Effect of the Number of Samples on Bayesian and Non-linear Least-squares Individualization: A Study of Cyclosporin Treatment of Haematological Patients with Multidrug Resistance

G. WU, F. PEA, P. COSSETTINI AND M. FURLANUT

Institute of Clinical Pharmacology and Toxicology, Chair of Special Pharmacology, Medical School, University of Udine, Udine, Italy

Abstract

We have studied whether the prediction of drug concentrations improves as the number of samples used for individualization is increased, and whether the Bayesian method of individualization is superior to the non-linear least-squares method. Data were obtained from ten adult haematological patients with multidrug resistance who were treated with cyclosporin. The predictions of blood–cyclosporin concentrations were made using the Abbott PKS program. The number of samples used for individualization was increased from 1 to 30 for the Bayesian method and from 4 to 30 for the non-linear least-squares method. Linear regression, percentage prediction error, and absolute and relative predictive performance were used to evaluate the predictions.

The results show that the Bayesian method affords greater precision than the non-linear least-squares method, but that the non-linear least-squares method is more accurate and results in less bias. Whereas for linear regression predictions improve as the number of samples is increased, other evaluations show improvement in the range from 5 to 11 samples; linear regression, percentage prediction errors and prediction bias support the opinion that the Bayesian method progressively becomes the non-linear least-squares method as the number of samples used for individualization is increased, but the accuracy and precision of prediction do not support this opinion.

The study supports the statement that Bayes' law requires parameters from an infinite population, otherwise the advantage of the Bayesian method might be marginal.

Several sophisticated computer software packages are currently used for clinical optimization of drug regimens and prediction of blood–drug concentrations. The pharmacokinetic parameters implemented by most software packages are obtained either from the literature or from the software provider's database. Because the pharmacokinetic parameters implemented might be well suited for some populations of patients, but not for a patient from other populations, individualization is used so that the pharmacokinetic parameters implemented suit patients from other populations.

Computer software packages perform individualization by using a pharmacokinetic model to fit a patient's blood–drug concentrations and the implemented pharmacokinetic parameters as the initial

values from which are obtained modified pharmacokinetic parameters which are more suitable for the given patient; it then uses the modified implemented pharmacokinetic parameters to optimize drugs regimens and predict blood–drug concentrations. Any previous blood–drug concentrations from a patient can be used for individualization: blood–drug concentrations from a previous course of treatment or blood–drug concentrations from the current course of treatment. Blood–drug concentrations from a previous course of treatment might provide more exact pharmacokinetic estimates because more samples can usually be obtained from a previous course of treatment; if the interval between courses of treatment is too long, however, the patient situation might change, leading to altered pharmacokinetic parameters. Such a change might even lead to less exact pharmacokinetic estimates because although fewer blood samples can be obtained from the current course of

Correspondence: G. Wu, Istituto di Farmacologia Clinica e Tossicologia, Cattedra di Farmacologia Speciale, Facoltà di Medicina e Chirurgia, Università degli Studi di Udine, P. le S. Maria della Misericordia, 33100 Udine, Italy.

treatment the patient situation is less likely to change.

Both Bayesian and non-linear least-squares methods can be used for individualization on the basis of the concentrations obtained. During individualization several questions might occur: At which time-point is a sample needed? How many samples are sufficient? Which of the two methods is better?

The first two issues are generally considered the optimum of sampling strategy. The answer to the first issue has been addressed in some detail (D'Argenio 1981; García et al 1994; Rodvold et al 1994; Wu et al 1995c; Wu 1997). The answer to the second issue seems simple, i.e. the more samples, the better the predictions. However, to the best of our knowledge, no study has been conducted to examine the second issue in great detail.

The answer to the third issue also seems simple, i.e. the Bayesian method is superior to the non-linear least-squares method when the number of samples is limited and the pharmacokinetic parameters from a population are available. This is because the Bayesian method incorporates into the pharmacokinetic model not only samples from a patient but also pharmacokinetic parameters from the patient population (Abbottbase 1992), whereas the non-linear least-squares method uses only the samples from a patient. However, a fundamental condition has so far been ignored in most Bayesian method applications, and that is that Bayes' law requires parameters from an infinite population (Mood et al 1988). Clearly, no pharmacokinetic parameters can be obtained from an infinite population of patients and so the Bayesian method might not work as well as expected.

In our previous studies (Wu et al 1995a, c, 1996; Wu & Furlanut 1996) we have addressed several issues concerning the prediction of blood–cyclosporin concentrations in haematological patients with multidrug resistance. However, we have not yet analysed the issue of different numbers of samples for individualization of predictions obtained using the Bayesian and non-linear least-squares methods. In this study we have attempted to address this issue by predicting blood–cyclosporin concentrations in haematological patients with multidrug resistance using a popular pharmacokinetic software package. In practice we have used different numbers of blood samples from the first course of treatment to individualize the implemented pharmacokinetic parameters in the computer software for each patient using the Bayesian and non-linear least-squares methods, then used the individualized implemented pharmacokinetic parameters with the second-course dosage to predict the second-course blood concentrations for each patient. Finally we have compared the measured and predicted second-course blood concentrations to make statistical inferences and draw conclusions.

Materials and Methods

Patients, cyclosporin dosage and blood–cyclosporin concentrations

Ten adult haematological patients with multidrug resistance (one acute lymphocytic leukaemia, seven acute non-lymphocytic leukaemia, one chronic myeloid leukaemia and one non-Hodgkin lymphoma) took two courses of treatment (Table 1). Table 2 gives details of patient liver function. The patients' renal functions were stable, i.e. the fluctuation of serum creatinine was smaller than 0.5 g L^{-1} from the start of treatment (Rodvold et al 1994). Cyclosporin treatment was conducted by continuous intravenous infusion for several days (Table 1), and blood–cyclosporin concentrations were monitored approximately four times a day during infusion (0000, 0600, 1200 and 1800 h) and 11 times after infusion (after 0, 0.5, 1, 2, 3, 5, 7, 9, 12, 24 and 36 h). The whole-blood samples were immediately analysed by a fluorescence polarization immunoassay method (TDx, Abbott Laboratories, Diagnostic Division, Irving, TX; Moyer et al (1991)).

Table 1. Patient demographics and cyclosporin dosage.

Course of treatment	First course	Second course
Age (years)	44 ± 16	44 ± 16
Weight (kg)	66 ± 11	64 ± 12
Dose (mg kg^{-1} day^{-1})	11 ± 1	12 ± 2
Number of samples	23 ± 6	26 ± 4
Time of continuous intravenous infusion (days)	3.71 ± 1.05	3.74 ± 0.95

The data are expressed as means ± s.d. (n = 10); there were eight male and two female patients. The time interval between two courses was 77 ± 73 days (means ± s.d.). No statistical difference was found between two courses (paired Student's t-test).

Table 2. Patient liver function.

Course of treatment	First course	Second course
SGPT (units L^{-1})	21.10 ± 9.45	40.30 ± 46.44
SGOT (units L^{-1})	22.50 ± 10.30	22.10 ± 10.35
γ-GT (units L^{-1})	60.70 ± 41.16	62.40 ± 43.46
Albumin (g L^{-1})	36.93 ± 5.68	37.12 ± 6.18

SGPT, serum glutamic pyruvic transaminase; SGOT, serum glutamic oxalacetic transaminase; γ-GT, gamma glutamyl transpeptidase. The data are expressed as means $\pm$ s.d. (n = 10); there were eight male and two female patients. No statistical difference was found between the two courses (paired Student's t-test).

Predictive program and pharmacokinetic parameters

The PKS program (Abbottbase Pharmacokinetic System, version 1.10, Abbott Laboratories, IL) was used to predict the second-course blood-cyclosporin concentrations. The Bayesian and non-linear least-squares methods are operated in the PKS program by use of a two-compartment model with volume of distribution in central compartment $(Vd_c = 0.70 \pm 0.26$ L $kg^{-1})$, clearance $(CL = 0.25 \pm 0.08$ L h^{-1} $kg^{-1})$ and inter-compartment rate constants $(k_{12} = 0.52 \pm 0.31$ and $k_{21} = 0.074 \pm 0.018$ $h^{-1})$. These implemented pharmacokinetic parameters are based upon 24 adult haematological patients with multidrug resistance treated with cyclosporin, and log-normally distributed (Wu et al 1995a, c, 1996). The weighting method is the inverse-squared predicted concentration (Abbottbase 1992).

Samples for individualization

The first-course blood cyclosporin samples from each patient were used for individualization: the first sample was that taken at the first sampling time; the second sample was that taken at the last sampling time; the third sample was that taken at the midpoint between the first and second sampling times; the fourth sample was that taken at the midpoint between the first and third sampling times; the fifth sample was that taken at the midpoint between the second and third sampling times; and so on.

Because, when using the non-linear least-squares method (Abbottbase 1992), individualization requires that the number of samples is greater than the number of parameters in the model, we used a minimum of four samples for the non-linear least-squares method, because there are four parameters in the two-compartment model, and the beginning of infusion can be considered as a sample at which the drug concentration is zero.

Statistics

Data calculation. The measured and predicted second-course blood–cyclosporin concentrations

were used to calculate percentage prediction errors from the equation (Wu 1995a; Wu et al 1995a, b, c, 1996; Wu & Furlanut 1996, 1997): percentage prediction error = [(predicted concentration – measured concentration)/measured concentration] $\times$ 100. Outliers (3 $\times$ s.d.) were detected according to the method of Healy (1979). The absolute and relative performances were calculated according to a method described elsewhere (Sheiner & Beal 1981; Wu 1995b). The absolute predictive performance includes: mean prediction error as bias; mean absolute error as accuracy; and mean-squared prediction error as precision. The relative predictive performances were differences in bias, accuracy and precision. Linear regression was used to calculate the correlation coefficient between measured and predicted concentrations.

Data presentation. The Shapiro-Wilk's W-test was used to determine the distribution of the data. For normal distribution, the data were presented as means $\pm$ s.d. For non-normal distribution, the data were presented as median with interquartile range.

Statistical inference. The paired Student's t-test, the Wilcoxon matched-pairs test and Kruskal–Wallis analysis of variance were used. $P < 0.05$ was considered to be indicative of statistical significance. Statistica for Windows (release 4.0 A, Statsoft 1993) was used to perform all the statistical tests.

Results

Figure 1 shows the correlation coefficients for comparison of predicted and measured concentrations in relation to the number of samples from the first course of treatment used for individualization. The correlation coefficient increases as the number of samples from the first course of treatment is increased. Generally, Bayesian method predictions are better than those obtained by the non-linear least-squares method (17 compared with 9), however no difference is found between the results
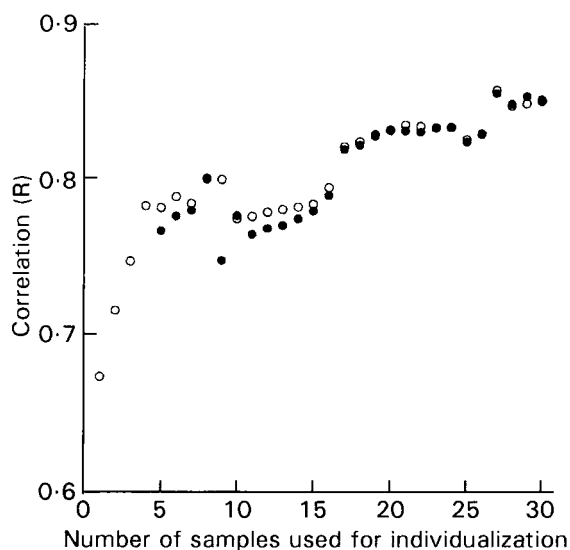
Figure 1. The correlation coefficient (y-axis) for comparison of predicted second-course concentrations with measured second-course concentrations in relation to the number of samples from the first course of treatment used for individualization (x-axis). ○, Prediction by the Bayesian method; ●, prediction by the non-linear least-squares method. The number of comparisons is reported in Figure 2.

from the Bayesian and non-linear least-squares methods when the number of samples is larger than 20. Also, in most instances differences between the results from the Bayesian and non-linear least-squares methods are not large enough to be significant.

Figure 2 shows the dependence of percentage prediction errors on the number of samples used for individualization. The predictions are better when the number of samples is in the range 10 to 16, because the prediction medians are closer to 0% prediction error. Among all the comparisons, three Bayesian methods are better than their corresponding non-linear least-squares methods, and six non-linear least-squares methods are better than their corresponding Bayesian methods.

Figure 3 shows the bias of predictions in relation to the number of samples used for individualization. The predictions are less biased when the number of samples is in the range 10 to 16. Among all the comparisons, five Bayesian methods have smaller bias than their corresponding non-linear least-squares methods, and six non-linear least-squares methods have smaller bias than their corresponding Bayesian methods.

Figure 4 shows the accuracy of predictions in relation to the number of samples used for individualization. The Bayesian method predictions are most accurate when the number of samples is four. Among all the comparisons, twelve non-linear least-squares methods are more accurate than their corresponding Bayesian methods, and five Baye-
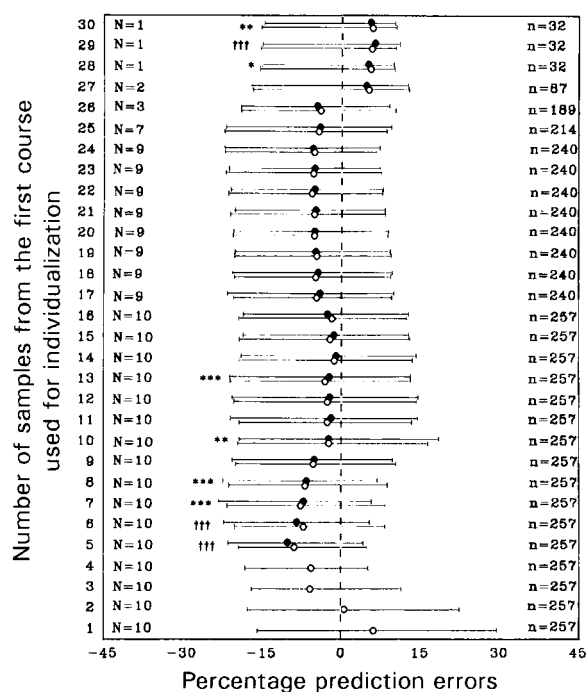

Figure 2. The percentage prediction errors (x-axis) for the Bayesian (○) and the non-linear least-squares (●) methods in relation to the number of samples from the first course used for individualization (y-axis). The data are expressed as median with interquartile range. The larger the interquartile range, the worse the predictions. When a percentage prediction error is larger than 0%, over-prediction occurs, that is, a predicted concentration is higher than its corresponding measured concentration. When a percentage prediction error is smaller than 0%, under-prediction exists, i.e. a predicted concentration is lower than its corresponding measured concentration. N is the number of patients and n the number of percentage prediction errors. $*P < 0.05$, $**P < 0.01$ and $***P < 0.001$, the predictions of the non-linear least-squares method are significantly better than those of the Bayesian method. $\dagger P < 0.05$, $\dagger\dagger P < 0.01$ and $\dagger\dagger\dagger P < 0.001$, the predictions of the Bayesian method are significantly better than those of the non-linear least-squares (Wilcoxon matched-pairs test). Taken together, the figure can be read, for example, five first-course samples (y-axis labels) from each of 10 patients (N = 10) were used for individualization in each of these 10 patients; 257 measured and predicted second-course concentrations (n = 257) were obtained from these 10 patients using the Bayesian and non-linear least-squares methods, respectively; 257 percentage prediction errors were calculated for Bayesian and non-linear least-squares methods, respectively; the medians of 257 percentage prediction errors were −8.84% and −10.20% for Bayesian and non-linear least-squares methods, respectively; the interquartile ranges were from −19.43% to 4.83% for the Bayesian method and from −21.34% to 4.10% for the non-linear least-squares method; the Bayesian method predictions are better than those of the non-linear least-squares method at the $P < 0.001$ significance level ($\dagger\dagger\dagger$). The Kruskal–Wallis analysis of variance was used to test whether the medians of percentage prediction errors are equal when the individualization from 1 to 30 first-course samples (y-axis labels) for Bayesian and from 5 to 30 first-course samples for non-linear least-squares methods.

sian methods are more accurate than their corresponding non-linear least-squares methods.

Figure 5 shows the precision of predictions in relation to the number of samples used for individualization. Bayesian method predictions are most
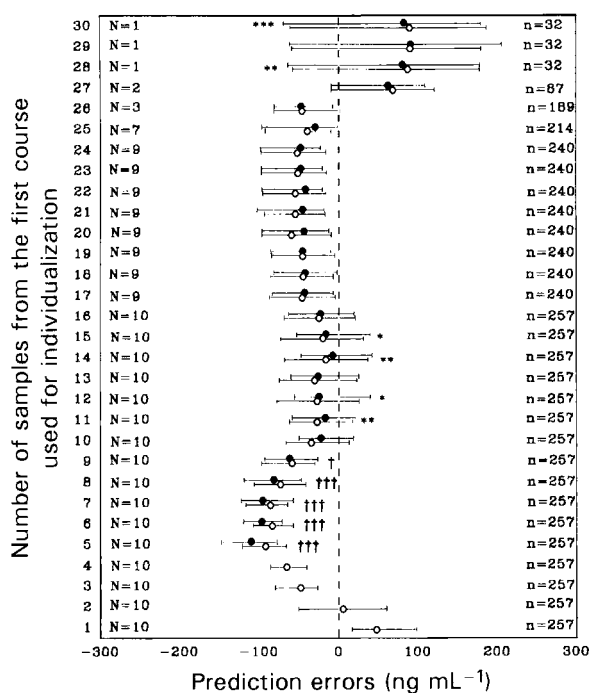
Figure 3. Bias of predictions obtained by use of the Bayesian (O) and non-linear least-squares (●) methods (x-axis) in relation to the number of samples from the first course of treatment used for individualization (y-axis). For details, see the caption to Figure 2.

Figure 4. The prediction accuracy obtained by use of the Bayesian (O) and non-linear least-squares (●) methods (x-axis) in relation to the number of samples from the first course of treatment used for individualization (y-axis). For details, see the caption to Figure 2.

precise when the number of samples is four. Among all the comparisons, fourteen Bayesian methods are more precise than their corresponding non-linear least-squares methods, and four non-linear least-squares methods are more precise than their corresponding Bayesian methods.

We also used Kruskal–Wallis analysis of variance to test whether the medians of percentage prediction errors, bias, accuracy and precision were equal among all predictions using different numbers of samples for individualization. The results show that they are not equal at the $P < 0.001$ level. This means that the number of samples from the first course of treatment for individualization indeed affects the predictions of the Bayesian and non-linear least-squares methods.

Discussion

We have used a data-rich sample schedule to analyse the effect of the number of samples used for individualization and the difference between Bayesian and non-linear least-squares methods, although such a data-rich sample schedule is not feasible for clinical settings.

This data-rich study was conducted for two reasons. Until now, too few samples have been used in most studies comparing the effect of the number of samples on the method of individualization used.
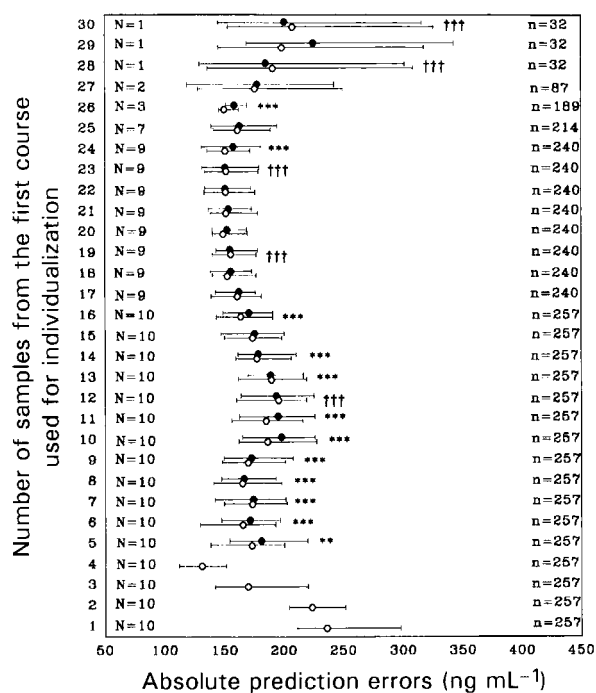

Figure 5. The prediction precision obtained by use of the Bayesian ((O) and non-linear least-squares (●) methods (x-axis) in relation to the number of samples from the first course of treatment used for individualization (y-axis). For details, see the caption to Figure 2.
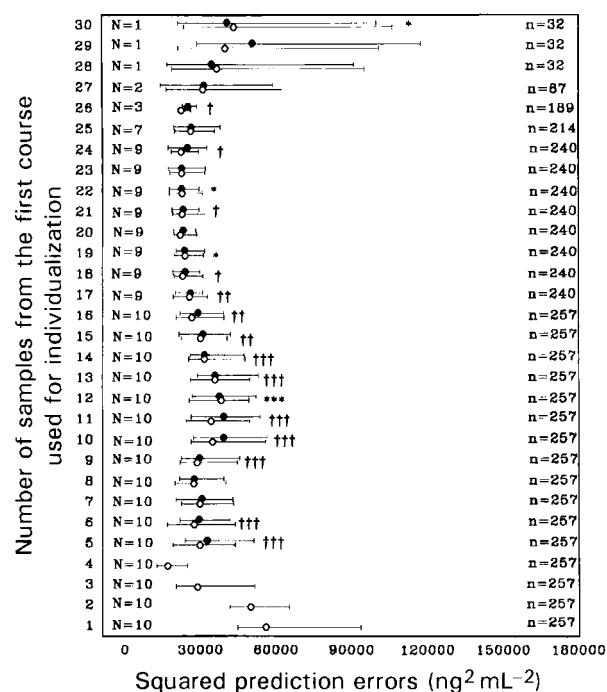
One can obtain some indication from these studies, but not a systematic understanding. The second reason is that the Bayesian method is widely regarded as being overwhelmingly superior to the non-linear least-squares method, but the fundamental condition for the Bayesian method is ignored, as indicated above. On the basis of these viewpoints, our study is a technique-oriented study rather than a clinical-oriented study.

Although the results of this study show the tendency, i.e. the more the samples, the better the prediction, this tendency is not very clear, because the results of linear regression show such a tendency in the range 1 to 30 samples whereas other evaluations demonstrated such a tendency only in the range 5 to 11 samples.

The results of this study did not show the Bayesian method to be overwhelmingly superior to the non-linear least-squares method, indeed several statistical evaluations indicate that the non-linear least-squares method is better. This might support the statement that the Bayes' law requires the use of parameters from an infinite population (Mood et al 1988), otherwise, the advantage of the Bayesian method might be marginal.

Current opinion is that the Bayesian method progressively becomes the non-linear least-squares method as the number of samples used for individualization increases (Kelman et al 1982; Jelliffe et al 1991). So far no publication has documented when this change occurs. The linear regression, percentage prediction error and prediction bias results supported this opinion, but prediction accuracy and precision results did not.

In this study the number of patients was not large; because statistical power analysis showed that the number of patients was sufficient to find the difference, we did not include more patients.

Until recently, numerous optimum sampling strategies have been developed in various situations (e.g. D'Argenio 1981; Wu 1997); it is, however, difficult to follow each of them to schedule our sample selection. One significant factor is that most optimum sampling strategies are concerned with using as few samples as possible. However, because we have so many samples from each patient, we hope to use a somewhat random samples schedule.

The results also show that predictions were dramatically improved as the number of samples was increased from one to two. This should encourage the taking of more than one sample for individualization and prediction.

In conclusion, the results show that the Bayesian method affords greater precision than the non-linear least-squares method, but use of the non-

linear least-squares method results in less bias and greater accuracy. Although linear regression shows that predictions improve as the number of samples is increased, other evaluations show improvement in the range 5 to 11 samples only. The linear regression, percentage prediction error and prediction bias results supported the opinion that the Bayesian method progressively becomes the non-linear least-squares method as an increasing number of samples is used for individualization, but prediction accuracy and precision did not support this opinion.

References

Abbottbase (1992) Pharmacokinetic System Operations Manual version 1.10, Abbott Laboratories, Abbott Park, IL, 6: 45–62

D'Argenio, D. Z. (1981) Optimal sampling times for pharmacokinetic experiments. J. Pharmacokinet. Biopharm. 9: 739–756

García, M. J., Gavira, R., Santos Buelga, D., Dominguez-Gil, A. (1994) Predictive performance of two phenytoin pharmacokinetic dosing programs from non-steady-state data. Ther. Drug Monit. 16: 380–387

Healy, M. J. R. (1979) Outliers in clinical chemistry quality-control schemes. Clin. Chem. 25: 675–677

Jelliffe, R. W., Lglesias, T., Hurst, A. K., Foo, K. A., Rodrinquez, J. (1991) Individualizing gentamicin dosage regimens: a comparative review of selected models, data fitting methods and monitoring strategies. Clin. Pharmacokinet. 21: 461–478

Kelman, A. W., Whiting, B., Bryson, S. (1982) OPT: a package of computer programs for parameters optimization in clinical pharmacokinetics. Br. J. Clin. Pharmacol. 14: 47–56

Mood, A. M., Graybill, F. A., Boes, D. C. (1988) Introduction to the Theory of Statistics. 3rd edn, McGraw-Hill, London, p. 36

Moyer, T. P., Winkels, J., Krom, R., Wiesner, R. (1991) Evaluation of Abbott TDx monoclonal assay for monitoring cyclosporine in whole blood [Tech. Brief]. Clin. Chem. 37: 1120–1121

Rodvold, K. A., Rotschafer, J. C., Gilliand, S. S., Guay, D. R. P., Vance-Bryan, K. (1994) Bayesian forecasting of serum vancomycin concentrations with non-steady-state sampling strategies. Ther. Drug Monit. 16: 37–41

Sheiner, L. B., Beal, S. L. (1981) Some suggestions for measuring predictive performance. J. Pharmacokinet. Biopharm. 9: 503–512

Wu, G. (1995a) Prediction of uptake of methyl mercury by rat erythrocytes using a two-compartment model. Arch. Toxicol. 70: 34–42

Wu, G. (1995b) Calculating predictive performance: a user's note. Pharmacol. Res. 31: 393–399

Wu, G. (1997) An explanation for failure to predict cyclosporine area under the curve using a limited sampling strategy: a beginner's second note. Pharmacol. Res. 35: 547–552

Wu, G., Furlanut, M. (1996) Prediction of blood–cyclosporine concentrations in haematological patients with multidrug resistance by means of total, lean and different adipose dosing body weight using Bayesian and non-linear least-squares methods. Int. J. Clin. Pharmacol. Res. 16: 89–97

Wu, G., Furlanut, M. (1997) Prediction of serum vancomycin concentrations by means of six different equations for calculation of creatinine clearance. Int. J. Clin. Pharmacol. Res. 17: 1–10

Wu, G., Furlanut, M., Baraldo, M., Pea, F., Damiani, D., Baccarani, M. (1995a) Prediction of cyclosporine blood concentrations in hematological patients with multidrug resistance. Pharmacol. Res. 31 (Suppl.): 44

Wu, G., Baraldo, M., Furlanut, M. (1995b) Calculating percentage prediction error: a user's note. Pharmacol. Res. 32: 241–248

Wu, G., Baraldo, M., Pea, F., Cossettini, P., Furlanut, M. (1995c) Effects of different sampling strategies on predictions of blood–cyclosporine concentrations in haematological patients with multidrug resistance by Bayesian and non-linear least-squares methods. Pharmacol. Res. 32: 355–362

Wu, G., Cossettini, P., Furlanut, M. (1996) Prediction of blood-cyclosporine concentrations in haematological patients with multidrug resistance by one-, two- and three-compartment models using Bayesian and non-linear least-squares methods. Pharmacol. Res. 34: 47–57