Limited Sampling Strategies to Estimate the Area under the Concentration-time Curve
Biases and a Proposed More Accurate Method

H. Tsuruta1; M. Fukumoto2; L. Bax3; A. Kohno4; Y. Morishita4

1Department of Medical Informatics, School of Allied Health Sciences, Kitasato University, Kanagawa, Japan;
2Division of Toxicology, Center for Clinical Pharmacy and Clinical Sciences, School of Pharmaceutical Sciences, Kitasato University, Tokyo, Japan;
3Pharsight Corporation, Sunnyvale, CA, USA;
4Department of Hematology and Oncology, Japan Agricultural Cooperatives, Aichi Konan Kosei Hospital, Aichi, Japan

Keywords
Limited sampling strategy, AUC, busulfan, nonlinear curve fitting, compartment model

Summary
Background: Over 100 limited sampling strategies (LSSs) have been proposed to reduce the number of blood samples necessary to estimate the area under the concentration-time curve (AUC). The conditions under which these strategies succeed or fail remain to be clarified.

Objectives: We investigated the accuracy of existing LSSs both theoretically and numerically by Monte Carlo simulation. We also proposed two new methods for more accurate AUC estimations.

Methods: We evaluated the following existing methods theoretically: i) nonlinear curve fitting algorithm (NLF), ii) the trapezium rule with exponential curve approximation (TZE), and iii) multiple linear regression (MLR). Taking busulfan (BU) as a test drug, we generated a set of theoretical concentration-time curves based on the identified distribution of pharmacokinetic parameters of BU and re-evaluated the existing LSSs using these virtual validation profiles. Based on the evaluation results, we proposed a new estimation method in which the most likely curve was selected from a set of pre-generated theoretical concentration-time curves.

Results: Our evaluation, based on clinical profiles and a virtual validation set, revealed: i) NLF sometimes overestimated the absorption rate constant Ka, ii) TZE overestimated AUC over 280% when Ka is small, and iii) MLR underestimated AUC over 30% when the elimination rate constant Ke is small. These results were consistent with our mathematical evaluations for these methods. In contrast, our two new methods had little bias and good precision.

Conclusions: Our investigation revealed that existing LSSs induce different but specific biases in the estimation of AUC. Our two new LSSs, a modified TZE and one using model concentration-time curves, provided accurate and precise estimations of AUC.

1. Introduction

Many anticancer agents, such as anthracyclines, busulfan (BU), carboplatin, docetaxel, etoposide, fluorouracil, irinotecan, and vinorelbine, as well as immunosuppressants, such as cyclosporin, mycophenolic acid, sirolimus, and tacrolimus, have narrow therapeutic windows between efficacy and toxicity. Moreover, their pharmacokinetics, i.e., the effect of an individual patient’s body on the drug, exhibit wide inter-patient variability. Therefore, dosing that takes account of individual differences in pharmacokinetics is necessary to avoid systemic toxicities and secure the target concentrations required for efficacy [1, 2].

An example of this may be seen in conditioning regimens preceding allogeneic or autologous bone marrow transplantation using BU, if the systemic exposure of BU is overtherapeutic, then the regimen-related toxicity, like in a veno-occlusive disease, may be lethal to the patient even if the transplantation is successful [3]. On the other hand, if the exposure is insufficient, then the undertherapy leads to failure of the transplantation.

In addition, the pharmacokinetics of BU varies by several factors, including the patient’s age, underlying disease, hepatic function, concomitant administration of other drugs, and daily therapeutic and bodily variations. BU pharmacokinetic studies suggest that the agent should not be dosed simply by body weight and/or body surface area, but that a pharmacokinetic-guided dosing is necessary to maintain tar-
get concentration and to minimize the incidence of dose-related toxicities [4–7].

The pharmacokinetic parameter that best assesses the total drug exposure to the body in most drugs is the area under the plasma drug concentration-time curve (AUC). However, estimation of AUC requires multiple blood samples, often more than 10 times over a single dosing interval, which is not feasible in usual clinical settings. Therefore, various limited sampling strategies (LSSs) have been proposed to reduce the number of blood samples necessary to estimate the AUC and other pharmacokinetic indices [4, 5, 7–13].

In the development of an LSS, pharmacokinetic profile data with multiple samples are obtained from patients and usually divided into the index set and the validation set. The data from the index set are used to develop an LSS, a formula to predict AUC, and the validation set is used for cross validation to assess accuracy and precision of the developed LSS. Many LSSs have been developed, with this procedure, for the various anticancer agents and immunosuppressants mentioned above; and many investigators of multiple studies have claimed that LSSs showed sufficient prediction performance in cross validation.

1.1 Background

We had been studying BU pharmacokinetics in Japanese patients and proposed an AUC estimation formula. In this study, we first performed a clinical study to assess the BU pharmacokinetics. Based on the pharmacokinetic analysis of the samples in this study, we developed an LSS (Fukumoto’s LSS) to estimate the AUC from four blood samples (0.5, 1.0, 2.0, and 6.0 hours):

\[
AUC_{\text{pred}} = 0.5 \cdot C(0.5) + 0.75 \cdot C(1) + 2.5 \cdot C(2) + 2.0 \cdot C(6) + C(6)/K_{\text{pred}},
\]

(1)

where

\[
K_{\text{pred}} = (\log(C(2) – \log(C(6)))/4.0
\]

(2)

and \(C(t)\) is blood BU concentration at time \(t\) (hour) after dose (►Appendix 1). Though the patients’ AUCs varied in a wide range from 2,409 to 12,030 (μg · hour/L), \(AUC_{\text{pred}}\) showed good agreement with the AUC based on 11 blood samples (\(r = 0.94\)).

Because we did not have a sufficient number of sample profiles for cross validation, we subsequently generated 300 various BU concentration-time curves based on the identified pharmacokinetic model. We then compared the estimated AUC for these theoretical time curves with each true AUC.

The results showed good agreement in 289 of 300 simulated cases; however, our LSS overestimated the AUC or took a negative value when the assumption that the absorption phase had ended before the last two sampling points was not valid. These results present the question, are the LSSs truly accurate even though they gave excellent estimations in cross validations.

1.2 Objectives

We first investigated and classified major methods to develop an LSS. We subsequently examined the existing LSSs from a theoretical point of view and clarified under which conditions the estimations failed. Taking BU as a test drug, we then verified these results by Monte Carlo simulation based on the virtual profiles in which the change of BU blood concentration follows the identified pharmacokinetic model and the identified distribution of within-patient errors. We finally investigated two new methods for accurate AUC estimations.

2. Methods

2.1 Classification and Theoretical Evaluation of the Methods Used to Estimate AUC

We searched for articles proposing LSSs using the PubMed interface with the keywords: “limited sampling strategy”, “limited sampling model”, and “limited sampling method”. We subsequently classified them based on the targets of the estimations and methods for developing LSSs. We then theoretically evaluated the accuracy of the major methods and clarified when the estimations would most likely fail.

2.2 Clinical Study and Analysis of BU Pharmacokinetics

BU is an alkylating agent of the methyl sulfonyl group with a strong cytotoxic reaction discovered by Haddow et al. [14]. In combination with cyclophosphamide, BU is widely used in conditioning regimes before allogeneic or autologous bone marrow transplantation instead of total body irradiation [15]. To evaluate LSSs by a validation set based on an actual drug, we studied the pharmacokinetics of BU. Following the clinical study plan approved by ethical committees of our hospitals, the blood samples were taken 11 times after the oral BU administration (1 mg/kg) in nine patients who had undergone bone marrow transplantation. Blood samples were also taken from the other 69 patients five times after the administration. These samples were then analyzed by the negative ion chemical ionization mode-GC/MS method, which we had developed earlier [16].

To identify which pharmacokinetic model best fits our observation, we analyzed the first 9 profiles with 11 blood samples using Akaike Information Criteria (AIC) [17] with WinNonlin [18]. We also estimated the distribution of the within-patient errors based on these nine cases.

Applying the identified model, we then analyzed all of the 78 observations by pharmacokinetic population analysis using NONMEM [19] to obtain the distribution of the following pharmacokinetic population parameters: first-order absorption rate constant \(K_a\), elimination rate constant \(K_e\), and volume of distribution \(V_d\).

2.3 Evaluation of LSSs Using a Virtual Validation Set with Within-patient Errors

In addition to the theoretical evaluation of the accuracy of LSSs, we also evaluated the accuracy and precision of each LSS exhaustively with the virtual profiles in which the change of BU blood concentration follows the identified pharmacokinetic model and...
the identified distribution of within-patient errors. For this purpose, we made the following Monte Carlo simulation, which we had developed to calculate the confidence interval for the AUC estimate by an LSS [20].

1. We identified a pharmacokinetic model that best describes the measured data and obtained the distribution of pharmacokinetic parameters by the pharmacokinetic population analysis described above in 2.2.

2. We then generated 208 BU theoretical concentration-time curves based on the identified pharmacokinetic model and by varying its parameters so as to cover all 78 cases we had observed with sufficient small steps for the evaluation of LSSs. We call this set of curves the “virtual validation set.”

3. We estimated the distribution of the within-patient errors.

4. For each of the concentration-time curves in our virtual validation set, we added random errors based on the identified distribution of the within-patient errors. We then estimated the AUC from the combination of generated concentrations using each LSS and compared the result to the true AUC of the original concentration-time curve before the within-patient errors were added. This process simulates the estimation by an LSS for an actual patient’s profile, which includes within-patient errors such as observation errors.

5. We repeated this procedure 5,000 times for each of the 208 concentration-time curves and obtained the distribution of estimated AUCs for each profile, from which we evaluated the accuracy and precision of the LSS.

6. The ratios of the mean of the estimated AUCs to the true AUC of each curve were depicted in 3D graphs so that the changes of accuracy could easily be observed as a function of the two rate constants $K_a$ and $K_e$.

All of the programs for the above-described Monte Carlo simulation were written in Standard Pascal using Embarcadero Delphi XE [21].

Table 1 shows the number of studies proposing new limited sampling strategies classified by targets and methods.

<table>
<thead>
<tr>
<th>Target</th>
<th>Method for developing LSS</th>
<th>MLR</th>
<th>TZE</th>
<th>Bayes</th>
<th>NNW</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUC</td>
<td></td>
<td>86</td>
<td>6</td>
<td>18</td>
<td>1</td>
</tr>
<tr>
<td>CL</td>
<td></td>
<td>4</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$C_{max}$</td>
<td></td>
<td>11</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PK</td>
<td></td>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>101</td>
<td>6</td>
<td>36</td>
<td>1</td>
</tr>
</tbody>
</table>

Note: MLR, multiple linear regression; TZE, trapezium rule with exponential curve approximation; NNW, neural network; AUC, area under the concentration curve; CL, clearance; $C_{max}$, maximum concentration; PK, other pharmacokinetic parameters

1. We generated a sufficient number of BU concentration-time curves assuming the identified pharmacokinetic model holds and varying its parameters so as to cover the range of each BU pharmacokinetic parameter. The step sizes were selected so that the estimated pharmacokinetic parameters had two significant digits.

2. When patient data were given, we searched for a curve that best fit the observed data from this set of model cases using the weighted least squares method and regarded its AUC as the patient’s AUC.

In the present study, we evaluated this new scheme of AUC estimation using our virtual validation set and the Monte Carlo simulation of within-patient errors described above in 2.3.

3. Results

3.1 Methods Used to Develop LSSs

LSSs have been developed and validated for various agents, such as BU, vinblastine, carboplatin, etoposide, fluorouracil, vinorelbine, etc. We found more than 150 studies each of which proposes or evaluates LSSs. We also found reviews [1, 2, 22, 23] and evaluation studies [24–26] on LSSs. These LSS targets are AUC, clearance, maximum concentration, and other pharmacokinetic indices. Table 1 shows the number of studies proposing new LSSs.

The standard method for the estimation of AUC is nonlinear curve fitting, which is often used when there are a sufficient number of observations. The major methods for cases with a limited number of observations were the trapezium rule with exponential curve approximation (TZE), multiple linear regression (MLR) or simple linear regression, and the Bayesian method. The linear trapezium rule without exponential curve approximation is also used to estimate the AUC in a finite time interval [12]. We also found an article proposing a neural network method [26].

© Schattauer 2012
3.2 Results of Clinical Study of BU Pharmacokinetics

AIC showed that the one-compartment model with lag time and first-order absorption was used to describe the data. Body weight was incorporated as a random effect for the volume of distribution. Results of the standard two-stage method for busulfan from 9 cases with 11 blood samples.

Table 2

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ka (hour)</td>
<td>1.72</td>
<td>0.76</td>
<td>0.334</td>
<td>5.616</td>
</tr>
<tr>
<td>Ke (hour)</td>
<td>0.267</td>
<td>0.043</td>
<td>0.195</td>
<td>0.423</td>
</tr>
</tbody>
</table>

\[
A = \frac{K_a \cdot F \cdot D}{Vd \cdot (K_a - K_e)},
\]

where

- \(K_e\) = elimination rate constant,
- \(K_a\) = absorption rate constant,
- \(\tau\) = time delay of the absorption phase,
- \(A\) = integrated concentration,
- \(Vd\) = volume of distribution,
- \(F\) = bio-availability,
- \(D\) = amount of dose

and \(\exp()\) is the exponential function \([27]\).

Integrating \(C(t)\), we obtain

\[
\text{AUC} = \int_0^\infty C(t)dt = \frac{F \cdot D}{K_a \cdot Vd} \cdot \frac{K_a - K_e}{K_a \cdot Ke}
\]

Applying Equation 3 as a pharmacokinetic model, we performed a population pharmacokinetic analysis for 78 cases (379 samples). Table 2a shows the means and standard deviations (SDs) of the two rate constants \(K_e\) and \(K_a\). These results were accordant with the analysis for nine cases (11 samples) using the traditional two-stage method, in which each patient’s data were analyzed individually (Stage 1); and, subsequently, the pharmacokinetic parameters were summarized by calculating the mean and SD (Stage 2) \([28]\).

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{example.png}
\caption{An example of clinical profile with 11 blood samples. Observed values are depicted by circles. The concentration-time curve is estimated assuming that one-compartment model with a lag time and first-order absorption describes the data.}
\end{figure}

Fig. 1 is an example of a clinical profile with eleven blood samples. The difference between the observed concentration \(x(t)\) (\(\mu g/L\)) (circles) and the theoretical concentration-time curve \(C(t)\) estimated by nonlinear curve fitting denotes within-patient error. We investigated the within-patient errors of BU in nine cases (83 observations), and found that the log-scale errors \(\epsilon\) were independent of time and nearly followed a normal distribution:

\[
\log(x(t)) \sim \log(C(t)) + \epsilon, \quad \epsilon \sim \mathcal{N}(0, 0.2014^2)
\]

We adopted this log-scale error model and used the empirical distribution of the within-patient errors (\(\text{Appendix 3}\)) to add within-patient errors in our Monte Carlo simulation.

3.3 Generation of a Virtual Validation Set

Based on the distribution of pharmacokinetic parameters in \(\text{Table 2}\), we varied the pharmacokinetic parameters as follows.

\[
K_a = 0.3 \text{ to } 6.0 \text{ (/hour)}
\]

\[
K_e = 0.15 \text{ to } 0.6 \text{ (/hour)}
\]

The value of \(F \cdot D / Vd\) was set to 1,000/0.58 to adjust the scales of concentration and within-patient error. The time lag \(\tau\) was set to zero to investigate the simple cases in the present study. Setting sufficiently small steps to \(K_a\) and \(K_e\) in \(\text{Table 3a}\), we then obtained 208 combinations of pharmacokinetic parameters, each of which corresponds to a patient’s pharmacokinetic state and yields a unique concentration-time curve by \(\text{Equation 3}\). We call this set of concentration-time curves the ‘virtual validation set’. Because the virtual validation set covers the pharmacokinetic states of the 78 patients in the present study, and includes no extraordinary cases, any LSS should accurately estimate the AUC of each virtual profile.

We also generated a more detailed set of 2,668 concentration-time curves for our concentration curve database search method based on the steps in \(\text{Table 3b}\).
3.4 Evaluation and Comparison of Methods Used to Estimate AUC

3.4.1 The Nonlinear Curve Fitting Method

Nonlinear curve fitting (NLF) is a method used to find a curve that best fits the observed data in which the weighted sum of squares of errors for the concentration estimation is minimized by varying the parameters of a preselected theoretical concentration-time curve. This method requires an iterative reweighting process because not the observed concentration but the model predicted concentration is often used as the weight for error summation.

This method is reliable if we have a sufficient number of observations, and it has been regarded as a standard method for evaluating a patient’s pharmacokinetic state. Many studies proposing a new LSS adopted the results estimated by NLF as a golden standard for the evaluation of a new estimation method.

However, the results of NLF may deviate when the number of samples is insufficient. Figure 2 shows what happens when the Gauss-Newton method is used, a typical method for NLF, in 69 cases with five samples. Though the standard two-stage method applied to nine cases with 11 samples resulted in very similar and reasonable results with a log-normal like distribution for the first-order absorption rate constant Ka, NLF resulted in impossible values over 20 (/hour) for seven cases. This is because NLF tries to fulfill its minimum error criterion by overfitting the model even to the observation errors and is the reason the NLF is not applicable to cases with a limited number of blood samples, and we need an LSS for these cases.

3.4.2 The Trapezium Rule with Exponential Curve Approximation

The trapezium rule is a simple method of numerical integration that works by approximating a segment under the curve as a trapezoid and summing these areas. TZE is based on the application of the trapezium rule at the absorption phase and the exponential approximation of the concentration-time curve at the elimination phase. Here is a typical three-sample formula proposed by Chattergoon et al. [8]:

\[ AUC = 0.75 \cdot C(1) + 0.25 \cdot C(1.5) + C(1.5)/k \]  

where

\[ k = (\log(C(1.5)) - \log(C(6)))/4.5 \]

and \( C(t) \) is the patient’s BU plasma concentration obtained after \( t \) hours of BU dose. This formula can theoretically be derived as follows (Fig. 3).

\[ AUC = A_1 + A_2 + A_3 \]

\[ = \frac{1}{2} \cdot 1.0 \cdot C(1) + \frac{1}{2} \cdot 0.5 \cdot (C(1) + C(1.5)) + \frac{C(1.5)}{e^{1.5k}} \int_{1.5}^{\infty} e^{-kt} dt \]

\[ = 0.75 \cdot C(1) + 0.25 \cdot C(1.5) + C(1.5)/k \]

where

\[ k = \frac{- \log(C(60)) - \log(C(1.5))}{6.0 - 1.5} = \frac{-1}{4.5} \log \left( \frac{C(60)}{C(1.5)} \right) \]

\( k \) is an estimate of Ke and can be determined so that the curve \( A e^{-kt} \) passes both \((1.5, C(1.5))\) and \((6.0, C(6.0))\).

In Equation 5, the coefficients 0.75 and 0.25 are derived mathematically from the trapezium rule and the rate constant \( k \) is determined directly from the observation. Therefore, TZE can be constructed not statistically but mathematically only if the optimal sampling times are selected.

To evaluate the accuracy of TZE, consider the case where the change of BU concentration follows the identified pharmacokinetic model, Equation 3. The discussion on estimation error in TZE can be divided into two parts, \( A_1 + A_2 \) and \( A_3 \). The error in the early stage is due to the trapezium rule and yields little error if the sampling times are selected appropriately.

On the contrary, the error caused by exponential approximation is critical in some cases. If the absorption phase is truly ended before the second sampling time (1.5 hour in this example) and the term \( -A_1 \cdot \exp [-\text{Ke}(t - \tau)] \) is negligible, the elimination phase can be approximated by an exponential curve \( A \cdot \exp [-\text{Ke}(t - \tau)] \) and

![Fig. 2](https://example.com/fig2.png)

**Fig. 2** Histogram of the estimated value of the first-order absorption rate constant Ka for 69 cases with five samples using the Gauss-Newton method. It resulted in impossible values over 20 (/hour) for seven cases.

| Table 3 Pharmacokinetic parameter values for generation of virtual profile set | a) for validation of LSSs, b) for candidates of concentration-time curves for our concentration curve database search method. |
|---|---|---|---|
| a) | Ka (hour) | Max | \( \Delta \) |
| | 0.30 | 6.00 | 0.35–0.5 |
| b) | Ke (hour) | Max | \( \Delta \) |
| | 0.15 | 0.60 | 0.03 |

© Schattauer 2012

Methods Inf Med 5/2012
Ke can be estimated from $C(1.5)$ and $C(6.0)$.

Because AUC of the elimination phase is approximated by $C(1.5)/k_e$ in Equation 5, the estimator can be appraised by evaluating the accuracy of $k_e$ as an estimation of the elimination rate constant $K_e$.

To evaluate the error caused by the slow absorption, we compared the estimated elimination rate constant $K_e$ with the true value for various values of the absorption rate constant $K_a$. Figure 4 shows the results: if $K_a$ is fast enough e.g., 2.0 (/hour), the estimation is accurate. On the other hand, $K_e$ is underestimated lower than half of the true value when $K_a$ is slow (0.5/hour). If $K_a$ and $K_e$ are both slow, as $K_a = 0.5$ and $K_e = 0.2$, the estimation fails and causes a more than four times overestimation of the area $A_3$. Because $K_e$ ranged from 0.195 to 0.423 (/hour) and $K_a$ ranged from 0.334 to 5.62 (/hour) in our patients, this overestimation will actually happen.

We observed this phenomenon from another point of view. Because $K_e$ was estimated uniquely by the ratio $C(6)/C(1.5)$ in this exponential approximation scheme, we compared the ratio $C(6)/C(1.5)$ and the estimated $K_e$. The results shown in Table 4 imply that the larger ratio of 0.5 will yield an unrealistic value for $K_e$. This provides a hint of when particular care should be given regarding the application of TZE to actual data.

We evaluated Chattergoon’s LSS by the use of our virtual validation set and Monte Carlo simulation of within-patient error. Figure 5a depicts the results, in which the ratio of the average of 5,000 estimated AUC values to the true AUC are depicted as a function of two rate constants $K_a$ and $K_e$. If the ratio is 1.0, the estimation is accurate and there is no bias. When $K_a \geq 1.5$ (/hour) and $K_e \geq 0.18$ (/hour), the maximum error was 20.2%. On the contrary, when $K_a \leq 1.0$, the TZE overestimated AUC by about four times in the worst case. It took negative values in eight cases, which were replaced by zero in the figure.

We made the same validation for another LSS based on TZE. Figure 5b shows the results for our previous LSS (Fukumoto’s LSS) which is defined by Equations 1 and 2. The shape of the bias surface was very similar to the one for Chattergoon’s LSS (Fig. 5a) and estimation also failed when $K_a \leq 1.0$. These results are consistent with our theoretical evaluation made in the beginning of this section.

### 3.4.3 Multiple Linear Regression

Stepwise MLR (multiple linear regression) is frequently used to construct LSSs for many anticancer agents, immunosuppressants, and other drugs. Using the PubMed database, we found more than 80 articles proposing an original LSS using MLR to estimate the AUC of these agents (Table 1).

If we have four observations $C(t_1), C(t_2), C(t_3),$ and $C(t_4),$ and can assume that the concentration changes strictly follow
the identified pharmacokinetic model, we can identify four pharmacokinetic parameters: $K_a$, $K_e$, $\tau$, and $A$ by solving simultaneous equations obtained from $\text{Equation 3}$. Then AUC can be calculated from $\text{Equation 4}$.

If we can assume that $\tau = 0$, we can calculate AUC from three observations. Because $\text{Equation 3}$ is highly nonlinear, the estimation process of AUC becomes nonlinear, and any LSS using MLR is nothing but a first-order approximation of a nonlinear procedure and inevitably accompanies bias to some extent.

To investigate under what condition the bias happens, compare the estimation formula of MLR:

$$ AUC = \sum_{i=1}^{n} a_i \cdot C(t_i) $$

with a TZE:

$$ AUC = \sum_{i=1}^{n} b_i \cdot C(t_i) + \frac{C(t_n)}{k} $$

$\text{Equation 7}$ approximates the area $A_n$ after $t_n$,

$$ A_n = \int_{t_0}^{t_n} C(t)dt \equiv \int_{t_0}^{t_n} Ae^{-kt}dt = \frac{C(t_n)}{k}, $$

and $A_n$ varies depending on the value of $K_e$. Though the last term $a_n \cdot C(t_n)$ in $\text{Equation 6}$ is thought to contribute to estimate the area around $t_n$, it has the same value for the same concentration $C(t_n)$ regardless of the value of $K_e$. Therefore, $a_n \cdot C(t_n)$ will underestimate $A_n$ if $K_e$ is small, and subsequently $\text{Equation 6}$ will underestimate AUC and vice versa.

To verify this assumption, we evaluated the typical MLR for BU by Hassan et al. [5]:

$$ AUC = 1.78 \cdot C(1) + 1.44 \cdot C(3) + 7.35 \cdot C(6) $$

using our virtual validation set and Monte Carlo simulation of within-patient errors. $\text{Figure 5c}$ shows the results. The means and SDs for the ratios of the average of 5,000 estimated AUCs to the true AUC were $0.984 \pm 0.057$ ($\text{Table 5}$). Though there were no large errors, this MLR underestimated the AUC from $8.6\%$ to $30.3\%$ when $K_e$ was less than or equal to $0.18$ (/hour).

### 3.5 Two New Methods for Accurate Estimation of AUC

To solve the bias problems revealed in the above results, we proposed two new methods for more accurate AUC estimation.

#### 3.5.1 Improvement of Our Previous LSS and Its Evaluation

Because the main cause of the estimation error in our TZE was due to the original definition of the elimination rate constant $K_{pred}$ we replaced $\text{Equation 2}$ with

$$ K_{pred} = \max(0.15, (\log(C(2) - \log(C(6))))/4.0) $$

to prevent $K_{pred}$ reaching an unrealistic small value because of slow absorption or observation error.

$\text{Figure 5d}$ shows the results of our newly revised method using the virtual validation set and Monte Carlo simulation of within-patient errors. The mean and SD...
for the ratios of the average of the estimated values to the true value were 1.063 ± 0.050 (Table 5). In addition to the accuracy being improved, the SD was improved dramatically from 0.416 to 0.050. It should be emphasized that the range was (0.820, 1.196), and there was no case in which the error was over 20%.

### 3.5.2 Evaluation of the Concentration Curve Database Search Method

To solve the bias problem, we proposed the concentration curve database search method described in 2.4.2. We first generated a detailed set of 2,668 concentration-time curves based on the identified pharmacokinetic model and shorter steps in Table 3b. These cases are thought to cover possible BU concentration changes except for the within-patient error. We subsequently searched for a curve that best approximates the given data from these pre-generated 2,668 profiles using the weighted least squares method.

Figure 5e shows the results of the evaluations using our virtual validation set and Monte Carlo simulation of within-patient errors. The mean and SD for the ratios of the average of the estimated values to the true value were 1.002 ± 0.026. The range was (0.938, 1.055), and there was no case in which the error was over 6.2% (Table 5).

### 3.6 Precision of LSSs

In addition to the accuracy of various limited sampling strategies, we were able to evaluate the precision of an estimate of AUC simultaneously. The right part of Table 5 is the summary for the coefficient of variation for 208 profiles, each of which was calculated from 1,000 Monte Carlo simulations of within-patient errors. Our two new methods showed better precision compared with conventional LSSs. It should be noted that Hassan’s MLR had over 30% bias in some cases, although that precision was close to that afforded with our new methods.

### 4. Discussion

In the present study, we evaluated various LSSs by the use of a virtual validation set of BU concentration-time curves and Monte Carlo simulation of within-patient errors. Whether the results are reliable or not will depend on the following questions:

1. Was the pharmacokinetic model we assumed correct?
2. Were the values of pharmacokinetic parameters we used for validation correct?
3. Were the error model and the parameter of the within-patient error correct?

Regarding the model correctness, AIC showed that the one-compartment model with absorption from the gut compartment best fits the observed data for all nine cases. Hassan et al. [5], Chattergoon et al. [8], and Sandstrom et al. [11] also reported that this model fits the observed data best and used this model for population pharmacokinetic analysis.

In addition to the population analysis using all the cases, we separately investigated 69 cases with five samples. The results were similar to those in Table 2b for nine cases with 11 samples that were calculated using the two-stage method. To compare with another pharmacokinetic study of BU, we calculated the ranges of the two rate constants from Hassan et al.’s study for 20 patients [5]. The results were: Ka ranged from 0.50 to 6.93 (1/hour); and Ke ranged from 0.167 to 0.825 (1/hour). Because these three analyses used independent sets of profiles and resulted in similar parameter values, the ranges of pharmacokinetic parameters in Table 2a will be reliable except that there remains a possibility that the maximum value for Ke is as much as 0.8 (1/hour) larger.

As for the last error model problem, we investigated 83 within-patient errors in nine cases with 11 samples and found the log-scale errors were nearly normally distributed and independent of time (Appendix 3). Moreover, the log-normal error model is the most common model for plasma concentration-time changes of various agents. Hence, to the best of our knowledge, we believe that the error model we adopted is adequate and provides the most accurate results.

We, therefore, concluded that each step in our Monte Carlo simulation most accurately emulated the application process of an LSS to actual observations with within-patient errors for an actual patient who has various and different pharmacokinetic parameter values. To assess the accuracy of LSSs, we had previously developed another simpler method in which we evaluated an LSS with the same set of theoretical concentration-time curves as in the present study [20]. The difference was that the previous model did not incorporate within-patient errors at all and only tested average cases. Though the method was able to point out similar problems in accuracy of LSSs, it could not investigate the precision of LSSs.
One of the advantages of our new evaluation method, presented in this study, is that with it we can simultaneously evaluate both the accuracy and precision of an LSS. Another advantage of our validation method is that we can apply it to any LSS if we can identify the pharmacokinetic model of the agent, the ranges of the model parameters, and the model of within-patient errors.

NLF is a standard and reliable method to estimate AUC when sufficient numbers of observations are available. However, when the numbers of observations are insufficient, we cannot use this method because it overfits the observations and yields unrealistic parameter values. Our estimation applying NLF to 69 clinical profiles with five samples resulted in a deviated estimation of the absorption-rate constant.

TZE is one of the major methods for the estimation of BU AUC [8–10, 20, 25]. Our evaluation revealed that TZE induced three times the overestimation of AUC when absorption is slow both in Chattergoon’s LSS and Fukumoto’s LSS. Though this feature of TZE could be predicted by our theoretical investigation as well as the MLR underestimation problem (vide infra), to our knowledge, no study on TZE seems to have pointed out this clinically important problem.

MLR is the most common method to develop LSS (Table 1) and many studies including independent evaluation researches usually resulted in the conclusion that MLR has a satisfactory estimation performance [1, 2, 5, 7, 13, 24, 25]. However, we showed that Hassan’s MLR induces a systematic bias over 30% when the elimination phase is slow. Panetta et al. [23] showed that MLR will generally give poor results in the estimation of clearance for an atypical patient. Though MLR showed good precision, we conclude that MLR is insufficient to accurately estimate AUC.

In the development of computer software, it is well known that a software test using random data can never assure software correctness [29]. Though systematic development, and step-by-step logical checks, may be the best way to prevent software bugs, exhaustive testing by all combinations of possible input data, if at all possible, improves the assurance levels of the software. A cross validation by an independent set of observations is similar to a software test by random data because it tests a restricted number of patients compared with the target population. Therefore, a cross validation does not guarantee the correctness of the estimation scheme for every case. If we want to develop a new clinical estimation scheme or validate one, we should evaluate its logic as theoretically as possible, and evaluate its performance as exhaustively as possible, as we did in the present study.

Although we could not make an evaluation in this study, the Bayesian method is another major method used to develop LSSs, and other investigators have discussed the merits and demerits of the Bayesian approach [1, 2, 11, 24, 26]. In the Bayesian method to predict drug concentration-time change or AUC, the estimation is made assuming that the simultaneous probability distribution of population pharmacokinetic parameters is already known and trying to find the most likely combination of pharmacokinetic parameters for a subject based on the information of population parameter distribution and the observed data [30].

The parameters, or priors, to be identified in the one-compartment model with first-order absorption are: 1) population mean of four parameters (Ka, Ke, τ, and Vd); 2) their variance-covariance matrix (10 parameters); and 3) variance of within-patient error. Therefore, in the present case, we need to identify 15 pharmacokinetic parameters in advance to apply the Bayesian method. This information is not always available in most LSS studies. Even if the priors can be estimated from the index set, the uncertainty of the prediction depends on the priors, and estimation errors could be large if the index set is small, as in the present case [1].

Moreover, to develop and use an LSS applying the Bayesian method, a specialized computer program for Bayesian analysis is required, and users need extensive training to operate the programs and interpret the results compared with other LSSs. However, if these conditions are cleared up, the Bayesian method will provide an unbiased estimation of the AUC [1, 2].

Another typical and important problem of the estimation of AUC is for ROC (receiver operating characteristic) curves. Hilgers pointed out that the point estimators derived from the empirical distribution functions using the trapezium rule are misleading if they are based on small samples [31]. We have proven that a parametric method for underlying distributions of diagnostic variables provides an unbiased estimator of the AUC [32]. However, not only the trapezium rule, TZE, but also the parametric approach, NLF, are biased in the AUC problem for the concentration-time curve. Our two new methods ought to solve the bias problem.

The concentration curve database search method had better performance in both accuracy and precision compared with conventional methods, and this method is applicable to other agents that require an AUC estimation as well. The limitation of this method is that it is necessary to know the pharmacokinetic model that the agent follows and the ranges of the model parameters in advance. If the wrong model is chosen, e.g., a two-compartment model with absorption in our example of oral BU dose, the estimation will deviate slightly. The time lag of the absorption phase also produces a deviation in the estimation. Hence, in the development of a new LSS for another drug using this concentration curve database search method, special care should be taken as to what pharmacokinetic model holds. However, if the pharmacokinetic model can be fixed, our method will most accurately give the best estimation of AUC.

Modified TZE had also better performance in accuracy and precision compared with conventional methods. The concentration curve database search method and the Bayesian method can only be applied when the pharmacokinetic model can be fixed. MLR needs an index set that is large enough to develop a regression equation. The merit of the modified TZE is that it can be constructed only if the minimum value of the elimination rate constant and the time when the absorption phase ends are known, both of which are relatively not difficult to estimate for most drugs. It should be emphasized that the exact pharmacokinetic model and the distribution of its parameters need not be known in advance. Furthermore, to the extent that we evalu-
lated, modified TZE showed better performance than did MLR and conventional TZE. Therefore, we recommend using a modified TZE instead of the conventional TZE and MLR when there is not sufficient information to develop a concentration curve database search method or to use the Bayesian method.

Finally, regarding the first clinical problem of BU dose adjustment, the target AUC of BU is from 4,200 to 5,400 (μg · hour/L) in our treatment protocol, which indicates that the permissible error is 12.5% of the mean value 4,800 (μg · hour/L). The dosing is repeated 16 times, once every 6 hours, which ends up to being a 4-day period. To achieve the target AUC, we measured five samples after the first dose, estimated the AUC by an LSS, and subsequently changed the dose so that the AUC became 4,800 (μg · hour/L) in subsequent doses. The evaluation results in Table 5 suggest that LSSs other than our concentration curve database method has a possibility that the bias is over 12.5% in some cases. However, even if there is no bias, every estimate has a variance, and we must take both bias and precision into account. If we want 90% of the patients to satisfy the request that the adjusted AUC is in the target range from 4,200 to 5,400 (μg · hour/L) when the bias is zero, the coefficient of variation of the estimate must be under 7.6% (= 12.5%/1.645). Our validation clarifies when there is a possibility that the adjusted dose is out of the target range; and, at that time, we should monitor the patient especially carefully. Though there is no perfect LSS, our new LSSs and validation method can be instrumental in achieving better dose individualization.

5. Conclusions

We evaluated the accuracy and precision of the conventional LSSs for the estimation of AUC both theoretically and using a set of simulated concentration-time curves. Our evaluation revealed:

1. The nonlinear curve fitting algorithm tended to overestimate the absorption rate constant Ka when the number of observations was limited.
2. The trapezium rule with exponential curve approximation overestimated the AUC more than 280% when the Ka was small.
3. The multiple linear regression method underestimated AUC more than 30% when the elimination rate constant Kε was small.

We conclude that these methods must be applied cautiously so as not to go beyond the limitation of each estimation scheme.

Based on these evaluation results, we improved the trapezium rule approach so as not to reach unrealistic values. We also proposed another estimation scheme in which the most likely curve was selected from a set of pre-generated theoretical concentration-time curves. The validation showed that both methods provided more accurate and precise estimations of the AUC than did conventional methods.

Acknowledgments

This work was partially supported by Grant-in-Aid for Scientific Research (C) 22500264. We thank Shoko Fukazawa for her assistance in making the graphs and Robert E. Brandt, Founder, CEO, and CME, and his insightful comments and suggestions. The figure in Appendix 3 is reprinted from the Japan Journal of Medical Informatics, in: “Development of an accurate estimation method for busulfan AUC and its confidence interval” [20] with permission of the Japan Association for Medical Informatics.

References

Appendix 1: Derivation of our Previous LSS Based on the Trapezium Rule

Let

\[ t = \text{time after dose (hours)}, \]
\[ C(t) = \text{plasma busulfan concentration at time } t, \]
\[ A_1 = \text{area under the line which ends } (0, 0) \text{ and } (0.5, C(0.5)), \]
\[ A_2 = \text{area under the line which ends } (0.5, C(0.5)) \text{ and } (1.0, C(1.0)), \]
\[ A_3 = \text{area under the line which ends } (1.0, C(1.0)) \text{ and } (2.0, C(2.0)), \]
\[ A_4 = \text{area under the line which ends } (2.0, C(2.0)) \text{ and } (6.0, C(6.0)), \]
\[ A_5 = \text{area under an exponential curve in the interval } (6.0, \infty) \text{ which passes } (2.0, C(2.0)) \text{ and } (6.0, C(6.0)). \]

Then

\[ \text{AUC} = A_1 + A_2 + A_3 + A_4 + A_5 \]

\[ = \frac{1}{2} \cdot 0.5 \cdot C(0.5) + \frac{1}{2} \cdot 0.5 \cdot (C(0.5) + C(1.0)) \]
\[ + \frac{1}{2} \cdot 1.0 \cdot (C(1.0) + C(2.0)) \]
\[ + \frac{1}{2} \cdot 4.0 \cdot (C(2.0) + C(6.0)) \]
\[ + \frac{C(6.0)}{e^{-6.0k}} \int_{6.0}^{\infty} e^{-kt} dt \]
\[ = 0.5 \cdot C(0.5) + 0.75 \cdot C(1.0) + 2.5 \cdot C(2.0) + 2.0 \cdot C(6.0) + C(6.0)/k, \]

where

\[ k = -\frac{\log C(6.0) - \log C(2.0)}{6.0 - 2.0} = -\frac{1}{4.0} \log \left( \frac{C(6.0)}{C(2.0)} \right), \]

\( k \) is an estimate of \( Ke \) and can be determined so that the curve \( Ae^{-kt} \) passes both \((2.0, C(2.0))\) and \((6.0, C(6.0))\).
Appendix 2: Block Diagram of the One-compartment Model with Lag Time and First-order Absorption

Variables. \( t \): time after dose; \( C_a(t) \): drug concentration in the gut compartment; \( C(t) \): drug concentration in the central compartment; \( D \): dose; \( K_a \): absorption rate constant; \( K_e \): elimination rate constant.

![Block Diagram](image)

Appendix 3: Results of the Clinical Study of Busulfan

Table 1 Demographic data of 78 patients. Note: group A, nine patients with 11 blood samples; group B, 69 patients with five blood samples.

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>Group A</th>
<th>Group B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acute lymphocytic leukemia</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Acute myelogenous leukemia</td>
<td>5</td>
<td>33</td>
</tr>
<tr>
<td>Chronic myelogenous leukemia</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>Myelodysplastic syndrome</td>
<td>2</td>
<td>27</td>
</tr>
<tr>
<td>Unknown</td>
<td>0</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 2 Comparison of model fitness using Akaike Information Criteria (mean ± SD). Note: one-compartment model, a one compartment model with absorption from the gut compartment; two-compartment model, a two-compartment model with absorption from the gut compartment; \( \hat{y} \), estimated concentration from the model.

<table>
<thead>
<tr>
<th>Weight for error summation</th>
<th>One-compartment model</th>
<th>Two-compartment model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No time lag</td>
<td>With time lag</td>
</tr>
<tr>
<td>( 1/\hat{y} )</td>
<td>58.91 ± 10.44</td>
<td>50.74 ± 9.31</td>
</tr>
<tr>
<td>( 1/(\hat{y} - \hat{y}) )</td>
<td>−3.95 ± 10.99</td>
<td>−11.70 ± 8.66</td>
</tr>
</tbody>
</table>

Fig. 1 Distribution of within-patient errors in nine patients with 11 samples