

# Predicting Total Drug Clearance and Volumes of Distribution Using the Machine Learning-Mediated Multimodal Method through the Imputation of Various Nonclinical Data

Hiroaki Iwata,\* Tatsuru Matsuo, Hideaki Mamada, Takahisa Motomura, Mayumi Matsushita, Takeshi Fujiwara, Kazuya Maeda, and Koichi Handa\*



Cite This: *J. Chem. Inf. Model.* 2022, 62, 4057–4065



Read Online

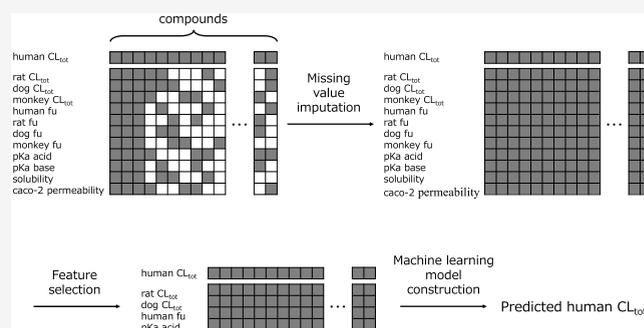
ACCESS |

Metrics & More

Article Recommendations

Supporting Information

**ABSTRACT:** Pharmacokinetic research plays an important role in the development of new drugs. Accurate predictions of human pharmacokinetic parameters are essential for the success of clinical trials. Clearance (CL) and volume of distribution (Vd) are important factors for evaluating pharmacokinetic properties, and many previous studies have attempted to use computational methods to extrapolate these values from nonclinical laboratory animal models to human subjects. However, it is difficult to obtain sufficient, comprehensive experimental data from these animal models, and many studies are missing critical values. This means that studies using nonclinical data as explanatory variables can only apply a small number of compounds to their model training. In this study, we perform missing-value imputation and feature selection on nonclinical data to increase the number of training compounds and nonclinical datasets available for these kinds of studies. We could obtain novel models for total body clearance ( $CL_{tot}$ ) and steady-state Vd ( $Vd_{ss}$ ) ( $CL_{tot}$ : geometric mean fold error [GMFE], 1.92; percentage within 2-fold error, 66.5%;  $Vd_{ss}$ : GMFE, 1.64; percentage within 2-fold error, 71.1%). These accuracies were comparable to the conventional animal scale-up models. Then, this method differs from animal scale-up methods because it does not require animal experiments, which continue to become more strictly regulated as time passes.



## INTRODUCTION

Pharmacokinetic evaluations play an important role in the development of new drugs throughout the entire process.<sup>1</sup> Clinical trials are particularly important in drug development, and improving the success rate of these requires the estimation of effective clinical dosages that produce the best drug effect profile. Therefore, it is necessary to accurately predict human pharmacokinetic parameters from nonclinical experimental data before transitioning to human clinical trials.<sup>2</sup> In general, the parameters that have a large effect on the blood concentration profile of a drug during intravenous administration are the volume of distribution (Vd), which quantifies the distribution of the drug inside the human body, and total body clearance ( $CL_{tot}$ ), which shows the drug processing capacity within the body as a whole.

Vd is determined by the physical properties of the drug, such as protein binding and membrane permeability, and predictions from machine learning models using chemical structures (CS) have been relatively accurate.<sup>3</sup> When nonclinical animal experimental values are available, the difference between the predicted and experimental values is maintained within approximately a 2-fold error using animal scale-up methods.<sup>4</sup> However, since this kind of highly accurate

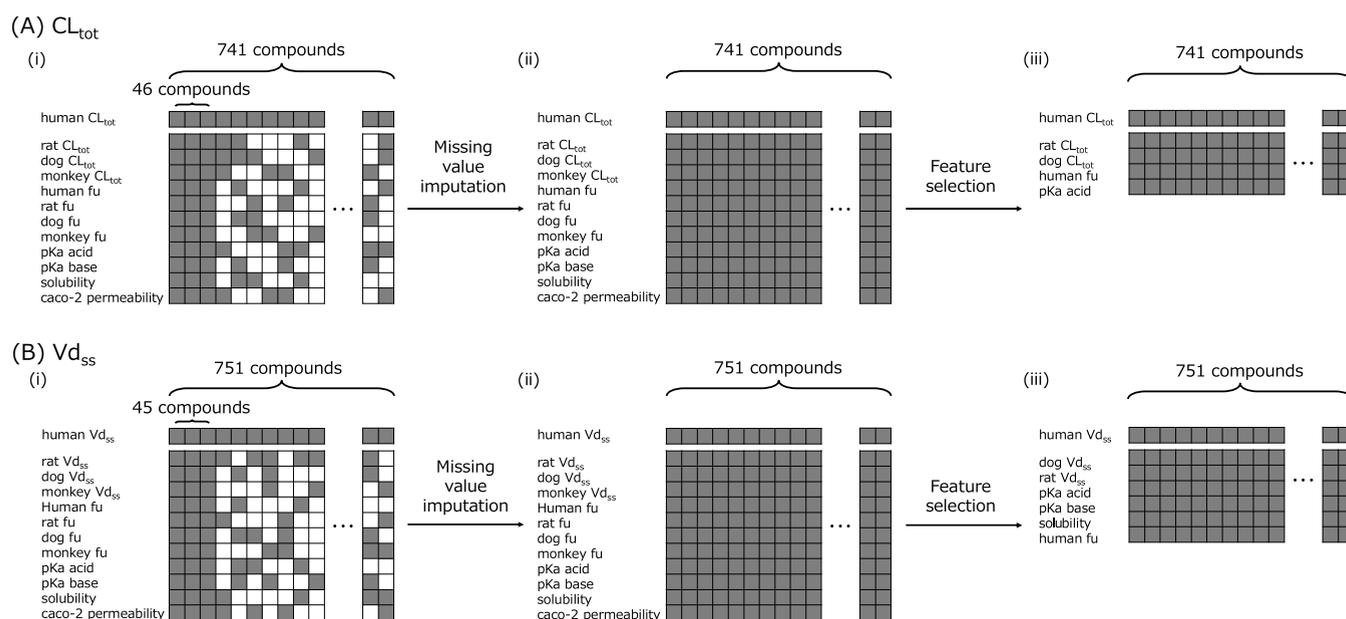
prediction method uses data from large animals such as dogs and monkeys, it is difficult to use this approach because of their high cost and ethical implications of large animal models.<sup>5</sup>

Predicting  $CL_{tot}$  is much more difficult than Vd because there are multiple drug CL pathways, including metabolism mainly by the liver and gastrointestinal tract, bile excretion of the unmetabolized drug, and its excretion in the urine. In one method, the intrinsic CL obtained in *in vitro* studies using human hepatocytes and microsomes was scaled up to determine hepatic CL. However, in many cases, the data cannot be accurately scaled because of issues around differences in the experimental systems and variations in lots between human specimens. Furthermore, there are currently no suitable *in vitro* experimental systems for other organs.<sup>6</sup> The empirical  $CL_{tot}$  predictions method using the animal weight

Received: March 23, 2022

Published: August 22, 2022





**Figure 1.** Workflow of our novel human  $CL_{tot}$  and  $Vd_{ss}$  prediction method. (A)  $CL_{tot}$  analysis flow. (i) There were 741 compounds with human  $CL_{tot}$  data and 46 that had values for all 11 features. (ii) All feature values were estimated via prediction using ADMETWORKS. (iii) Feature extraction was performed using XGBoost or Random Forest, and a prediction model was constructed. (B)  $Vd_{ss}$  analysis flow. (i) There were 751 compounds with human  $Vd_{ss}$  data and 46 that had values for all 11 features. (ii) All feature values were estimated via prediction using ADMETWORKS. (iii) Feature extraction was performed using XGBoost or Random Forest, and a prediction model was constructed.

power law showed accurate at an average of approximately 2-fold error. However, verification has not been performed on external datasets.<sup>7,8</sup>

For predicting  $CL_{tot}$ , several studies have already investigated using machine learning.<sup>9</sup> Then, some reports used related experimental values to  $CL_{tot}$  as explanatory variables.<sup>10</sup> We proposed using a machine learning method based on multimodal learning that takes the CS and nonclinical data for predicting human  $CL_{tot}$ .<sup>11</sup> The main point of this method to note is that the human  $CL_{tot}$  prediction accuracy is increased using both CS data and animal experimental data, suggesting that it may be possible to further improve human  $CL_{tot}$  prediction accuracy using not only rat  $CL_{tot}$  but also the  $CL_{tot}$  values from various animals (e.g., dogs and monkeys) and *in vitro* experimental values such as the protein binding ratio for each animal species as explanatory variables. However, these experimental values are often missing from the compound datasets.

Missing-value imputation is a well-known method for resolving this issue. Methods that use machine learning, such as kNNimputation,<sup>12</sup> multivariate imputation by chained equations,<sup>13</sup> and missForest,<sup>14</sup> are established imputation methods known for their ability to provide improved accuracy in these types of applications. The prediction of drug repositioning with high accuracy has been made possible by the addition of missing-value imputation based on the similarity of compound structures.<sup>15</sup> Missing data related to activity values for different compound targets were predicted using the Random Forest method, and a QSAR model was constructed, which uses the data from these imputed missing values as explanatory variables.<sup>16</sup>

In this study, first, we constructed machine learning models for predicting the missing nonclinical data using chemical compound descriptors. Next, we predicted the missing nonclinical data and then we constructed our machine learning model for predicting human  $CL_{tot}$  and  $Vd_{ss}$ , which used the

nonclinical data with imputed missing values and CS data as explanatory variables. XGBoost and Deep Tensor,<sup>17</sup> a deep learning method that can learn graph data, were used as the bases for these machine learning models. As a result, the prediction accuracy of this method is comparable with the many animal scale-up methods. Different from the conventional methods, since these models do not need new experimental data, it seems to be appropriate for predicting the human parameters in not only the clinical stage but also the early drug discovery stage.

## MATERIALS AND METHODS

**Workflow.** The workflow used in this study consisted of the following three steps: (i) gathering of the chemical compound and nonclinical data; (ii) imputing the missing values in the nonclinical data using ADMETWORKS (<https://www.fujitsu.com/jp/solutions/business-technology/tc/sol/admeworks/index.html>); and (iii) selecting features by XGBoost or Random Forest and constructing the prediction model (Figure 1).

**Gathering Chemical Compounds, Nonclinical Data, and Data Preprocessing.** We obtained 741 sets of human  $CL_{tot}$  data with CS data, and 751 sets of human  $Vd_{ss}$  data with CS data from JCP2013<sup>4,7</sup> and ChEMBL23.<sup>18</sup> We also obtained various sets of animal experimental data ( $CL_{tot}$ ,  $Vd_{ss}$ , and fraction-unbound data for rats, dogs, and monkeys) and human fu data for each of these compounds.<sup>4,7</sup> In addition, we collected the  $pK_a$  acid,  $pK_a$  base, solubility, and caco-2 permeability data including the calculated values for each compound from PubChem<sup>19</sup> and DrugBank.<sup>20</sup> Caco-2 permeability is a positive/negative binary value, and the values denoted as predicted values were also collected (Table 1). This left us with 46  $CL_{tot}$  and 45  $Vd_{ss}$  compounds that recorded all 11 data items for  $CL_{tot}$ : rat  $CL_{tot}$ , dog  $CL_{tot}$ , monkey  $CL_{tot}$ , human fu, rat fu, dog fu, monkey fu,  $pK_a$  acid,  $pK_a$  base,

Table 1. Details of the Compound Data

feature	number of compounds	source
human $CL_{tot}$	741	JCP2013, ChEMBL23
rat $CL_{tot}$	387	JCP2013, ChEMBL23
dog $CL_{tot}$	284	JCP2013, ChEMBL23
monkey $CL_{tot}$	129	JCP2013, ChEMBL23
human $Vd_{ss}$	751	JCP2013, ChEMBL23
rat $Vd_{ss}$	351	JCP2013, ChEMBL23
dog $Vd_{ss}$	274	JCP2013, ChEMBL23
monkey $Vd_{ss}$	125	JCP2013, ChEMBL23
human fu	577	JCP2013, ChEMBL23
rat fu	237	JCP2013, ChEMBL23
dog fu	179	JCP2013, ChEMBL23
monkey fu	88	JCP2013, ChEMBL23
$pK_a$ acid	334	Pubchem, DrugBank
$pK_a$ base	335	Pubchem, DrugBank
solubility	339	Pubchem, DrugBank
caco-2 permeability	307	Pubchem, DrugBank

solubility, and caco-2 permeability; and  $Vd_{ss}$ : rat  $Vd_{ss}$ , dog  $Vd_{ss}$ , monkey  $Vd_{ss}$ , human fu, rat fu, dog fu, monkey fu,  $pK_a$  acid,  $pK_a$  base, solubility, and caco-2 permeability (hereinafter defined as “nonclinical data”). These sets of compounds were labeled the “evaluation dataset” and the sets of compounds not in the “evaluation dataset” were labeled the “training dataset” for human  $CL_{tot}$  or  $Vd_{ss}$  (dataset.xlsx). In addition, the set of compounds formed by removing the “evaluation dataset” from the set of compounds that had rat data were labeled “training dataset (rat)” for the human  $CL_{tot}$  and  $Vd_{ss}$  data. Note that although the excretion pathways have not been identified in most of the compounds used in this study, when the data from Lombardo et al., who gathered drug data,<sup>7</sup> and Varma et al.,<sup>21</sup> who investigated human kidney excretion, were compared, kidney excretion of 50% or less of the  $CL_{tot}$  was found in 157 of the 231 remaining compounds. Furthermore, although data were gathered during intravenous administration of eptalprost, there are errors in the cited references, and it has been shown that the human data are for oral administration.

The gathered data were then preprocessed as follows before applying each of the machine learning calculations described in the next section. First, the  $CL_{tot}$ ,  $Vd_{ss}$ , and solubility data underwent a logarithmic transformation and then all of the data except the caco-2 permeability values were standardized. It is worth noting that the data used in the animal scale-up method evaluations were the raw data without data preprocessing.

**Missing-Value Imputation Using ADMEWORKS.** Some of the 11 features of the nonclinical data were missing for some of the compounds in the “training dataset” (Table 1). Therefore,

we created machine learning models using existing data for each item and predicted the unknown nonclinical values for these features using these models. ADMEWORKS was used to complete the descriptor calculations, descriptor extractions, prediction model construction, and prediction. First, compound descriptors with a total of 1465 dimensions were calculated from descriptors with 555 dimensions identified from their atom/bond-related parameters, topology-related parameters, and physiochemical parameters, and descriptors with 910 dimensions were obtained by counting a partial structure search of CS. Next, descriptors containing missing values, calculation errors, and descriptors with a correlation coefficient of 0.9 or higher (the default setting in ADMEWORKS) were excluded. In addition, high-level feature extraction (particle swarm optimization)<sup>22</sup> was performed, and a model was constructed using the remaining descriptors. The machine learning model that maximized the percentage within a 2-fold error in 5-fold cross-validation (maximized the two-class accuracy of caco-2 permeability only) was then adopted for downstream analyses. The nonclinical data for each compound were then predicted using each prediction model and the machine learning methods used for each item and the learning model accuracies are listed in Table S1.

**Feature Selection and Prediction Model Construction.**

**Feature Selection.** Given the possibility that nonclinical data may not be that useful for prediction and the possibility of inappropriate missing-value imputation having an adverse effect on prediction, specific items among the 11 nonclinical data categories were selected as explanatory variables. This selection was performed based on the importance of the explanatory variable as determined during the construction of the prediction model using the XGBoost or Random Forest method.<sup>23</sup> First, the prediction models for human  $CL_{tot}$  and  $Vd_{ss}$  were constructed with “training dataset” by the XGBoost or Random Forest method using all 11 items from the nonclinical data as explanatory variables. This allowed us to evaluate the importance of each of the 11 variables within the model. Next, the top  $k$  nonclinical data of importance were selected as explanatory variables. The best  $k$  was determined by the search using the initial  $k$  value where the total of the top  $k$  importance exceeded 0.5 with each  $k$  evaluated in single-value increments. For each  $k$ , 5-fold cross-validation using the “training dataset” was performed using the multimodal model described below, whose explanatory variables included the CS and the top  $k$  nonclinical data of importance. This was evaluated using the geometric mean fold error (GMFE) and percentage within a 2-fold error, which are described below. When either the GMFE or percentage within a 2-fold error became worse than the previous value of  $k$  (i.e.,  $k - 1$ ), the search was complete. Finally, the  $k$  that gave the best GMFE

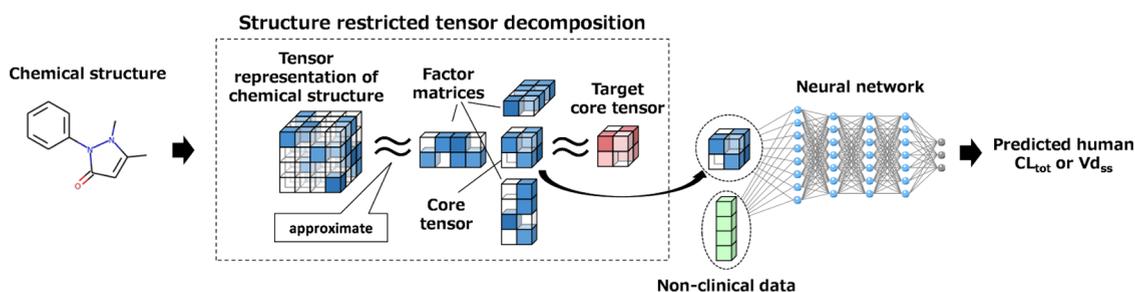


Figure 2. Overview of the multimodal Deep Tensor model.

and percentage with a 2-fold error within the search range was considered the best  $k$ , and the nonclinical data at the best  $k$  were selected for further evaluation.

**Deep Tensor Model.** To evaluate the effectiveness of the missing-value imputation, a prediction model that uses only the CS data as explanatory variables was constructed using Deep Tensor for comparison. The “training dataset” was used to construct this prediction model and the other conditions were set to the same values as those used in the multimodal Deep Tensor model described below.

**Multimodal Deep Tensor Model.** Prediction models for human  $CL_{tot}$  and  $Vd_{ss}$  were constructed using the CS and nonclinical data as explanatory variables using the previously described method<sup>11</sup> for Deep Tensor,<sup>17</sup> a deep learning technology for structured graph data (Figure 2). For  $CL_{tot}$  and  $Vd_{ss}$ , four combinations of the nonclinical data were used as explanatory variables with the training dataset used to construct the prediction model: (1) rat data only + “training dataset (rat)” (as CS + rat  $CL_{tot}$  or  $Vd_{ss}$  in Table S2); (2) rat data only + “training dataset” (as CS + rat  $CL_{tot}$  or  $Vd_{ss}$  imputed in Table S2); (3) all 11 nonclinical data points + “training dataset” (as CS + 11 features in Table S3); and (4) nonclinical data selected by the feature selection described above + “training dataset” (as CS + selected features in Table S3). The core tensor size was set at  $50 \times 50$  and the neural network structure consisted of two intermediate layers, 1000 neurons in each layer, and one neuron in the output layer. The ReLU function<sup>24</sup> was used as the activation function, and batch normalization<sup>25</sup> with a decay rate of moving average = 0.9, epsilon value =  $2 \times 10^{-5}$ , and dropout<sup>26</sup> at a rate of 0.5 was applied to produce the intermediate layers. Then, during training, the number of epochs was set to 50, and the minibatch size was set to 100.

**XGBoost Model.** To evaluate the effectiveness of the missing-value imputation, we performed XGBoost calculation as a traditional machine learning model. XGBoost was implemented using scikit-learn of Python language.<sup>27</sup> Prediction models that used only the CS data and the CS data and nonclinical data as explanatory variables were constructed. We transformed the CS data into the extended connectivity fingerprint with bond diameter four (ECFP4). The ECFP4 compound descriptor was calculated using RDKit with parameters of radius 2 and 2048 dimensions.

**Animal Scale-Up and Conventional Machine Learning Methods.** The single-species allometric scaling (SSS) method for  $CL_{tot}$ , which uses the  $CL_{tot}$  values of any single model, rats, dogs, or monkeys, the simple allometry (SA) method, which uses all three species, and the fraction-unbound corrected intercept method (FCIM)<sup>28</sup> are often implemented as the conventional method of human  $CL_{tot}$  prediction. For the construction of each model of SSS, the compounds that have the value needed for each model are used from “training dataset.” Then, each model is evaluated by the “evaluation dataset.” For SA and FCIM that have no training process by the training dataset, the parameters are tuned by the data of each compound to be predicted. The number of compounds in the training process for parameter selection for each method and the equations for each method are shown in Table S4.

The SSS and SA method for  $Vd_{ss}$  uses the  $Vd_{ss}$  values similar to the  $CL_{tot}$  prediction models described above. Then, the  $\text{Oie-Tozer}$ <sup>29</sup> method was also used as the conventional method of human  $Vd_{ss}$  prediction. The process of construction and evaluation of each model of SA is similar to  $CL_{tot}$ . For the

SA and  $\text{Oie-Tozer}$  method that have no training process by the training dataset, the parameters are tuned by the data of each compound to be predicted. The number of compounds in the training dataset for each parameter and each method and the equations for each method are shown in Table S5.

**Performance Evaluation.** GMFE and percentage within a 2-fold error were used as indicators for evaluating the prediction accuracy of the method. When  $GMFE = X$ , the mean error between the measured and predicted values can be interpreted as an X-fold error. GMFE is expressed by the following equation

$$GMFE = 10^{\text{mean}|\log_{10}(\text{predicted}/\text{observed})|}$$

where GMFE values closer to 1 indicate improved accuracy. Furthermore, a percentage within a 2-fold error indicates the proportion of data that are within a 2-fold error ( $1/2 \times \text{correct value} \leq \text{predicted value} \leq 2 \times \text{correct value}$ ). Values of percentage within a 2-fold error closer to 100% indicate better accuracy. When the evaluation results did not match for both indicators, GMFE was used as the primary predictor of accuracy as this is the more comprehensive indicator value.

## RESULTS

**Evaluation of the Usefulness of Missing-Value Imputation.** To clarify the effectiveness of missing-value imputation, the accuracy was evaluated using rat data, which had the fewest missing variables. More specifically, three models were created: a model trained using only CS data, a multimodal model using CS and rat  $CL_{tot}$  or rat  $Vd_{ss}$  (CS + rat  $CL_{tot}$  or rat  $Vd_{ss}$ ) data, and a multimodal model using CS and rat  $CL_{tot}$  or rat  $Vd_{ss}$  imputed data using predicted values (CS + rat  $CL_{tot}$  imputed or rat  $Vd_{ss}$  imputed), and the effectiveness of the missing-value imputation was evaluated (Table S2). Evaluation was performed using the evaluation dataset that included established values for human, rat, dog, and monkey  $CL_{tot}$  or  $Vd_{ss}$  for 45 compounds for the CL prediction or 46 compounds for the  $Vd_{ss}$  data predictions. Note that variation was inhibited in these evaluations due to the limited number of compounds in the evaluation dataset; construction of the models and evaluation using the evaluation dataset were completed five times, and the mean values were used for the evaluation.

Table 2 shows the results of the missing-value imputation for  $CL_{tot}$  prediction. First, the accuracies were compared for the model trained using only CS and the multimodal model using CS and rat  $CL_{tot}$  (CS + rat  $CL_{tot}$ ). The training data for the model using only CS consisted of 695 compounds excluding those present in the evaluation dataset (only CS), the GMFE was XGBoost: 2.53 and Deep Tensor: 2.44, and the percentage within the 2-fold error was XGBoost: 45.7% and Deep Tensor: 45.7%. For the multimodal model including CS and rat  $CL_{tot}$ , taking the 343 compounds that had a rat  $CL_{tot}$  value from the 695 compounds as training data (CS + rat  $CL_{tot}$ ), the GMFE was XGBoost: 2.15 and Deep Tensor: 2.15, and the percentage within 2-fold error was XGBoost: 52.2% and Deep Tensor: 54.8%. This finding confirmed that the accuracy was improved by introducing rat  $CL_{tot}$  values to the  $CL_{tot}$  prediction. This result is consistent with our previous report.<sup>11</sup> Next, although accuracy is generally increased by increasing the amount of training data, there were fewer compounds for which nonclinical data, such as rat  $CL_{tot}$  values, were measured. Therefore, we performed prediction using ADMEWORKS for the compounds with no rat  $CL_{tot}$  value data (see the Materials

**Table 2. Results of the Accuracy Evaluations for Imputations of Rat  $CL_{tot}$  Data**

	method	GMFE	% of 2-fold error
$CL_{tot}$	695 compounds	2.53	45.7
	XGBoost: only CS		
	343 compounds	2.15	52.2
	XGBoost: CS + rat $CL_{tot}$		
	695 compounds	2.09	54.3
	XGBoost: CS + rat $CL_{tot}$ imputed		
	695 compounds	2.44	45.7
	Deep Tensor: only CS		
	343 compounds	2.15	54.8
	Deep Tensor: CS + rat $CL_{tot}$		
$Vd_{ss}$	706 compounds	1.66	82.2
	XGBoost: only CS		
	306 compounds	1.72	75.6
	XGBoost: CS + rat $Vd_{ss}$		
	706 compounds	1.73	68.9
	XGBoost: CS + rat $Vd_{ss}$ imputed		
	706 compounds	1.85	62.2
	Deep Tensor: only CS		
	306 compounds	1.89	56.9
	Deep Tensor: CS + rat $Vd_{ss}$		
	706 compounds	1.75	64.4
	Deep Tensor: CS + rat $Vd_{ss}$ imputed		

and Methods for more details). This meant that we could then train the multimodal model using the CS and rat  $CL_{tot}$  imputed data with 695 compounds (CS + rat  $CL_{tot}$  imputed). This model produced a GMFE of XGBoost: 2.09 and Deep Tensor: 2.09, and the percentage within 2-fold error was XGBoost: 54.3% and Deep Tensor: 54.3%, making it the most accurate model. These results suggest that model accuracy can be improved by increasing the data used during training, even if this is imputed data.

Next, Table 2 also shows the results of the missing-value imputation for  $Vd_{ss}$ . When a comparison was performed using training data for 706 compounds, excluding the evaluation dataset, the model trained using only CS (only CS) and the multimodal model using CS and rat  $Vd_{ss}$  (CS + rat  $Vd_{ss}$ , using 306 compounds) had GMFEs of XGBoost: 1.66 and 1.72, Deep Tensor: 1.85 and 1.89, and percentages within 2-fold errors of XGBoost: 82.2% and 75.6% and Deep Tensor: 62.2% and 56.9%, respectively. This means that the model using only CS (only CS) was more accurate than the multimodal model using CS and actual nonclinical data (CS + rat  $Vd_{ss}$ ). It is known that the structure–activity relationship is stronger for  $Vd_{ss}$  than  $CL_{tot}$  for prediction from structure information,<sup>3</sup> and this suggests that the increase in structure alone in the training data increases the accuracy of the prediction model. Furthermore, when we evaluated the model produced using the training data of 706 compounds with imputed nonclinical data where there was no rat  $Vd_{ss}$  value, the multimodal model using CS and rat  $Vd_{ss}$  imputed using predicted values (CS + rat  $Vd_{ss}$  imputed) had a GMFE of XGBoost: 1.73 and Deep Tensor: 1.75, and a percentage within 2-fold error of XGBoost: 68.9% and Deep Tensor: 64.4%, making it the most accurate model.

**Improving Accuracy Using Feature Selection.** Improved accuracy was achieved using multimodal machine learning models and imputed values for the missing rat  $CL_{tot}/Vd_{ss}$  data. We then went on to evaluate the addition of multimodal machine learning models designed to select specific features from the 11 items of nonclinical data in each dataset (Table S3).

We first created a set of learning models for each of the 11 items in the nonclinical data using ADMEWORKS, and missing-value imputation was performed by prediction. Details of the prediction model and model accuracy for each item of the nonclinical data are shown in Table S1. Multiple models were constructed for each item, and missing-value imputation was performed by predicting the unknown clinical data using the model with the highest accuracy.

We went on to complete feature selection using these 11 items with their imputed data for any missing values. The importance of each item in the nonclinical data is shown in Figures S1–S4. We then used their importance to select the four items (rat  $CL_{tot}$ , dog  $CL_{tot}$ , human fu, and  $pK_a$  acid) critical for  $CL_{tot}$  prediction (Figure S1 and Table S6) and the five items (rat  $Vd_{ss}$ , dog  $Vd_{ss}$ ,  $pK_a$  acid,  $pK_a$  base, and human fu) identified for  $Vd_{ss}$  prediction (Figure S2 and Table S7) from the nonclinical data using the XGBoost algorithm. We then used their importance to select the four items (rat  $CL_{tot}$ , dog  $CL_{tot}$ , human fu, and  $pK_a$  acid) critical for  $CL_{tot}$  prediction (Figure S3 and Table S8) and the six items (dog  $Vd_{ss}$ , rat  $Vd_{ss}$ ,  $pK_a$  acid,  $pK_a$  base, solubility, and human fu) identified for  $Vd_{ss}$  prediction (Figure S4 and Table S9) from the nonclinical data using the Random Forest algorithm.

Finally, multimodal machine learning models were constructed using the selected nonclinical and CS data. The accuracies of this model, five different conventional methods, the model using only CS, and the multimodal model using CS and all 11 items of nonclinical data were evaluated using the “evaluation dataset.” We then repeated the dataset evaluations five times for both the multimodal machine learning model and the model using CS only, and the mean value was used for evaluation, similar to the evaluation of the usefulness of missing-value imputation. Table 3 shows the results of the  $CL_{tot}$  prediction using eight different models. Among the five types of conventional models, models that use monkey  $CL_{tot}$  data, such as SSS monkey (GMFE: 1.93, percentage within 2-fold error: 58.7%) and FCIM (GMFE: 1.99, percentage within 2-fold error: 52.2%) were the most accurate. Among the multimodal models using nonclinical data with missing-value imputation and CS as proposed in this research, the model using all 11 items (CS + 11 features) presented with a GMFE of XGBoost: 2.06 and Deep Tensor: 2.11 and a percentage within 2-fold error of XGBoost: 58.7% and Deep Tensor: 52.2% that was equivalent to the conventional methods, while the model using feature selection (CS + selected features) was shown to be the most accurate with a GMFE of XGBoost: 1.98 and Deep Tensor: 1.92 and a percentage within a 2-fold error of XGBoost: 50.0% and Deep Tensor: 66.5%. These results indicate that predictive models can be improved by increasing the number of compounds used for training and that these can be enhanced by first imputing any missing data using predictive values. In addition, these results suggest that a better model can be constructed by performing feature selection and training using only the important features from the nonclinical data.

Table 3. Results of Accuracy Evaluations

	method <sup>a</sup>	GMFE <sup>b</sup>	% of 2-fold error	
CL <sub>tot</sub>	SSS rat	2.36	43.5	
	SSS dog	2.30	39.1	
	SSS monkey	1.93	58.7	
	SA	2.33	45.7	
	FCIM	1.99	52.2	
	XGBoost: only CS	2.40	50.0	
	XGBoost: CS + 11 features	2.06	58.7	
	XGBoost: CS + selected features	1.98	50.0	
	Deep Tensor: only CS	2.44	45.7	
	Deep Tensor: CS + 11 features	2.11	52.2	
	Deep Tensor: CS + selected features	1.92	66.5	
	Vd <sub>ss</sub>	SSS rat	1.91	62.2
		SSS dog	1.93	71.1
		SSS monkey	1.60	80.0
SA		2.07	68.9	
Øie–Tozer		1.46	84.4	
XGBoost: only CS		1.70	77.8	
XGBoost: CS + 11 features		1.64	71.1	
XGBoost: CS + selected features		1.66	71.1	
Deep Tensor: only CS		1.85	62.2	
Deep Tensor: CS + 11 features		1.75	69.8	
Deep Tensor: CS + selected features		1.74	74.2	

<sup>a</sup>SSS: single-species allometric scaling; SA: simple allometry; FCIM: fu-corrected intercept method; CS: chemical structure. <sup>b</sup>GMFE: geometric mean fold error.

Table 3 also summarizes the results from the Vd<sub>ss</sub> predictions. First, the Øie–Tozer method (GMFE: 1.46, percentage within 2-fold error: 84.4%) was shown to be the most accurate of the conventional methods evaluated and was closely followed by SSS monkey (GMFE: 1.60, percentage within 2-fold error: 80.0%). However, there were no significant differences among the multimodal models using nonclinical data with missing-value imputation and CS data, with CS + 11 features producing a GMFE of XGBoost: 1.64 and Deep Tensor: 1.75 and a percentage within 2-fold error of XGBoost: 71.1% and Deep Tensor: 69.8% and the feature selection model (CS + selected features) producing a GMFE of XGBoost: 1.66 and Deep Tensor: 1.74 and a percentage within 2-fold error of XGBoost: 71.1% and Deep Tensor: 74.2%, despite a small improvement in the overall accuracy as described by the percentage within 2-fold error value. Unlike in CL<sub>tot</sub> prediction, Vd<sub>ss</sub> prediction was still more accurate using the conventional Øie–Tozer method, which is based on animal scale-up data.

## DISCUSSION

**Missing-Value Imputation (NA Imputation).** This research demonstrates that high-accuracy predictions of CL<sub>tot</sub> and Vd<sub>ss</sub> can be achieved via data extension facilitated by missing-value imputation. Since this method does not require new experimental values, it can be used from the initial stages of drug development.

Although the effectiveness of imputation using this method has been confirmed for rat data, it had not been evaluated for other nonclinical data. Therefore, nonclinical data probably exists where imputation is not effective, but rather reduces prediction accuracy. Furthermore, nonclinical data may also exist where imputation is appropriate but not effective for

prediction. With these possibilities in mind, we were careful to select the explanatory variables in this study. Given this, it was necessary to evaluate all combinations of the 12 candidate explanatory variables, consisting of the CS data and 11 items of nonclinical data, to select the explanatory variables that are truly optimal for developing a multimodal Deep Tensor model. However, since this would take a huge amount of time, this study used a simplified version of these evaluations to select the explanatory variables. More specifically, the CS data were taken as always selected, and the required items from among the 11 nonclinical data items were differentially applied. When selecting the nonclinical data features, we determined that the number of explanatory variables that should be selected were evaluated based on the importance of the explanatory variables (i.e., nonclinical data), as obtained from the XGBoost or Random Forest evaluations (Figures S1–S4 and Tables S6–S9). This allowed us to reduce the number of combinations of explanatory variables that needed to be evaluated. Note that the importance of the nonclinical data obtained from the XGBoost or Random Forest method assumes that a prediction model is constructed using the XGBoost or Random Forest method. As a result, it is possible that these do not match the importance of the multimodal Deep Tensor model. Furthermore, in the XGBoost or Random Forest method, the importance of each of the explanatory variables was calculated using only the nonclinical data. As a result, when these are used together with the CS data in the multimodal Deep Tensor model, it is possible that their importance may change. Therefore, we cannot definitively say whether we selected the best explanatory variables.

**Selected Explanatory Variables.** The explanatory variables selected for CL<sub>tot</sub> prediction (CS + selected features), which had the highest prediction accuracy were CL<sub>tot</sub> rat, CL<sub>tot</sub> dog, human fu, and pK<sub>a</sub> acid in both XGBoost and Deep Tensor (Figures S1 and S3 and Tables S6 and S8). We believe that the selection of CL<sub>tot</sub> for multiple species may help to accurately reflect the inherent metabolic differences between species.<sup>30</sup> However, the fact that monkey CL<sub>tot</sub> was not selected could possibly be due to problems with imputation accuracy since the number of datasets used for missing-value imputation for the monkey values was significantly smaller than that of the other species. In addition, we believe that these selections are valid as human fu was selected and this is a common consideration for human CL<sub>tot</sub> prediction.<sup>7</sup> The addition of the pK<sub>a</sub> acid variable may help to accurately reflect the allocation of the metabolism/excretion pathways depending on the compound's physical properties.<sup>31</sup>

The explanatory variables selected for the most accurate prediction of Vd<sub>ss</sub> (CS + selected features) included rat Vd<sub>ss</sub>, dog Vd<sub>ss</sub>, human fu, pK<sub>a</sub> acid, and pK<sub>a</sub> base in XGBoost (Figure S2 and Table S7) and rat Vd<sub>ss</sub>, dog Vd<sub>ss</sub>, human fu, pK<sub>a</sub> acid, pK<sub>a</sub> base, and solubility in Deep Tensor (Figure S4 and Table S9). The Vd of multiple species is likely to reflect the inherent differences in the Vd pathways between species<sup>32</sup> and the validity of these selections was similarly supported by the addition of the human fu parameters.<sup>4</sup> However, significantly more physical property parameters (e.g., pK<sub>a</sub> acid, pK<sub>a</sub> base, solubility) were included in the Vd<sub>ss</sub> evaluations; this is likely designed to reflect the fact that the Vd of various compounds is determined by interactions between the compound and the constituent components of the tissue (e.g., lipids, phospholipids, acidic glycoprotein).<sup>33</sup>

**Comparison with Animal Scale-Up Methods.** Conventional animal scale-up methods are known to be useful, offer good accuracy, and are often used for human pharmacokinetic parameter prediction during drug development.<sup>4,7</sup> Here, we showed that SSS monkey, which uses monkey data, had the next highest accuracy for  $CL_{tot}$  prediction after applying a multimodal model with missing-value imputation + feature selection (Deep Tensor: CS + selected features). This was followed by the FCIM method, which uses data from rats, dogs, and monkeys and XGBoost: CS + selected features. Among the  $Vd_{ss}$  prediction methods, the  $\emptyset_{ie}$ –Tozer method, which is calculated based on the plasma and tissue binding rates in rats, dogs, and monkeys, had the highest accuracy, followed by an animal scale-up method using monkey (SSS monkey) data and the proposed multimodal models with missing-value imputation and features (XGBoost: CS + 11 features and Deep Tensor: CS + selected features), which had slightly worse accuracy (Table 3). Given this, the prediction accuracy of human  $CL_{tot}$  and  $Vd_{ss}$  using monkey data can be said to be valid for humans given the close genetic relationship between these species.<sup>34,35</sup> However, when the experimental costs and ethical aspects of performing animal experiments are considered, since monkey experiments tend to be completed at later stages in the nonclinical development stage, employing these in the initial stages of drug development is difficult.<sup>34,36</sup> Therefore, the proposed multimodal model that includes missing-value imputation and feature selection using existing data can be applied in the initial stages of drug development and is expected to contribute substantially to efficient drug development.

## CONCLUSIONS

This study constructed a set of high-accuracy  $CL_{tot}$  and  $Vd_{ss}$  prediction models using missing-value imputation and feature selection for nonclinical data. Previous evaluations using nonclinical data as explanatory variables were shown to be less effective as the number of missing data points meant that the final number of evaluated compounds was too small for accurate machine learning. Therefore, we confirmed that the accuracy of these models is improved as a result of increasing the number of compounds used for training and increasing the number of (nonclinical data) explanatory variables that can be used by performing missing-value imputation on nonclinical data. This method differs from animal scale-up methods in that it does not require animal experiments, which have become more strictly regulated in recent years. Although we used XGBoost and Deep Tensor algorithms in this research, the other machine learning algorithms could be applied because this proposed method of imputation has no preference in machine learning algorithms. The increased accuracy of the  $CL_{tot}$  and  $Vd_{ss}$  predictions produced by this method are expected to facilitate the evaluation and identification of candidate structures with improved pharmacokinetic properties at the earlier stages of drug discovery.

## ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jcim.2c00318>.

The data used in this paper are listed in Supporting\_information\_dataset.xlsx (PDF)

## AUTHOR INFORMATION

### Corresponding Authors

Hiroaki Iwata – Graduate School of Medicine, Kyoto University, Kyoto 606-8507, Japan; [orcid.org/0000-0001-9791-0008](https://orcid.org/0000-0001-9791-0008); Email: [iwata.hiroaki.3r@kyoto-u.ac.jp](mailto:iwata.hiroaki.3r@kyoto-u.ac.jp)

Koichi Handa – Toxicology & DMPK Research Department, Teijin Institute for Bio-medical Research, Teijin Pharma Limited, Tokyo 191-8512, Japan; Email: [ko.handa@teijin.co.jp](mailto:ko.handa@teijin.co.jp)

### Authors

Tatsuru Matsuo – Fujitsu Ltd., Kawasaki-shi, Kanagawa 211-8588, Japan

Hideaki Mamada – DMPK Research Laboratories, Central Pharmaceutical Research Institute, Japan Tobacco Inc., Takatsuki, Osaka 569-1125, Japan; [orcid.org/0000-0002-5433-7042](https://orcid.org/0000-0002-5433-7042)

Takahisa Motomura – Central Pharmaceutical Research Institute, Japan Tobacco Inc., Takatsuki, Osaka 569-1125, Japan

Mayumi Matsushita – Fujitsu Ltd., Kawasaki-shi, Kanagawa 211-8588, Japan

Takeshi Fujiwara – Graduate School of Medicine, Kyoto University, Kyoto 606-8507, Japan

Kazuya Maeda – Graduate School of Pharmaceutical Sciences, Department of Molecular Pharmacokinetics, The University of Tokyo, Tokyo 113-0033, Japan

Complete contact information is available at: <https://pubs.acs.org/10.1021/acs.jcim.2c00318>

### Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

### Funding

This paper is based on the partial results obtained from a project commissioned by the New Energy and Industrial Technology Development Organization (NEDO) and a grant-in-aid of The Fugaku Trust For Medicinal Research.

### Notes

The authors declare no competing financial interest. Software used: ADMEWORKS (<https://www.fujitsu.com/jp/solutions/business-technology/tc/sol/admeworks/index.html>).

## ACKNOWLEDGMENTS

This research was conducted as part of the activities of the Life Intelligence Consortium (LINC). The authors thank Dr. Yasushi Okuno from the Graduate School of Medicine at Kyoto University for supporting their research activities at the LINC.

## ABBREVIATIONS USED

$CL$ , clearance;  $CL_{tot}$ , total clearance;  $Vd$ , volume of distribution;  $Vd_{ss}$ , steady-state volume of distribution; SSS, single-species allometric scaling; SA, simple allometry; FCIM, fraction-unbound corrected intercept method; CS, chemical structure; GMFE, geometric mean fold error; CYP, cytochrome P450

## REFERENCES

- (1) Ballard, P.; Brassil, P.; Bui, K. H.; Dolgos, H.; Petersson, C.; Tunek, A.; Webborn, P. J. The right compound in the right assay at the right time: an integrated discovery DMPK strategy. *Drug Metab. Rev.* **2012**, *44*, 224–252.
- (2) Andrade, E. L.; Bento, A. F.; Cavalli, J.; Oliveira, S. K.; Schwanke, R. C.; Siqueira, J. M.; Freitas, C. S.; Marcon, R.; Calixto, J. B. Non-clinical studies in the process of new drug development - Part II: Good laboratory practice, metabolism, pharmacokinetics, safety and dose translation to clinical studies. *Braz. J. Med. Biol. Res.* **2016**, *49*, No. e5646.
- (3) Lombardo, F.; Jing, Y. In Silico Prediction of Volume of Distribution in Humans. Extensive Data Set and the Exploration of Linear and Nonlinear Methods Coupled with Molecular Interaction Fields Descriptors. *J. Chem. Inf. Model.* **2016**, *56*, 2042–2052.
- (4) Lombardo, F.; Waters, N. J.; Argikar, U. A.; Dennehy, M. K.; Zhan, J.; Gunduz, M.; Harriman, S. P.; Berellini, G.; Rajlic, I. L.; Obach, R. S. Comprehensive assessment of human pharmacokinetic prediction based on in vivo animal pharmacokinetic data, part 1: volume of distribution at steady state. *J. Clin. Pharmacol.* **2013**, *53*, 167–177.
- (5) Russell, W. M. S.; Burch, R. L. *The Principles Of Humane Experimental Technique*; Methuen, 1959.
- (6) Shiran, M. R.; Proctor, N.; Howgate, E.; Rowland-Yeo, K.; Tucker, G.; Rostami-Hodjegan, A. Prediction of metabolic drug clearance in humans: in vitro–in vivo extrapolation vs allometric scaling. *Xenobiotica* **2006**, *36*, 567–580.
- (7) Lombardo, F.; Waters, N. J.; Argikar, U. A.; Dennehy, M. K.; Zhan, J.; Gunduz, M.; Harriman, S. P.; Berellini, G.; Liric Rajlic, I.; Obach, R. S. Comprehensive assessment of human pharmacokinetic prediction based on in vivo animal pharmacokinetic data, part 2: clearance. *J. Clin. Pharmacol.* **2013**, *53*, 178–191.
- (8) (a) Crouch, R. D.; Hutzler, J. M.; Daniels, J. S. A novel in vitro allometric scaling methodology for aldehyde oxidase substrates to enable selection of appropriate species for traditional allometry. *Xenobiotica* **2018**, *48*, 219–231. (b) Mahmood, I. A Single Animal Species-Based Prediction of Human Clearance and First-in-Human Dose of Monoclonal Antibodies: Beyond Monkey. *Antibodies* **2021**, *10*, 35. (c) Sasabe, H.; Koga, T.; Furukawa, M.; Matsunaga, M.; Kaneko, Y.; Koyama, N.; Hirao, Y.; Akazawa, H.; Kawabata, M.; Kashiwama, E.; Takeuchi, K. Pharmacokinetics and metabolism of brexpiprazole, a novel serotonin-dopamine activity modulator and its main metabolite in rat, monkey and human. *Xenobiotica* **2021**, *51*, 590–604.
- (9) (a) Wang, Y.; Liu, H.; Fan, Y.; Chen, X.; Yang, Y.; Zhu, L.; Zhao, J.; Chen, Y.; Zhang, Y. In Silico Prediction of Human Intravenous Pharmacokinetic Parameters with Improved Accuracy. *J. Chem. Inf. Model.* **2019**, *59*, 3968–3980. (b) Gombar, V. K.; Hall, S. D. Quantitative structure-activity relationship models of clinical pharmacokinetics: clearance and volume of distribution. *J. Chem. Inf. Model.* **2013**, *53*, 948–957. (c) Demir-Kavuk, O.; Bentzien, J.; Muegge, I.; Knapp, E. W. DemQSAR: predicting human volume of distribution and clearance of drugs. *J. Comput. Aided Mol. Des.* **2011**, *25*, 1121–1133.
- (10) (a) Kosugi, Y.; Hosea, N. Direct Comparison of Total Clearance Prediction: Computational Machine Learning Model versus Bottom-Up Approach Using In Vitro Assay. *Mol. Pharm.* **2020**, *17*, 2299–2309. (b) Miljković, F.; Martinsson, A.; Obrezanova, O.; Williamson, B.; Johnson, M.; Sykes, A.; Bender, A.; Greene, N. Machine Learning Models for Human In Vivo Pharmacokinetic Parameters with In-House Validation. *Mol. Pharm.* **2021**, *18*, 4520–4530.
- (11) Iwata, H.; Matsuo, T.; Mamada, H.; Motomura, T.; Matsushita, M.; Fujiwara, T.; Kazuya, M.; Handa, K. Prediction of total drug clearance in humans using animal data: proposal of a multimodal learning method based on deep learning. *J. Pharm. Sci.* **2021**, *110*, 1834.
- (12) Troyanskaya, O.; Cantor, M.; Sherlock, G.; Brown, P.; Hastie, T.; Tibshirani, R.; Botstein, D.; Altman, R. B. Missing value estimation methods for DNA microarrays. *Bioinformatics* **2001**, *17*, 520–525.
- (13) Schafer, J. L.; Olsen, M. K. Multiple imputation for multivariate missing-data problems: A data analyst's perspective. *Multivariate Behavioral Res.* **1998**, *33*, 545–571.
- (14) Stekhoven, D. J.; Bühlmann, P. MissForest–non-parametric missing value imputation for mixed-type data. *Bioinformatics* **2012**, *28*, 112–118.
- (15) Sawada, R.; Iwata, H.; Mizutani, S.; Yamanishi, Y. Target-Based Drug Repositioning Using Large-Scale Chemical-Protein Interactome Data. *J. Chem. Inf. Model.* **2015**, *55*, 2717–2730.
- (16) Martin, E. J.; Polyakov, V. R.; Zhu, X. W.; Tian, L.; Mukherjee, P.; Liu, X. All-Assay-Max2 pQSAR: Activity Predictions as Accurate as Four-Concentration IC50s for 8558 Novartis Assays. *J. Chem. Inf. Model.* **2019**, *59*, 4450–4459.
- (17) Maruhashi, K.; Todoriki, M.; Ohwa, T.; Goto, K.; Hasegawa, Y.; Inakoshi, H.; Anai, H. Learning Multi-way Relations Via Tensor Decomposition with Neural Networks, In *Thirty-Second AAAI Conference on Artificial Intelligence*, 2018.
- (18) Gaulton, A.; Bellis, L. J.; Bento, A. P.; Chambers, J.; Davies, M.; Hersey, A.; Light, Y.; McGlinchey, S.; Michalovich, D.; Al-Lazikani, B.; Overington, J. P. ChEMBL: a large-scale bioactivity database for drug discovery. *Nucleic Acids Res.* **2012**, *40*, D1100–D1107.
- (19) Kim, S.; Thiessen, P. A.; Bolton, E. E.; Chen, J.; Fu, G.; Gindulyte, A.; Han, L.; He, J.; He, S.; Shoemaker, B. A.; Wang, J.; Yu, B.; Zhang, J.; Bryant, S. H. PubChem Substance and Compound databases. *Nucleic Acids Res.* **2016**, *44*, D1202–D1213.
- (20) Wishart, D. S.; Feunang, Y. D.; Guo, A. C.; Lo, E. J.; Marcu, A.; Grant, J. R.; Sajed, T.; Johnson, D.; Li, C.; Sayeeda, Z.; Assempour, N.; Iynkkaran, I.; Liu, Y.; Maciejewski, A.; Gale, N.; Wilson, A.; Chin, L.; Cummings, R.; Le, D.; Pon, A.; Knox, C.; Wilson, M. DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Res.* **2018**, *46*, D1074–D1082.
- (21) Varma, M. V. S.; Feng, B.; Obach, R. S.; Troutman, M. D.; Chupka, J.; Miller, H. R.; El-Kattan, A. Physicochemical determinants of human renal clearance. *J. Med. Chem.* **2009**, *52*, 4844–4852.
- (22) Jahandideh-Tehrani, M.; Bozorg-Haddad, O.; Loaiciga, H. A. Application of particle swarm optimization to water management: an introduction and overview. *Environ. Monit. Assess.* **2020**, *192*, 281.
- (23) (a) Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. (b) Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V. Scikit-learn: Machine learning in Python. *J. Machine Learning Res.* **2011**, *12*, 2825–2830.
- (24) Glorot, X.; Bordes, A.; Bengio, Y. In *Deep Sparse Rectifier Neural Networks*. Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics 2011; pp 315–323.
- (25) Ioffe, S.; Szegedy, C. *Batch normalization: Accelerating deep network training by reducing internal covariate shift*. arXiv preprint arXiv:1502.03167 **2015**.
- (26) Srivastava, N.; Hinton, G.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R. Dropout: a simple way to prevent neural networks from overfitting. *J. Machine Learning Res.* **2014**, *15*, 1929–1958.
- (27) Varoquaux, G.; Buitinck, L.; Louppe, G.; Grisel, O.; Pedregosa, F.; Mueller, A. Scikit-learn: Machine learning without learning the machinery. *GetMobile: Mobile Comput. Commun.* **2015**, *19*, 29–33.
- (28) Tang, H.; Mayersohn, M. A novel model for prediction of human drug clearance by allometric scaling. *Drug Metab. Dispos.* **2005**, *33*, 1297–1303.
- (29) (a) Jones, R. D.; Jones, H. M.; Rowland, M.; Gibson, C. R.; Yates, J. W.; Chien, J. Y.; Ring, B. J.; Adkison, K. K.; Ku, M. S.; He, H.; et al. PhRMA CPCDC initiative on predictive models of human pharmacokinetics, part 2: comparative assessment of prediction methods of human volume of distribution. *J. Pharm. Sci.* **2011**, *100*, 4074–4089. (b) Obach, R. S.; Baxter, J. G.; Liston, T. E.; Silber, B. M.; Jones, B. C.; Macintyre, F.; Rance, D. J.; Wastall, P. The prediction of human pharmacokinetic parameters from preclinical and in vitro metabolism data. *J. Pharmacol. Exp. Ther.* **1997**, *283*, 46–58.

(30) Słoczyńska, K.; Gunia-Krzyzak, A.; Koczurkiewicz, P.; Wojcik-Pszczola, K.; Zelaszczyk, D.; Popiol, J.; Pekala, E. Metabolic stability and its role in the discovery of new chemical entities. *Acta. Pharm.* **2019**, *69*, 345–361.

(31) Wakayama, N.; Toshimoto, K.; Maeda, K.; Hotta, S.; Ishida, T.; Akiyama, Y.; Sugiyama, Y. In Silico Prediction of Major Clearance Pathways of Drugs among 9 Routes with Two-Step Support Vector Machines. *Pharm. Res.* **2018**, *35*, 197.

(32) Poulin, P.; Dambach, D. M.; Hartley, D. H.; Ford, K.; Theil, F. P.; Harstad, E.; Halladay, J.; Choo, E.; Boggs, J.; Liederer, B. M.; Dean, B.; Diaz, D. An algorithm for evaluating potential tissue drug distribution in toxicology studies from readily available pharmacokinetic parameters. *J. Pharm. Sci.* **2013**, *102*, 3816–3829.

(33) (a) Poulin, P.; Theil, F. P. A priori prediction of tissue: plasma partition coefficients of drugs to facilitate the use of physiologically-based pharmacokinetic models in drug discovery. *J. Pharm. Sci.* **2000**, *89*, 16–35. (b) Poulin, P.; Schoenlein, K.; Theil, F. P. Prediction of adipose tissue: plasma partition coefficients for structurally unrelated drugs. *J. Pharm. Sci.* **2001**, *90*, 436–447. (c) Poulin, P.; Krishnan, K. A biologically-based algorithm for predicting human tissue: blood partition coefficients of organic chemicals. *Hum. Exp. Toxicol.* **1995**, *14*, 273–280. (d) Berezhkovskiy, L. M. Volume of distribution at steady state for a linear pharmacokinetic system with peripheral elimination. *J. Pharm. Sci.* **2004**, *93*, 1628–1640. (e) Rodgers, T.; Rowland, M. Physiologically based pharmacokinetic modelling 2: predicting the tissue distribution of acids, very weak bases, neutrals and zwitterions. *J. Pharm. Sci.* **2006**, *95*, 1238–1257. (f) Schmitt, W. General approach for the calculation of tissue to plasma partition coefficients. *Toxicol. In Vitro* **2008**, *22*, 457–467.

(34) Tess, D. A.; Eng, H.; Kalgutkar, A. S.; Litchfield, J.; Edmonds, D. J.; Griffith, D. A.; Varma, M. V. S. Predicting the Human Hepatic Clearance of Acidic and Zwitterionic Drugs. *J. Med. Chem.* **2020**, *63*, 11831–11844.

(35) Miyamoto, M.; Iwasaki, S.; Chisaki, I.; Nakagawa, S.; Amano, N.; Hirabayashi, H. Comparison of predictability for human pharmacokinetics parameters among monkeys, rats, and chimeric mice with humanised liver. *Xenobiotica* **2017**, *47*, 1052–1063.

(36) Russell, W.; Burch, R. *The Principles of Humane Experimental Technique*. Wheathampstead, Universities Federation for Animal Welfare: UK; 1959.