Title of article:

Long Short-Term Memory Algorithm for Personalized Tacrolimus Dosing: A Simple and Effective Time Series Forecasting Approach Post-Lung Transplantation

Authors:

Haruki Choshi,[1], Kentaroh Miyoshi[1], Maki Tanioka[2], Hayato Arai[3], Shin Tanaka[1], Kazuhiko Shien[1], Ken Suzawa[1], Mikio Okazaki[1], Seiichiro Sugimoto[1], Shinichi Toyooka[1,2]

Affiliations of authors:

[1]Department of General Thoracic Surgery and Breast and Endocrinological Surgery, Okayama University Graduate School of Medicine, Dentistry and Pharmaceutical Sciences, Okayama, Japan

[2]Department of Medical Data Science Innovator Training Program, Okayama University Graduate School of Medicine, Dentistry and Pharmaceutical Sciences, Okayama, Japan

[3]Department of Nephrology, Rheumatology, Endocrinology and Metabolism, Okayama University Graduate School of Medicine, Dentistry and Pharmaceutical Sciences, Okayama, Japan.

Corresponding author email:

Kentaroh Miyoshi MD, PhD,

Department of General Thoracic Surgery and Breast and Endocrinological Surgery, Okayama University Graduate School of Medicine, Dentistry and Pharmaceutical Science, 2-5-1, Shikata-cho, kita-ku, Okayama 700-8558, Japan.

Tel: +81-86-235-7265 Fax: +81-86-235-7269

E-mail address: kmiyoshi@okayama-u.ac.jp

Running title:

LSTM can predict TTL with only previous TRF

Abbreviations

AI, artificial intelligence; Alb, albumin; ALP, alkaline phosphatase; ALT, alanine aminotransferase; AST, aspartate aminotransferase; BUN, blood urea nitrogen; CHDF, continuous hemodiafiltration; CPB, cardiopulmonary bypass; Cre, creatinine; CRP, C-reactive protein; CYP, cytochrome P450; D-BIL, direct bilirubin; ECMO, extracorporeal membrane oxygenation; FFP, fresh frozen plasma; G-GT, γ-glutamyl transpeptidase; Hb, hemoglobin; Hct, hematocrit; LDH, lactate dehydrogenase; LSTM, long short-term memory; LTx, lung transplantation; MSE, mean squared error; NO, nitric oxide; PC, platelet concentrate; POD, postoperative day; PT-INR, prothrombin time-international normalized ratio; RCC, red cell concentrate; SHAP, Shapley Additive exPlanations; T-BIL, total bilirubin; TRF, tacrolimus-related factors; TTL, tacrolimus trough level; WBC, white blood cell

Word count: main body, 3687 words; abstract, 243 words

Abstract

Background: Management of tacrolimus trough levels influences morbidity and mortality after lung transplantation. Several studies have explored pharmacokinetic and artificial intelligence models to monitor tacrolimus levels. However, many models depend on a wide range of variables, some of which, like genetic polymorphisms, are not commonly tested for in regular clinical practice. This study aimed to verify the efficacy of a novel approach simply utilizing time series data of tacrolimus dosing, with the objective of accurately predicting trough levels in the variety of clinical settings.

Methods: Data encompassing 36 clinical variables for each patient were gathered, and a multivariate long short-term memory algorithm was applied to forecast subsequent tacrolimus trough levels based on the selected clinical variables. The tool was developed using a dataset of 87,112 data points from 117 patients and its efficacy was confirmed using six additional cases.

Results: Shapley Additive exPlanations revealed a significant correlation between trough levels and prior dose-concentration data. By using simple trend learning of dose, administration route, and previous trough levels of tacrolimus, we could predict values within 30% of the actual values for 88.5% of time points, which facilitated the creation of a tool for simulating tacrolimus trough levels in response to dosage adjustments. The tool exhibited the potential for rectifying clinical misjudgments in a simulation cohort.

Conclusions: Utilizing our time series forecasting tool, precise prediction of trough levels is attainable independently of other clinical variables, through the analysis of historical tacrolimus dose-

61 concentration trends alone.

62

Introduction

Tacrolimus has been used as a primary agent after lung transplantation (LTx) among several immunosuppressive agents[1-2] and the monitoring of tacrolimus trough level (TTL) correlated with the area under the concentration-time curve has been the standard methods in clinical practice.[3] Tacrolimus dosing is practically customized for each patient by physicians, considering time series data for monitored TTLs and other clinical conditions to be controlled within the target range. However, TTLs often deviate from the target range because of the influence of various misleading clinical factors.[4-7]

In recent years, advancements in machine learning technology have led to the practical application of advanced artificial intelligence (AI) techniques across various fields utilizing diverse algorithms. Several studies have reported on the prediction method of tacrolimus concentrations after solid organ transplantation; however, nearly all extant prediction models have been constructed utilizing a multitude of variables, encompassing data not commonly available in clinical practice, such as patient-specific genetic polymorphisms, including cytochrome P450 (CYP) 3A4/5 genotypes.[8-16] This suggests that the practical use of these models may be hindered by the ease with which data gaps can occur, or by the impact of clinical factors not included in the inputs, leading to inconsistencies in predictive accuracy and limitations in applicability to patients with diverse clinical backgrounds. Conversely, the application of time series forecasting methodologies with machine learning technique remain unreported in the literature. Of these, long short-term memory (LSTM) algorithms have been used in various medical fields, such as drug concentration prediction and arrhythmia detection.[17-18]

Outside of the medical field, LSTM is also used in technical analysis for stock price forecasting, which predicts future stock prices based on trends and patterns identified only from preceding price movements.[19-21] The advantage of the time series analyses lies in its simplicity: it utilizes learning data for the predictive factors, which inherently incorporates the outcomes of all clinical factors. Based on these concepts, we hypothesized that an AI model utilizing LSTM algorithm could predict TTLs effectively from trends in a small number of clinical data obtained during routine medical care. This study aimed to validate the efficacy of a clinically applicable simulation tool for dose-to-TTL prediction using a novel time series forecasting approach, which utilizes minimal clinical factors commonly obtained during routine care.

Materials and methods

Participants

This retrospective observational study of tacrolimus concentration was performed at a single center, Okayama University Hospital. Flow diagram of the subject selection is shown in Figure 1. We assigned 117 cases up to March 2022 as the study cohort to create the AI tool and allocated the subsequent six cases as the simulation cohort for testing the tool. Within the study cohort, using the train_test_split function provided by the Scikit-learn library, 80%, 10%, and 10% of the cases were randomly allocated to the training, validation, and test datasets, respectively.

We extracted data on the 36 clinical factors from the medical records of all eligible patients at multiple

time points (Table 1). For postoperative factors, tacrolimus-related factors, and laboratory data, data were collected at each observation point, including detailed time-series data on the medication interfering tacrolimus metabolism such as azole antifungal drugs. In our protocol, antifungal medication typically starts with intravenous micafungin and is switched to oral itraconazole once the patient can take oral medications. We did not use any other medications that could significantly influence tacrolimus metabolism in a routine manner. This dataset was composed of data based on the specific postoperative protocols. This study was approved by the Institutional Review Board of the Okayama University Hospital (No.2208-059).

Management of tacrolimus at Okayama University Hospital

Maintenance immunosuppression therapy consisted of tacrolimus, mycophenolate mofetil, and prednisolone. Tacrolimus was initiated with bolus infusion via intravenous administration every 12 hours immediately after transplantation.[22] After overall stability of the postoperative systemic condition and normal intestinal peristalsis were confirmed, the administration was switched from the intravenous to oral route. Serum tacrolimus concentration was measured using the antibody-conjugated magnetic immunoassay with Dimension Tacrolimus Flex® Reagent Cartridge (TAC) from Siemens Healthcare Diagnostics. TTL measurements were initiated twice daily at fixed times; once the patient's condition stabilized, the frequency was reduced to once daily. The observation period of the study was defined as within the timeframe of measuring TTLs twice a day, and clinical data were collected over a maximum period of 4 weeks. This timeframe referred to the period of the most

stringent tacrolimus medication management before the stabilization phase, during which TTL measurements were taken twice daily. We targeted tacrolimus concentrations in the first month after LTx at 8–13 ng/mL from April 2011 to March 2021 and at 10–15 ng/mL from April 2022.

Data preparation

There is a difference in dosage between intravenous and oral administration of tacrolimus, and previous studies have reported conversion dose ratios of 2.78–7.4 for switching from intravenous to oral tacrolimus administration.[23-26] In the data from this study, the median dosage was 1.2 and 0.2 mg for oral and intravenous administration, respectively. Therefore, the intravenous dosage was set using data obtained by multiplying the actual prescribed dose by six as the input value to handle data in the model development.

Model development

All model generation processes were performed using Python version 3.10.12. A multivariate LSTM model constructed based on a three-layered structure (see Supplementary Figure S1) was utilized to predict TTLs, as illustrated in Figure 2. Figure 2 showed that TTL predictions were made using data from the previous three time points. In this context, complete information for the prior three time points does not exist before time point 4. Therefore, due to the lack of preceding data for the first three time points (1, 2, and 3), we expediently used the mean value of each factor from the study cohort as imputation data for the non-existent values at time points -2, -1, and 0.

Figure 3 illustrates a flowchart detailing the model creation process. We examined the appropriate

combination of inputs to create a multivariate LSTM model that predicts TTL as a development model for the dose-TTL simulation tool. Shapley Additive exPlanations (SHAP) was implemented to select potential factors that could effectively predict TTL. SHAP assigns each feature an importance value called the SHAP value by calculating the marginal contribution of features to the model output from all features.[27-28] SHAP values represent the relative contribution of each input factor to the predicted value increase or decrease from the baseline value in the AI model.[29] The mean absolute values of the SHAP values indicate the importance of the SHAP feature that affects the model. In this study, SHAP was performed to select the top 15 influential factors with high SHAP feature importance, using 117 cases from the study cohort. By simulating multiple combinations of factors with high SHAP feature importance, we investigated and selected the optimal combination of factors that predicted the TTL with the highest accuracy.

The mean squared error (MSE) was introduced to assess accuracy.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2$$

where $\hat{y}_i$ is the predicted value, y_i is the actual value, and n is the series length. The model with the lowest MSE loss in the test dataset was adopted to create a dose-TTL simulation tool that could predict TTLs corresponding to tacrolimus dosages. To evaluate the overall performance of the ultimate model, we plotted the correlation between the predicted and actual values for the study cohort and calculated the R^2 value on the test dataset to demonstrate how accurately the predictive model explains the

observed data.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$$

where $\hat{y}_i$ is the predicted value, y_i is the actual value, $\bar{y}$ is the mean value of the actual value, and n is the series length. Finally, simulated validation was conducted using the simulation cohort by using our dose-TTL simulation tool.

Results

Patient characteristics

The patient characteristics in the dataset for the simulation tool development are shown in Supplementary Table S1. The observation time points ranged from 10 (5 days) to 56 (28 days), with a median of 26 (13 days). All 87,112 clinical data points were collected by capturing 36 clinical factors at each observation time point. The TTLs were measured at 3,327 time points. The percentage of TTLs within the target concentration range was 52.2% during the first postoperative week and 56.6% during the entire period. Among the study population, the 5-year chronic lung allograft dysfunction-free survival rate was 73.3% (95% confidence interval: 0.633–0.810), and the 5-year overall survival rate was 77.6% (95% confidence interval: 0.684–0.845) (Supplementary Figure S2).

Analyses of the impact of each clinical parameter using SHAP

Figure 4a shows the 15 clinical factors with high SHAP feature importance. The SHAP implied that the preceding tacrolimus dose and TTL had the greatest influence on the prediction of subsequent TTL

176 in this model. The SHAP feature importance of these two factors was more than three times higher
177 than that of other clinical factors. In addition to tacrolimus-related factors (TRF), relatively high SHAP
178 feature importance was observed for prednisolone dose, albumin (Alb), and azole antifungals.
179 SHAP summary plots were drawn, as shown in Figure 4b, which present a clear overview of the
180 variable importance along with their direction of impact on 15 clinical factors with high SHAP feature
181 importance. This figure highlights that the distribution of SHAP values was positively skewed for the
182 preceding tacrolimus dose and TTL, emphasizing that these factors were positively correlated with the
183 next TTL values. The SHAP feature importance and summary plots for all factors are shown in
184 Supplementary Figure S3.
185 **Selection of input combination to optimize dose-TTL simulation algorithm**
186 First, we established a benchmark LSTM model using the tacrolimus dose and TTL as inputs, which
187 showed overwhelmingly high SHAP feature importance. Next, we set up the models by sequentially
188 adding other candidate clinical factors to the input layer, along with dose and TTL, calculating and
189 comparing the MSE loss for each on the test dataset. Imputation was excluded from the candidates
190 because it is an essential flag in implementing the model. Consequently, it was demonstrated that
191 adding tacrolimus administration route information to the benchmark model resulted in the lowest
192 MSE loss value (Supplementary Table S2). As we did not observe a reduction in MSE loss with the
193 addition of other factors, we ultimately opted for a simple LSTM model to predict tacrolimus
194 concentration using only TRF, which included TTL, dose, and route (a total of 10,184 data) as inputs.

Figure 5a shows a positive correlation between the predicted and actual values, resulting in the R^2 value of 0.67 on the test dataset. Regarding the actual approximation, the percentage of predicted TTLs within ± 20% and ± 30% of actual TTLs was calculated to be 74.3% and 88.5%, respectively. Stratifying the results by the time periods after LTx revealed that the predictions were even more accurate one week after LTx (Figure 5b). Additionally, the route of administration on the test dataset did not significantly affect the prediction accuracy (Figure 5c). Figure 6 depicts examples of the temporal evolution of the predicted TTL in the test dataset of the study cohort, along with the actual measured TTL. These patients underwent LTx at various times and had diverse backgrounds. However, the predicted TTL values approximate the actual measured values throughout the observation period. Examination of the intravenous administration period (yellow zone) and oral administration period (purple zone) also revealed that the predicted TTL values closely matched the actual measured values irrespective of the administration route. The LSTM model, which predicts the next TTL using the TRF as the input, was ultimately adopted as the most accurate and straightforward TTL prediction tool.

Validation case analyses

The TTL prediction simulations were conducted using the LSTM model with six new simulation cases (Figure 7). Patient characteristics of the simulation cohort are shown in Supplementary Table S3. The mean absolute error between the actual and predicted value was 4.78, 2.20, and 1.51 ng/mL at time point 1, 2, and 3, respectively, indicating that the predicted values quickly approached the actual values. In cases 1 and 2, the TTL reached and stabilized within the target range after several doses.

The TTL values predicted using the simulation tool closely approximated these values. In contrast, the other four cases deviated from the target TTL early after the initiation of tacrolimus administration. At time point 4 in case 3 and time point 6 in case 4, the actual administered dose resulted in abnormally high TTLs, which can be predicted by the simulation tool. Instead, it suggested that an appropriate TTL can be achieved by reducing the dose. Conversely, at time point 5 in case 5 and time point 7 in case 6, the TTL did not reach the target range. The AI tool accurately predicted that the previous doses of tacrolimus would result in TTLs below the target value. It suggested that an increase in medication dosage was needed to achieve the target TTL. Validation using these simulation cohorts suggested that the LSTM-based simulation tool developed in this study has practical utility in proposing an appropriate tacrolimus dosage to achieve the target TTL, and the contribution of this simulation tool to dose optimization seems promising.

Discussion

The blood concentration of tacrolimus can be influenced by various clinical factors, making it challenging to precisely predict the TTL from the dosage, especially in the early post-transplant period.[4, 6-7] Indeed, 70% of TTLs were reportedly outside the target range in the first week after LTx.[5] In addressing this issue, we have been employing bolus intravenous administration during the early post-transplant acute phase, resulting in a relatively high achievement rate of target trough levels.[22] However, approximately half of the medications fail to reach the target value. This study demonstrates

that accurate predictions of the next tacrolimus concentration can be achieved using only a simple input of previous tacrolimus dosing information using a novel time series approach with an LSTM algorithm. The predicted TTL values calculated using this model approximated the actual TTL values with 88.5% accuracy even with a limited size of training dataset. This result is comparable to the accuracy reported in a kidney transplantation study (64.87% - 91.33%) that utilized 10 commonly employed machine learning algorithms, which required a large number of input variables, specifically 25 clinical factors.[12] Furthermore, in the simulation of the validation cases, this simple LSTM model demonstrated the ability to accurately predict actual TTL values and proposed more appropriate medication dosages at points where the actual values deviated from the target. Overall, the AI methodology proposed in this study has the potential to provide practical tacrolimus recommendations after LTx. In the field of LTx, there have been no reports on tacrolimus management using deep learning, and this study provides new insights.

LSTM is useful in forecasting time series data because it has been improved from conventional recurrent neural networks with regard to long-term memory and the vanishing gradient problem.[30-31] This algorithm has attracted much attention in recent years.[27, 32-33] It has been commonly used in technical analysis for stock forecasting[19-21]; and, we believe that it could be applied to trend analysis of various time series data. In this study, we employed a multivariate LSTM, which can predict the next TTL using a certain time step with multiple input clinical factors, and identified several features of this AI algorithm. Interestingly, the prediction accuracy of the proposed LSTM model was not

252 significantly affected by various background clinical conditions. The cohort used to create the TTL
253 prediction model included LTx patients spanning 11 years, providing a relatively broad historical range
254 of clinical backgrounds. However, the TTL prediction performance of this tool remained stable across
255 different periods. Additionally, the prediction accuracy of TTL did not change even during the switch
256 of drug administration routes in the early post-LTx period. Furthermore, this LSTM model was capable
257 of regaining prediction accuracy between several subsequent drug administrations and measurements,
258 even in the face of rapid TTL fluctuations owing to unpredictable clinical conditions. The LSTM model
259 can flexibly and relatively accurately predict future TTLs using minimal information regarding
260 tacrolimus medication and its trends.

261 We utilized SHAP to extract clinical factors as effective input features for creating a TTL prediction
262 tool. SHAP is a type of explainable AI that can be used to detect factors that contribute to the model.
263 An explanation of AI considerations may provide transparency for clinical applications.[18, 27, 34] When
264 assessing the importance of factors for an AI model, such as LSTM, there is a possibility that factors
265 identified by other multivariate analyses, such as elastic net, may not truly be crucial for the model.
266 Furthermore, in settings such as this study, where there are numerous confounding factors to be
267 evaluated, multivariate analysis methods can encounter limitations in statistical precision owing to
268 issues related to sample size. Therefore, we employed SHAP to examine the contribution of the overall
269 data to the TTL results in our LSTM model. The results showed that the TTL and tacrolimus dose were
270 the most significant predictors, showing a positive correlation with subsequent TTL values, which

markedly eclipsed the other variables in terms of feature importance. The validity of these input factor selections was substantiated through validation analysis.

Several other clinical factors with high feature importance were identified using SHAP. Previous clinical studies have suggested that these factors are associated with tacrolimus metabolism and blood concentration. There are many reports of a relationship between TTL and Alb because tacrolimus distributes into proteins in the blood.[7, 35-39] Additionally, high-dose or continuous prednisolone administration may induce CYP enzymes and leads to high tacrolimus clearance.[5, 40] Similarly, azole antifungal drugs share CYP3A4 as a metabolic enzyme with tacrolimus.[41-43] However, these factors with high feature importance had no additive effect of accuracy improvement in the LSTM model. The lack of improvement in accuracy is related to overfitting and the negligible impact of these factors on the algorithm settings. The study results indicate that the simple relationship between the previous tacrolimus-related information and the subsequent trough concentration is a conclusive factor that incorporates all the various elements influencing tacrolimus levels. This implies that our dose-TTL simulation tool is universal, irrespective of the variety of clinical conditions and the transplant physician's policy on a case-by-case basis. It would be practical to adapt the simplest model to flexibly available and limited clinical factors when applied to patients under various conditions.

This study has some limitations. First, the size of data was limited. Since this study was conducted at a single facility, it can be asserted that the algorithm developed from our dataset by itself does not possess generalizability. However, the aim of this study was to validate the methodology of time series

forecasting, and it can be claimed that the efficacy of the method involving the learning of time series data through LSTM for the prediction of TTL has been demonstrated. Theoretically, training the model with a larger dataset should enable more accurate predictions. Furthermore, by learning from datasets specific to certain clinical contexts (such as medical facilities, eras, or administration methods), this LSTM model would likely be able to make predictions that are both flexible and precise in similar conditions. Although there are some differences between facilities, such as dosing methods and TTL measurement intervals, if tacrolimus dosing management follows certain rules at a facility, it is possible to create a simulation tool. Trend analysis with LSTM algorithm is a suitable analytical approach, and it is worth validating in multiple facilities. The second limitation pertains to the tacrolimus medication protocol. In this study, there was no significant difference in prediction accuracy between intravenous and oral administration. This is likely because the twice-daily bolus intravenous administration that we adopted did not cause instability in the absorption of tacrolimus immediately after surgery, allowing us to obtain regular data with peak and trough values similar to those seen with oral administration. Although the twice-daily bolus intravenous administration is not common, additional data from facilities that start with oral administration from the first dose may benefit those centers. However, there remains the possibility that prediction results may vary due to the variability in absorption during initial oral administration. Also, because the medication management in this study was based on fixed time schedules, the accuracy may decrease when used in outpatient settings where the timing of medication intake and concentration measurement is likely to shift. In such situations, it would be

309 necessary to add the data on the interval between tacrolimus administration and blood concentration
310 measurement. Third, there is insufficient information about all factors affecting tacrolimus metabolism
311 and absorption. We did not consider the diversity of genetic polymorphisms present in the study cohort.
312 This information could potentially contribute to the identification of individual TTL patterns. In
313 particular, the mean absolute error at time point 1 was large in the simulation cohort, which might be
314 improved with information on genetic polymorphisms. Additionally, all subjects in this cohort were
315 ethnically Japanese, and the number of cystic fibrosis patients was limited due to the rarity of the
316 disease in Japan. Regarding concomitant drugs, there is a lack of cases with azole antifungal drugs
317 other than itraconazole and medications affecting TTL, such as rifabutin. It may not be realistic to
318 directly apply our model to cohorts with diversity in ethnicity and primary disease. However, since
319 there are reports that LSTM-based models are capable of making accurate predictions even in cohorts
320 with diversity,[45-47] it may be possible to create a widely usable model by training on a large, diverse
321 dataset.

322 In conclusion, the LSTM model allows for a dose-TTL simulation with high accuracy solely by
323 inputting previous dose-concentration data. This deep learning technique for time series analysis is
324 capable of returning accurate predictions of TTL based on simple inputs, tailored to the learning
325 environment. The effectiveness of tacrolimus concentration prediction and dosage determination
326 assistance tools using time series approach warrants further prospective validation in a variety of
327 clinical transplant cohorts.

328

329 **Author contributions**

330 HC, KM and MT designed and performed the experiments and interpreted the data; HC and KM wrote

331 the manuscript; ST and SS collected the clinical data; HA assisted in the design of the programming;

332 All co-authors approved the paper.

333

334 **Acknowledgments**

335 We would like to thank Megumi Ishihara and We would like to thank Megumi Ishihara and the other

336 recipient coordinators at the Organ Transplant Center, Okayama University Hospital.

337

References

1. Grimm M, Rinaldi M, Yonan NA, et al. Superior prevention of acute rejection by tacrolimus vs. cyclosporine in heart transplant recipients--a large European trial. *Am J Transplant.* 2006;6(6):1387–1397.
2. Bedair B, Hachem RR. Management of chronic rejection after lung transplantation. *J Thorac Dis.* 2021;13(11):6645–6653.
3. Schiff J, Cole E, Cantarovich M. Therapeutic monitoring of calcineurin inhibitors for the nephrologist. *Clin J Am Soc Nephrol.* 2007;2(2):374–384.
4. Sikma MA, van Maarseveen EM, van de Graaf EA, et al. Pharmacokinetics and toxicity of tacrolimus early after heart and lung transplantation. *Am J Transplant.* 2015;15(9):2301–2313.
5. Sikma MA, Hunault CC, van Maarseveen EM, et al. High variability of whole-blood tacrolimus pharmacokinetics early after thoracic organ transplantation. *Eur J Drug Metab Pharmacokinet.* 2020;45(1):123–134.
6. Kotecha S, Ivulich S, Snell G. Review: immunosuppression for the lung transplant patient. *J Thorac Dis.* 2021;13(11):6628–6644.
7. Sikma MA, Hunault CC, Huitema ADR, De Lange DW, van Maarseveen EM. Clinical pharmacokinetics and impact of hematocrit on monitoring and dosing of tacrolimus early after heart and lung transplantation. *Clin Pharmacokinet.* 2020;59(4):403–408.
8. Cheng F, Li Q, Cui Z, Wang Z, Zeng F, Zhang Y. Tacrolimus concentration is effectively predicted using combined clinical and genetic factors in the perioperative period of kidney transplantation and associated with acute rejection. *J Immunol Res.* 2022;2022:3129389. doi: 10.1155/2022/3129389
9. Min S, Papaz T, Lambert N, et al. An integrated clinical and genetic prediction model for tacrolimus levels in pediatric solid organ transplant recipients. *Transplantation.* 2022;106(3):597–606.
10. Luo X, Zhu L, Cai N, Zheng L, Cheng Z. Prediction of tacrolimus metabolism and dosage requirements based on CYP3A4 phenotype and CYP3A5(*)3 genotype in Chinese renal transplant recipients. *Acta Pharmacol Sin.* 2016;37(4):555–560.
11. Storset E, Holford N, Hennig S, et al. Improved prediction of tacrolimus concentrations early after kidney transplantation using theory-based pharmacokinetic modelling. *Br J Clin Pharmacol.* 2014;78(3):509–523.
12. Zhang Q, Tian X, Chen G, et al. A prediction model for tacrolimus daily dose in kidney transplant recipients with machine learning and deep learning techniques. *Front Med (Lausanne).* 2022;9:813117. doi: 10.3389/fmed.2022.813117
13. Srinivas L, Gracious N, Nair RR. Pharmacogenetics based dose prediction model for initial tacrolimus dosing in renal transplant recipients. *Front Pharmacol.* 2021;12:726784. doi:

10.3389/fphar.2021.726784
14. Tang J, Liu R, Zhang Y, et al. Application of machine-learning models to predict tacrolimus stable dose in renal transplant recipients. *Sci Rep.* 2017;7:42192. doi: 10.1038/srep42192
15. Miano TA, Feng R, Griffiths S, et al. Development and validation of a population pharmacokinetic model to guide perioperative tacrolimus dosing after lung transplantation. *medRxiv.* 2023. doi: 10.1101/2023.06.26.23291248
16. Hong E, Carmanov E, Shi A, et al. Application of physiologically based pharmacokinetic modeling to predict drug-drug interactions between elexacaftor/tezacaftor/ivacaftor and tacrolimus in lung transplant recipients. *Pharmaceutics.* 2023;15(5):1438. doi: 10.3390/pharmaceutics15051438
17. Khusial R, Bies RR, Akil A. Deep learning methods applied to drug concentration prediction of olanzapine. *Pharmaceutics.* 2023;15(4):1139. doi: 10.3390/pharmaceutics15041139
18. Petmezas G, Stefanopoulos L, Kilintzis V, et al. State-of-the-art deep learning methods on electrocardiogram data: systematic review. *JMIR Med Inform.* 2022;10(8):e38454. doi: 10.2196/38454
19. Roondiwala M, Patel H, Varma S. Predicting stock prices using LSTM. *Int J Sci Res (Raipur).* 2017;6(4):1754–1756.
20. Nelson DMQ, Pereira ACM, de Oliveira RA. Stock market's price movement prediction with LSTM neural networks. *2017 International Joint Conference on Neural Networks (IJCNN)*. 2017.1419–26.
21. Moghar A, Hamiche M. Stock market prediction using LSTM recurrent neural network. *Procedia Comput Sci.* 2020;170:1168–1173.
22. Hirano Y, Sugimoto S, Mano T, et al. Prolonged administration of twice-daily bolus intravenous tacrolimus in the early phase after lung transplantation. *Ann Transplant.* 2017; 2:484–492.
23. Yang C, Xi Y, Chen W, et al. Conversion ratio of tacrolimus switching from intravenous infusion to oral administration after lung transplantation. *J Thorac Dis.* 2020;12(8):4292–4298.
24. Suetsugu K, Ikesue H, Miyamoto T, et al. Analysis of the variable factors influencing tacrolimus blood concentration during the switch from continuous intravenous infusion to oral administration after allogeneic hematopoietic stem cell transplantation. *Int J Hematol.* 2017;105(3):361–368.
25. Pasternak AL, Marcath LA, Li Y, et al. Impact of pharmacogenetics on intravenous tacrolimus exposure and conversions to oral therapy. *Transplant Cell Ther.* 2022;28(1):19 e1–19 e7.
26. Kanamitsu K, Yorifuji T, Ishida H, et al. Clinical factors affecting the dose conversion ratio from intravenous to oral tacrolimus formulation among pediatric hematopoietic stem cell transplantation recipients. *Ther Drug Monit.* 2020;42(6):803–810.

27. Butnariu D. STABILITY AND SHAPLEY VALUE FOR AN n-PERSONS FUZZY GAME. *Fuzzy Sets Syst.* 1980; 4: 63–72
28. Lundberg SM, Lee S. A Unified Approach to Interpreting Model Predictions. *31st Conference on Neural Information Processing Systems (NIPS 2017)*. 2017. Long Beach, CA, USA. doi: 10.48550/arXiv.1705.07874
29. Yang E, Zhang H, Guo X, Zang Z, Liu Z, Liu Y. A multivariate multi-step LSTM forecasting model for tuberculosis incidence with model explanation in Liaoning Province, China. *BMC Infect Dis.* 2022;22(1):490. doi: 10.1186/s12879-022-07462-8
30. Hochreiter S, Schmidhuber J. Long short-term memory. *Neural Comput.* 1997; 9(8): 1735–80
31. van Houdt G, Mosquera C, Nápoles G. A review on the long short-term memory model. *Artif Intell Rev.* 2020;53(8):5929–5955.
32. Gul S, Khan GM, Yousaf S. Multi-step short-term $PM_{2.5}$ forecasting for enactment of proactive environmental regulation strategies. *Environ Monit Assess.* 2022;194(5):386. doi: 10.1007/s10661-022-10029-4
33. Ji R, Shi S, Liu Z, Wu Z. Decomposition-based multi-step forecasting model for the environmental variables of rabbit houses. *Animals (Basel).* 2023;13(3):546. doi: 10.3390/ani13030546
34. Greff K, Srivastava RK, Koutnik J, Steunebrink BR, Schmidhuber J. LSTM: A Search Space Odyssey. *IEEE Trans Neural Netw Learn Syst.* 2017;28(10):2222–2232.
35. Cheng F, Li Q, Wang J, et al. Genetic Polymorphisms Affecting Tacrolimus Metabolism and the Relationship to Post-Transplant Outcomes in Kidney Transplant Recipients. *Pharmacogenomics Pers Med.* 2021;14:1463-1474.
36. Zhu L, Zhang J, Song H, et al. Relationships of related genetic polymorphisms and individualized medication of tacrolimus in patients with renal transplantation. *Int J Clin Exp Med.* 2015;8(10):19006–19013.
37. Andrews LM, Hesselink DA, van Schaik RHN, et al. A population pharmacokinetic model to predict the individual starting dose of tacrolimus in adult renal transplant recipients. *Br J Clin Pharmacol.* 2019;85(3):601–615.
38. Francke MI, Visser WJ, Severs D, van Egmond AME, Hesselink DA, de Winter BCM. Body composition is associated with tacrolimus pharmacokinetics in kidney transplant recipients. *Eur J Clin Pharmacol.* 2022;78(8):1273–1287.
39. Van Looy S, Verplancke T, Benoit D, et al. A novel approach for prediction of tacrolimus blood concentration in liver transplantation patients in the intensive care unit through support vector regression. *Crit Care.* 2007;11(4):R83. doi: 10.1186/cc6081
40. Passey C, Birnbaum AK, Brundage RC, Oetting WS, Israni AK, Jacobson PA. Dosing equation for tacrolimus using genetic variants and clinical factors. *Br J Clin Pharmacol.* 2011;72(6):948–957.
41. Groll AH, Townsend R, Desai A, et al. Drug-drug interactions between triazole antifungal

agents used to treat invasive aspergillosis and immunosuppressants metabolized by cytochrome P450 3A4. *Transpl Infect Dis.* 2017;19(5):e12751. doi: 10.1111/tid.12751

42. Klatt ME, Eschenauer GA. Review of pharmacologic considerations in the use of azole antifungals in lung transplant recipients. *J Fungi (Basel).* 2021;7(2):76. doi: 10.3390/jof7020076

43. Zhao Y, Xiao C, Hou J, et al. The effect of voriconazole on tacrolimus in kidney transplantation recipients: a real-world study. *Pharmaceutics.* 2022;14(12):2739. doi :10.3390/pharmaceutics14122739

44. Abu-Elmagd KM, Fung J, Draviam R, et al. Fourhour versus 24-hour intravenous infusion of FK 506 in liver transplantation. *Transplant Proc*. 1991; 23(6): 2767–2770.

45. Carbune V, Gonnet P, Deselaers T, et al. Fast multi-language LSTM-based online handwriting recognition. *Int J Doc Anal Recognit.* 2020; 23 (2): 89-102.

46. Sheikhi F, Kowsari Z. Time series forecasting of COVID-19 infections and deaths in Alpha and Delta variants using LSTM networks. *PLoS ONE*. 2023; 18(10): e0282624. https://doi.org/10.1371/journal.pone.0282624

47 Gautam Y. Transfer Learning for COVID-19 cases and deaths forecast using LSTM network. *ISA Trans*. 2022; 124: 41-56.

Figure legends

Figure 1. Flow diagram of the subject selection

A total of 139 patients underwent lung transplantation at the Okayama University Hospital between April 2011 and May 2023. Tacrolimus was administered intravenously via a bolus or orally to 129 patients. Three cases were excluded because of short administration periods (< 5 days) and two infants with body weights of less than 10 kg were excluded. One patient was excluded because of tacrolimus trough level (TTL) with significant outliers (probably owing to sampling error).

Figure 2. Multivariate three-step long short-term memory model used in this study

The input layer consisted of multivariate factors, and the output layer was set to the tacrolimus trough level (TTL). The three-step means that the output layer is predicted using the input factors at the prior three time points. The preceding three-step (t-3, t-2, t-1) time series data were used to predict the TTL at a specific time point (t). All clinical data were standardized using Z-score normalization, and imputation data were defined as the mean values of the respective factors. To distinguish the imputation data from other data, flags were set on these data and used in the analysis. The imputation flag reveals the point at which the TTL was unmeasured (red parts in the figure).

Figure 3. Flowchart of model development in the study cohort

First, 15 candidate clinical factors associated with tacrolimus trough level (TTL) were detected using Shapley Additive exPlanations (SHAP) for long short-term memory (LSTM) with all clinical factors. Subsequently, the mean squared error (MSE) loss was compared between LSTM models with various

combinations of candidate factors. Finally, a dose-TTL simulation tool was developed using the LSTM model, leading to the lowest MSE loss.

Figure 4. The results of Shapley Additive exPlanations

Alb, albumin; AST, aspartate aminotransferase; Cre, creatinine; D-BIL, direct bilirubin; Hct, hematocrit; PC, platelet concentrate; PT-INR, prothrombin time-international normalized ratio; TTL, tacrolimus trough level

The graph shows the top 15 absolute value of mean Shapley Additive exPlanations (SHAP) value (a) and SHAP summary plots (b). The mean (|SHAP value|) indicates the average impact of the model output magnitude. The higher the importance of the SHAP feature, the more likely that a factor contributes to the model. The SHAP summary plots contributed to comprehending the variable importance along with their direction of impact on the 15 clinical factors with high SHAP feature importance. The horizontal axis represents the SHAP value; positive SHAP values affect the predicted TTL in an upward direction, whereas negative SHAP values affect it in a downward direction. In each dot, red and blue indicate high and low feature values, respectively. In other words, the red dots for the positive SHAP values and blue dots for the negative SHAP values indicate that the factor was positively correlated with the predicted values.

Figure 5. Evaluation of the overall performance of the model and examination of accuracy based on the stratification of the test dataset

Plotting the relationship between the predicted and actual values for the study cohort showed a positive

509 correlation (a). The test dataset was divided into three periods: up to day 3, from day 4 to day 7, and
510 beyond day 8. The accuracy for each period was calculated, and the results showed that more accurate
511 prediction was observed with longer post-transplant periods (b). The results, stratified by the route of
512 administration, showed that the predictions for oral and intravenous administration had equivalent
513 accuracy (c).

514 **Figure 6. The predicted result examples of tacrolimus trough level using a three-step long short-**
515 **term memory model utilizing only tacrolimus-related factors**

516 In the upper left of the graph, the period of lung transplantation is indicated. Regardless of patient
517 background, the predicted tacrolimus trough levels (TTLs) were similar to the actual TTLs. Distinction
518 by the route of administration in the prediction curve for TTLs. The predicted TTLs are similar to the
519 actual TTLs after intravenous (yellow) and oral (purple) administration

520 **Figure 7. Results of a dose-TTL simulation tool for the latest six lung transplantation cases**

521 Cases 1 and 2 were successfully managed within the target range of the actual tacrolimus trough level
522 (TTL). The actual and predicted TTLs at the final time point in cases 3 and 4 were outside the target
523 range, with an intravenous dose of 0.30 mg. Our tool indicated that the dose to bring the predicted TTL
524 into the target range was 0.18 mg in case 3 and 0.15 mg in case 4. In case 5, 0.27 mg was administered
525 intravenously at time point 4 and the actual TTL was less than 10.0 ng/mL. The predicted TTL was
526 also 9.0 ng/mL at 0.27 mg, and 0.37 mg was needed to derive 10.0 ng/mL. In case 6, by intravenous
527 administration of 0.30 mg, the actual and predicted TTL at time point 7 were 8.4 and 8.7 ng/mL,

528 respectively. The dose required to bring the predicted TTL to the target range was 0.45 mg.

529

Supplementary information

Supplementary Figure S1. The three-layer long short-term memory structure

c, long-term memory; $\tilde{c}_t$, new information from x_t and h_{t-1}; f_t, forget gate vector; h, hidden state (short-term memory); i_t, input gate vector; o_t, output gate vector; σ, sigmoid function; tanh, hyperbolic tangent function; x_t, input; y_t, output

Unique gates, called "forget gate," "input gate," and "output gate" allow to hold on the long-term memory.

Supplementary Figure S2. Chronic lung allograft dysfunction-free survival curve (a) and overall survival curve (b) in the study population

The median observation period is 5 years and 0 months. The 5-year chronic lung allograft dysfunction-free survival rate was 73.3% (95% confidence interval: 0.633–0.810), and the 5-year overall survival rate was 77.6% (95% confidence interval: 0.684–0.845).

Supplementary Figure S3. Shapley Additive exPlanations for all factors

Alb, albumin; ALP, alkaline phosphatase; ALT, alanine aminotransferase; AST, aspartate aminotransferase; BUN, blood urea nitrogen; CHDF, continuous hemodiafiltration; CPB, cardiopulmonary bypass; Cre, creatinine; CRP, C-reactive protein; D-BIL, direct bilirubin; ECMO, extracorporeal membrane oxygenation; FFP, fresh frozen plasma; G-GT, γ-glutamyl transpeptidase; Hb, hemoglobin; Hct, hematocrit; LDH, lactate dehydrogenase; NO, nitric oxide; PC, platelet concentrate; POD, postoperative day; PT-INR, prothrombin time-international normalized ratio; RCC,

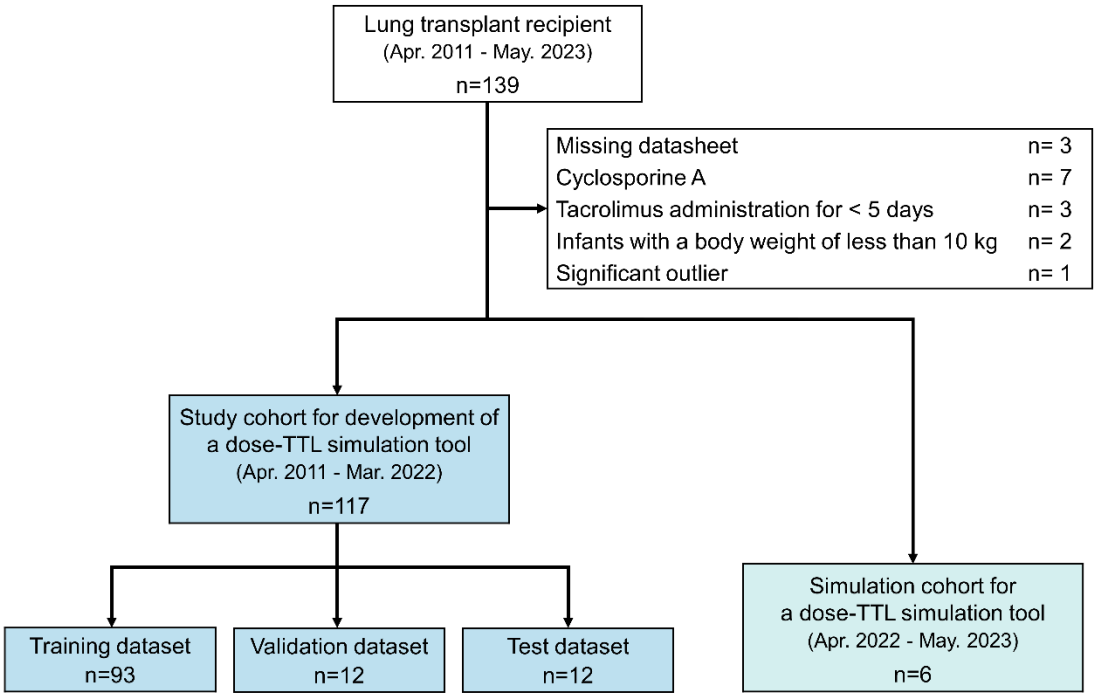
549 red cell concentrate; T-BIL, total bilirubin; TTL, tacrolimus trough level; WBC, white blood cell

550 The graph shows the Shapley Additive exPlanations (SHAP) feature importance (a) and SHAP

551 summary plots (b) for all factors.

552

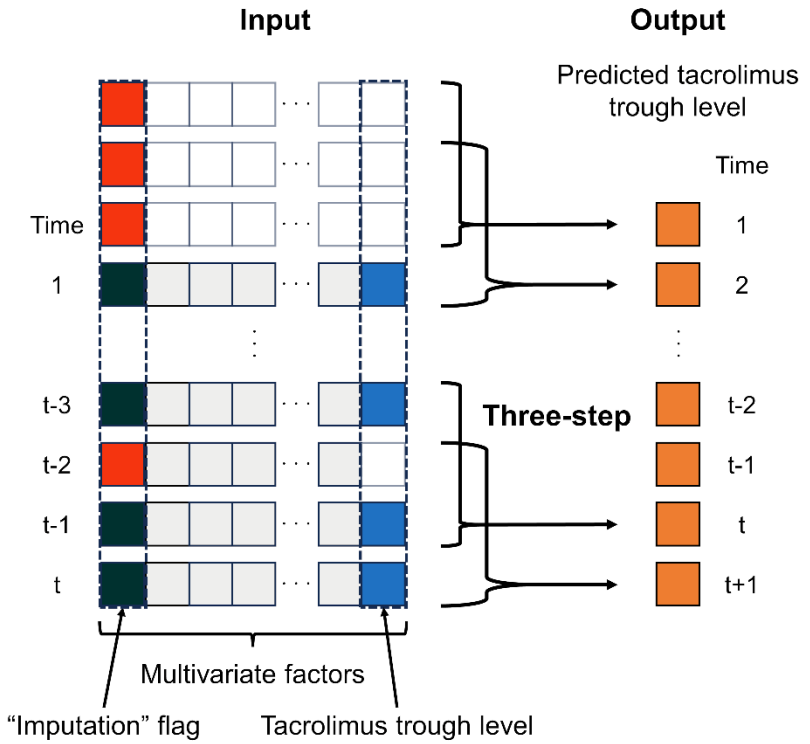
553 Figure 1



554

555

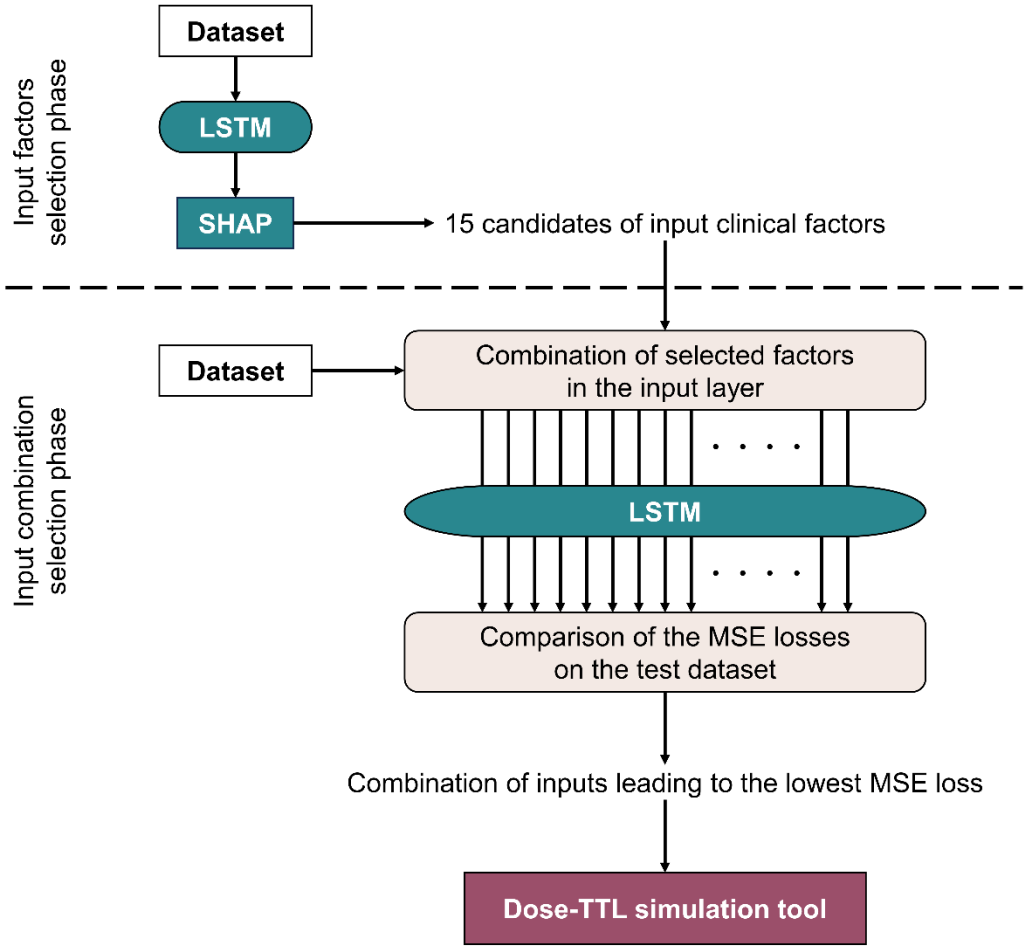
556 Figure 2



557

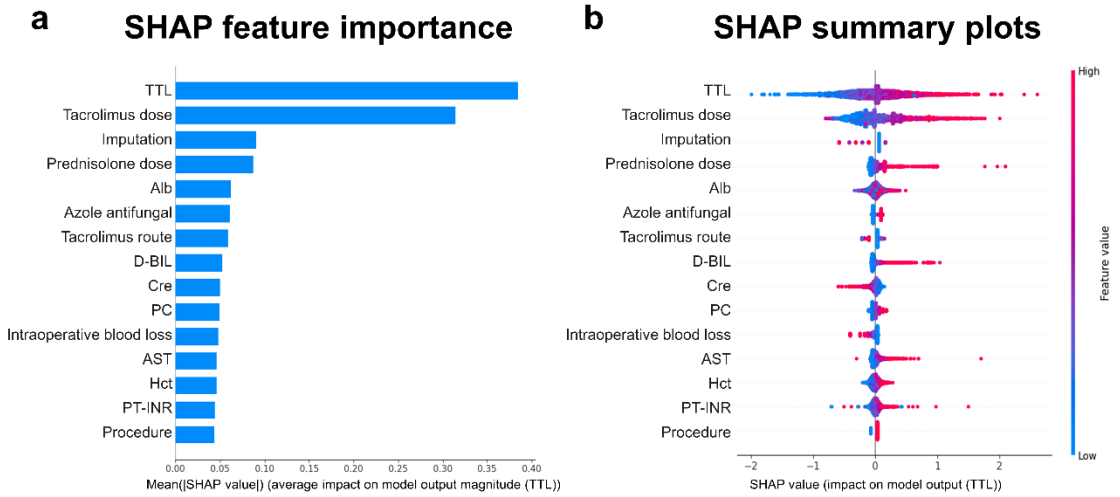
558

559 Figure 3



560
561

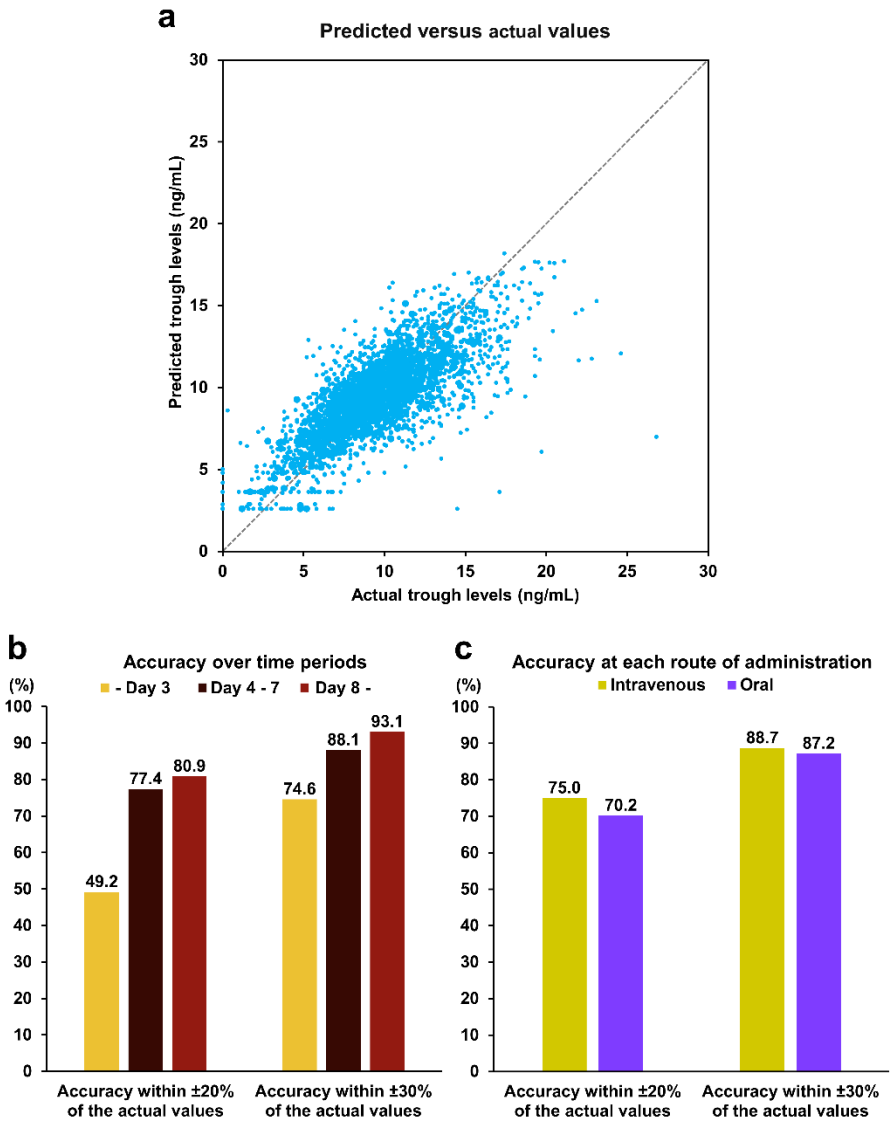
562 Figure 4



563

564

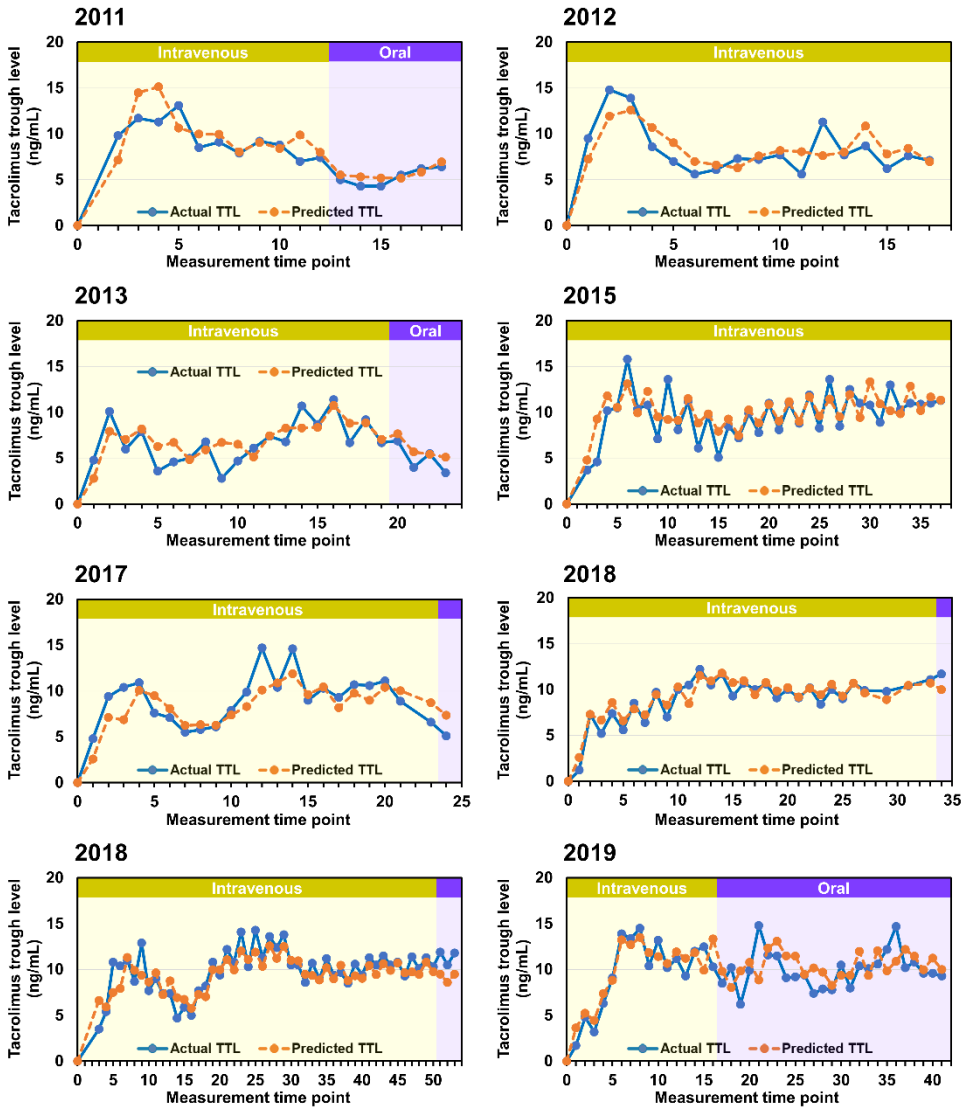
565 Figure 5



566

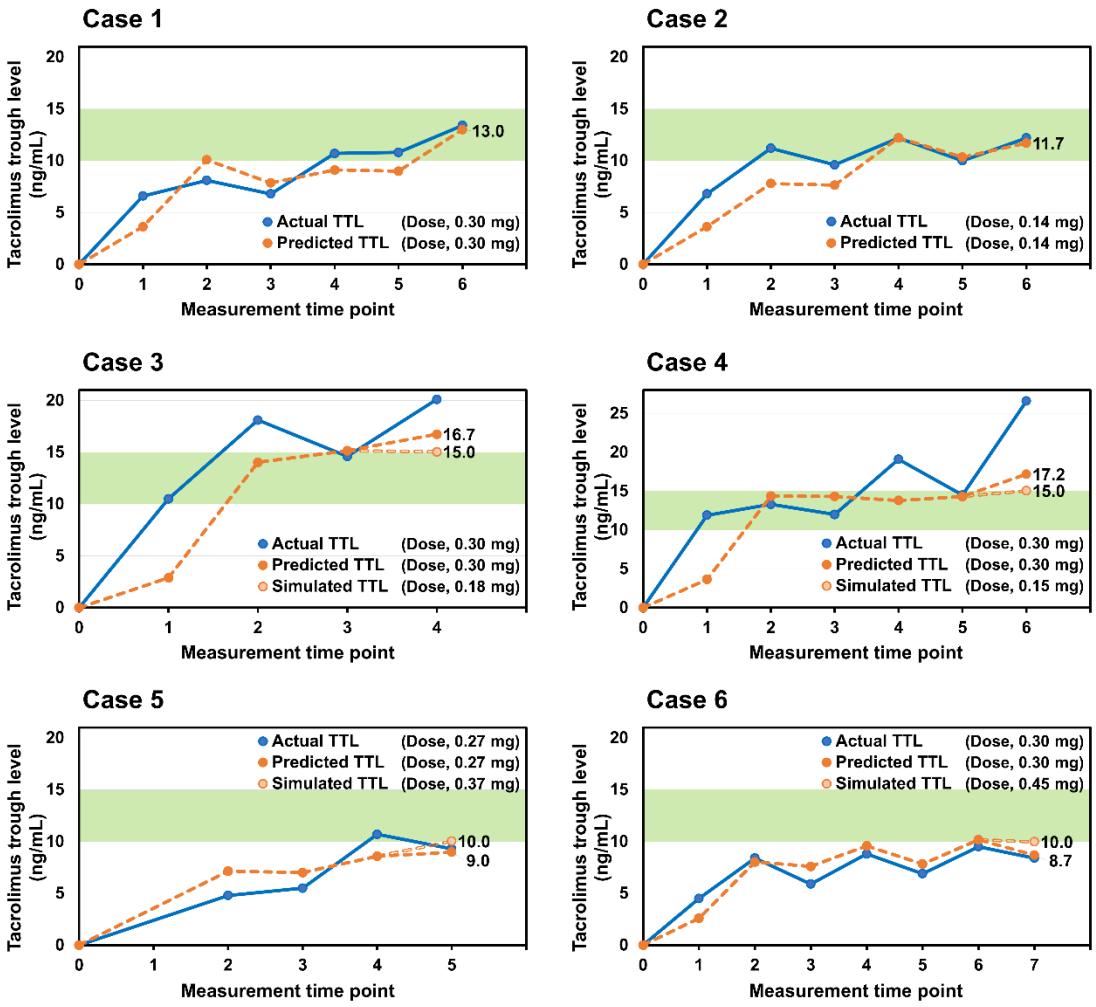
567

568 Figure 6



569
570

571 Figure 7



572

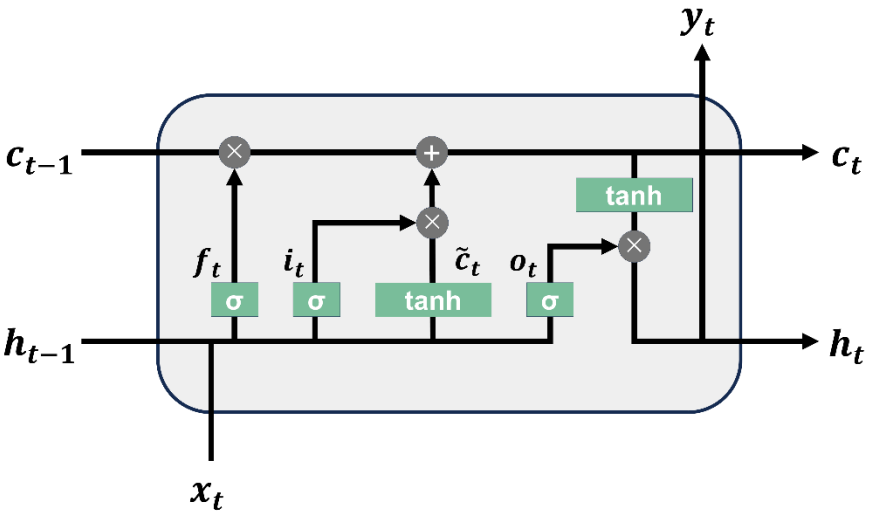
573

574 **Table 1. Collected 36 clinical factors for the input factors**

Patient characteristics	**Tacrolimus-related factors**
Age	Tacrolimus dose
Sex	Tacrolimus route (route of tacrolimus administration)
Height	TTL
Weight	
Intraoperative factors	**Laboratory data**
Procedure (single/bilateral LTx)	BUN
Operation time	Cre
CPB	AST
Intraoperative blood loss	ALT
Intraoperative blood transfusion volume	LDH
- RCC	ALP
- FFP	T-BIL
- PC	D-BIL
Postoperative factors	G-GT
Intubation period	Alb
ECMO/CHDF period	PT-INR
Azole antifungal	Hb
Inhalation of NO period	Hct
POD	WBC
Prednisolone dose	CRP
Mycophenolate mofetil dose	

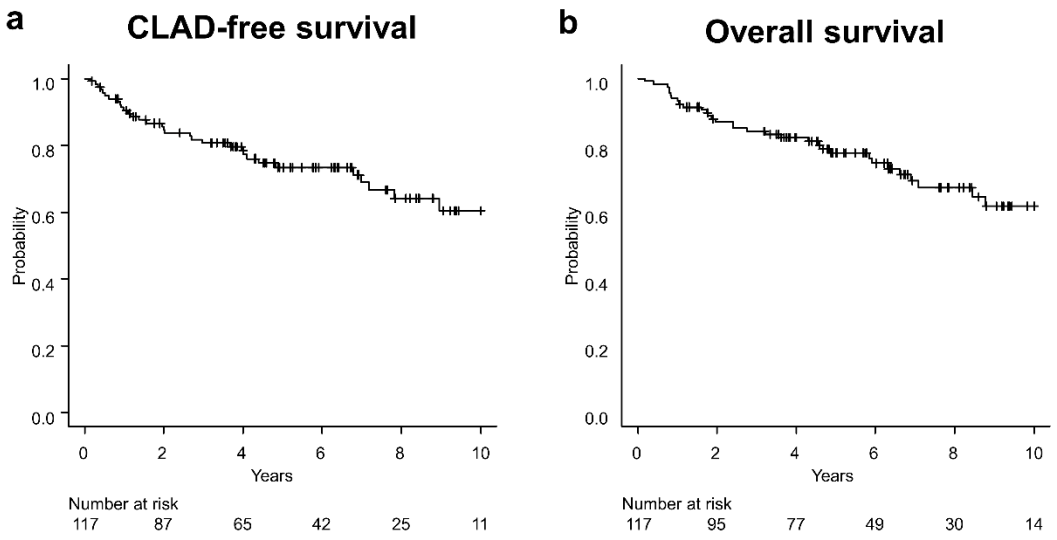
575 Alb, albumin; ALP, alkaline phosphatase; ALT, alanine aminotransferase; AST, aspartate
576 aminotransferase; BUN, blood urea nitrogen; CHDF, continuous hemodiafiltration; CPB,
577 cardiopulmonary bypass; Cre, creatinine; CRP, C-reactive protein; D-BIL, direct bilirubin; ECMO,
578 extracorporeal membrane oxygenation; FFP, fresh frozen plasma; G-GT, γ-glutamyl transpeptidase;
579 Hb, hemoglobin; Hct hematocrit; LDH, lactate dehydrogenase; LTx, lung transplantation; NO, nitric
580 oxide; PC, platelet concentrate; POD, postoperative day; PT-INR, prothrombin time-international
581 normalized ratio; RCC, red cell concentrate; T-BIL, total bilirubin; TTL, tacrolimus trough-level;
582 WBC, white blood cell.
583

584 Supplementary Figure S1



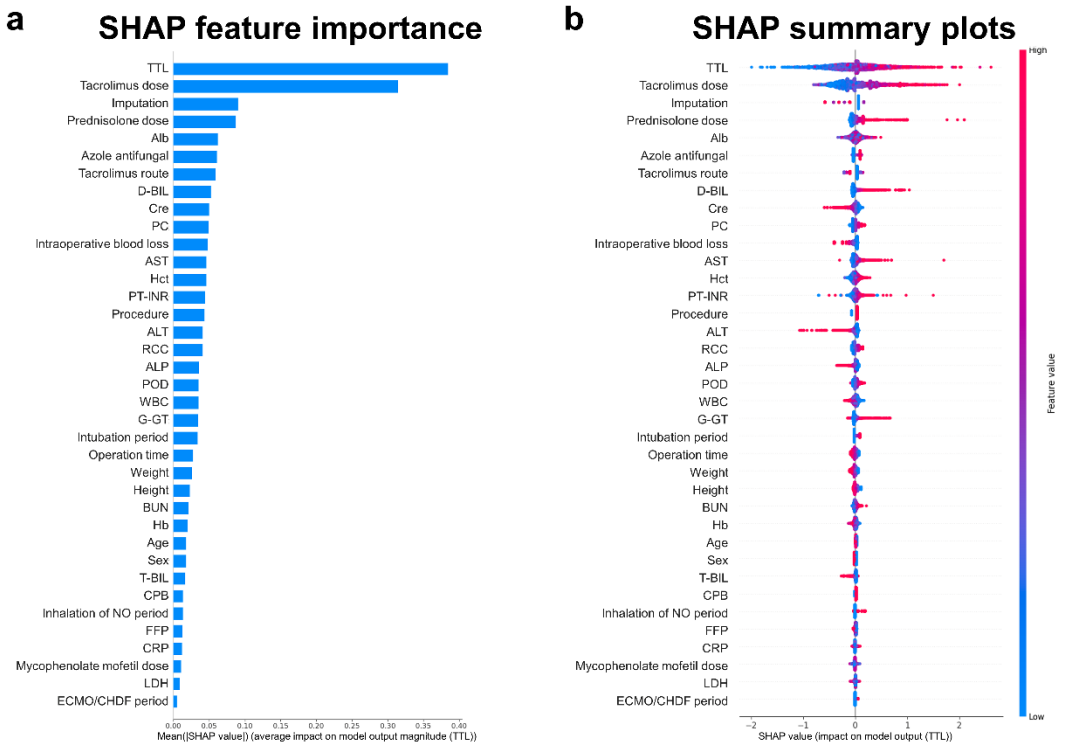
585
586

587 Supplementary Figure S2



588
589

590 Supplementary Figure S3



591

592

Supplementary Table S1. Patient characteristics

Characteristics	Overall n=117
Age (median)	2–68 (42)
Male (%)	59 (50.4)
Preoperative information	
Height (median) (cm)	90–181 (161.3)
Weight (median) (kg)	11.2–95.5 (45.2)
Intraoperative information	
Bilateral lung transplantation (%)	82 (70.1)
Lobar transplantation (%)	46 (39.3)
CPB use (%)	80 (68.4)
Operation time (median) (min)	219–845 (466)
Intraoperative blood loss (median) (mL)	10–15,860 (830)
Observation points (median)	10–56 (26)

CPB, cardiopulmonary bypass

The observation period is defined as the timeframe for measuring tacrolimus trough levels twice daily, with a maximum period of 4 weeks.

Supplementary Table S2. Mean squared error loss of the model for each additional factor to tacrolimus dose and trough level in the input layer

Input factors	Additional factor	Loss
Tacrolimus dose + TTL		1.69
	+ Intraoperative blood loss	2.09
	+ Prednisolone dose	1.75
	+ PC	1.71
	+ Azole antifunga	1.70
	+ Hct	1.70
	+ Procedure	1.69
	+ AST	1.68
	+ Cre	1.68
	+ PT-INR	1.66
	+ D-Bil	1.66
	+ Alb	1.65
	+ **Route of tacrolimus administration**	**1.58**

Alb, albumin; AST, aspartate aminotransferase; Cre, creatinine; D-BIL, direct bilirubin; Hct, hematocrit; PC, platelet concentrate; PT-INR, prothrombin time-international normalized ratio; TTL, tacrolimus trough level

Supplementary Table S3. Patient characteristics of the simulation cohort

Case	Age	Sex	Disease	Procedure	CPB	Operation time	Intraoperative blood loss
1	50	M	ILD	Single	-	375 min	310 mL
2	57	M	BOS	Single	+	514 min	3530 mL
3	17	M	PAH	Bilateral	+	685 min	555 mL
4	48	M	ILD	Single	+	456 min	830 mL
5	55	F	DPB	Bilateral	+	544 min	1460 mL
6	16	M	BOS	Bilateral	+	561 min	2120 mL

BOS, bronchiolitis obliterans syndrome; CPB, cardiopulmonary bypass; DPB, diffuse panbronchiolitis; ILD, interstitial lung disease; PAH, pulmonary arterial hypertension