL MODEL

Regarding the Hgb response to EPO dosage, an example of clinical data record is shown in Figure 1. Hgb level is recorded around every 2 weeks, patients with late-stage renal disease receive EPO treatment 1 to 3 times per week [12].

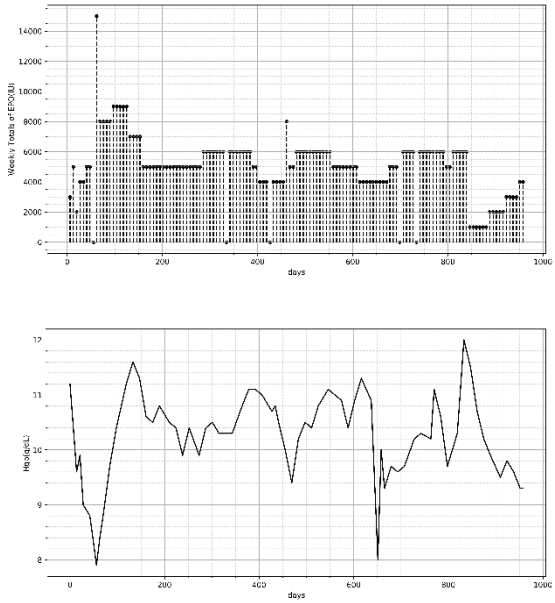


Figure 1. A clinical example data of EPO dosages and Hgb responses

Paper [13] has proposed a physiological erythropoiesis model to describe the Hgb-EPO relationship. The model consists of PK and PD model equations. The PK model describes how the body affects specific exogenous chemical substances through absorption and distribution mechanisms after drug administration, as well as changes in the metabolism of substances in the body, and the effects and excretion pathways of drug metabolites [14]. The PD model explains pharmacological effects on living systems, including reactions and binding to cellular components, and the biochemical and physiological consequences of these effects [15].

For the PK model, it can be described as:

$$\frac{dE(t)}{dt} = -\frac{V \cdot E(t)}{K_m + E(t)} - \alpha \cdot E(t) + dose(t) \quad (1)$$

$$E_p(t) = E(t) + E_{en} \quad (2)$$

$$k_{in}(t) = \frac{S \cdot E_p(t)}{C + E_p(t)} \quad (3)$$

$$\mathrm{E_{en}} = \frac{\mathrm{C \cdot H_{en}}}{\mu \cdot \mathrm{K_H \cdot S - H_{en}}} \quad (4)$$

The PD model is defined as bellow:

$$\frac{dR(t)}{dt} = k_{in}(t - D) - \frac{4x_1(t)}{\mu^2} \quad (5)$$

$$\frac{dx_1(t)}{dt} = x_2(t) \quad (6)$$

$$\frac{dx_2(t)}{dt} = k_{in}(t - D) - \frac{4x_1(t)}{\mu^2} - \frac{4x_2(t)}{\mu} \quad (7)$$

$$Hgb(t) = K_H \cdot R(t) \quad (8)$$

In the PK model equations, $E(t)$ denotes the amount of exogenous recombinant human EPO, E_{en} denotes the endogenous EPO, $E_p(t)$ is the total EPO of the dynamic pool in plasma, $k_{in}(t)$ is the red blood cells (RBC) production rate, and $dose(t)$ is the EPO dosing in international unit (IU) which is modeled as a train of impulses.[13] Additionally, the model contains some parameters: H_{en} is the Hgb level due to endogenous EPO, μ represents the mean RBC life span, V is the maximum exogenous EPO clearance rate, K_m stands for the exogenous EPO level that produces half-maximum clearance rate, α is the linear clearance constant, S represents the maximal RBC production rate stimulated by EPO, C is the amount of EPO that produces half-maximum RBC production rate [13].

In the PD model, states $R(t)$ represent the population of red blood cells (RBC), states $x_1(t)$ and $x_2(t)$ are internal states that aid in calculating $R(t)$, $Hgb(t)$ is the hemoglobin level which can be detected clinically, parameters D is the time required for EPO-stimulated RBCs to start forming, K_H is the average amount of Hgb per RBC (mean corpuscular hemoglobin, or MCH, in a complete blood count) which takes value of $K_H = 29.5pg/cell$ [13].

The initial conditions can be determined as below

$$R_0 = \frac{Hgb_0}{K_H} \quad (9)$$

$$x_{10} = \frac{\mu \cdot (H_{en} - \mu \cdot K_H \cdot \dot{R}_0)}{4K_H} \quad (10)$$

$$x_{20} = \frac{K_H \cdot R_0 - H_{en} + \mu \cdot K_H \cdot \dot{R}_0}{K_H} \quad (11)$$

Based on the above physiological model, eight unknown model parameters $\alpha, C, D, H_{en}, K_m, \mu, S, V$ can be estimated using collected data for each patient. In this work, we use the inverse PINN for the parameter estimation.

PHYSICS-INFORMED MACHINE LEARNING

Paper [11] proposed the PINN which is a type of neural network trained to solve supervised learning tasks while following given physical law described by partial differential equations. It is shown by [16] that the method performs well to identify the unknown model parameters.

The PINN structure [17] used in this work is shown in Figure 2. Time is the input. States of the physiological system are output. The hidden layers perform nonlinear transformations on the data [18]. It is similar to a fully connected neural network but adds three extra layers to accelerate convergence. Input-scaling layer is designed to shrink the input time domain through a linear scaling function. When differential equations solution has a certain pattern, for example, the solution follows periodicity or attenuation, feature layer can be set as $sin(kt)$ or e^{-kt} respectively [16]. If states have different magnitudes, the output-scaling layer can be used to scale them.

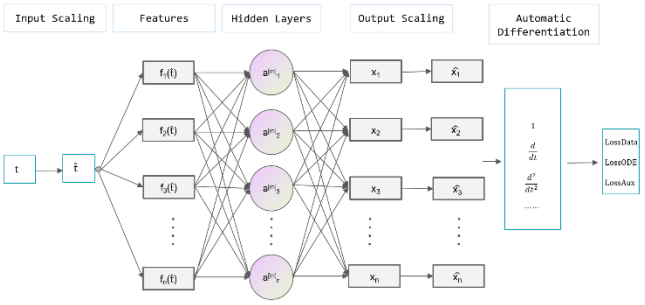


Figure 2. Physics informed neural networks architecture

The main idea of PINN is to incorporate the physical model equation residual (error) into the loss function of the neural network training. During the training process, the network model parameters and the physical model parameters can be estimated simultaneously. Consider a set of ODE equations

$$\frac{dx_s}{dt} = f_s(x_s, t; p) \qquad s = 1, \dots, S \tag{12}$$

The loss function is composed of 3 parts as follows.

$$Loss(\theta, p) = Loss^{data}(\theta) + Loss^{ode}(\theta) + Loss^{aux}(\theta) \tag{13}$$

where

$$Loss^{data}(\theta) = \sum_{m=1}^{M} w_m^{data} Loss_m^{data}$$

$$= \sum_{m=1}^{M} w_m^{data} \left[\frac{1}{N^{data}} \sum_{n=1}^{N^{data}} \left(y_m(t_n) - \widehat{x_m}(t_n;\theta)\right)^2\right] \tag{14}$$

$$\mathrm{Loss}^{\mathrm{ode}}(\theta) = \sum_{s=1}^{S} \mathrm{w}_s^{\mathrm{ode}} \mathrm{Loss}_s^{\mathrm{ode}}$$

$$= \sum_{s=1}^{S} \mathrm{w}_s^{\mathrm{ode}} \left[\frac{1}{\mathrm{N}^{\mathrm{ode}}} \sum_{n=1}^{\mathrm{N}^{\mathrm{ode}}} \left(\frac{\mathrm{d}\widehat{x_s}}{\mathrm{dt}}|_{\tau_n} - \mathrm{f_s}(\widehat{x_s}(\tau_n;\theta), \tau_n; \mathrm{p})\right)^2\right] \tag{15}$$

$$Loss^{aux}(\theta) = \sum_{s=1}^{S} w_s^{aux} Loss_s^{aux}$$

$$= \sum_{s=1}^{S} w_s^{aux} [x_s(T_0) - \widehat{x_s}(T_0;\theta)]^2 \tag{16}$$

$\mathrm{N}^{\mathrm{data}}$ is the number of sample data points where both the input (time and additional control input) and the output response are available. N^{ode} is the number of collocation points used to evaluate the model residual. Note that there is no response data needed for those collocation points. $Loss^{data}$ is difference values between measurements of $y_1, y_2, \dots, y_M$ and network outputs $\widehat{x_1}, \widehat{x_2}, \dots, \widehat{x_M}$ at time $t_1, t_2, \dots, t_{N^{data}}$. $Loss^{aux}$ is similar to $Loss^{data}$, but it specifically considers the start time point T_0 as an additional source. $Loss^{ode}$ is the key point of PINN. By automatic differentiation, the derivative of output states $\widehat{x_1}, \widehat{x_2}, \dots, \widehat{x_S}$ concerning input t at the time point $\tau_1, \tau_2, \dots, \tau_{N^{ode}}$ can be obtained. Then we can calculate the residual error according to the differential equations and use it as a part of the loss function. In this way, differential equations are integrated into the neural network, which attaches physical constraints to machine learning. The weighting coefficients $(w_1^{data}, w_2^{data}, \dots, w_M^{data})$, $(w_1^{ode}, w_2^{ode}, \dots, w_S^{ode})$ and $(w_1^{aux}, w_2^{aux}, \dots, w_S^{aux})$ are used to balance the loss terms. Finally, by minimizing the loss function, the parameters θ of the neural network and unidentified parameters p of differential equations are optimized together.

MODIFIED PK/PD MODEL FOR PINN

To incorporate the physiological model into the PINN framework, we face two challenges from the original PK/PD model, which are explained below.

Impulse input sequences in PK equations

Eq. 1 is a differential equation with impulse input sequence. This equation describes the decay process of exogenous EPO in the human body. Based on a simulation of this differential equation with parameter V, Km, α being set as 1660, 76.5, 0.25, respectively. Black dash-dot curve in Figure 3 illustrates the trajectory of EPO in human body during 10 days after receiving 7000 IU EPO medications on the second day. In practice, dose(t) is a train of impulses. This causes two issues when PINN is used to incorporate this physiological equation. First, the width of the impulse tends to be zero and the derivative $\frac{dE(t)}{dt}$ goes to infinity at the dosing time. It is impossible to directly evaluate the differential equation residual $\frac{dE(t)}{dt} - \left[-\frac{V \cdot E(t)}{K_m + E(t)} - \alpha \cdot E(t) + dose(t)\right]$. Secondly, the profile of $E(t)$ is not smooth under an impulse sequence input as shown in the figure. It is not very efficient to approximate this nonsmooth function through the neural network. To address this issue, we propose a method to approximate this differential equation based on the following observations.

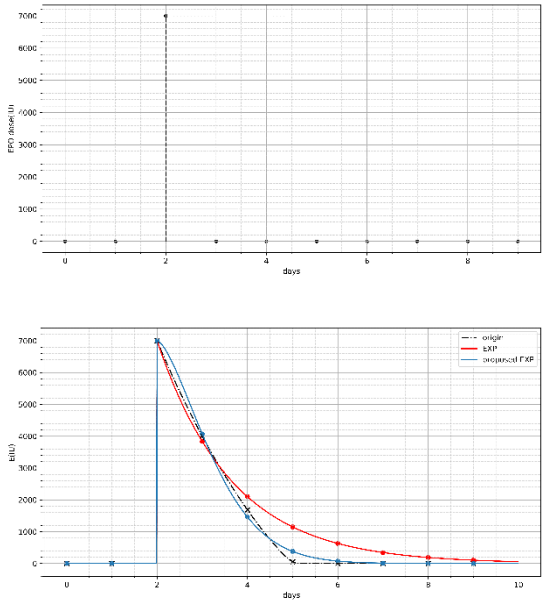


Figure 3. Trajectory of E(t) under a single EPO dosage

According to the differential equation 1, when $E(t)$ is much bigger than K_m, the equation can be simplified as

$$\frac{\mathrm{d}E(t)}{\mathrm{d}t} \approx -V - \alpha \cdot E(t) + dose(t) \tag{17}$$

For this equation, the solution trajectory of $E(t)$ is an exponential function as shown by the red line in Figure 3 with $\alpha = 0.6$. Compared with the exponential function, the curve of differential equation solution $E(t)$ decreases more quickly.

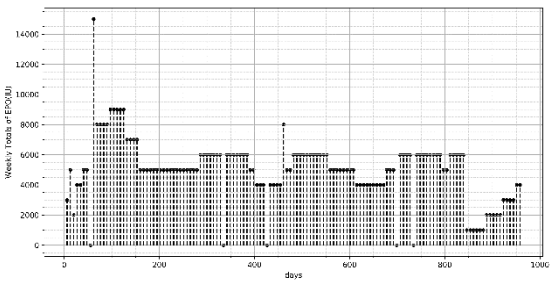
To improve the approximation accuracy, we propose the following exponential function Eq. 18 to approximate $E(t)$,

$$E(t) = \sum_{j=1}^{N(t)} dose_j \cdot exp\left[-\left(a_0 \cdot e^{\frac{-dose_j}{a_1}} + a_2\right) \cdot \left(t - t_j\right)^{a_3}\right] \tag{18}$$

where a_0, a_1, a_2, a_3 are four undetermined parameters, t_j and $dose_j$ correspond to the j-th EPO administration time and dosage value, respectively. $N(t)$ is the total number of dosing times up to time t. For example, if the patient receives 5000 IU EPO treatment and 10000 IU EPO treatment on the 20th day and 60th day, the corresponding $(t_j, dose_j)$ are (20, 5000), (30, 10000). Using exponentiation of time difference, this proposed exponential function can match the differential equation solution better at the later stage.

To demonstrate the performance of the proposed model equation, we simulate the original equation 1 to get the profile of $E(t)$ under the EPO dose sequence as shown in the top part of Figure 3. Afterwards, we sample data from the true solution (as shown by the black dash-dot curve in Figure 3) and then use least squares method to estimate the parameters in the proposed model equation 18. The estimated parameter values are a0 = 1.87, a1 = 3640, a2 = 0.269, a3 = 1.53. The $E(t)$ trajectory calculated by the proposed exponential function Eq. 18 is drawn in Figure 3 by blue line, which approximates the true response curve (black dash-dot curve) very well.

In addition to the single impulse input study, the accuracy of the proposed model equation is also tested over a sequence of EPO dosages which was obtained from clinical data. The top part of Figure 4 is the EPO treatment record. The solution of the differential equation and the approximate exponential function are shown in the middle part of Figure 4, respectively. Notice that the bottom one is the zoomed version of the red box in the middle figure to show more details. R^2 of $E(t)$ prediction is equal to 99.76%. This result verifies that the proposed model equation approximates the original exponential differential equation very well.



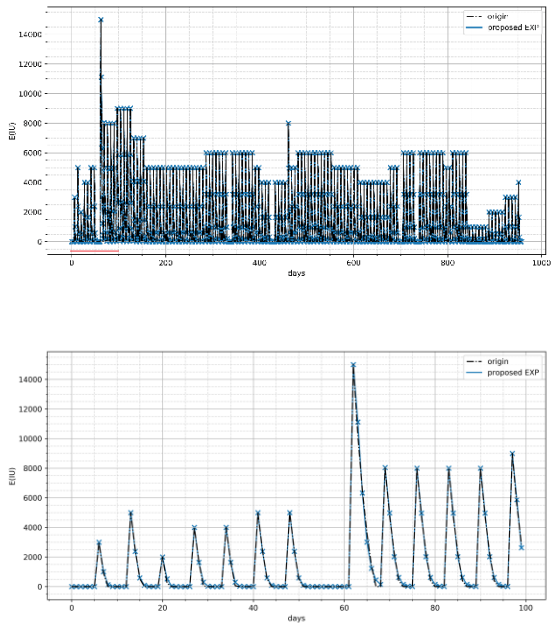


Figure 4. Trajectories of EPO in the body based on clinical data

TIME DELAY IN PD EQUATIONS

The other issue comes from Eq. 5 and Eq. 7. These two equations are delay differential equations. The delay item D is the parameter to be estimated. However, in the neural network, it is hard to calculate the partial derivative of the loss function with respect to the delay parameter.

Paper [13] has studied the low-pass filter nature of the RBC pool. As shown in Figure 5, a twice-weekly dosing sequence is simulated and it generates pulsatile and periodic EPO levels E_P and corresponding production rate k_{in}; but the periodic dynamics are largely smoothed out by the low-pass nature of the RBC pool filter [13].

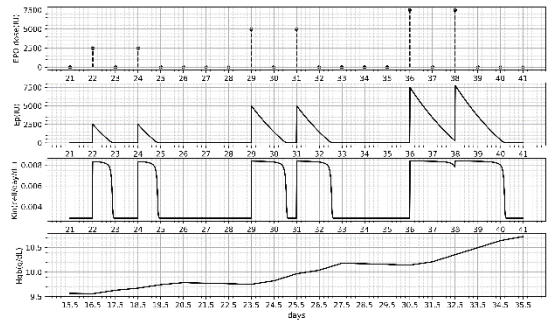


Figure 5. Low-pass nature of the RBC pool filter

Therefore, during the therapy, PK and cell production PD is relevant to the mean value of the production rate, which is denoted as $\overline{k_{\text{in}}}$ in Eq. 19 where $[iT, (i+1)T]$ is a single dose period. The Eq. 3 can be reconsidered as a memoryless nonlinear relationship between EPO doses and mean production rate

$\overline{k_{\text{in}}}$,[13] which means a similar mean production rate profile $\overline{k_{\text{in}}}$ will lead to similar Hgb trajectory.

$$\overline{k_{\text{in}}}(dose_i, T) = \frac{1}{T}\int_{iT}^{(i+1)T} k_{\text{in}}(t)dt \tag{19}$$

Based on the above analysis, we propose to convert the delay differential equations into a different form which makes the estimation easier. The method is based on a weighting function and the new equations are defined as follows:

$$\frac{dR(t)}{dt} = \lambda_1 k_{in}(t - D_1) + \lambda_2 k_{in}(t - D_2) - \frac{4x_1(t)}{\mu^2} \tag{20}$$

$$\frac{dx_2(t)}{dt} = \lambda_1 k_{in}(t - D_1) + \lambda_2 k_{in}(t - D_2) - \frac{4x_1(t)}{\mu^2} - \frac{4x_2(t)}{\mu} \tag{21}$$

$$\lambda_1 + \lambda_2 = 1 \tag{22}$$

The term $k_{in}(t - D)$ is replaced by the weighting function $\lambda_1 k_{in}(t - D_1) + \lambda_2 k_{in}(t - D_2)$, where λ_1 and λ_2 are parameters to be determined, and D_1 and D_2 are fixed as 4 and 7, respectively. This is based on the fact that the time required for progenitor cells to be stimulated by EPO and finally become reticulocytes ready to mature into RBCs is 4 - 7days [3]. The original delay parameter D can be estimated as

$$D = \lambda_1 D_1 + \lambda_2 D_2 \tag{23}$$

Figure 6 shows the RBC production rate k_{in} and the average weekly production rate $\overline{k_{\text{in}}}(T = 7)$ of the original form and the proposed weighting function respectively. Here, λ_1 and λ_2 are both set as 0.5. It illustrates that during every dose period, there is some difference between the original form $k_{in}(t - D)$ and the proposed weighting function $\lambda_1 k_{in}(t - D_1) + \lambda_2 k_{in}(t - D_2)$. For example, in the second dose period (day 22 to 29), the original model solution includes two pulses whereas the proposed weighting function produces three pulses with a smaller magnitude. However, the average weekly production rate $\overline{k_{\text{in}}}$ of the original form gets superimposed by the $\overline{k_{\text{in}}}$ of the proposed weighting function. The trajectories of Hgb level obtained from these two methods in this short term are similar, as shown in Figure 7.

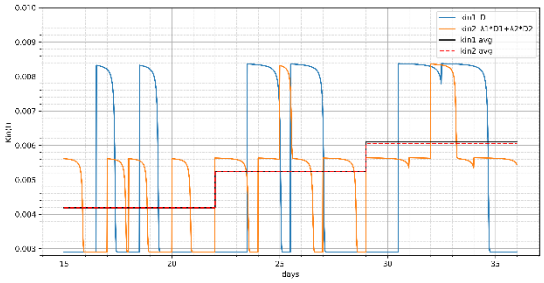


Figure 6. Comparison between the average weekly production rate $\overline{k_{\text{in}}}$ and k_{in}

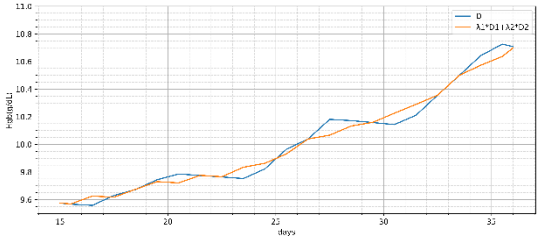


Figure 7. Hgb value of two methods in short term

Table 1: PARAMETERS FOR PK/PD MODELS

α	K_m	V	C	D	H_{en}	μ	S
0.25	46.5	2800	22.45	5.5	7.9	92.2	0.0084

Finally, we check the approximation performance over a long horizon. With parameters set as Table 1, the original model and the approximated model are both simulated. Figure 8 shows these two Hgb trajectories of the original form and delay differential equations with a weighting function. Root mean square error (RMSE) between two curves equals 0.0712.

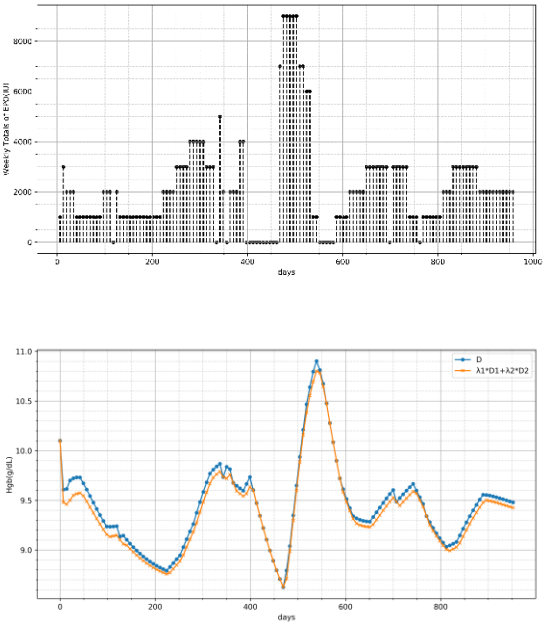


Figure 8. Long term Hgb responses: original model and approximation model

Above results show that the proposed PK/PD model modification approximate the original model very well. They provide a basis for the PINN modeling as described in the next section.

PINN USING THE MODIFIED PK/PD MODEL

According to the proposed approximation model explained, the overall physiological model used in the PINN is based on equations 18, 2,3,4,20, 6, 21, 8.

Table 2: PINN1 (FROM NOISE-FREE DATA), PINN2 (FROM NOISY DATA

	α	K_m	V	C	D	H_{en}	μ	S
True	0.1	120	600	30	5.2	9	90	0.004
PINN1	0.0959	127	607	27.6	5.36	8.89	95.0	0.00378
PINN2	0.0966	127	616	30.2	5.38	9.07	84.6	0.00422

During the PINN training process, the neural network parameters and the parameters in the physiological model are simultaneously estimated. The loss term corresponding to the model residual is based on equation 20, 6, 21.

Based on the parameters a_0, a_1, a_2, a_3, original parameters V, K_m, α can be further estimated through least squares method. Besides, the delay parameter D can be evaluated using equation 23.

Test on simulated data

To demonstrate the proposed erythropoiesis modeling based on PINN, true parameter values as listed in Table 2 and a set of EPO input sequence as shown in Figure 9 are chosen to generate a series of Hgb data by solving this PK/PD model. Then Gaussian noise with zero mean and the standard deviation of $\sigma_\epsilon = c\mu$ is added to Hgb data to simulate measurement noise, where μ is the standard deviation of original Hgb data and c is equal to 5%. According to the noise-free Hgb data and the noise-containing Hgb data, we use PINN to identify these parameters in the differential equations separately and compare the results. The algorithm is implemented in Python with the open-source library DeepXDE [19]. The neural network is formed from 5 hidden layers and each one has 64, 128, 256, 128, 64 neurons. The feature layer adopts $t, sin(t), sin(2t), sin(3t), sin(4t), sin(5t)$. The swish function is set as the activation function. In addition, we use the Adam optimizer [20] and 500000 iterations with a learning rate equal to 10^{-4}.

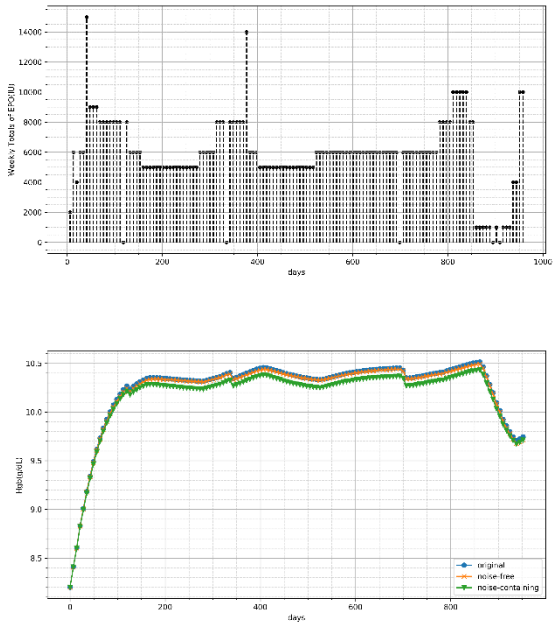


Figure 9. Simulated true Hgb responses and PINN model predictions

Based on noise-free data and noise-containing data, we can estimate the parameters for erythropoiesis modeling in Table 2. The fitting results based on the two cases are shown in Figure 9. Corresponding RMSE are 0.0160 (no noise) and 0.0755 (with noise), respectively. The result shows these inferred parameters have a higher degree of accuracy. The agreement between the Hgb solution based on the estimated parameters and exact dynamics is good considering the noise in the training data.

CONCLUSION

In this paper, we applied PINN technique to model the Hgb response under EPO treatment. This method combined physiological PK/PD model and neural network learning technology to estimate the parameters of PK/PD model. During the training of the neural network, physical laws describing the physiological model are enforced by adding the model residuals to the loss function. To address the problem that the PINN cannot easily handle the residual of the differential equation at the time instants with impulse inputs, we proposed an approximate model to replace the PK model equation. In addition, to handle the time delay in the PD differential equation, we proposed a weighting function-based formulation so that the delay parameter can be estimated by training the PINN. Tests on simulated data show that the proposed method has good prediction accuracy and is robust to noise.

The proposed modeling technique can help build individualized model for patients with renal disease. Physicians can rely on this modeling technique to develop patient-specific EPO dosing strategy to optimally manage the Hgb level of different patients. Future work can be done by integrating the PINN model into feedback control strategies to achieve this objective.

ACKNOWLEDGEMENTS

This work was supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC).

REFERENCES

1. Lankhorst CE, Wish JB. Anemia in renal disease: diagnosis and management. *Blood rev.* 24(1):39-47 (2010).
2. Hayat A, Haria D, Salifu MO. Erythropoietin stimulating agents in the management of anemia of chronic kidney disease. *Patient prefer. adher.* 2: 195 (2008).
3. Guyton A C, Hall J E. Textbook of medical physiology. *Philadelphia: Saunders* 548 (1986).
4. Bradbury BD, Danese MD, Gleeson M, Critchlow C. Effect of epoetin alfa dose changes on hemoglobin and mortality in hemodialysis patients with hemoglobin levels persistently below 11 g/dL. *Clin. J. Am. Soc. Nephrol.* 4(3):630-637 (2009).

5. Uehlinger DE, Gotch FA, Sheiner LB. A pharmacodynamic model of erythropoietin therapy for uremic anemia. *Clin. Pharm. Therap.* 51(1):76-89 (1992).
6. Ramakrishnan R, Cheung WK, Wacholtz MC, Minton N, Jusko WJ. Pharmacokinetic and pharmacodynamic modeling of recombinant human erythropoietin after single and multiple doses in healthy volunteers. *J. clin. pharmacol.* 44(9):991-1002 (2004).
7. Woo S, Krzyzanski W, Jusko WJ. Pharmacokinetic and pharmacodynamic modeling of recombinant human erythropoietin after intravenous and subcutaneous administration in rats. *J. Pharmacol. Exp. Ther.* 319(3):1297-1306 (2006).
8. Gaweda AE, Jacobs AA, Brier ME, Zurada, JM. Pharmacodynamic population analysis in chronic renal failure using artificial neural networks—a comparative study. *Neural Netw.* 16(5-6):841-845 (2003).
9. Gaweda AE, Muezzinoglu MK, Aronoff GR, Jacobs AA, Zurada JM, Brier ME. Individualization of pharmacological anemia management using reinforcement learning. *Neural Netw*. 18(5-6):826-834 (2005).
10. Shukla K, Jagtap AD, Karniadakis GE. Parallel physics-informed neural networks via domain decomposition. *J. Comput. Phys.* 447:110683 (2021).
11. Raissi M, Perdikaris P, Karniadakis GE. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J. Comput. Phys.* 378:686-707 (2019).
12. Robinson S, Conteh FS, Oppong AY, Yellowhair TR, Newville JC, Demerdash NE, Shrock CL, Maxwell JR, Jett S, Northington FJ, Jantzie LL. Extended combined neonatal treatment with erythropoietin plus melatonin prevents posthemorrhagic hydrocephalus of prematurity in rats. *Front. Cell. Neurosci.* 12:322 (2018).
13. Chait Y, Horowitz J, Nichols B, Shrestha RP, Hollot CV, Germain MJ. Control-relevant erythropoiesis modeling in end-stage renal disease. *IEEE Trans. Biomed. Eng.* 61(3):658-664 (2013).
14. Chen H X, Mooney M, Boron M, Vena D, Mosby K, Grochow L, Jaffe C, Rubinstein L, Zwiebel J, Kaplan RS. Phase II multicenter trial of bevacizumab plus fluorouracil and leucovorin in patients with advanced refractory colorectal cancer: an NCI Treatment Referral Center Trial TRC-0301. *J. Clin. Oncol.* 24(21):3354-3360 (2006).
15. Duffus J. Glossary for chemists of terms used in toxicology (IUPAC Recommendations 1993). *Pure appl. chem.* 65(9):2003-2122 (1993).
16. Karniadakis GE, Kevrekidis IG, Lu L, Perdikaris P, Wang S, Yang L. Physics-informed machine learning. *Nat. Rev. Phys.* 3(6):422-440 (2021).
17. Yazdani A, Lu L, Raissi M, Karniadakis GE. Systems biology informed deep learning for inferring parameters and hidden dynamics. *PLoS comput. Biol.* 16(11):e1007575 (2020).
18. Bishop C M. Neural networks for pattern recognition. *Oxford university press* (1995).
19. Lu L, Meng X, Mao Z, Karniadakis GE. DeepXDE: A deep learning library for solving differential equations. SIAM Rev. 63(1):208-228 (2021).
20. Kingma DP, Ba J. Adam: A method for stochastic optimization. *arXiv preprint* arXiv:1412.6980 (2014).

This conference proceeding has not been peer reviewed.