CLINICAL PHARMACOLOGY & THERAPEUTICS
VOLUME 39 NUMBER 6 JUNE 1986

ORIGINAL ARTICLES

An evaluation of population pharmacokinetics in therapeutic trials. Part I. Comparison of methodologies

NONMEM, a computer program that uses the method of extended least-squares analysis, has been advocated as a means of obtaining estimates of population pharmacokinetic parameters when only fragmentary information can be obtained from subjects. To assess the performance of this program, we compared NONMEM with traditional methods for the estimation of population pharmacokinetic parameters with data collected during a phase III clinical trial of alprazolam. NONMEM estimates of the population mean clearance and its coefficient of variation were identical to the estimates obtained with traditional pharmacokinetic techniques. Moreover, NONMEM estimates of these parameters remained stable even when as few as three data points were available per subject. NONMEM estimates of the mean volume of distribution and its coefficient of variation appear to be overestimated, apparently because of the sampling scheme used to generate data for the NONMEM analysis. Suggestions for the effective use of NONMEM in clinical trials, to maximize the benefits of this approach, are provided. Our results lend further support for the use of NONMEM to estimate population pharmacokinetic parameters of a drug from data generated during phase III clinical trials. (CLIN PHARMACOL THER 1986;39: 605-12.)

Thaddeus H. Grasela, Jr., Pharm.D., Edward J. Antal, Ph.D., Raymond J. Townsend, Pharm.D., and Randall B. Smith, Ph.D.
Ann Arbor and Kalamazoo, Mich.

From the University Hospital, Department of Pharmacy Services and College of Pharmacy, University of Michigan, Ann Arbor, and the Pharmacokinetic/Pharmacodynamic Research Unit, The Upjohn Company, Kalamazoo.

Supported in part by Grant No. BRSG2S07RR05571-19 from the Biomedical Research Support Program, Division of Research Resources, National Institutes of Health and by a grant from The Upjohn Company.
Received for publication July 22, 1985; accepted Jan. 26, 1986.

Reprint requests to: Dr. Thaddeus H. Grasela, Jr., The Clinical Pharmacokinetics Laboratory, Millard Fillmore Hospitals, 3 Gates Circle, Buffalo, NY 14209.

In the past, pharmacokinetic studies have focused on the individual, in that studies were designed to yield the maximum information on the disposition of drugs in individual subjects. More recently, there has been increased interest in the determination of population pharmacokinetic parameters, i.e., parameters that define the typical pharmacokinetic behavior of a drug in a large group of subjects or patients.

These population pharmacokinetic parameters include fixed effect parameters, which quantify a population's average kinetics (including average relationships between physiology and pharmacokinetics), and random effect parameters, which quantify the typical magnitude of random interindividual kinetic variability and the typical magnitude of the residual variability as a result of random intraindividual kinetic variability, drug level measurement, and model specification error. Estimates of these parameters have proved useful for a number of clinically relevant purposes, including the development of dosing guidelines for specific populations of patients and the revision of dosing regimens by the use of measured drug concentrations.[1]

Current procedures for the estimation of these parameters involve performing traditional pharmacokinetic studies, after single or multiple doses, in normal volunteers or in patients with mild degrees of a disease of interest. The major problem with this approach concerns the representativeness of the information obtained, because drug disposition in patients who receive a drug for a therapeutic effect may be significantly different from drug disposition in volunteers. Unfortunately, traditional pharmacokinetic studies can be difficult to perform in the clinical setting. Serious ethical problems arise when one attempts to perform these studies in critically ill, pediatric, and elderly patients who may not be able to tolerate the rigors of such a study. This can result in a paucity of clinically relevant information and significant delays before problems are recognized.

Sheiner et al.[2] have advocated an alternative approach to the problem of estimating population pharmacokinetic parameters by the use of data generated during the routine clinical care of patients. This approach, implemented in the computer program NONMEM, has been shown to provide accurate and precise estimates of population pharmacokinetic parameters from such data in both simulation studies and in analysis of clinical data.[3-8]

It has been proposed that this data analysis approach be applied to data collected during phase III and phase IV clinical trials to identify more quickly populations that are at risk for toxicity because of altered pharmacokinetics. To validate this new methodology as applied to phase III clinical trials, a series of studies are currently being conducted to evaluate the applicability and performance of NONMEM in a variety of clinical study settings. In addition to our present report addressing the comparability of methodologies, other multicenter, limited sampling studies will evaluate (1) the ability of NONMEM to detect potential drug-drug interactions, (2) the applicability and practicality of conducting a NONMEM analysis as an addendum to a large-scale clinical efficacy study, and (3) the feasibility and performance of NONMEM when used as a true pharmacokinetic screen in a long-term general patient population. These results will be the subject of future reports. Although a marketed drug will be used in all cases, the results should provide insights into premarketing situations.

The objectives of this study are twofold: (1) to compare NONMEM empirically with two standard methods for the pharmacokinetic analysis of data obtained during a multiple-dosing trial of alprazolam, and (2) to evaluate the ability of NONMEM to use fragmentary amounts of data per individual. This was accomplished by repeating the analysis of the above data with progressively fewer data per subject.

METHODS

Ten healthy adult men with a mean ($\pm$SD) weight of 79.9 $\pm$ 44.4 kg were initially given alprazolam, 1 mg po, followed by 0.5 mg po every 8 hours for 7 days.

Eighteen alprazolam plasma concentrations were measured after the first and last dose at 0.25, 0.5, 0.75, 1, 1.25, 1.5, 2, 3, 4, 6, 8, 10, 12, 16, 24, 32, 40, and 48 hours. In addition, 12 trough levels were measured from each subject during the multiple-dosing period (Fig. 1).

Population pharmacokinetic parameters of alprazolam were obtained by use of the following data analysis procedures.

Standard two-stage method. The Standard two-stage (STS) method, as the name implies, proceeds in two stages. In the first stage each individual's data are separately analyzed to obtain estimates of the individual's pharmacokinetic parameters. For the purposes of our study, total body clearance (CL) was estimated after the first and last dose by dividing the dose by the appropriate plasma AUC. The elimination rate constant (k_e) was estimated by least-squares regression analysis of the terminal log-linear decay phase. The apparent volume of distribution (V_{area}) was then calculated as: $V_{area} = CL/k_e$. The individual's estimated CL, k_e, and

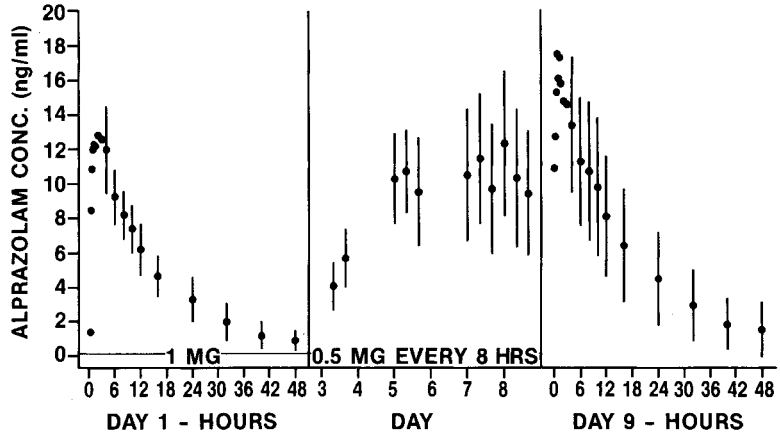

Fig. 1. Mean (± SD) alprazolam plasma concentration-time profile observed with the multiple-dose regimen. Alprazolam concentrations measured after the first and last doses were used for STS approach. Trough levels measured on days 3 through 8 were used for analysis with the accumulation model and NONMEM.

V_{area} was taken to be the mean of the values obtained after the first and last dose for the respective parameters.

In the second stage, the mean population parameters were estimated by combining the corresponding individual estimates. In light of the pharmacostatistical model used for NONMEM analysis (see below), the geometric mean of the individual estimates was used to estimate the population mean parameter. Estimates of the interindividual random effect parameters were obtained by calculating the standard deviations of the logarithms of the corresponding fixed effect parameters. A confidence interval for each fixed effect parameter is obtained from the exponentials of the end points of the usual 95% confidence interval for a mean computed from the logarithms of the individual estimates. There is no appropriate method to compute 95% confidence intervals for random interindividual effect parameters.

Accumulation model. The 12 trough levels (Fig. 1) measured from each subject during the multiple-dosing regimen were fit to Eq. 1:

$$C_{min,i} = C^{ss}_{min,i} \cdot (1 - e^{-k_{e,i}t}) \quad (1)$$

where $C_{min,i}$ is the trough concentration measured at time t in the ith subject, $C^{ss}_{min,i}$ is the average steady-state trough concentration for the ith subject, t is the time (in hours) of the sampling of $C_{min,i}$ measured from the start of the multiple-dosing regimen, where $0 < t <$ time of the last measured concentration, and $k_{e,i}$ is the k_e in the ith subject.

NONMEM. The trough levels (Fig. 1) measured from each subject during the multiple-dosing trial were combined and analyzed by NONMEM. To evaluate the ability of NONMEM to estimate population pharmacokinetic parameters from fragmentary data, the number of trough levels for each subject was progressively reduced by the random removal of trough levels from each individual. In this fashion, three data sets were created: the original data set, containing 12 trough levels per subject (Data-1); a second data set, containing six trough levels per subject (Data-2); and a third data set, containing three trough levels per subject (Data-3). Ten subjects were included in each data set. Population pharmacokinetic parameters of alprazolam were estimated for each data set by NONMEM and were compared with the results of the STS approach and the accumulation model.

Pharmacokinetic model. The pharmacokinetics of alprazolam in plasma are described by a one-compartment open model with first-order elimination. Drug absorption is modeled as a zero-order infusion over a 1.5-hour period, the average time to peak reported in a previous study.[9] This simple pharmacokinetic model was required because of the nature of the sampling schedule (only trough concentrations were available; see Discussion). To assess the influence of the value selected for the duration of infusion, several values between 0.5 and 2.5 hours were selected and the fits to the data were compared. Bioavailability of the drug in all data analysis methods is assumed to be unity.[9] The pharmacokinetic model is cast in a recursive form, as described previously,[7] because of the repetitive nature of the dosing history. The pharmacokinetic parameters to be estimated are CL and V_{area}.

Statistical model. Unlike the standard approach

Table I. Comparison of methods for the estimation of population pharmacokinetic parameters of alprazolam

	STS	*Accumulation model*	*NONMEM*		
			Data-1	*Data-2*	*Data-3*
No. of subjects	10	10	10	10	10
No. of samples per subject	36	12	12	6	3
CL (L/hr/kg)	0.060*	NA	0.060	0.059	0.0596
95% CI	(0.03-0.12)		(0.047-0.073)	(0.050-0.068)	(0.049-0.070)
Coefficient of variation (%)	29	NA	29	27	28
95% CI	NA		(16-39)	(17-34)	(17-36)
V_{area} (L/kg)	1.05*	NA	1.56	1.37	1.53
95% CI	(0.84-1.33)		(1.2-1.9)	(1.1-1.6)	(1.1-1.9)
Coefficient of variation (%)	14	NA	45	47	84
95% CI	NA		(16-62)	(0-82)	(0-145)
k_e (%)	0.058*†	0.054*	0.038†	0.044†	0.039†
95% CI	(0.031-0.108)	(0.022-0.134)	(0.028-0.048)	(0.035-0.053)	(0.031-0.046)
Coefficient of variation (%)	28	40	54	55	89
95% CI	NA	NA	—	—	—

NA = Not available; CI = confidence interval; — = not calculated.
*Geometric mean.
†Calculated from estimates of CL and V_{area}.

to pharmacokinetic analysis, NONMEM requires that an explicit statistical model be supplied. This statistical model must account for interindividual variation in pharmacokinetic parameters—in this case, CL and V_{area}—and for residual error. The latter represents uncertainty in the relationship between the plasma concentration predicted by the pharmacokinetic model and the respective measured values, and is modeled by Eq. 2:

$$\ln(C_{ij}) = \ln(\hat{C}_{ij}) + \epsilon_{ij} \quad (2)$$

where $\ln(C_{ij})$ is the logarithm of the ith measured plasma concentration in the jth individual and $\ln(\hat{C}_{ij})$ is the logarithm of the corresponding predicted concentration resulting from the pharmacokinetic model. The ϵ_{ij} values are independent, identically distributed statistical errors with mean zero and variance σ^2. By modeling the logarithms of the drug concentrations, we state that the error intervening between the observed and predicted concentrations increased in proportion to the measured concentration, a phenomenon frequently observed in practice. The model in Eq. 2 should be regarded as only an approximation to what is undoubtedly a more complex error model, because ϵ_{ij} must represent all uncertainty caused by intraindividual time variation in CL and V_{area}, pharmacokinetic model misspecification, analytic errror in measurement of plasma concentrations, and errors in the reported time of dosing or sampling for drug level determination.

For interindividual variation, we assume that $\ln(CL_j) = \ln(CL) + \eta^{CL_j}$ and $\ln(V_{area,j}) = \ln(V_{area}) + \eta^{V_{area,j}}$, where CL and V_{area} are the population mean values and the η_j values are individual random pertubations from these predictions that are independent and identically distributed, with mean zero and variances equal to ω_{CL}^2 and $\omega_{V_{area}}^2$, respectively. These statistical models are written in logarithmic terms so that (1) the individual parameters must be greater than zero, and (2) if a symmetric distribution is assumed for the η_j, distribution of individual parameters is skewed to the right.

Under the additional assumption of zero covariances among the individual parameters, there are five population parameters for this pharmacostatistical model: CL, V_{area}, ω_{CL}^2, $\omega_{V_{area}}^2$, and σ_ϵ^2.

RESULTS

Table I summarizes our results. The estimate of the population mean CL by the STS approach is 0.06 L/hr/kg (95% confidence interval 0.03 to 0.12 L/hr/kg), which is identical to the estimate of CL by NONMEM, namely 0.06 L/hr/kg (95% confidence interval approximately 0.05 to 0.07 L/hr/kg). Note that the NONMEM estimates remain stable, along with the 95% confidence interval, even as the number of samples is reduced from 12 to six to three per subject.

Estimates of the coefficient of variation for CL are also similar for the two methods, approximately 29% for both. The 95% confidence interval for this parameter by NONMEM is 16% to approximately 39%. An estimate for this interval is not available with the STS approach.

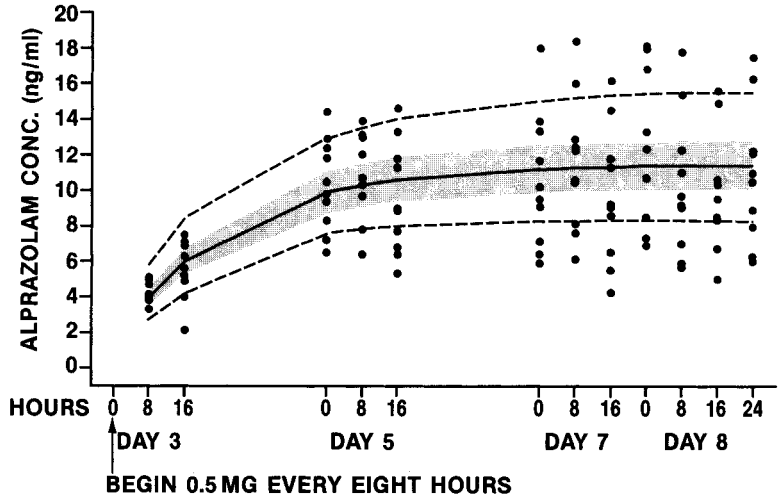

Fig. 2. NONMEM fit to Data-1 (12 samples per subject, 10 subjects). The *solid line* connects trough levels predicted for the average individual in this study (79.9 kg man) from population pharmacokinetic parameters obtained with NONMEM. The *shaded area* represents ±SD of intraindividual variability. The area bounded by the *dashed lines* represents ±SD of both inter- and intraindividual variability.

The estimate of the population mean V_{area} by the STS approach is 1.05 L/kg (95% confidence interval 0.84 to 1.33 L/kg), whereas the corresponding NONMEM estimates are somewhat higher, 1.56 L/kg (95% confidence interval 1.2 to 1.9 L/kg), 1.37 L/kg (95% confidence interval, 1.1 to 1.6 L/kg), and 1.53 L/kg (95% confidence interval 1.1 to 1.9 L/kg) for Data-1, Data-2, and Data-3, respectively. Moreover, the estimates of the coefficient of variation of V_{area} by NONMEM are also higher (45%, 47%, and 84% for Data-1, Data-2, and Data-3, respectively) than the corresponding estimate by the STS approach (14%). Also note that the NONMEM 95% confidence interval for this parameter is rather large and increases as the amount of data decreases.

The estimate of the k_e by the accumulation model (Eq. 1) is 0.054 $hours^{-1}$, as compared with the value of 0.058 $hours^{-1}$ calculated from the STS estimates of CL and V_{area}. The values for k_e as calculated by NONMEM estimates of CL and V_{area} are 0.038, 0.044, 0.039 $hours^{-1}$ for Data-1, Data-2, and Data-3, respectively.

To determine the consequences of the use of a zero-order infusion process to model what is typically a first-order absorption process, and to assess the influence of a variable time to peak concentration, the data sets were reanalyzed with the time to peak concentration fixed at 0.5, 1.0, 1.5, and 2.5 hours. No differences in the parameter estimates or the goodness of fit were found.

Figs. 2 to 4 show the fit to the three data sets from the population pharmacokinetic parameters obtained with NONMEM.

DISCUSSION

Our results suggest that estimates of desired population pharmacokinetic parameters can be obtained directly in patients of interest by use of the fragmentary information generated during phase III clinical trials or obtained as part of the routine clinical care of patients.

The estimates of CL by either the STS method or NONMEM are identical even when only three trough levels per subject are used. In a study of the performance of NONMEM in the analysis of simulated routine clinical data, Sheiner and Beal[5] demonstrated that NONMEM was capable of accurately estimating the population mean CL of a drug when as few as two concentrations, measured under a variety of circumstances, were available per individual. Our results suggest that the above findings can be extended to the fragmentary data often collected as part of a phase III clinical trial.

The estimates of V_{area} are somewhat higher with NONMEM analysis as compared with the STS method, and the precision of the estimate is worse with NONMEM as indicated by the larger 95% confidence interval. This is undoubtedly the result of the sampling scheme. Most of the alprazolam levels used for the NONMEM analyses were measured at steady state, and because the dosing interval was less than the $t_{1/2}$ of the

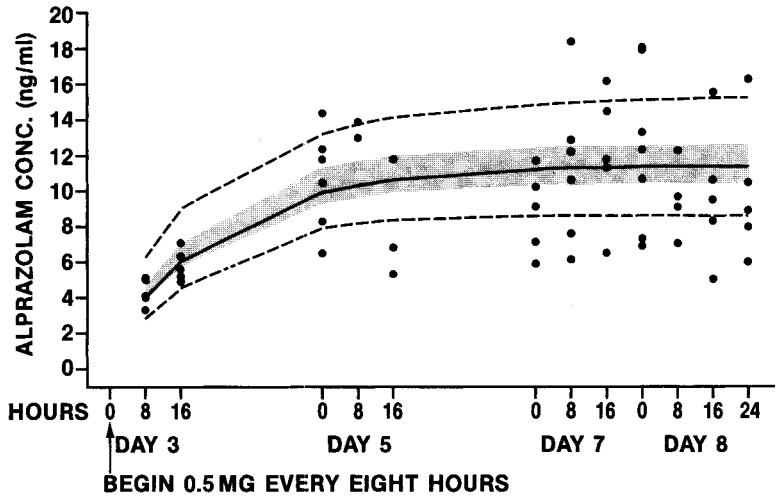

Fig. 3. NONMEM fit to Data-2 (six samples per subject, 10 subjects). The *solid line* connects trough levels predicted for the average individual in this study (79.9 kg man) from population pharmacokinetic parameters obtained with NONMEM. The *shaded area* represents ± SD of intraindividual variability. The area bounded by the *dashed lines* represents ± SD of both inter- and intraindividual variability.

drug, these levels contained very little information regarding V_{area}. Moreover, very few levels were measured during the accumulation stage and no peak-trough combinations were available to improve the amount of information regarding this parameter. Identical results were obtained by Sheiner and Beal[5] in simulations with the trough-only sampling design; the NONMEM estimate of the population mean V_{area} was approximately 20% higher than the true value. This bias can be reduced, however, by obtaining measurements according to a more random sampling pattern.

Although NONMEM is not able to estimate the population V_{area} very accurately from the data available, it is important to note that NONMEM does provide information that can be used to assess the precision of the estimate, namely the standard error of the estimate. NONMEM provides these values for both fixed and random effect parameters, and these can be used to estimate 95% confidence intervals for the population parameters. The larger the 95% confidence intervals, the poorer the precision of the estimate. Given the rather large estimates for the 95% confidence interval for V_{area}, it is clear that there is little information regarding V_{area} available in this data set and the estimate of V_{area} provided by NONMEM should not be trusted.

To evaluate the results obtained herein properly, one should ideally know the true values of the population pharmacokinetic parameters. Unfortunately, these true values are rarely known, except, of course, when simulations are performed. However, in a study to compare the performance of the STS approach with NONMEM in the analysis of simulated experimental data, Sheiner and Beal[4] found that both approaches yield acceptable (nonbiased and precise) estimates for the fixed effect parameters. Because a large quantity of experimental data was used to obtain the STS estimates in the current study, one can assume the estimates of the fixed effect parameters obtained with this method represent reasonable estimates of the true values and can serve as a standard against which the NONMEM estimates can be compared.

This same conclusion, however, cannot be made with regard to the estimates of interindividual variability. In the same study, Sheiner and Beal[4] found that the STS approach systematically overestimates these parameters. This is because each parameter is estimated from the original drug concentration-time data with some error. This error adds variability to the parameter estimates that is not biologic in origin, resulting in an upward-biased estimate of interindividual variability. NONMEM estimates of interindividual variability tend to be more reliable because they are not contaminated with the error involved in the estimation of individual parameters. Although NONMEM estimates of these parameters have been shown to be relatively unbiased, they are highly imprecise. This problem has been attributed to the small number of subjects generally included in such studies, because estimates of variability

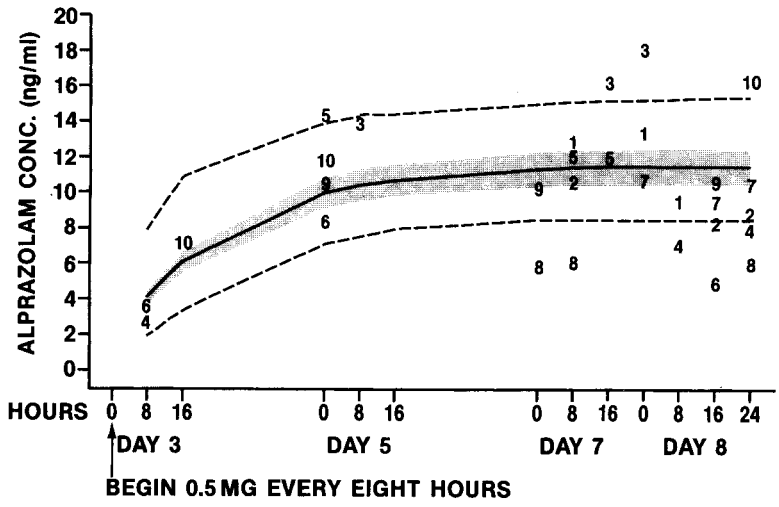

Fig. 4. NONMEM fit to Data-3 (three samples per subject, 10 subjects). The *solid line* connects trough levels predicted for the average individual in this study (79.9 kg man) from population pharmacokinetic parameters obtained with NONMEM. The *shaded area* represents ±SD of intraindividual variability. The area bounded by the *dashed lines* represents ±SD of both inter- and intraindividual variability. *Numbers* identify samples from specific individuals.

are considerably less precise, given a small number of subjects, than are estimates of means. Thus it is not known to what extent the estimates of interindividual variability obtained by either approach we used reflect the true values for these parameters.

When analyzing data measured during the accumulation to steady state with a multiple-dosing regimen, one can use the accumulation model as presented in Eq. 1. Indeed, the estimate of k_e obtained with this approach is very similar to the estimate expected on the basis of the results of the STS analysis of data obtained after the first and last dose. Use of the accumulation model, however, is limited because it allows only the estimation of k_e, and independent estimates of CL and V_{area} are not possible under this approach. The NONMEM estimate of k_e is (presumably) downward biased secondary to the estimate of V_{area}. As discussed previously, this is probably the result of the trough-only sampling scheme.

NONMEM was originally developed to be used in the analysis of population pharmacokinetic data, data consisting of a dosing history and only a few measured drug concentrations from a large number and variety of individuals. Inasmuch as the data analyzed herein were obtained from a relatively small and homogeneous population, our results should not be construed as population pharmacokinetic parameters. We have shown, however, that the data as typically collected during phase III and phase IV trials can be used for purposes other than to document compliance. To use these data most efficiently, however, the traditional approaches to data collection in this setting, i.e., sampling only trough levels, must be discarded.

Our results, specifically the difficulty encountered in the estimation of V_{area}, and the work performed by Sheiner and Beal[5] with simulated routine clinical data suggest that the following guidelines be used in setting up a population pharmacokinetic "study design."

First, samples should be obtained at random time points from each individual and *not* according to a rigid experimental protocol. As we have shown, the common method of measuring only trough levels is unnecessarily restrictive and limits the information that can be extracted from the data. Second, a minimum of two to four samples should be obtained from each subject, depending on the number of pharmacokinetic parameters to be estimated. However, data that consist of only one sample per subject are capable of supplying additional information when combined with more extensive data.[5] Third, a minimum of 50 to 100 subjects should be included and the population should be composed of patients representative of the population who will be receiving the drug for therapeutic purposes. This will ensure that a representative sample of patients is included and improve the estimates of interindividual variability. Finally, subjects can be receiving a variety of concomitant medications and diet should not be restricted; careful analysis of this data can provide in-

formation on a variety of possible drug-drug and drug-food interactions.

In the past, data analysts have been reluctant to draw conclusions based on the analysis of nonexperimental data because of the widely recognized problems associated with the analysis of such data. Thus dosing guidelines for newly released drugs and regulatory agency decisions are frequently established, in part, on the basis of traditional pharmacokinetic studies performed in normal subjects. However, recent experience with the drug benoxaprofen[10] suggest that these rather limited studies are inappropriate for the establishment of dosing guidelines for patients and can have tragic consequences.

Our results lend further support for the use of NONMEM in the analysis of data generated during phase III clinical trials and in other clinical settings. The incorporation of a NONMEM-like approach into phase III and phase IV clinical trials, in addition to the well-designed pilot studies currently performed, could speed the identification of populations at risk for toxicity caused by altered pharmacokinetics, and provide at least an initial quantification of the magnitude of this alteration. In this way we may reduce the risk of injury to patients exposed to new drugs.

References

1. Sheiner LB, Beal SL. Bayesian individualization of pharmacokinetics: simple implementation and comparison with non-Bayesian methods. J Pharm Sci 1982;71: 1344-8.
2. Sheiner LB, Rosenberg B, Marathe VV. Estimation of population characteristics of pharmacokinetic parameters from routine clinical data. J Pharmacokinet Biopharm 1977;5:445-79.
3. Sheiner LB, Beal SL. Evaluation of methods for estimating population pharmacokinetic parameters. I. Michaelis-Menten model: routine clinical pharmacokinetic data. J Pharmacokinet Biopharm 1980;8:553-71.
4. Sheiner LB, Beal SL. Evaluation of methods for estimating population pharmacokinetic parameters. II. Biexponential model and experimental pharmacokinetic data. J Pharmacokinet Biopharm 1981;9:635-51.
5. Sheiner LB, Beal SL. Evaluation of methods for estimating population pharmacokinetic parameters. III. Monoexponential model: routine clinical data. J Pharmacokinet Biopharm 1983;11:303-19.
6. Grasela TH, Sheiner LB, Rambeck B, et al. Steady-state pharmacokinetics of phenytoin from routinely collected patient data. Clin Pharmacokinet 1983;8:355-64.
7. Grasela TH, Sheiner LB. Population pharmacokinetics of procainamide from routine clinical data. Clin Pharmacokinet 1984;9:545-54.
8. Grasela TH, Donn SM. Neonatal population pharmacokinetics of phenobarbital derived from routine clinical data. Dev Pharmacol Ther 1985;8:374-83.
9. Smith RB, Kroboth PD, Vanderlugt JT, et al. Pharmacokinetics and pharmacodynamics of alprazolam after oral and IV administration. Psychopharmacology 1984; 84:452-6.
10. Grahame R. The rise and fall of benoxaprofen [Editorial]. Rheumatol Rehabil 1982;21:191-2.