Ann Rech Vét (1990) 21, suppl 1, 87s–92s

Statistical analysis of pharmacokinetic data with special applications to bioequivalence studies

J Powers, T Powers

Departments of Statistics and Veterinary Clinical Sciences, The Ohio State University, 1958 Neil Avenue, Columbus, OH 43210, USA

(Pharmacokinetics of Veterinary Drugs, 11–12 October 1989, Fougères, France)

Summary — The objectives of this investigation are: 1) to describe techniques for determining the validity of the assumptions; 2) to suggest data transformations which may validate the use of parametric procedures; and 3) to describe a non-parametric alternative to the analysis of variance for crossover designs. Two assumptions common to all parametric procedures include the underlying normal distribution of the observations and equality of variances across treatment groups. Normal probability plots and/or stem and leaf plots are good diagnostic techniques to address the assumption of normality, while Bartlett's test is the most common method of determining equality of variances. To evaluate bioequivalence data, the Food and Drug Administration suggests the use of analysis of variance for crossover designs. If the underlying assumptions are valid, the appropriate statistical models are well known. On the other hand, if the assumptions are not valid, the investigator has one of two choices: 1) transform the data in such a way as to satisfy the assumptions, or 2) use a non-parametric procedure. Square root or logarithmic transformations are commonly used in this situation. However, if a suitable transformation cannot be found, then a non-parametric procedure should be used. Koch (*Biometrics* (1972) 28, 577-584) developed a non-parametric crossover test, which is relatively easy to apply, but the corresponding power calculations required by the FDA are less obvious.

statistical analysis / bioequivalence / parametric and non-parametric tests

Résumé — Analyse statistique des données pharmacocinétiques et application aux essais de bioéquivalence. *Les objectifs de ce rapport sont : 1) de décrire les techniques utilisées pour vérifier les conditions d'application des tests, 2) de suggérer des transformations de données en vue de l'utilisation de tests paramétriques, et 3) de décrire une approche non paramétrique de l'analyse de variance pour les essais de bioéquivalence. Deux conditions sont requises pour la mise en œuvre des tests paramétriques : la normalité des distributions et l'homogénéité des variances pour les différents groupes. La normalité des distributions peut être vérifiée par différents tests graphiques. Le test de Bartlett est le plus généralement utilisé pour vérifier l'homogénéité des variances. Pour évaluer une bioéquivalence la* Food and Drug Administration *suggère d'utiliser une analyse de variance pour les dessins expérimentaux croisés. Si les conditions d'utilisation sont réunies, le modèle statistique à utiliser est bien connu. En revanche, si les conditions d'utilisation ne sont pas vérifiées, l'investigateur a le choix entre transformer ses données pour satisfaire les conditions d'utilisation du test ou utiliser une approche non paramétrique. Les transformations les plus généralement utilisées sont les transformations racine carrée et logarithmique. Si une transformation appropriée ne peut pas être trouvée, une approche non paramétrique doit être retenue. Koch (*Biometrics *(1972) 28, 577–584), a développé un test non paramétrique pour les dessins croisés. Ce test est relativement facile à appliquer mais le calcul de la puissance qui lui est associée est moins évident.*

analyse statistique / bioéquivalence / tests paramétriques / tests non paramétriques

INTRODUCTION

The use of statistics in the area of pharmacokinetics comprises, among others, two broad categories: 1) estimation techniques to obtain the coefficients of a specified model and/or the techniques necessary for non-compartmental analysis, and 2) tests of hypotheses to compare the estimated pharmacokinetic parameters for different treatment groups, as required in bioequivalence studies. The subject of model fitting and estimation of the coefficients of a given model is part of the more general area of non-linear regression. Regression theory almost exclusively relies upon the estimation procedure known as least squares, which is a special case of maximum likelihood with certain traits. On the other hand, non-compartmental analysis uses an estimation procedure known as method of moments. Both of these approaches lead to estimates of the pharmacokinetic parameters *per se*. Often the investigator wants to compare these estimates for different drugs, different formulations of the same drug or different routes of administration. When the researcher is investigating the properties of a new drug, an objective may be to determine if the kinetics of the drug are dose-dependent. In each of these areas, the pharmacokineticist finds solutions to his questions in the area of statistics.

STATISTICAL ASPECTS OF COMPARTMENTAL MODELING LINEAR MODEL

The one compartmental model is a straightforward problem of fitting a linear function. The method of least squares is the most commonly used estimation procedure. This model can be represented by (1):

$$\text{Ln } C = \text{Ln } a + bt + \varepsilon \qquad (1)$$

where Ln is the natural logarithm, C is the drug concentration and t is the time.

The objective of the method of least squares estimation is to minimize the function :

$$f(t) = \sum_{i=1}^{n} (\text{Ln } C_i - \text{Ln } a - bt_i)^2 \qquad (2)$$

This method is appealing not only because of the ease of obtaining the estimates of the slope and intercept but also because the resulting estimates are asymptotically normally distributed.

Once the estimates have been obtained, the investigator should use some criterion to determine if the resulting model 'adequately' describes the data. A plot of the residuals, the difference between the observed and predicted concentrations, against either time or observed concentrations is a very helpful diagnostic criterion. If this plot indicates the residuals are random, the model is adequate, *ie*, can be used as a predictor. Another fit criterion, Akaike's information criterion (AIC), will be discussed with non-linear models.

One additional point to be addressed is the assumption of constant variances. Basically this means, if repeated samples are taken at each time, the variances at each of these times are all equal. In most pharmacokinetic investigations, this is not the case. One way to correct this problem is to modify (2) by employing a system of weights:

$$f(t) = \sum_{i=1}^{n} W_i (\text{Ln } C_i - \text{Ln } a - bt_i)^2 \qquad (2a)$$

Non-constant variances are a problem not only in the linear case under discussion,

but also in non-linear models. The choice of the appropriate set of weights has been the subject of much statistical research. Judge *et al* (1985) presented an in depth discussion of this problem.

NON-LINEAR MODELS

Most of drug concentration–time profiles seen by the pharmacokineticist are represented by non-linear models. These models are not just an 'obvious' extension of the single exponential case to the sum of two or more exponentials, but instead the non-linear case contains a set of statistical challenges unknown to the linear model. The general non-linear model is represented by (3) :

$$C = \sum_{i=1}^{K} A_i e^{\alpha_i t} \qquad (3)$$

Non-linear models of the form given in eqn 3 require an iterative procedure to estimate the parameters of the model. Such procedures are widely available to the pharmacokineticist. In addition to all of the packages individually constructed for use specifically in kinetic investigations (*eg,* Metzler *et al* (1974), Beal and Sheiner (1984)), more general packages such as SAS contain iterative procedures with the capability of fitting data to such models (SAS Institute, 1985).

The problem of heteroscedacity (unequal variances across time) is persistent in this non-linear case as it was in the linear model. The incorrect choice of weights can lead to serious bias and/or lack of precision. Peck *et al* (1984) described a procedure called extended least squares (ELS) regression which minimized the expression:

$$\sum_{j=1}^{n} \frac{(C_j \sum_{i=1}^{K} A_i e^{\alpha_i t_j})^2}{\text{var } j} + \text{Ln (var } j) \qquad (4)$$

This function is in contrast to the ordinary least squares (OLS) or weighted least squares (WLS) expressions defined earlier. In addition, when iterative procedures are used, weights are recalculated with each iteration yielding an iteratively reweighted least squares (IRLS) function to be minimized.

The researcher must have an algorithm for choosing a model. This algorithm should contain criteria for selecting a particular model. It must be pointed out that no selection procedure can choose the 'correct model' but only selects the specific model(s) from a given set which best satisfies the selection criteria. Changing the selection criteria, the weighting strategy and/or set of models from which to make a choice can result in choosing different 'best' models. Thus, the problem of selecting a best model and best weighting strategy does not necessarily have a unique solution.

Two very commonly used selection criteria are the residual sum of squares (5) and Akaike's information criterion, AIC (6):

$$R_e = \sum_{i=1}^{n} W_i (\hat{C}_i - C_i) \qquad (5)$$

$$AIC = n \text{ Ln } R_e + 2p \qquad (6)$$

where W_i is the weighting strategy, n = number of data points and $\hat{C}_i$ and C_i are the predicted and observed concentrations, respectively, and p is the number of parameters estimated in the model (Yamaoka *et al,* 1978).

To select a 'best' model from a set of models under consideration, the investigator can choose that model with the minimum R_e or the minimum AIC (MAIC). By comparing eqns (5) and (6), it is noted that if the R_e for two models are approximately the same, MAIC selects the model which has fewer parameters to estimate.

COMPARTMENTAL *VERSUS* NON-COMPARTMENTAL ANALYSIS

Without discussing the statistical considerations of non-compartmental analysis, let us contrast the two methodologies. First, it has been stated many times that non-compartmental analysis is not model-independent. In fact, DiStefano and Landaw (1984) pointed out that non-compartmental analysis is based on a model with a far more restricted structure than multicompartmental models. Under the constraints of the non-compartmental model, they demonstrate the relationships of the parameters to their counterparts from the multiexponential models. Next, Landaw and DiStefano (1984) continue their comparisons in the context of data analysis and statistical considerations. These authors present a comprehensive discussion of modeling in general with good examples of the specific points being addressed.

HYPOTHESIS TESTING OF PHARMACOKINETIC PARAMETERS

The area of inference as related to the coefficients of a compartmental model is statistically straightforward. Least squares regression guarantees that the resulting estimates are asymptotically normally distributed. Because of this property, all parametric inference procedures are valid, including the *t* test for two populations and all analysis of variance models. Further, the property of asymptotic normality is valid for the weighted least squares procedure and also when it is necessary to use an iterative procedure, such as Newton Raphson or Marquardt.

In the case of bioequivalence or bioavailability studies, the guidelines issued by the regulatory agency request comparisons of pharmacokinetic parameters, such as half-life, volume of distribution, *AUC* and characteristics, such as time to maximum concentration and maximum concentration *per se*. Estimates of these parameters are not normally distributed and indeed their distributions are unknown. Typically, the sample size in this type of investigation is not large, so that there are serious doubts posed by using large sample theory.

Hypothesis testing, or equivalently the construction of confidence intervals, for coefficients of a compartmental model, may be accomplished by using parametric procedures such as the *t* test or an analysis of variance model. This is true if the estimates of the coefficients are obtained from an estimation procedure such as least squares. Estimates obtained for example using the trapezoidal rule, do not have this property; thus, unless the sample sizes are big enough to rely upon large sample theory, it is not valid to use parametric procedures. Winer (1971) has given detailed descriptions and numerical examples for a wide range of analysis of variance models encountered in pharmacokinetic studies. Typical examples include multifactor, repeated measures and crossover models. SAS has the capability of handling all of these models (SAS Institute, 1985).

Westlake (1972) has suggested an approach to evaluating bioequivalency using confidence intervals instead of a test of hy-

pothesis. Schuirmann (1987) has proposed a test procedure which employs two one-sided tests for assessing bioequivalency as compared to the standard 'power approach' (a test with a level of significance of 0.05 and estimated power of at least 0.80). The criticism of the power approach is that the type II error is large. If the test for coincidence of the interaction profiles, as described by Gill and Hafs (1971), were to be incorporated into the analysis of variance table for crossover designs, the power approach may be superior.

Let us suppose a researcher wishes to compare the half-life of two drugs in two groups of animals. The appropriate non-parametric test is the Mann–Whitney procedure. On the other hand, if estimates of the half-life of the two drugs are obtained from the same group of animals, then the appropriate test is Wilcoxon's rank sum. Hollander and Wolfe (1973) describe these tests in detail and provide the reader with numerical examples. Further, let us suppose the investigator wishes to compare the half-lives from '$k > 2$' drugs obtained from distinct groups: now Kruskal–Wallis is the test to be used. This test is the non-parametric counterpart to a one way analysis of variance and there are post tests available to the user. However, if these k half-lives were estimated from one group of subjects, then the Friedman test should be used to evaluate the data. Again, the necessary post tests are available. Hollander and Wolfe (1973) not only illustrate these tests but also have all of the necessary tables available.

The application of these non-parametric tests is not limited to the half-life of a drug, but was used only for illustrative purposes. All of the estimates of pharmacokinetic parameters, including *AUC*, volume of distribution, and *AUMC*, to name a few, should be evaluated using non-parametric techniques.

The FDA guidelines for bioequivalency tests include crossover designs to measure equivalency as measured by the pharmacokinetic parameters. The underlying assumptions necessary for the crossover analysis of variance are not satisfied, rendering this test invalid for small samples. Koch (1972) has developed a non-parametric method for two period change-over design. The availability of this method in a statistical computer package, such as SAS, is not known to this author at this time. However, Koch presents a numerical example to illustrate the procedure and it is clear that an automated package is not essential. More recently, Elswick and Uthoff (1989) developed a non-parametric test for the two treatments, two period, four sequence design.

Westlake (1972) has suggested using confidence intervals in evaluating bioequivalence data. Because of the lack of normality, the conventional techniques for construction of confidence intervals for estimates of pharmacokinetic parameters are not valid. Lam *et al* (1985) have used a jackknife technique to estimate the variance of the harmonic mean for half-lives. Using this estimate of the variance, they constructed a confidence interval for the harmonic mean of half-lives.

Using the property of normality of the estimates of the coefficients of a compartmental model, Bartoszynski and Powers (1990) constructed a confidence interval for the half-life of a drug. This confidence interval was shown to be of minimum length of all such intervals.

Much work needs to be done in developing confidence intervals for other pharmacokinetic parameters. Also, the extension of the Friedman test to a multi-way layout would be very helpful to the pharma-

cokineticist. The statistical community needs to direct some of its efforts toward determining the distributions of the estimates of pharmacokinetic parameters. From these efforts, precise confidence intervals and powerful inference procedures will be forthcoming.

REFERENCES

Bartoszynski R, Powers JD (1990) Estimation of confidence intervals for half-lives. *J Vet Pharmacol Ther* 13, 1-6

Beal SL, Sheiner LB (1984) NONMEM users guide, part I. University of California, San Francisco

DiStefano JJ III, Landaw EM (1984) Multiexponential multicompartmental and noncompartmental modeling I. Methodological limitations and physiological interpretations. *Am J Physiol* 246, R651-R664

Elswick RR, Uthoff VA (1989) *Biometrics* 45, 663-667

Gill JL, Hafs HD (1971) Analysis of repeated measurements of animals. *J Anim Sci* 33, 381-336

Hollander M, Wolfe DA (1973) *Nonparametric Statistical Methods.* J Wiley & Sons, New York

Judge GG, Griffiths WE, Hill RC, Lutkepohl H (1985) *The Theory and Practice of Econometrics,* 2nd edn. J Wiley & Sons, New York

Koch GG (1972) The use of nonparametric methods in the statistical analysis of the two-period change-over design. *Biometrics* 28, 577-584

Lam FC, Hung CT, Perrier DG (1985) Estimation of variance for harmonic mean half-lives. *J Pharm Sci* 74, 229-231

Landow EM, DiStefano JJ III (1984) Multiexponential, multicompartmental and noncompartmental modeling II. Data analysis and statistical considerations. *Am J Physiol* 246, R665-R677

Metzler CM, Elfring GL, McEwen AJ (1974) A users manual for NONLIN and associated programs. Upjohn, Kalamazoo, MI

Peck CC, Beal SL, Sheiner LB, Nichols AI (1984) Extended least squares nonlinear regression: a possible solution to the 'choice of weights' problem in analysis of individual pharmacokinetic data. *J Pharmacokinet Biopharm* 12, 545-558

SAS Institute (1985) *SAS User's Guide: Statistics,* 5th edn. SAS Institute, Inc, Cary, NC

Schuirmann DJ (1987) A comparison of the two one-sided tests procedure and the power approach for assessing the equivalence of average bioavailability. *J Pharmacokinet Biopharm* 15, 657-680

Westlake WJ (1972) Use of confidence intervals in analysis of comparative bioavailability trials. *J Pharm Sci* 61, 1340-1341

Winer BJ (1971) *Statistical Principles in Experimental Design,* 2nd edn. McGraw–Hill, New York

Yamaoka K, Nakagawa T, Uno T (1978) Application of Akaike's information criterion (AIC) in the evaluation of linear pharmacokinetic equations. *J Pharmacokinet Biopharm* 6, 165-175