ISSN 2618-1584 Vol. 7, Issue 4: 47-55, DOI: 10.25236/FMSR.2025.070406

# Machine Learning Approaches Combined with Power Analysis to Improve the Prediction of Population Pharmacokinetic Modeling

**Wen Huang[1,2], Ying Li[2], Jiangyan Lu[2], Nansong Wang[2]**

*[1]School of Life Sciences and Biopharmaceutical Science, Shenyang Pharmaceutical University, Shenyang, 110016, China*
*[2]Department of Pharmacy, The University of Hong Kong-Shenzhen Hospital, Shenzhen, Guangdong, 518000, China*

***Abstract:*** *This study aims to determine if integrating machine learning (ML) approaches with power analysis for correlated covariates enhances predictive performance in modeling. Two drugs were selected to apply population pharmacokinetics methodology. Parameter-covariate relationships were estimated using stepwise covariate modeling (SCM). Power analysis identified correlated true covariates for each parameter chosen by SCM. If correlated true covariates were identified, ML methods further selected the most significant covariate. The calibrated model utilizing this significant covariate was compared to the SCM model, evaluating performance using relative error. Non-compartment analysis (NCA) calculated individual reference AUC, enabling comparison of AUCs from calibrated and SCM models with the reference AUC to assess predictive performance. Results showed that calibrated models outperformed SCM models, with $R^2$ values exceeding 80%. Overall, the calibrated models predicted AUC more accurately than the SCM models. Power analysis helped eliminate false covariates among correlated ones, while ML methods provided criteria for selecting covariates from the identified true covariates. Thus, combining these methods can enhance population pharmacokinetic model predictions.*

***Keywords:*** *machine learning, calibrated models, pharmacokinetic*

## 1. Introduction

Model Informed Drug Discovery and Development (MID3) aims to enhance the quality, efficiency, and cost-effectiveness of decision-making through a predictive and extrapolative quantitative framework[1]. Within this context, the population pharmacokinetic model is pivotal in the research and development of new drugs[2]. Identifying parameter-covariate relationships is crucial for elucidating inter-individual variability, thereby improving the predictive performance of the population pharmacokinetic model[3]. To date, potential covariates have primarily been identified using a stepwise forward-addition and backward-deletion strategy, which relies on pre-specified P-values as criteria for inclusion and exclusion. However, a significant limitation of this approach is the selection bias inherent in choosing covariates, which may overstate the importance of the selected variables, particularly in cases of high correlation among covariates[4]. Consequently, addressing the issue of correlated covariates has emerged as a pressing challenge that necessitates resolution.

In certain instances, a spurious but strongly correlated covariate may substitute for the true covariate, thereby increasing the risk of a type II error (failing to select a true covariate)[5]. While stringent criteria ($p<0.05$ or $p<0.01$) are implemented to mitigate type I error, a fundamental drawback of the stepwise covariate modeling (SCM) method is its inability to effectively reduce type II error in statistical analyses. Therefore, it is essential to employ effective methodologies to reduce type II error in the covariate model. Power analysis of covariates provides robust algorithms for this purpose. The parametric power estimation (PPE) algorithm uses Monte Carlo simulation principles to compute the unknown non-centrality parameter from a limited dataset[6]. Meanwhile, the Fisher information matrix-based power calculation (FIM-PC) generates the power curve based on the expected information matrix and the Wald test[7,8]. Lastly, the Monte Carlo Mapped Power (MCMP) method offers a simulation-based approach for power calculation using the likelihood ratio test[9,10].

In this study, stepwise covariate modeling (SCM) was employed to estimate the relationships

ISSN 2618-1584 Vol. 7, Issue 4: 47-55, DOI: 10.25236/FMSR.2025.070406

between parameters and covariates. The SCM operates as a stepwise procedure, which is iteratively applied based on likelihood ratio tests to evaluate the advantages of incorporating additional covariates[11]. The parameter-covariate relationships of correlated covariates identified through the SCM method were further examined through power analysis. This power analysis of correlated covariates was conducted using Monolix and Python software, following the principles of the MCMP method.

The power analysis of covariates offers the advantage of mitigating the inclusion of spurious covariates arising from correlated covariates[5]. However, it is important to note that the analysis cannot address correlated true covariates. Tang et al. reported that machine learning (ML) models are capable of assessing significant covariates[12]. Consequently, the ML approach can provide a comprehensive array of algorithms for estimating the impact of covariates on parameters.

ML is increasingly used in the medical field for regression and classification tasks[13]. Multi-classification challenges, for example, are becoming more common in disease diagnosis, high-throughput virtual screening for drug discovery[14], and identifying targeted biomarkers for specific diseases[15]. An artificial neural network model outperforms the population pharmacokinetic model developed via structural equation modeling (SCM) in predictive performance[16]. ML can bridge the gap between big data and pharmacometrics, enabling the development of models for parameter estimation and analysis of correlated true covariates[17].

In this study, we employed the ML method to construct our model. Traditional feature importance metrics indicate the significance of covariates but only highlight important ones. In contrast, the SHAP (Shapley Additive explanation) method clearly shows how each feature affects the ML model's output and is recognized as the most consistent approach for feature attribution[18]. Additionally, permutation importance assesses feature significance by permuting the response vector multiple times to evaluate the relevance of all features[19].

Research on power analysis of correlated covariates has shown it can effectively exclude false covariates. However, a key challenge is selecting the right covariate from correlated true covariates. ML analysis offers an advantage by identifying the most relevant true covariate among these. Therefore, this study aims to investigate whether combining power analysis with ML analysis can enhance the predictive performance of the population pharmacokinetic model.

## 2. Methods

The study was conducted in four distinct phases: the initial screening of covariates utilizing the SCM method, conducting a power analysis for correlated covariates, performing an ML analysis of correlated true covariates, and a comparative assessment of predictive performance. Key information pertaining to each phase has been encapsulated in Figure 1.

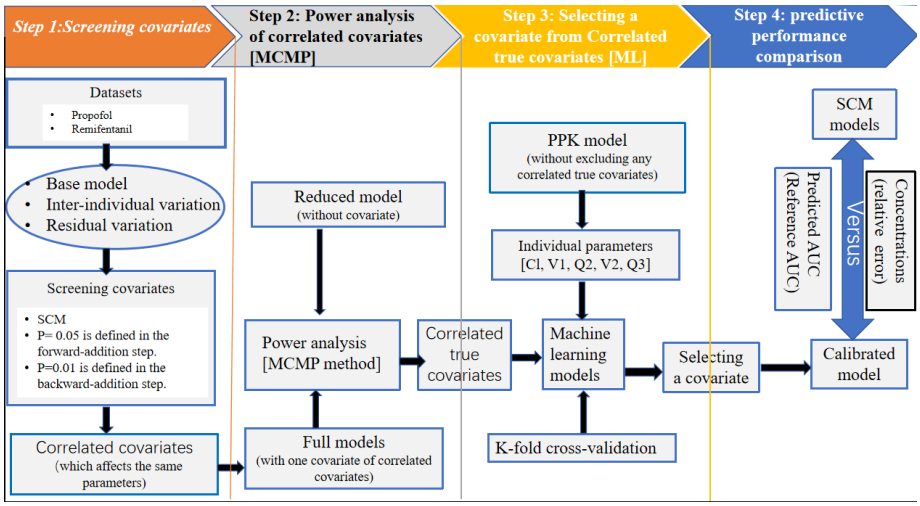


*Fig. 1 The different steps of this study for both two drugs. PK pharmacokinetics, AUC area under the plasma concentration-time curve, NCA non-compartment analysis.*

ISSN 2618-1584 Vol. 7, Issue 4: 47-55, DOI: 10.25236/FMSR.2025.070406

### *2.1 Data*

Pharmacokinetic data were sourced from Open-TCI and categorized into two distinct datasets: one for propofol and another for remifentanil. These pharmacokinetic data have already been utilized in pharmacokinetic studies[20-22]. For remifentanil, ideal body weight (IBW) was calculated by $2.396 \times \exp(0.01863 \times \mathrm{HT})$ in children[23], and by $45.4 + 0.89 \times (\mathrm{HT[cm]} - 152.4) + 4.5(\text{if male})$ in adult[24], for which HT was height of subject. Adjusted body weight (ABW) was calculated by $\mathrm{IBW} + 0.4 * (\mathrm{TBW} - \mathrm{IBW})$, for which TBW was total body weight[24]. For propofol, the IBW, ABW, and TBW were calculated as propofol, and they were also calculated by $2.396 \times \exp(0.01863 \times \mathrm{HT})$ if the HT was less than 152.4. The propofol data from Schnider et al on Open-TCI had 24 subjects, each were measured on two occasions. Therefore, the analysis did treat the separate occasions as new subjects, and only the infusion data were collected. Subject characteristics and previous studies using these data were presented in Supplementary Material.

### *2.2 Population Pharmacokinetic Analysis*

For propofol, the dataset integrated multiple study data, so the sampling methods (artery or vein) and physical condition (healthy or patients) were analyzed as binary covariate. For remifentanil, the physical condition ((healthy or patients) were also analyzed as binary covariate. Covariate analysis followed a standard forward and backward selection process. The categorical covariates were evaluated by $\theta_i = \theta \times e^{cat_i}$, and the continuous covariates were evaluated by $\theta_i = \theta \times e^{con_i}$, where the $\theta_i$ represents the individual parameter value, $\theta$ represents the typical individual parameter value, and the covariate WT was converted by $WT/70$. In this study, during the forward-addition process, only covariates that demonstrated significant changes in the objective function value (OFV) were retained, adhering to a predefined significance threshold of $p<0.05$. Correspondingly, in the backward-addition phase, only covariates that resulted in significant changes in the OFV were retained, with a predefined significance level of $p<0.01$. These criteria were established within the SCM methodology for the purpose of covariate screening.

#### *2.2.1 True covariates selecting*

Several covariates associated with the same parameter, as identified through the SCM method, exhibited a high degree of correlation. Consequently, power analysis was employed to evaluate these correlated covariates for each parameter, with the aim of selecting those covariates that demonstrated significant power values. The power analysis utilized a substantial simulation dataset generated via Simulx software, based on a predefined model and effect size. A total of 1,000 individuals were simulated for this study.

The principle of the MCMP approach involved substituting the overall OFV with the summation of individual objective function values (iOFV). By employing both a comprehensive model (incorporating one covariate) and a reduced model (excluding the covariate), a substantial pool of iOFVfull and iOFVReduced values could be derived from the extensive simulation dataset. The summation of the differences between iOFVfull and iOFVReduced ($\sum \Delta$iOFV) for each individual was utilized in lieu of the overall $\Delta$OFV for statistical inference in the likelihood ratio test (LRT). To delineate the relationship between power and sample size comprehensively, we randomly sampled the $\Delta$iOFV 1,000 times, subsequently calculating the sum of $\Delta$iOFV for comparison against the critical $\chi 2$ OFV for each sample size.

In this study, considering that the covariates included first could affect the inclusion of later covariates, the iOFVfull was calculated by a full model, which included only one covariate at a time. This was to explore which covariate has a more significant impact on the parameter without covariate influence. With the increase of the sample size, if the power value can reach more than 90%, the covariate would be regarded as a true covariate. On the contrary, if the power value was always at a very low level, the covariate would be regarded as a false covariate.

#### *2.2.2 Covariate selecting*

After analyzing the correlated covariates, we found that some parameters still exhibited high correlation. To address this, we employed machine learning techniques to select a single covariate from the correlated true covariates, including gradient boosting regressor (GB), decision tree regressor (DT), extra trees regressor (ET), ada boost regressor (AB), and random forest regressor (RF), developing ML

ISSN 2618-1584 Vol. 7, Issue 4: 47-55, DOI: 10.25236/FMSR.2025.070406

models that utilized k-fold cross-validation.

Kamiński B reported using a directed graph $G = (V, E), E \in V^2$, to construct a decision tree model, for which nodes $V$ represents three disjoint sets (decision, chance, and terminal nodes). And for each edge $e \in E$, the first element ($e_1 \in V$) denote its parent node, the second element ($e_2 \in V$) denote its child node. The decision tree model uses feature to split the parent node into child nodes based on minimizing the mean-square error or Gini Coefficient. We used the covariate and mean-square error to develop decision tree regression model (Figure 2). Besides the decision tree regression model, RF was also considered. The RF is an ensemble learning method consisted of n decision tree. Svetnik V et al[26] reported given a dataset, $D = \{(X_1, Y_1), \ldots, (X_n, Y_n)\}$, bootstrap n sample sets, where $X_i$ ($i = 1, 2, 3, \ldots, n$) is a vector of descriptors, and $Y_i$ ($i = 1, 2, 3, \ldots, n$) is the corresponding labels. For each bootstrap sample grow a decision tree. In regression, outputs of all trees are aggregated to produce one average prediction (average of the individual tree predictions) presented in Figure 2. For the AB algorithm, firstly, imputing the training dataset D whose samples had equal weight to obtain the weak learner 1. Similarly, the error rate ($e_1$) and learner coefficient ($\alpha_1$) were calculated to update weight which was used to weight the training dataset D to train the model again. Secondly, according to the weight calculated by the $e_1$ and $\alpha_1$, the weak learner 2 can be developed, and then the $e_2$ and $\alpha_2$ were calculated to update the weight again which was also used to weight the training dataset D to obtain the weak learner 3. Based on the above process, n weak learner can be developed and integrated into a strong learner (Figure S1). For the GB algorithm, developing the first decision tree ($T_1(x)$) to calculate residual error ($r_{1,i}$), then developing the second decision tree ($T_2(x)$) to fit the $r_{1,i}$, and calculate the $r_{2,i}$. Commonly, the third decision tree ($T_3(x)$) was developed to fit the $r_{2,i}$ and calculate the $r_{3,i}$. Finally, n decision trees were developed to integrate into a strong learner (Figure S2). For the ET algorithm, the algorithm is similar to the RF algorithm, the difference is that the ET algorithm uses same training dataset D to train each decision tree and randomly split the parent node into child nodes.

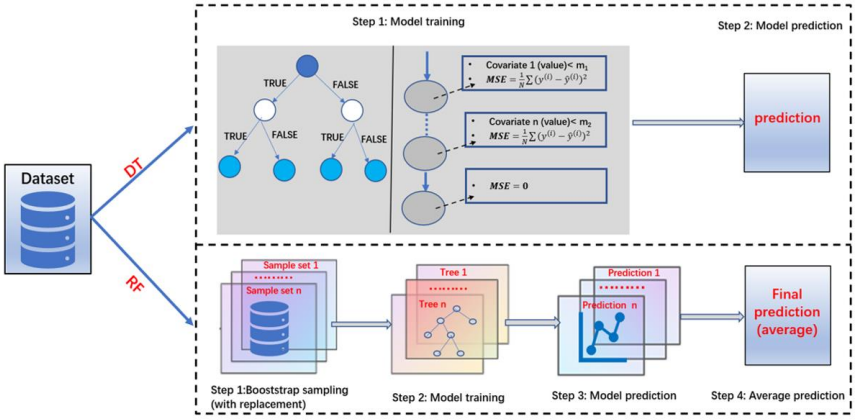


*Fig. 2 Illustrating decision tree (DT) and random forest (RF).*

Permutation importance can identify significant covariates among correlated true covariates, while SHAP values gauge their influence on parameter predictions. This study applied permutation importance to assess covariate significance and calculated SHAP values to determine which covariate had a greater impact. The SHAP method identified a covariate with a stronger influence, which was then validated by permutation importance. Ultimately, the covariate with the highest importance and SHAP value was integrated into the model. Both SHAP value and permutation importance computations were conducted using the same machine learning model.

### *2.2.3 Calibrated model*

Following the screening of covariates using the SCM methodology, a calibrated model was developed through the application of power analysis and machine learning techniques to select a covariate from the correlated covariates. Conversely, the SCM model was formulated by intentionally excluding a covariate that was disregarded in the power analysis and machine learning processes from the set of correlated covariates. The calibrated covariate model demonstrates not only a reduced likelihood of Type I error but also a diminished probability of Type II error.

ISSN 2618-1584 Vol. 7, Issue 4: 47-55, DOI: 10.25236/FMSR.2025.070406

### *2.3 Comparison*

In this section, we conducted a comparative analysis of the predictive performances of the SCM and the calibrated model, focusing on two key aspects.

Firstly, we examined the predicted plasma drug concentrations. The dataset was partitioned into training and test datasets in an 80:20 ratio. The training datasets were utilized to perform power analysis, machine learning analysis, and to develop population pharmacokinetic models. The plasma concentrations predicted from the test datasets were subsequently employed to calculate the mean square error (MSE) and the coefficient of determination ($R^2$), which facilitate the evaluation of model performance.

Secondly, we assessed the area under the plasma concentration-time curve (AUC). For propofol and remifentanil, individual AUC estimates were derived from the population pharmacokinetic models, and non-compartmental analysis (NCA) was applied to compute the AUC for each individual, designated as the "reference AUC." The goodness-of-fit plots of AUCprediction and AUCreference were presented to distinguish the performance of these models. And these ratios of AUCprediction to AUCreference for each individual were presented as scatter plots to compare predictive performance.

Finally, the impact of correlated true covariates on inter-individual Variability (IIV) was estimated to compare which covariate was more able to reduce the IIV.

### *2.4 Software*

The Monolix software (version 2021R1) was employed to develop the population pharmacokinetic model. The machine learning component was implemented using Python (version 3.7) in conjunction with the scikit-learn library for model development. The mean squared errors (MSE) and R-squared ($R^2$) values were computed using R and RStudio (version 4.1.2).

## 3. Results

### *3.1 Population Pharmacokinetic Analysis*

A total of 151 individuals contributed to the analysis, resulting in 3,335 sampling points for propofol. Additionally, 97 individuals were included for remifentanil, yielding 2,167 sample points. The base model utilized for covariate screening was a three-compartment model characterized by zero-order absorption from the depot to the central compartment, combined with linear elimination from the central compartment for both drugs. Exponential models were used to account for IIV in the structural parameters of two drugs. And combined 1 ($y=f+(a+bfc)\varepsilon$) was selected to residual model in propofol; combined 2 ($y=f+\sqrt{a^2+b^2(f^c)^2}\varepsilon$) was selected to residual model in remifentanil.

#### *3.1.1 Screening results of covariates*

For propofol, the covariate related to clearance of the central (CL) was covariate WT screened by the SCM method. The volume of the central (V1) was not included any covariate. The covariates ABW, BSA, and WT were related to the highly perfused compartment (Q2), the covariates ABW and BSA were related to the scarcely perfused compartment (Q3), and the covariate AGE was related to the highly perfused compartment (V2). For remifentanil, the covariates AGE and WT were related to the parameter CL, the covariates ABW and BSA were related to parameter V1, the covariates AGE and WT were related to the covariate Q2, the covariate WT was related to the parameter V2, and the covariate AGE was related to the covariate Q3.

#### *3.1.2 Correlation analysis*

In this study, Kendall's tau-b correlation coefficients were calculated between the main continuous covariates. Table S2 showed the correlation between the main covariates of propofol. The covariates BSA, ABW, and WT showed a high correlation, and the covariates ABW, BSA, and HT also showed a high correlation. So only one covariate should be selected for the parameters Q2 and Q3. Table S3 showed the correlation between the main covariates of remifentanil. The covariates ABW and BSA also had a high correlation, so only one covariate should be selected on the parameter V1.

ISSN 2618-1584 Vol. 7, Issue 4: 47-55, DOI: 10.25236/FMSR.2025.070406

#### *3.1.3 Power analysis of highly correlated covariates*

Figure 3 showed the power value of correlated covariates of propofol and remifentanil. Figures 3A, 3B, and 3C showed the power value of the covariates WT, ABW, and BSA included in the model of the parameter Q2 was more than 90% with the increase of the sample size, so the covariates WT, BSA, and ABW should not be excluded. While with the increase of the sample size, the power of the covariate HT included in the model of the parameter Q3 was less than 20%, so the covariate HT should be excluded. However, the power value of the covariates ABW and BSA included in the model of the parameter Q3 was more than 90%, so they both should be included (Figures 3E and 3F). Similarly, with the increase of the sample size, the power of the covariate BSA and ABW included in the model of the parameter V1 was up to 90%, so the covariates BSA and ABW should be included.

#### *3.1.4 ML analysis of correlated true covariates*

In this section, the correlated true covariates and parameters obtained by including correlated true covariates were used to develop ML models. Five ML methods were used and the SHAP values and permutation importance were calculated based on the same model. The performance of five ML models was presented in Table 2. For the parameters Q2 and V1, the GB model had a better performance.

Figure S3 showed that a covariate was selected by the SHAP method from these correlated true covariates. For the GB model of the parameter Q2 in propofol, the covariate ABW had a more significant impact on model output in original data. Therefore, the covariate ABW should be selected from correlated true covariates, and then be included in the pharmacokinetic model of the parameter Q2. Similarly, based on a more significant impact on the GB model output of the parameter Q3, the covariate ABW should be selected from correlated true covariates, and then be included in the pharmacokinetic model of the parameter Q3 in propofol. For remifentanil, the covariate BSA should be selected as a more significant covariate to be included in the pharmacokinetic model of the parameter V1. Similarly, the DT, ET, and RF models showed the same result of selecting the more significant covariate presented in Supplementary Material. However, the SHAP values can't be obtained by the AB model in this study.

The permutation importance was applied to verify the selected results of the SHAP method, and the result were presented in Supplementary Material. The selected results of the correlated true covariates gained by the permutation importance method were similar to the SHAP method for all machine learning models.

#### *3.1.5 Final model*

Based on the power analysis and ML analysis, for the calibrated model, the included finally covariates were: WT on CL, ABW on Q2, AGE on V2, ABW on Q3 of propofol; AGE and WT on CL, BSA on V1, AGE and WT on Q2, WT on V2, AGE on Q3 of remifentanil.

To prove the performance of the calibrated model, the covariates excluded by the ML analysis from correlated true covariates were used to develop the SCM models. Therefore, the finally included covariates of propofol were: WT on CL, BSA on Q2, AGE on V2, BSA on Q3 (SCM model1); WT on CL, WT on Q2, AGE on V2, BSA on Q3 (SCM model 2). Similarly, the finally included covariates of remifentanil were: AGE and WT on CL, ABW on V1, AGE and WT on Q2, WT on V2, AGE on Q3 (SCM model)

Following the power and ML analysis, the parameters of the calibrated models were evaluated. The parameters of the SCM models were also assessed, as detailed in the same table. Goodness of fit plots and residual scatter plots for the final models of both drugs are included in the Supplementary Material. These plots demonstrate that all final models align well with the observed data, showing no systematic bias. Furthermore, residual scatter plots confirm the absence of systematic bias in the training datasets. The supplementary material similarly presents these plots for the test datasets. Compared to the calibrated remifentanil model, the SCM model shows a more pronounced deviation in the goodness of fit plots for the test dataset.

### *3.2 Comparison of Predictive Performance*

#### *3.2.1 Predictive Plasma Drug Concentration*

The performance of the various models was rigorously evaluated, with the results summarized in Table 3. In the training datasets, the calibrated model for propofol demonstrated superior performance compared to SCM model 1 and SCM model 2. The $R^2$ values derived from population predictions and

ISSN 2618-1584 Vol. 7, Issue 4: 47-55, DOI: 10.25236/FMSR.2025.070406

observations for these models exceeded 80%. Similarly, the calibrated models for remifentanil also outperformed the SCM model, with $R^2$ values surpassing 90%. In the test datasets, both for propofol and remifentanil, the Mean Squared Error (MSE) and $R^2$ values indicated that the calibrated models exhibited enhanced predictive performance relative to the SCM model.

#### *3.2.2 Predictive AUC*

All final models were employed to predict the Area Under the Curve (AUC) for each individual, with the goodness-of-fit results for both drugs illustrated as scatter plots in Figure 4. For propofol, the calibrated model achieved reliable predictions of individual AUC. Likewise, for remifentanil, the calibrated models also yielded accurate predictions of individual AUC.

The AUCprediction/AUCreference was calculated and the results were presented in figure S9. Figure S9A showed that most of the AUCprediction/AUCreference of three propofol models were between 0.8 and 1.2. Figure S9B also showed that most of the AUCprediction/AUCreference of two remifentanil models were between 0.8 and 1.2, but the bias of the SCM model was larger.

### *3.3 The influence of highly correlated covariates on inter-individual Variability*

Compared with the SCM model 1 and model 2, the calibrated model had lower IIV for the parameter Q2 and the same IIV for the parameter Q3 of remifentanil, which means that the covariate ABW on the parameters Q2 and Q3 were beneficial to explaining IIV.

To better understand which covariate of correlated true covariates can better explain IIV, the covariate was changed on the parameters Q2 or Q3 based on the calibrated model. The result was presented in Supplementary Material. For propofol, inclusion of the covariate ABW can decrease the IIV in Q2 by about 13.16%. This was a better covariate than BSA and WT, which were 10.53% and 13.16%, respectively. Inclusion of the covariates BSA and ABW can decrease the IIV in Q3 by about 7.55% and 7.55%, respectively. Therefore, inclusion of the covariate BSA was better than ABW. For remifentanil, inclusion of the covariate BSA and ABW can decrease the IIV in V1 by about 73.4% and 68.5%, so the BSA was a better covariate than ABW.

## 4. Discussion

To enhance the development of a robust population pharmacokinetic model, it is essential to conduct a power analysis of correlated covariates. This analysis aims to identify true correlated covariates while minimizing the risk of including spurious covariates. In the present study, several covariates associated with the same parameter, as identified through the SCM method, exhibited high correlation, necessitating confirmation regarding the inclusion of any false covariates. Consequently, a power analysis was imperative for the selection of true covariates. The MCMP method was employed in this study to facilitate the power analysis of covariates. Notably, the computational time required by the MCMP method to yield relevant power information was less than 1% of that required by the SSE method, while also providing a precise prediction of the relationship between power and sample size[10]. In contrast, the power calculation based on the Fisher information matrix (FIM-PC) involved a more intricate process than that of the MCMP method [10, 27].

For propofol, the covariates ABW, BSA, and HT on the parameter Q3 were selected by the SCM method. However, the power analysis showed that less than 20% power value was maintained of the covariate HT, as the sample size increased. This means that the covariate HT had a low power value to be included in the pharmacokinetic model without the influence of other covariates. Therefore, the covariate HT should be viewed as a false covariate. The covariates ABW, WT, and BSA have been identified as significant covariates for the parameter Q2 through the SCM method. Power analysis supports their classification as true covariates. In the case of remifentanil, the covariates ABW and BSA for the parameter V1 should also be considered true covariates, as the power exceeded 90% with an increasing sample size.

To determine a more effective covariate, it is essential to analyze the correlated true covariates pertaining to the same parameter. The ML method offers a robust criterion for covariate selection. The ML models employed for vancomycin, latamoxef, cefepime, azlocillin, ceftazidime, and amoxicillin have demonstrated substantial predictive accuracy in forecasting individual clearance[12]. Additionally, the application of artificial neural network analysis for predicting the plasma drug concentration of remifentanil has yielded high predictive performance[16].

ISSN 2618-1584 Vol. 7, Issue 4: 47-55, DOI: 10.25236/FMSR.2025.070406

Regarding propofol, the covariates WT, BSA, and ABW concerning the parameter Q2 were identified as true covariates. However, both SHAP and permutation importance methods consistently indicated that the covariate ABW is more critical and exerts a more significant influence on the prediction of parameter Q2 across all ML models. Similarly, the covariate ABW had a more significant impact and importance on the parameter Q3 than the covariate BSA. For remifentanil, the covariates ABW and BSA on the parameter V1 were viewed as true covariates, and the covariate BSA was viewed as more important based on the SHAP and permutation importance methods.

For propofol, the Schnider model[22] was used to calculate the range of parameters according to the range of covariates, which was within the range of 2.5%-97.5% of predicted parameters obtained by the calibrated model. Similarly, for remifentanil, the Minto model[28] was also used to calculate the range of parameters, which was also within the range of 2.5%-97.5% of predicted parameters obtained by the calibrated model. Therefore, the calibrated models had no distortion.

According to MSE and R2, the calibrated models had a better prediction performance than the SCM models. Similarly, the goodness-of-fit results of AUCprediction and AUCreference showed that the calibrated models had a better performance than the SCM models. And these AUCprediction/AUCreference were in the range of 0.8-1.25 for all models, but for remifentanil, several AUCprediction/AUCreference of the SCM model were mor than 1.6 or less than 0.7.

The ML method used to analyze correlated covariates had a drawback in that the significant covariate selected by the SHAP value or permutation importance may be not of clinical interest. However, the drawback may be offset by using the power analysis to exclude false covariate or excluding clinically unrelated covariate in advance.

In conclusion, the calibrated models had a better predictive performance than the SCM models. Similarly, the analysis of IIV also showed that the covariate selected to include in the pharmacokinetic model by all the ML models had more ability to decrease IIV. Therefore, combining the power analysis and ML method to analyze correlated covariates was beneficial to improving predictive performance.

## References

*[1] WORKGROUP E M, MARSHALL S F, BURGHAUS R, et al. Good Practices in Model-Informed Drug Discovery and Development: Practice, Application, and Documentation [J]. CPT Pharmacometrics Syst Pharmacol, 2016, 5(3): 93-122.*

*[2] MARSHALL S, MADABUSHI R, MANOLIS E, et al. Model-Informed Drug Discovery and Development: Current Industry Good Practice and Regulatory Expectations and Future Perspectives [J]. CPT Pharmacometrics Syst Pharmacol, 2019, 8(2): 87-96.*

*[3] ZOU Y, TANG F, NG C M. A Modified Hybrid Wald's Approximation Method for Efficient Covariate Selection in Population Pharmacokinetic Analysis [J]. The AAPS Journal, 2021, 23(2): 37.*

*[4] WAHLBY U, JONSSON E N, KARLSSON M O. Comparison of stepwise covariate model building strategies in population pharmacokinetic-pharmacodynamic analysis [J]. AAPS PharmSci, 2002, 4(4): E27.*

*[5] RIBBING J, JONSSON E N. Power, selection bias and predictive performance of the Population Pharmacokinetic Covariate Model [J]. J Pharmacokinet Pharmacodyn, 2004, 31(2): 109-34.*

*[6] UECKERT S, KARLSSON M O, HOOKER A C. Accelerating Monte Carlo power studies through parametric power estimation [J]. J Pharmacokinet Pharmacodyn, 2016, 43(2): 223-34.*

*[7] RETOUT S, COMETS E, SAMSON A, et al. Design in nonlinear mixed effects models: optimization using the Fedorov-Wynn algorithm and power of the Wald test for binary covariates [J]. Stat Med, 2007, 26(28): 5162-79.*

*[8] UECKERT S, HENNIG S, NYBERG J, et al. Optimizing disease progression study designs for drug effect discrimination [J]. J Pharmacokinet Pharmacodyn, 2013, 40(5): 587-96.*

*[9] KLOPROGGE F, SIMPSON J A, DAY N P, et al. Statistical power calculations for mixed pharmacokinetic study designs using a population approach [J]. Aaps j, 2014, 16(5): 1110-8.*

*[10] VONG C, BERGSTRAND M, NYBERG J, et al. Rapid sample size calculations for a defined likelihood ratio test-based power in mixed-effects models [J]. AAPS J, 2012, 14(2): 176-86.*

*[11] AYRAL G, SI ABDALLAH J-F, MAGNARD C, et al. A novel method based on unbiased correlations tests for covariate selection in nonlinear mixed effects models: The COSSAC approach [J]. CPT: Pharmacometrics & Systems Pharmacology, 2021, 10(4): 318-29.*

*[12] TANG B H, GUAN Z, ALLEGAERT K, et al. Drug Clearance in Neonates: A Combination of Population Pharmacokinetic Modelling and Machine Learning Approaches to Improve Individual*

ISSN 2618-1584 Vol. 7, Issue 4: 47-55, DOI: 10.25236/FMSR.2025.070406

*Prediction [J]. Clin Pharmacokinet, 2021, 60(11): 1435-48.*
*[13] LIANG H, TSUI B Y, NI H, et al. Evaluation and accurate diagnoses of pediatric diseases using artificial intelligence [J]. Nat Med, 2019, 25(3): 433-8.*
*[14] KHANDELWAL A, KRASOWSKI M D, RESCHLY E J, et al. Machine learning methods and docking for predicting human pregnane X receptor activation [J]. Chem Res Toxicol, 2008, 21(7): 1457-67.*
*[15] JEON J, NIM S, TEYRA J, et al. A systematic approach to identify novel cancer drug targets using machine learning, inhibitor design and high-throughput screening [J]. Genome Med, 2014, 6(7): 57.*
*[16] KANG S H, POYNTON M R, KIM K M, et al. Population pharmacokinetic and pharmacodynamic models of remifentanil in healthy volunteers using artificial neural network analysis [J]. Br J Clin Pharmacol, 2007, 64(1): 3-13.*
*[17] MCCOMB M, BIES R, RAMANATHAN M. Machine learning in pharmacometrics: Opportunities and challenges [J]. Br J Clin Pharmacol, 2022, 88(4): 1482-99.*
*[18] YANG J. Fast TreeSHAP: Accelerating SHAP Value Computation for Trees [J]. 2021.*
*[19] ALTMANN A, TOLOSI L, SANDER O, et al. Permutation importance: a corrected feature importance measure [J]. Bioinformatics, 2010, 26(10): 1340-7.*
*[20] BIDESHWAR K. KATARIA S A V H F E A. The Pharmacokinetics of Propofol in Children Using Three Different Data Analysis Approaches [J]. 1994, 80: 104–22.*
*[21] ROSS A K, DAVIS P J, DEAR GD G L, et al. Pharmacokinetics of remifentanil in anesthetized pediatric patients undergoing elective surgery or diagnostic procedures [J]. Anesth Analg, 2001, 93(6): 1393-401, table of contents.*
*[22] SCHNIDER T W, MINTO C F, GAMBUS P L, et al. The influence of method of administration and covariates on the pharmacokinetics of propofol in adult volunteers [J]. Anesthesiology, 1998, 88(5): 1170-82.*
*[23] TRAUB S L, KICHEN L. Estimating ideal body mass in children [J]. Am J Hosp Pharm, 1983, 40(1): 107-10.*
*[24] SAHINOVIC M M, STRUYS M, ABSALOM A R. Clinical Pharmacokinetics and Pharmacodynamics of Propofol [J]. Clin Pharmacokinet, 2018, 57(12): 1539-58.*
*[25] KAMINSKI B, JAKUBCZYK M, SZUFEL P. A framework for sensitivity analysis of decision trees [J]. Cent Eur J Oper Res, 2018, 26(1): 135-59.*
*[26] SVETNIK V, LIAW A, TONG C, et al. Random Forest: A Classification and Regression Tool for Compound Classification and QSAR Modeling [J]. Journal of Chemical Information and Computer Sciences, 2003, 43(6): 1947-58.*
*[27] RETOUT S D S, MENTRE F. Development and implementation of the population Fisher [J]. Comput Methods Programs Biomed, 2001, 65(2): 141–51.*
*[28] MINTO CHARLES F, SCHNIDER THOMAS W, EGAN TALMAGE D, et al. Influence of Age and Gender on the Pharmacokinetics and Pharmacodynamics of Remifentanil: I. Model Development [J]. Anesthesiology, 1997, 86(1): 10-23.*