2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Adaptive Individualized Drug-Dose Response Modeling from a Limited Clinical Data: Case of Warfarin Management

Affan Affan, Jacek M. Zurada, Michael E. Brier, and Tamer Inanc

***Abstract*— Administration of drugs requires sophisticated methods to determine the drug quantity for optimal results, and it has been a challenging task for the number of diseases. To solve these challenges, in this paper, we present the semi-blind robust model identification technique to find individualized patient models using the minimum number of clinically acquired patient-specific data to determine optimal drug dosage. To ensure the usability of these models for dosage predictability and controller design, the model (In)validation technique is also investigated. As a case study, the patients treated with warfarin are studied to demonstrate the semi-blind robust identification and model (In)validation techniques. The performance of models is assessed by calculating minimum means squared error (MMSE).**

***Clinical Relevance*— This work establishes a general framework for adaptive individualized drug-dose response models from a limited number of clinical patient-specific data. This work will help clinicians in decision-making for improved drug dosing, patient care, and limiting patient exposure to agents with a narrow therapeutic range.**

I. Introduction

Thromboembolic events occur when blood clots break loose and are stuck in narrow vessels further downstream in potentially harmful locations. The effect of warfarin in the human body is determined by testing the International Normalized Ratio (INR). For optimal therapeutic results, the INR value should be between 2-3 [1]. Due to inter and intra-variability of patient characteristics, warfarin management is one of the most common causes for accidental in-house hospitalization and the failure of a general drug dosing strategy would be costly to any care facility that employs it for a long enough period. In the case of improper warfarin dosage, problems related to atrial fibrillation, post-myocardial infractions, and deep vein thrombosis (DVT) could arise. Therefore, individualized predictive models reflecting each patient's unique dose-response characteristics would be ideal.

Population-based drug dosing strategies developed and proven to be effective through common Warfarin management protocols [2], [3], however, run a risk of being ineffective or harmful in cases of overdose for outlier patients. In [2], individualized models based on a population pharmacokinetic (PK) are created using knowledge of genetic markers that effectively dose patients, with good results for patients with low responsiveness to anticoagulant. Similarly, in [4], Warfarin Regimen using A Pharmacogenetics-guided Initiation Dosing (WRAPID) algorithm is proposed to incorporate the loading and maintenance of doses based on genetic information. The collection of genetic markers from patients is expensive, and time-consuming and requires months of measurements for each patient to generate their predictive models. [2], [4], [5]. In [6] individualized patient models based on a Bayesian forecasting method are created that do not utilize genetic markers and do not rely on a large amount of patient data to be effective. However, the Bayesian forecasting method requires large amounts of patient background information, which may not always be available. Warfarin dosage prediction models are also developed using supervised machine learning and recurrent neural network (RNN) in [7], [8]. These methods do not address the inter and intra-variability among patients as it does not develop individualized models and requires a higher number of data points for training and testing the model. The research work in [9] showed that it was possible to create an individualized approach to Warfarin dosing utilizing a reinforcement learning approach that kept a patient's INR in the therapeutic range. In this approach, it is known beforehand how a patient will respond to anticoagulation therapy, which is not realistic in some circumstances and is a common problem to be addressed in other research works [2], [5].

The classical methods such as auto-regressive with exogenous input (ARX) are useful to create the patient dose-response model to warfarin. The disadvantage of classical methods is the requirement of large data set for better model performance and the model structure is required to be defined beforehand. On the other hand, robust identification (RI) methods would be ideal to create an individualized patient model with fewer experimental data points [10].

In robust system identification, the information on the maximum gain of the system K, the stability margin of the system response r, and an upper bound on the noise are required[10], [11]. In the literature, l_1 robust identification and H_∞ robust identification techniques are used for time domain and frequency domain data, respectively. However, in these techniques, zero initial conditions are assumed which is not a valid assumption for the applications of personalized

*Research supported by NSF under grant 1722825 and this material is based upon work supported by the Department of Veterans Affairs, Veterans Health Administration, Office of Research and Development, Clinical Sciences Research.

Affan Affan is a Ph.D. student at Electrical and Computer Engineering department, University of Louisville, KY, USA (e-mail: affan.affan@louisville.edu). Jacek M. Zurada is a Professor of Electrical and Computer Engineering department, University of Louisville, KY, USA (e-mail: jacek.zurada@louisville.edu). Michael E. Brier is a Professor of Medicine, University of Louisville, and at the Robley Rex Veterans Administration Medical Center, Louisville, KY, USA (e-mail: mbrier@louisville.edu). *Tamer Inanc is an Associate professor of Electrical and Computer Engineering department, University of Louisville, KY, USA (correspondence phone: (502)852-7508; fax: (502)852-1577; e-mail: t.inanc@louisville.edu).

drug dosage as the patient may have been using medication prior to the model identification process. To consider the effect of the initial condition, the semi-blind robust identification technique provides appropriate tools for patient modeling with time-domain data [12], [13], [14].

This paper outlines the individualized dose-response models for patients treated with warfarin obtained using a semi-blind robust system identification technique and patient-specific time-domain clinical data [13], [15]. To adapt the models to the change in patient status, the identified models are checked for (In)validation using new patient-specific clinical data [12], [13].

The remainder of this paper is organized as follows. First, we discuss the semi-blind robust system identification in section-II. Section III discusses the model (In)validation and section-IV presents the individualized patient model results obtained by the one-step-ahead prediction along with the error analysis followed by the conclusion.

II. Semi-blind System Identification

The semi-blind robust identification technique incorporates the effect of initial conditions of the system (patient) [12], [13]. By considering, $G = \begin{bmatrix} A_g & B_g \\ C_g & D_g \end{bmatrix}$, as the state space representation of model, G, problem for semi-blind identification can be defined as follows:

Problem 1: *Given input sequence u, output sequence y, noise bound $\in \mathcal{N}$, maximum stability gain and characteristics of past input u^-, determine $G(z) = G_p(z) + G_{np}(z)$, which is compatible with priori and posteriori information, such that τ is a non-empty set, where τ is defined as:*

$$\tau(y) \doteq y_i = \sum_{i=0}^{N} g_i u_{N-i} + C_g A_g^{N-1} \left(\Gamma_g^N u^-\right)_{i=0} \quad (1)$$

where $g_0 = D_G; g_i = C_G (A_G)^{i-1} B_G$. The solution to (1) involves solving a Bi-Affine matrix, which is a non-convex, NP-hard problem. The above problem can be converted to the convex problem as mentioned in [13]. The convex problem can be defined as follows:

Problem 2: *Determine $G(z) = G_p(z) + G_{np}(z)$, which is compatible with a priori and a posteriori information, such that τ is non-empty set:*

$$\tau(y) = \left\{ G(z) \in S : y_i - \left(T_g^N u^+\right)_i + \left(\Gamma_g^N u^-\right)_i \right\} \quad (2)$$

where, $\left|\left(\Gamma_g^N u^-\right)_i\right| \leq \gamma K_u; i = 0,1,\cdots,N-1$ and T_g^N is the Toeplitz matrix and Γ_g^N is the Hankel matrix.

The first part of the τ set corresponds to the patient response for input u and the latter part provides information for system response for past inputs u^-. This problem can be solved by following LMIs as presented in [13]**:**

$$M(g) = \begin{bmatrix} KR^{-2} & \left(T_g^N\right)^T \\ T_g^N & KR^{-2} \end{bmatrix} \geq 0,$$

$$\left| y - (T_u^N p P + T_u^N g) - \Gamma_g^N u^- \right| \in \mathcal{N},$$

$$-\gamma K_u \leq \Gamma_g^N u^- \leq \gamma K_u,$$

where γ, K_u, p, P represents γ-ball in a normed space, bound on the norm of sequence u^-, affine parameters, and the parametric portion of the model, respectively.

As time passes, the patient's characteristics may change due to aging, change in food habits, and new medication for a new disease. The change in patient's status may reflect in the patient's drug-dose response model identified at an early stage, and the model may not match the patient's current dose-response characteristics. To solve this problem, the individualized patient-specific model needs to be adapted to reflect the current status of the patient. This is achieved via model (In)validation methodology discussed in the next section.

III. Model (In)Validation

The theory gives the tools to design the system model which is robust and stable in theory but no information regarding the stability and robustness of the model in practice. This issue rises due to the system uncertainties and unmodeled dynamics in the system. The model (In)validation techniques provide evidence about the usability of the model under these uncertainties for controller design and prediction to some extent by testing the identified model on a new experimental data set. The number of data points required for model (In)validation is equal to the order of the identified model. By assuming multiplicative and additive noise, the problem can be stated as follows:

Problem 3: *Given N_t experimental data points (y_i, u_i), the nominal model $G(z) \in S$, descriptions of admissible noise $\mathcal{N}$, uncertainty Δ and initial conditions x_0, determine if there exists at least one triple (η, Δ, x_0) that can reproduce the available experimental data by the following equation* [12], [13]*:*

$$y = (I + \Delta)\left(T_g u^+ + T_g^{ic} x_0\right) + \eta \quad (3)$$

where u^+ is the input after t=0, T_g maps the input to the output, whereas T_g^{ic} maps the initial conditions to the output.

The above problem has a term $T_g^{ic} x_0$, where x_0 is the initial condition. During the identification, this unknown term can be replaced with some term representing the effect of these initial conditions such as $u^- \in \mathcal{U}_-$.

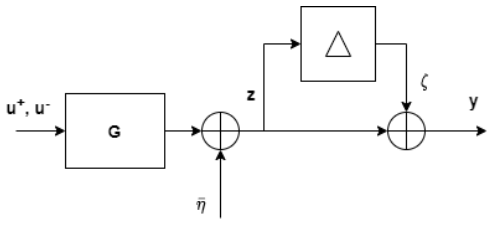

Fig. 1. The convex relaxed model (In)validation for semi-blind robust identification.

Fig. 1 shows the setup for the model (In) validation, where measurement noise is also affected by Δ, $\eta \doteq (1 + \Delta)\bar{\eta}$. Equation (3) can be modified as follows:

$$y = (I + \Delta)\left(T_g u + \Gamma_g^N u^- + \bar{\eta}\right) \quad (4)$$

where $\mathcal{U}_-$ and Γ_g^N denote the past admissible inputs and Hankel matrix. Equation (4) is satisfied if a triple $(u^-, \bar{\eta}, \Delta)$ exists and $\|\Delta\|_\infty < 1$. For further details of convex relaxation

of the model (In)validation process, please see [13].

IV. Simulation Results

To demonstrate the practical use of this approach, the clinical data of forty-six patients treated with warfarin has been collected at The Robley Rex Vetrans Administration Medical Center as part of a separate project by author MEB. Each patient has been dosed with warfarin with INR recorded on the daily basis. The data consists of patient ID numbers, daily warfarin dosage and INR values. Due to the limited space, only the results of the two patients with the most challenging cases from the available database are shown in this paper. In the following figures of semi-blind robust identification simulation results, the red line with square markers shows the actual clinical INR values, the solid blue line shows the model prediction results of the identified full order model, the green line with diamond markers shows the prediction results of the identified reduced-order model, the magenta vertical dashed line shows the number of data points, N_t, used in the identification process for the first and the cyan vertical dashed line represents the points where the model is invalidated and the model is updated. Finally, the vertical blue bars in Figs. 2 and 3 show the daily warfarin dosages.

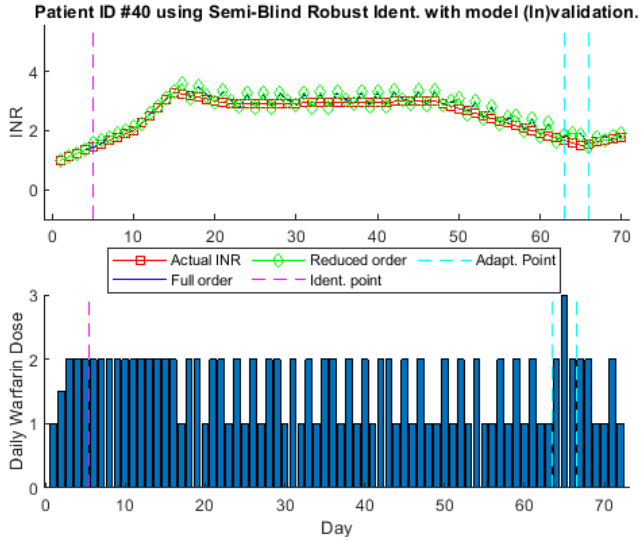

Fig. 2. Prediction results for patient-40 model obtained by semi-blind robust system identification technique.

It is important to mention that whenever the model is (In)validated, all the previously available data points are used to update the model. The full order model is equal to the $N_t + 2$, model consists of parametric and non-parametric portion.

TABLE I Model (In)validation for patient 40.

Time Step (days)	5-62	63-65	66-70
Model (In)validation. $\|\Delta\|_\infty$	0.1	0.88	0.99

The selection of reduced-order model is based on model (In)validation conditions, the model order which satisfies the model (In)validation condition of $\|\Delta\|_\infty < 1$ is selected as the final reduced-order model.

$$G_{40}(z) = \begin{cases} \dfrac{0.2z^3 - 0.2z^2 + 0.1z - 0.01}{z^3 - 1.7z^2 + 0.8z - 0.14} & 5 \le n \le 62 \\ \dfrac{0.1z^4 - 0.1z^3 - 0.02z^2 + 0.1z - 0.02}{z^4 - 2.2z^3 + 1.2z^2 + 0.5z - 0.4} & 63 \le n \le 65 \\ \dfrac{0.05z^4 - 0.06z^3 - 3e^{-3}z^2 + 0.03z - 4e^{-3}}{z^4 - 2.3z^3 + 1.6z^2 + 0.1z - 0.4} & 66 \le n \le 70 \end{cases} \quad (12)$$

The one-step-ahead prediction results for the patient with ID number 40 are shown in Fig. 2. The patient-40 is a challenging patient to find the model representing its characteristics as the clinical data of patient 40 shows a lot of variations, which strongly proves the need for individualized models instead of population-based models for drug dosing. For the model identification, the first five data points are selected. These five-time steps contain most of the patient dose-response information and can be seen in Fig. 2. Therefore, the model identified at the first identification point worked till the 62nd day. At the 63rd-time step, the model faced a change in warfarin dosage, which is not captured by the first model, therefore, the model is (In)validated, and a new model is identified which accommodates these variations. The model (In)validation results for patient-40 are shown in TABLE I. The mathematical expressions for the patient with ID number 147 are shown in (13).

$$G_{147}(z) = \begin{cases} \dfrac{0.1z^3 + 0.2z^2 + 0.1z - 1.5e^{-14}}{z^3 + 0.4z^2 - z - 0.4} & 5 \le n \le 6 \\ \dfrac{0.1z^3 - 0.1z^2 + 0.03z - 9.2e^{-5}}{z^3 - 1.9z^2 + 1.1z - 0.23} & n = 7 \\ \dfrac{0.12z^3 - 0.14z^2 + 0.05z - e^{-3}}{z^3 - 2.1z^2 + 1.4z - 0.3} & 8 \le n \le 72 \\ \dfrac{0.03z^4 + 0.01z^3 + 0.02z^2 + 0.02z - 2e^{-3}}{z^4 - 0.6z^3 + 0.6z^2 - 0.5z - 0.4} & 73 \le n \le 74 \\ \dfrac{0.03z^4 + 0.02z^3 + 0.03z^2 + 0.02z - 2e^{-3}}{z^4 - 0.4z^3 + 0.5z^2 - 0.6z - 0.4} & 75 \le n \le 84 \end{cases} \quad (13)$$

The prediction results for the patient with ID number 147 are shown in Fig. 3. Patient 147 is also challenging because due to a lot of variation even in the early days, the model is updated two times. This shows that the model identification for patient 147 is difficult. However, around the 9th day, the identification process can find the model which worked for a longer time and satisfied the conditions of model (In)validation throughout till the 72nd day. On 73rd and 75th day, the model was required to be updated to accommodate the changes in data for better prediction.

TABLE II Model (In)validation for patient 147.

Time Step (days)	5-6	7	8-72	73-74	75.84
Model (In)validation. $\|\Delta\|_\infty$	0.46	0.5	0.22	0.89	0.80

The model (In)validation results are shown in TABLE II, which shows that model (In)validation conditions are satisfied for each identified model, proving that the models are useful for controller design and prediction purposes. The MMSE values are 0.078 ± 0.044, which shows that the predicted values are close to the actual clinical data, and identified models were able to adapt according to the change in data and performed good predictions.

V. Conclusion

This paper presents the approach based on the individualized patient models that use a semi-blind robust identification

technique for patients treated with warfarin by considering non-zero initial conditions. The measurement data of forty-six patients from the Robley Rex Veterans Administration Medical Center are used to demonstrate the practical use of the algorithm.

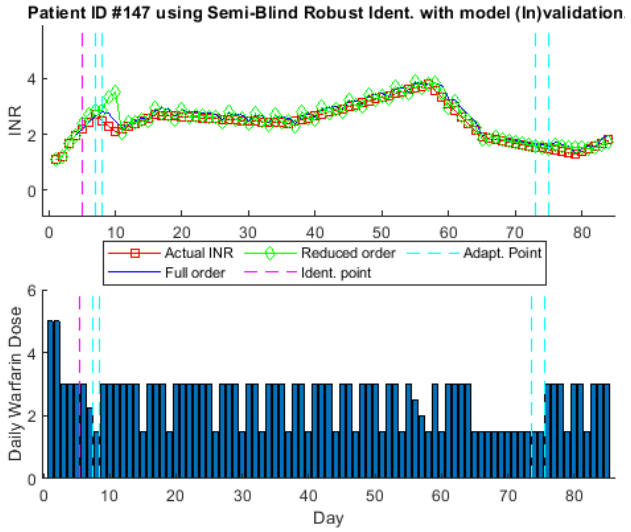

Fig. 3 Prediction results for patient-147 model obtained by semi-blind robust system identification technique.

For prediction and controller design using the identified models, the models are adaptive and change in time due to the progress of the disease, life changes. In this paper, we investigate the model (In)validation technique as a model adaptation strategy for the patient-specific individualized models. It is shown by the simulations using time-domain patient-specific clinical data that the model (In)validation algorithm can adapt the models efficiently to capture the varying model dynamics. The MMSE results show that models identified by semi-blind robust identification performed very well and predicted the INR values close to the actual data in response to warfarin dosage. The future aim is to design the controller based on the individualized models identified by the Semi-blind robust identification technique.

ACKNOWLEDGMENT

Funding from National Science Foundation (NSF) under grant 1722825 is gratefully acknowledged.

REFERENCES

[1] Food and Drug Administration (FDA), "Coumadin (warfarin sodium) tablets Label," 10 2011. [Online]. Available: https://www.accessdata.fda.gov/drugsatfda_docs/label/2011/009218s107lbl.pdf.

[2] Y. b. Zhu, X. H. Hong , M. Wei, J. Hu, X. Chen, S. k. Wang, J. r. Zhu, F. Yu and J. g. Sun, "Development of a novel individualized warfarin dose algorithm based on a population pharmacokinetic model with improved prediction accuracy for Chinese patients after heart valve replacement," *Acta Pharmacologica Sinica volume,* vol. 38, pp. 434-442, 2017.

[3] T. Sasaki, H. Tabuchi, S. Higuchi and I. leiri, "Warfarin-dosing algorithm based on a population pharmacokinetic/pharmacodynamic model combined with Bayesian forecasting," *Pharmacogenomics,* vol. 10, no. 8, 2009.

[4] I. Y. Gong, R. G. Tirona, N. Crown, G. K. Dresser, S. LaRue, N. Langlois, A. l. Langner, G. Zou, D. M. Roden, C. M. Stein, M. Rodger, M. Carrier, M. Forgie, P. S. Wells and R. B. Kim, "Prospective evaluation of a pharmacogenetics-guided warfarin loading and maintenance dose regimen for initiation of therapy," *Blood,* vol. 118, no. 11, pp. 3163-3171, 2011.

[5] A. K. Hamberg, "Pharmacometric Models for Individualisation of Warfarin in Adults and Children," Acta Universitatis Upsaliensis, Uppsala, 2013.

[6] D. F. B. Wright and S. B. Duffull, "A Bayesian dose-individualization method for warfarin," *Clin Pharmacokinet,* vol. 52, no. 1, pp. 59-68, 2013.

[7] Y. H. Hu, F. Wu, C. L. Lo and C. T. Tai, "Predicting warfarin dosage from clinical data: a supervised learning approach," *Artif Intell Med,* vol. 56, no. 1, pp. 27-34, Sep 2012.

[8] O. Tan, J. Z. Wu, K. Yeo, S. Lee, J. Hon, P. Chinag, Y. H. Lau, R. Moh and J. M. Fam, "Building a predictive model for warfarin dosing via machine learning," *European Heart Journal,* vol. 41, no. Supplement_2, 2020.

[9] M. Jacobs, "Personalized anticoagulant management using reinforcement learning.," ThinkIR: The University of Louisville's Institutional Repository , Louisville, 2014.

[10] T. Inanc, M. Sznaier, P. A. Parrilo and R. S. S. Peña, "Robust identification with mixed parametric/nonparametric models and time/frequency-domain experiments: theory and an application," *IEEE Transactions on Control Systems Technology,* vol. 9, no. 4, pp. 608-617, 2001.

[11] M. Sznaier and R. S. S. Peña, Robust System Theory and Applications, Wiley, 1998.

[12] W. Ma, M. Yilmaz, M. Sznaier and C. Lagoa, "SEMI-BLIND ROBUST IDENTIFICATION/MODEL (IN)VALIDATION WITH APPLICATIONS TO MACRO-ECONOMIC MODELLING," in *16th Triennial World Congress*, Prague, 2005.

[13] M. Wenjing, "Semi-blind Robust Identification and Model (In)Validation," The Pennsylvania State University, State College, Pennsylvania, 2006.

[14] E. Akabua, T. Inanc, A. Gaweda, M. E. Brier, S. Kim and J. M. Zurada, "Individualized model discovery: the case of anemia patients," *Computer Methods and Programs in Biomedicine,* vol. 118, no. 1, pp. 23-33, 2015.

[15] S. Toffner-Clausen, System Identification and Robust Control, Springer, 1996.