REVIEW ARTICLE

Use of Statistical Methods in Evaluation of In Vivo Performance of Dosage Forms

WILFRED J. WESTLAKE

CONTENTS

Interest in the evaluation of the *in vivo* performance of drug dosage forms has increased greatly in the last few years, as an examination of recent volumes of this Journal will readily attest. The pharmacokinetic aspects of dosage form performance have received particular emphasis. In a typical experiment, the dosage form is administered to a group of subjects and blood samples are drawn at various sampling times following administration. Assay of the blood samples for drug content gives a sequence of drug concentrations which may be used to characterize the performance of the formulation. Such questions as how much of the drug does the formulation make available and how rapidly are of obvious interest and importance. When assay methods for blood drug levels are not feasible, an alternative approach consists of collecting and assaying urine specimens.

Methods of statistical analysis are becoming more widely used in the evaluation of such experiments, and there appear to be two main areas of application. The first is in the characterization of the behavior of the formulation in the biological organism by some kind of mathematical model; the second is in the comparison of the behavior of different dosage forms. Both types of application lead naturally to the use of certain statistical techniques. The first leads to estimation problems, particularly those of estimating the various parameters in a mathematical model. The second leads to problems in hypothesis testing, that is, in testing the hypothesis that no difference exists between two formulations. As discussed later in this article, a different approach to the problem of comparison of formulations—*via* confidence intervals—was recently suggested.

Olson and Lee (1) previously reviewed the application of statistical methods to quality control in the pharmaceutical industry. The purpose of the present article is to provide a similar review of the evaluation of *in vivo* performance of dosage forms. No attempt is made to refer to every article in which statistical methods were used. Rather, this article provides a critical review of proposed methods, some of which are now in wide use, with reference to selected papers. The author's opinion is that many statistical techniques have been applied

uncritically, albeit enthusiastically, to this field. The main aim of this review is, therefore, to expose the statistical problems and proposed methods of solution in a critical examination.

CHARACTERIZATION OF A FORMULATION

To examine the *in vivo* behavior of a formulation, one administers the formulation to a subject, draws blood samples at various times following administration, and obtains a sequence of blood concentrations, or blood levels as they are often called. The sequence of blood levels is then used to characterize the formulation. A number of methods of characterization are possible and have been used by various workers. The most popular method, by far, is that of characterizing the formulation by one or more estimated parameters (*e.g.*, an absorption rate constant) obtained from fitting a biologically meaningful mathematical model to the blood level data. In the great majority of cases, compartmental models are employed in which the biological organism is represented as a series of connected compartments (*e.g.*, blood compartment and tissue compartment) and transfers of drug among compartments are usually first order.

A second, purely empirical, method of characterization consists of fitting some mathematical function to the blood levels. In this case, the mathematical function is not intended to be a meaningful model of the biological organisms but only to give a good fit to the observed blood levels. However, the possibility exists—and, indeed, has been utilized—of choosing some parameter of the mathematical function to characterize the formulation. Other methods of characterization of the performance of the formulation, such as choosing some feature of the blood level sequence (*e.g.*, peak blood level or time of peak level) or some computed characteristic of the blood level sequence (*e.g.*, area under the blood level curve as a measure of amount of drug absorbed) are possible and are widely used. The statistical methodology that has been employed in utilizing these methods of characterization will be discussed, and some of the statistical problems will be examined.

Compartmental Models—The overwhelming choice of most workers who are faced with the problem of using a biologically meaningful model to characterize blood levels produced by a drug formulation is the compartmental model with linear transfers. One-, two-, and three-compartment models [see, for example, Chapter 38 of *Reference 2*, Riegelman *et al.* (3), and Nagashima *et al.* (4)] figure prominently in the pharmaceutical literature. Taking the one-compartment open model as an example, one can derive the expression:

$$C(t) = \frac{Dk_a}{V(k_a - k_e)} (e^{-k_e t} - e^{-k_a t}) \qquad (Eq. 1)$$

for the concentration of drug in the blood at time t. The symbol D represents the amount of drug absorbed, V is an apparent volume of distribution of the blood, and k_a and k_e are absorption and elimination rate constants, respectively. The salient feature of this expression, as it is in all compartmental models, is that it is the sum of a number of exponential terms.

The statistical problem is that of finding the values of the unknown parameters D/V, k_a, and k_e which, in

some way, best "fit" the observed data. Again, in the overwhelming majority of cases, the technique of fitting used is that of "least squares," *i.e.*, the estimation of the parameters such that they minimize the sum of squares of differences between observed and predicted blood level values. The method has a number of desirable statistical features but is not the only method possible. Determination of this minimum sum of squares results in three simultaneous equations to be solved for the three parameters. If these equations were linear in the parameters, one could fall back on a well-developed theory for their solution; but in the case of the present example and all compartmental models, the equations are nonlinear in the parameters and little general theory is available for solution of the equations. Therefore, the equations are usually solved by iterative methods. That is, an initial rough estimate of the parameters is obtained and one then proceeds, iteratively, to find a sequence of solutions for which the residual sum of squares becomes progressively smaller. When no further substantial diminution in the sum of squares appears possible, the process is said to have converged to the required least-squares solution.

The initial estimates of the pharmacokinetic parameters are usually obtained by the familiar process of "peeling off" the exponential terms one at a time from a semilogarithmic plot of the blood levels. The technique is frequently mentioned in the pharmaceutical literature; for a systematic approach that has been programmed for a digital computer, reference may be made to Foss (5). Once initial tentative estimates of the pharmacokinetic parameters have been obtained, several techniques of proceeding to the least-squares solution are available. Most of these are based, at least in part, on the so-called Gauss Newton method, which involves linearizing the problem in the neighborhood of the initial estimates and computing by linear least-squares techniques the solution giving the "minimum" sum of squares. One then proceeds, iteratively, from the new solution through further improved solutions until convergence is attained.

Hartley (6) gave a modified version of this technique and Marquardt (7) proposed a method based both on the Gauss Newton method and the method of "steepest descent." Numerous programs based on these methods [*e.g.*, Metzler (8) and Marquardt (9)] are available and are widely used in fitting compartmental models to blood level data. Smith and Shanno (10) produced an improved version of Marquardt's original method, and Ackerman *et al.* (11) proposed a method of solution termed "iterative guessing." However, they gave no reasons for considering their method superior to the other methods already discussed. Lawton and Sylvestre (12) suggested the possibility of taking advantage of any linearities that occur in the functional equations. Since only one term, the D/V introduced earlier, appears linearly in compartmental models, their suggestion would appear to have little utility in the solution of the problems discussed here.

Other methods of fitting sums of exponentials to data are available. Two examples are Cornell's (13) method, which requires equal spacing of the observations in time, and Foss's (14) method, which gives a noniterative

scheme for sums of two or three exponentials not restricted to equal spacing.

An important concern is the reliability of the estimates of the parameters given by the least-squares solution. The usual way of approaching this statistical problem is through the construction of a "confidence region" for the parameter estimates. Briefly, a 95% confidence region, for example, is a region in the parameter space that is so constructed that in repeated trials the true values of the parameter will lie in the confidence region in 95% of the cases. As an example, consider the two-compartment open model in which the five parameters, D/V, k_a, k_{12}, k_{21}, and k_e—or, equivalently, D/V, k_a, k_{21}, α, and β (where $\alpha\beta = k_e k_{21}$ and $\alpha + \beta = k_e + k_{12} + k_{21}$)— are to be estimated. In this case, the confidence region is a five-dimensional ellipsoid and the practical difficulties of interpreting such regions lead, in most programs, to a series of confidence statements for the individual parameter estimates based on the concepts of tangent planes to the ellipsoid. In addition, one can make the straightforward confidence interval statements for each parameter estimate separately, as was done by Maas and Patel (15), but these statements ignore the substantial correlations that usually exist between parameter estimates whenever one is fitting a sum of exponential terms. One result of the strong correlation existing between the parameter estimates is that confidence ellipsoids or, equivalently, the ellipsoids corresponding to the contours of constant residual sum of squares are greatly elongated. In practical terms, this means that one may be able to change the parameter estimates substantially without changing the residual sum of squares or, in other words, that many very different sets of estimates of the pharmacokinetic parameters give about equally good fits of the model to the data.

It should be apparent from this brief discussion that, in most cases in which the equation for blood concentration obtained from a compartmental model is fitted to experimental data, very little faith should be placed in the estimated pharmacokinetic parameters *per se*. The problems inherent in fitting sums of exponential terms to data have been recognized by mathematicians and statisticians for a long time, and the well-known book on numerical analysis by Lanczos (16, chap. IV) contains an excellent discussion of the subject. In view of its great significance to would-be fitters of compartmental models, it is worth spending a little time on the example given by Lanczos.

Lanczos gives a set of 24 data points (accurate to the second decimal place) generated by a function that is the sum of a number of exponentials; the data are error free except for the round-off to the second decimal place. An exact method for estimating the coefficients (amplitudes) of the exponential term and the values of their exponents is used. Lanczos starts tentatively with a three-term exponential model but finds that the series of observations needs only two exponential terms to describe it. He then finds the values of the parameters of the two-exponential term model that best fit the data and is able to generate a near-perfect fit. When the residuals (the differences between the actual data value and the value predicted by the fitted model) are examined, only one of them exceeds 0.005—it is, in fact,

0.006—and the square root of the residual mean square error is 0.0026. One could hardly expect a more perfect fit. Then Lanczos drops his bombshell: not only is the fitted model incorrect, but it does not even have the correct number of terms! In fact the data were generated from an expression that was the sum of *three* exponential terms. Not only was the exponential term with the smallest coefficient lost in the fitting process, but the other two terms of the original function were very different from the two exponentials as actually fitted. Lanczos decides that the result is thus completely unsatisfactory and his brief summation deserves to be quoted in full (p. 279).

It would be idle to hope that some other modified mathematical procedure could give better results, since the difficulty lies not with the manner of evaluation but with the extraordinary sensitivity of the exponents and amplitudes to very small changes of the data, which no amount of least square or other form of statistics could remedy. The only remedy would be an increase of accuracy to limits which are far beyond the possibilities of our present measuring devices.

This quotation might serve as a suitable warning to all who embark on the task of fitting compartmental models to blood level data, especially since the pharmacokineticist is faced with a far more severe problem than that illustrated by Lanczos in his example. Not only is round-off error present, but assay error is also inevitable in determining the blood samples. However, there is worse to follow in view of the fact that no biological organism can be expected to behave as a theoretically perfect compartmental model. One should be prepared to think in terms of another component of error, namely, lack of fit of the actual organism to the theoretical model.

The pharmaceutical literature abounds with examples of blood level data to which the authors have fitted compartmental models by the method of least squares. The foregoing discussion suggests that rather than confining their attention solely to a least-squares solution, authors would be well advised to give confidence limits, alternative solutions with almost equal sums of squares and, additionally, solutions for alternative models having almost equally good fits. Only in this manner would it be possible to form some idea of the validity of the proposed solution from a purely statistical standpoint. Naturally, if extrastatistical arguments (*e.g.*, physiological) can be adduced, the nature of the problem changes.

Methods other than least squares have been proposed for fitting sums of exponentials to blood level data. Gardner *et al.* (17), for example, proposed a method based on Fourier transforms. To date, however, methods other than least squares have not been widely used in the characterization of blood levels by compartmental models. Ackerman *et al.* (11), discussing briefly the problems already raised in this paper and referring to the literature on multicompartment analysis of tracer experiments (18), stated: "Complex transform methods have been attempted but none appear to have proved especially helpful with real biological data."

One aspect of the characterization of a drug formulation by compartmental models that has received little emphasis to date is the possibility of arranging the blood sampling times so as to estimate only certain specified parameters with high precision. For example, the essential characteristic of an orally administered formulation may lie in one parameter only, namely, an absorption rate constant. If one is allowed only a certain number of blood samples from each subject, the best procedure might be to arrange them so as to give the maximum precision in the estimation of this absorption rate constant at the expense of poorer precision for the other parameters. The problem of designing experiments where the underlying model is nonlinear in the model parameter has attracted some attention in the statistical literature but little, as yet, in pharmaceutical applications. Box and Lucas (19) wrote perhaps the first paper on the topic, and since then a number of papers by associates of Box have appeared. The paper by Box (20) extended the theory and also provided useful additional references. A characteristic requirement for the construction of an optimum design in a nonlinear situation is the need for some idea of the magnitude of the parameters that one is attempting to estimate. In practice, pilot studies would normally be carried out from which such rough estimates could be obtained, and this field appears potentially ripe for further investigation.

Normally, one estimates the parameters in a compartmental model by fitting the model to blood level data since this is the only compartment that can be sampled at varying points in time. An interesting possibility is that of obtaining data from other compartments of the model by indirect means. Wagner *et al.* (21) demonstrated that for subjects receiving the drug LSD-25, there appears to be a simple linear relation between the scores obtained by the subjects on arithmetic tests and the predicted concentration of drug in the so-called tissue compartment of a two-compartment open model. Metzler (22) used this idea to fit a two-compartment model to both blood level and test score data simultaneously. In a two-compartment model, there are five parameters $(D/V, k_a, k_{12}, k_{21},$ and $k_e)$ to be estimated. If the assumption is made that the arithmetic test scores are of the form found in Eq. 2:

$$S(t) = 100 - F_1 C_T(t) \qquad \text{(Eq. 2)}$$

where $S(t)$ is the score at time t, $C_T(t)$ is the tissue concentration at time t, and F_1 is an unknown parameter, then there are six parameters to be estimated but the number of data points available is now the sum of the number of blood levels and the number of test scores. The model is fitted by the usual least-squares techniques, and no new statistical principles are involved. This procedure, when valid, may be viewed as a means of increasing the amount of experimental data available for the fitting process.

In this discussion of compartmental models, attention should be drawn to the paper of Rodda *et al.* (23) which breaks new ground in several aspects of parameter estimation. Their paper dealt, however, only with one of the simplest possible cases, namely, the one-compartment model with oral drug administration. By simulation techniques, they examined the effects of assay error and of random variation in the pharmacokinetic parameters (physiological variation) on the least-squares estimates of the parameters. Their results suggest that the least-squares estimates do, in fact, give unbiassed estimates of the true underlying parameters. However, their estimates of the sampling variance of these parameters are sufficiently different from those obtained by the usual linearization techniques used in most nonlinear least-squares programs to suggest that the confidence intervals given by such programs may be substantially in error. This paper also examined a favorite device of workers in this field, namely, "curve averaging"—the practice of fitting a compartmental model to blood level data that are taken not from one trial but are the mean values of a number of trials. Here, they found, not surprisingly, a distinct tendency for the estimates of the parameters to be biassed.

Finally, the authors (23) discuss an alternative estimation technique which appears to be less sensitive to "outliers" than the usual least-squares estimation. An outlier is an observation subject to considerably greater error than most of the other observations. As should be only too apparent fom the earlier discussion, an outlier in a series of blood level values can have a disastrous effect on the least-squares estimates of the pharmacokinetic parameters. Outright discarding of apparent outliers is open to the objection of arbitrariness and the charge of attempting to make the data fit preconceived theories; the outliers may, in fact, be legitimate observations.

The contents of this paper have been discussed in detail because it is one of the few attempts so far made to examine some of the severe statistical and computational problems that beset the user of compartmental models. However, since it dealt only with one of the simplest possible compartmental models, much statistical examination yet remains to be done to counter the uncritical manner in which workers in this field have embraced compartmental models and nonlinear least-squares programs that fit them to their data. In view of Lanczos's sobering example, an interesting simulation study would be that of fitting two-compartment models to data generated from one-compartment models (and vice versa), with an accompanying examination of the relationship between the true and estimated pharmacokinetic parameters. By using intravenous infusion to mimic known absorption rates, Loo and Riegelman (24) demonstrated that the choice of the wrong model can lead to substantial errors in the estimates of the absorption rate constant. The clear implication is that blood level data must be fitted to the "correct" model for accurate characterization of a dosage form. However, the earlier discussion and, in particular, the example given by Lanczos imply that statistical and mathematical manipulation of the data may provide little guidance in determining this model. To demand that a nonlinear least-squares program decide on the appropriate model and estimate its parameters solely on the basis of statistical criteria is, demonstrably, unrealistic. Moreover, even if nonstatistical criteria (for example, physiological factors) can be adduced to support the validity of a particular model, the severe problems associated with the fitting of the model

parameters to the blood level data remain. The question of the sensitivity of exponential curves to their parameters was discussed in a recent paper by Julius (25).

Empirical Characterization—A feature of the characterization of blood level data by the compartmental model is that the model is intended to represent, however imperfectly and approximately, the biological organism and its handling of the drug. One can, of course, characterize the blood level data in a purely empirical manner by any type of mathematical function that gives a good fit to the data. Westlake (26) suggested that if the characterization is required only to predict future blood levels (e.g., from various regimens), then a simple mathematical representation is adequate. Wold (27) proposed a more elaborate, but still empirical, characterization based on spline functions. Wold used a polynomial spline function to characterize blood levels of a group of subjects receiving various modifications of a semisynthetic penicillin. Briefly, a polynomial spline function is a function that is, piecewise, a polynomial of specified degree. The pieces join each other in points called knots and are subject at the knots to various continuity conditions on the function itself and its derivatives. After a logarithmic transformation of the blood levels, Wold was able to obtain a good fit to blood level data for the 60 subjects in his experiment using a spline function of two pieces, although he encountered numerical difficulties in trying to fit the sum of two exponential terms to the same data sets. Even though the use of spline functions is an empirical approach, Wold suggested taking this idea one step further and using certain features of the function (such as time of peak and final slope) as representative of "meaningful" biological parameters (such as absorption rate and elimination rate). He cited the greater stability of these estimates within subject groups (on the same modification of the drug formulation) in comparison with estimates of the traditional pharmacokinetic parameters as a great advantage. And he considered the disadvantage, namely, the lack of any direct relationship to the parameters of a compartmental model, to be unimportant and mainly a question of the experimenter's familiarization with this new idea.

COMPARISON OF FORMULATIONS

The previous section dealt with the problem of characterizing the in vivo performance of a formulation of a drug. In many cases, however, one is less concerned with characterizing a formulation per se than with comparing its performance with that of another formulation. This is particularly the case of an abbreviated new drug application (NDA) submitted to the Food and Drug Administration (FDA), where a part of the submission is a comparative blood level trial in which the new formulation is compared with an already approved "standard" formulation. The general problem is that of comparing the in vivo behavior of two or more formulations of the same drug by conducting trials in which the formulations are administered to subjects and blood samples (or urine samples) are taken at various times following administration.

A major problem in comparing the performance of the formulations is deciding which characteristics of the

blood level sequence should be compared. The following list of possibilities is suggested:

1. Comparison of the formulations with respect to estimated parameters (e.g., an absorption rate constant or area under the curve as a measure of the amount of drug absorbed) in a biologically meaningful model, such as a compartmental model.

2. Comparison of the formulations with respect to some empirical, but presumably meaningful, characteristic of the blood level sequence (e.g., peak blood level achieved or time of peak blood level).

3. Comparison of the formulations solely with respect to the sequences of blood levels that they generate. Essentially, the comparison of formulations is treated as a "repeated-measurements" experiment (the statistical terminology for a sequence of observations—in this case, blood levels at various sampling times—on each item in the trial). The essential feature of such a method of analysis is that one does not pick out any "meaningful" feature of the blood level sequence for special treatment.

Trials in which the in vivo performance of formulations are compared have come to be known as (comparative) bioavailability trials. The term bioavailability denotes different things to different workers. For a limited definition as "relative absorption efficiency," see Reference 2 (chap. 25). Metzler (28) suggested that, in its broadest use, "the term includes the study of the factors which influence and determine the amount of active drug which gets from the administered dose to the site of pharmacologic action as well as the rate at which it gets there." Metzler gave an excellent, brief summary of the statistical problems arising in comparative bioavailability trials. Underlying all comparative studies of the in vivo performance of different formulations lies a question that is of the utmost importance: If statistical analysis detects "significant" differences between formulations, are these differences of any clinical or therapeutic importance? Without an adequate answer to the question of what differences in blood level characteristics or estimated pharmacokinetic parameters correspond to significant clinical or therapeutic differences, all comparative blood level trials must, to a certain extent, be academic exercises.

Design of Comparative Blood Level Trials—The design of comparative blood level trials is based, at least in its statistical aspects, on the well-established principles laid out by Cochran and Cox (29), for example. Designs may employ parallel groups in which each subject of a group receives a given formulation, or they may be of the crossover type in which each subject receives all formulations, in some sequence, with a suitable wash-out period between administrations of the different formulations. The crossover design has achieved almost universal acceptance because, as in much biological experimentation, subject-to-subject variation is considerable. Use of the crossover design, as contrasted with the design in parallel groups, for example, enables one to eliminate subject-to-subject variation in the analysis, thereby achieving a more sensitive test of the formulations. The various ways of laying out the assignment of formulations to subjects were reviewed by Westlake (30). Briefly, one can assign all formulations to

each subject in a random sequence (randomized blocks), or the administration of formulation to subjects can be made in sequences that form a series of Latin squares. A further possibility, not widely used to date, is that of assigning formulations in a balanced incomplete block design. In this case, each subject does not receive all of the formulations (if more than two are involved in the study) but receives at least two of them. The analysis is only slightly more complicated than the more usual Latin-square design, and the crossover concept, which enables one to isolate the subject-to-subject variation, is preserved.

Once the appropriate design has been selected, the remaining problem is determining the appropriate number of subjects for the trial. Again, there is a standard statistical methodology for achieving this end. (See, for example, chap. 11, *Reference 29*.) One needs: (*a*) an estimate of the minimum difference between two formulations with respect to some characteristic that one wishes to detect, (*b*) a probability of detecting this difference should it occur, (*c*) a significance level (commonly 0.05) at which the test of the null hypothesis of no difference between the formulations is to be made, and (*d*) an estimate of the residual error standard deviation, which is a measure of the inherent variability in the data. Usually, such an estimate will be available from preliminary work or from a pilot experiment. An elementary discussion of the problem of determining the number of subjects required in the trial was given by Westlake (30).

Analysis of Variance in Comparative Blood Level Trials—The basic method of analyzing a crossover trial in a Latin-square design is an analysis of variance (ANOVA). For example, suppose that it is desired to analyze a characteristic for which there is one value per administration of a formulation to a subject. Examples are an estimated pharmacokinetic parameter, an area under the blood level curve, and a peak blood level. Then the basic ANOVA partitions the total sum of squares of the observations about their mean into a component due to the subjects, a component due to the different days of administration, a component due to the formulations, and a component termed residual error. The formulation sum of squares is tested against the residual error sum of squares to check for differences among formulations, and, of course, similar tests are available to test for differences between individual pairs of formulations.

Variations on the basic ANOVA are possible and correspond to various factors in the assignment of subjects to the respective sequences of formulations that they receive. Cochran and Cox (29, chap. IV) summarized the variations on the basic crossover design and its analysis. Wagner (31) mentioned three of these variations in connection with blood level trials. The variation that he designated as Type 1 is the same as that discussed above with a further component, due to sequence of administration, extracted from the sum of squares due to subjects. Type II, on the other hand, contains one essential point of difference from the basic design. Subjects are assigned to Latin-squares in a nonrandom manner; for example, one might split them up into fairly homogeneous groups with respect to weight

and assign each group to a Latin-square. One can then extract from the subjects' sum of squares a component due to Latin-squares which enables one to test whether the weight of the subjects was a factor of importance. One feature of the analysis is that instead of extracting one overall "day of administration" effect, a day effect for each Latin-square is generated. This has the effect of reducing the number of degrees of freedom available for the error term and it might be questioned whether one would really expect the day effect, if present, to vary from square to square. If the day effects are collapsed into one overall day effect, Type II becomes essentially the same as Type I. The Type III crossover design can be analyzed as in Type I, but it can also be subjected to a much more elaborate form of analysis which adjusts for potential "residual effects," that is, the possible carryover from one administration to the following administration. This type of design has potential application in certain clinical trials, but for blood level trials there is already a very direct check on residual effects, namely, the zero-hour blood level, which is customarily taken at the time of administration of the formulation.

Another variation based on an analysis given in Lindquist (32) was used by MacLeod *et al.* (33) in their study of the bioavailability of three brands of ampicillin. Their analysis extracts from the error term a component that can only be described as a "within subject interaction between formulations and days." One would not, perhaps, be greatly interested in testing this component, but its removal from the error term could possibly affect the test on formulations, since the error term would now be different. This component can only be obtained if the number of formulations compared in the Latin-square design is greater than two. The "among subjects interaction between formulations and days," which the authors extracted from the subjects' sum of squares, is the same thing as the sequence component referred to earlier. Finally, the analysis of a crossover trial arranged as a balanced incomplete block design was given by Westlake (30).

Validity of Analysis of Variance and Transformations—Although the basic analysis of the crossover design has been discussed, the question now arises of whether the data should be transformed in some manner before analysis. To answer this question, an understanding of the underlying structure implicit in the analysis is necessary, namely, that any observation is made up of the sum of an overall mean, a component due to the particular subject, a component due to the day, a component due to the formulation, and a random error term. The error terms for the various administrations are assumed independent and normally distributed with zero mean and uniform variance. A characteristic that is often analyzed in a crossover trial is an estimated pharmacokinetic parameter (*e.g.*, an absorption rate constant), the object being to determine if the formulations differ with respect to this parameter. In numerous examples in the literature, the analyses were carried out in terms of the basic ANOVA. If each experiment yielded an observed value of the parameter, the analysis might be appropriate since it could reasonably be assumed that the observed value was equal to the true value for that subject–week–formulation combination plus a random

error term. In actual practice, of course, the parameter has not been observed but estimated from data; consequently, a second error term applies to the estimated parameter. This can be expressed as: estimated parameter value equals true mean parameter value plus e_1 plus e_2, where e_1 is the error term expressing random fluctuation of the parameter value about its mean value for that subject–day–formulation combination, and e_2 is a sampling error term which is dependent on the goodness-of-fit of the model to the data. An estimate of the variance of e_2 can be obtained directly when the parameter is estimated, and frequently this variance is far from uniform.

The foregoing discussion suggests that the use of ANOVA may not be appropriate and that a better course of action lies in analyzing the data with a test that does not require the assumption of uniform variance. A number of nonparametric tests are possible candidates. Westlake (30) suggested that, in a crossover trial, the simple sign test would be suitable for testing estimated pharmacokinetic parameters. Bernard *et al.* (34), in comparing half-lives and peak serum concentrations of minocycline in different groups of patients, used another nonparametric test, the Mann–Whitney U-test. An elementary reference to these and several other nonparametric tests is Siegel (35). An interesting possibility[1], which does not yet appear to have been studied, is that of comparing formulations with respect to the variance of some estimated pharmacokinetic parameter— not with respect to the mean value of the parameter as is usually done. A formulation that provided a much more variable value of the absorption rate constant, k_a, for example, might well be regarded as an inferior, because more variable, product.

The area under the blood level curve from zero to infinity is one of the most frequently analyzed characteristics of the blood level curve. Under the assumptions that elimination of the drug occurs only from the blood compartment and that the elimination is first order, Eq. 3 can be readily demonstrated:

$$\text{area} = FD/k_eV \qquad \text{(Eq. 3)}$$

where F is the proportion of the dose D of drug absorbed, while k_e and V are the elimination rate constant and apparent volume of distribution, respectively. Wagner (2, chap. 25) discussed several ways of transforming the areas prior to analysis. He suggested that, in a crossover design, one can obtain the ratio of the proportions of drug absorbed by taking the ratios of the areas. In this case, D cancels out of the ratio, as does the product k_eV if this is assumed constant for a given subject. However, the ensuing analysis must then ignore the "day effect," since this was absorbed into the ratio. An alternative approach, suggested by Westlake (30), is to take the logarithms of the areas which gives the relation found in Eq. 4:

$$\log \text{area} = \log F + \log D - \log (k_eV) \qquad \text{(Eq. 4)}$$

The log of the area may now be analyzed by the standard ANOVA, since the term $\log (k_eV)$ can be treated as an

additive effect due to the subject. It is not essential that k_eV remains exactly constant for a given subject but only that it be distributed around a mean value for that subject with either a relatively small variance or with a standard deviation proportional to the mean.

Wagner (2) suggested performing the analysis of the data after a transformation which involves correction for the volume of distribution and elimination rate constant. My understanding of this method is as follows. The area is multiplied by the product of W (body weight) and the decay constant, as estimated from the final exponential decay of the blood levels, and then divided by the dose. For a one-compartment model, this transformation results in Eq. 5:

$$\frac{(\text{area})(Wk_e)}{D} = \frac{(FD/Vk_e)(Wk_e)}{D} = \frac{F}{V/W} \qquad \text{(Eq. 5)}$$

Thus, on the assumption that V/W is constant and k_e has been correctly estimated, the transformed area can be analyzed as a measure of fraction of dose absorbed. However, if the behavior of the subject is more closely modeled by a two-compartment model, one sees that the transformation gives not $F/(V/W)$ but $F/(V/W)$ multiplied by k_{21}/α or, equivalently, by β/k_e. The approximate constancy of the mean value of the factor k_{21}/α in the group of subjects does not, of course, imply that this factor therefore cancels out in the analysis since it is multiplicative and not additive. In fact, the situation seems quite similar to the analysis of unadjusted areas proposed first. In the former case, one is analyzing FD/k_eV; with a logarithmic transformation and the assumption that k_eV for a given subject is approximately constant or distributed with relatively small variance, one can proceed to the standard ANOVA. In the case where areas are adjusted in the manner just described, one is analyzing $[F/(V/W)](k_{21}/\alpha)$ and, again, a logarithmic transformation and the assumption that both V/W and k_{21}/α are distributed with a relatively small variance about a mean value allow one to proceed to the ANOVA. It is thus difficult to see, on statistical grounds, what the adjustment of areas has achieved, especially since one is also making the assumption that β (the final exponential decay constant) has been adequately estimated from the data. However, the adjustment of areas might be an appropriate procedure if it could be demonstrated that the product k_eV were considerably more variable in a given subject than the quotient of the pharmacokinetic parameters k_{21}/α.

Analysis of the area under the blood level curve, when feasible, is likely to be of major importance in any comparative blood level trial in view of its direct relation to the amount of drug absorbed. Other parameters that are not dependent on any model of the drug distribution, such as peak blood level and time at which it occurs, are sometimes studied since they characterize the performance of the drug formulation. One either picks the peak blood level occurring among the samples or attempts to interpolate it in the sequence of sample values. Little appears to have been written on the problems associated with interpolation of maximum blood levels, although the spline-function approach, suggested by Wold (27), should be of value in this regard.

In addition, the question can be asked: Should such

---

[1] Suggested by Dr. K. R. Heimlich, personal communication.

data be transformed before subjecting them to an analysis of variance? Again, little seems to have been done on this subject, although Westlake (30) tentatively suggested that *any* concentration data can be regarded as possible candidates for a logarithmic transformation. The reason put forward is that for any pharmacokinetic model the apparent volume of distribution will appear in the denominator of the expression for drug concentration in the blood. Thus, if volume of distribution is roughly constant for a subject and is to be treated as an additive effect in the ANOVA, then the data cannot be expected to conform to the ANOVA model without some prior transformation. Logarithmic transformation will certainly have the effect of bringing the volume of distribution into the model additively, as required, but the question is a complicated one and no completely satisfactory answer appears forthcoming.

Analysis of Comparative Blood Level Trials as "Repeated-Measurements" Experiments—In the final type of analysis of blood level data to be considered, no particular features of the sequence of blood levels are picked out for special treatment. For each subject-day-formulation combination—in other words, for each administration of the drug—one is simply faced with a sequence of blood levels taken at various sampling times. The statistical term for such an experiment is a repeated-measurements experiment, the repeated measurements, of course, being the sequence of blood levels taken at repeated times. Numerous methods of statistical analysis are available and are based on no assumptions (implicit or otherwise) concerning the nature of the biological organism or the relative importance of features of the blood level sequence. Such methods have a certain appeal, particularly inasmuch as it usually is not clear which features of the blood level sequence are important or have a direct bearing on the therapeutic effect of the drug formulation.

The method most often used is that of performing the basic ANOVA at each sampling time, that is, a series of n ANOVA's where n is typically in the 6-12 range. Countless literature references could be given so that there is little point in singling out any one as an example. The problem, statistically speaking, is that one is forced to adjust the significance level to make this series of tests meaningful. For example, to test a null hypothesis (the hypothesis of no difference between formulations) at the 0.05 level means that one is accepting a 0.05 probability of deciding that the formulations are different when they are, in fact, the same. However, if there are, say, 10 sampling times, then it can be shown that if one rejects the null hypothesis on the grounds that any individual ANOVA shows a significant difference between the formulations at the 0.05 level, the probability of rejecting the null hypothesis even when true is about 0.40 (assuming independence of sampling times). The reason is, of course, that there are 10 opportunities to reject the null hypothesis; thus, to make sense of the series of tests, one must scale the significance level for individual tests down to a very much lower level. This procedure also has some unsatisfactory features, and the problem was discussed at length by Westlake (30).

A method of analyzing comparative blood level trials, which is in many ways more satisfactory, consists of viewing the trial as a "split-plot" design or, more accurately stated, as a "strip-plot" design. Details of the analysis of variance of a repeated-measurements design treated in this way were given by Winer (36). Briefly, the ANOVA comprises two sections, each with its appropriate error term for testing the various effects. The so-called "main plot" part of the analysis is, essentially, an analysis of mean blood levels (the mean being taken over all sampling times) and has the same structure as that given earlier for the basic ANOVA, namely a partitioning into components due to days, subjects, formulations, and residual. The so-called "subplot" is comprised of a partitioning into components due to sampling times and the interactions of sampling times with days, subjects, and formulations, respectively, and a subplot error term. Two tests, one in the main plot and one in the subplot, are of particular interest in the comparison of formulations. The test on the formulation mean square in the main plot tests whether there are differences among formulations with respect to their mean blood levels over the sampling times. The test on the (sampling times X formulations) interaction in the subplot tests whether there are differences among formulations with respect to the *patterns* of blood levels to which they give rise. Tests for difference between mean blood levels and patterns of blood levels for any given pair of the formulations studied in the design can also be made.

This method of analysis has a number of advantages, not least of which is that it produces one simple test on the mean blood levels and one test on the blood level patterns over time. Another advantage is that the concepts of mean blood level and blood level pattern are readily grasped by the nonstatistical scientists who must make the final decision concerning equivalency. There is an important restriction in that the method of analysis is only valid if the variance–covariance matrix of residuals has a particular form known as "uniform" (that is, the variances at the different sampling times are equal and all covariances between different sampling times are equal). A practical consequence of this is that the correlation between blood levels at different sampling times must be independent of the distance in time between the samples. In his book, Winer (36) gave details of a test, first devised by Box (37), to be carried out to determine whether the assumption of uniform variance–covariance holds. The author's experience (see *Reference 38*) with many orally administered drugs—in which samples are usually taken at intervals approximately an hour apart—is that the assumption is often valid, especially when a transformation of the data is made to ensure equal variances. This paper (38) also gives an extension of the analysis, including Box's test, to the case of comparative blood level trials arranged as balanced incomplete block designs. For the case where the uniform variance–covariance matrix requirement is not satisfied, Greenhouse and Geisser (39) gave an approximate test which could be useful in blood level trials.

Other more complex methods of analysis are, of course, available. Cole and Grizzle (40) discussed an example in which the effects of intravenous morphine sulfate and intravenous trimethaphan on blood hista-

mine levels of dogs were studied. Sampling times were at 0, 1, 3, and 5 min. and, perhaps not surprisingly with samples taken so close together in time, the uniform variance–covariance assumption was not valid. They discussed the application of a more complex method of multivariate analysis, namely the multivariate analysis of variance (MANOVA). Apart from their example, few applications of these techniques to blood level data appear to have been published in the literature. One drawback to routine application of MANOVA to typical blood level trials in which formulations are to be compared is the difficulty of interpreting these more sophisticated concepts so that they are meaningful to the nonstatistical scientists who must make the decisions based on the analysis.

Another possible method of analysis, described by Church (41), is based on principal components. The sequence of, say, n blood levels over the sampling times is treated as a vector of n observations. One then seeks a linear transformation of these n variables into a new set of n orthogonal (uncorrelated) variables such that the first variable accounts for the maximum amount of variability in the data, the second component for the maximum amount of the remaining variability, and so on. The advantage is that the new variables are uncorrelated and can be ranked in order of importance in accounting for the variability of the data. The dimension of the vector may also be effectively reduced if it is found that fewer than n principal components account for the bulk of the variability. Redman (42) discussed the application of principal components to blood level experiments. Snee (43) proposed a modification of Church's method in which the variation of the data is partitioned into two parts, the first connected with the mean of the blood levels and the second with the pattern of the blood levels. Principal components are then used to analyze the pattern component. A possible problem with the use of methods based on principal components is, as with MANOVA, the potential difficulties of conveying the statistical ideas to a nonstatistical user of the results. A second problem is that, instead of one simple test for pattern difference among formulations, there are a number of tests dependent on how many principal components are thought to be sufficient to account for the variation of the data. Thus, there is a greater likelihood that some difference among patterns will be detected as statistically significant, even though it may not have any meaningful relation to clinical or therapeutic significance. So at the end of this discussion of possible methods of analysis, one returns to the problem that plagues the whole subject of comparative bioavailability, namely, the relationship, if any, of differences between the blood level characteristics of the formulations to differences in their therapeutic effects. Until much more is known about this fundamental aspect of *in vivo* testing, unsatisfactory features will be connected with most statistical analysis.

Use of Confidence Intervals—All of the methods of analysis discussed so far have been formulated in terms of the statistical theory of hypothesis testing. To illustrate the simplest possible case, suppose that a crossover trial is conducted in which two formulations are to be compared with respect to some single parameter, such as area under the blood level curve as a measure of drug absorbed. If the basic ANOVA is used to analyze the data, one is essentially testing the null hypothesis that $\mu_1 = \mu_2$, where μ_1 denotes the mean amount of drug absorbed when the first formulation is administered to an essentially infinite population of subjects and μ_2 has the same meaning for the second formulation. Now it appears reasonable to assert that, in actual fact, μ_1 and μ_2 will never be equal since there will always be some differences, however slight, in the composition and methods of manufacture of the formulations. Consequently, if a sufficiently large number of subjects is employed in the crossover trial, one should be able to detect a statistically significant difference between any pair of formulations. This observation suggests that the idea of testing a simple null hypothesis of the type $\mu_1 = \mu_2$ may not be a very realistic approach to testing whether two formulations provide equivalent amounts of drug absorbed. The key lies in the word "equivalent" which, in a practical situation, means therapeutically or clinically equivalent. This might mean, for example, that the second formulation could be considered equivalent to the first if it resulted in an area under the blood level curve (taken to be a measure of absorption) that was within $n\%$ of that achieved with the first formulation. How small n must be to indicate therapeutic equivalence is, of course, a decision for the clinician to make; and as Levy (44) pointed out, the nature of n depends on the particular drug under consideration.

Suppose that for a given drug the value of n is taken to be 20; that is, the mean area under the blood level curve for the second formulation must be within 20% of the mean area for the first formulation in order to be judged "practically" equivalent. Then a suitable null hypothesis to test would be ($0.80\ \mu_1 \leq \mu_2 \leq 1.20\ \mu_1$) rather than simply $\mu_1 = \mu_2$. However, construction of a test poses considerable problems and an alternative approach is to abandon hypothesis testing methods completely and replace them with the concept of confidence intervals. A 95% confidence interval statement, for example, resulting from the analysis of the example under discussion would be of the type ($\mu_1 - k_1 \leq \mu_2 \leq \mu_1 + k_2$). Note that this gives an indication of where μ_2 is located in relation to μ_1; when the result of the analysis is stated in this form, it gives the physician or clinical pharmacologist information in a form from which he or she can make a judgment as to the practical equivalence of the formulations. Metzler (28) discussed the use of confidence intervals as a means of summarizing a comparative bioavailability trial and gave a number of examples, for each of which various criteria of acceptance are proposed based on a confidence interval approach. He then showed how certain formulations meet the acceptance criteria while some do not. Some that meet the acceptance criteria would be rejected as being statistically significantly different from the standard formulation against which comparison was made; some that would pass the hypothesis testing do not meet the acceptance criteria based on confidence intervals.

Many instructive examples can be constructed to

show how the confidence interval approach, which can be made the basis of a decision by the physician or clinical pharmacologist, leads to quite different results from the hypothesis testing approach, which is based only on the statistician's more-or-less arbitrary selection of significance levels. Defenders of the hypothesis testing approach will probably reply that, in planning the trial, the number of subjects is selected to allow detection of a certain minimum difference between the bioavailability of the formulations based on the error variance (estimated from pilot studies), significance level of the test, and required probability of detecting the minimum difference. This is true, but much of this careful planning may be negated if the variability turns out to be different from that estimated from the pilot experiments. In this case, the trial may be either incapable of detecting the minimum difference as statistically significant or, on the other hand, may show supposedly unimportant differences to be statistically significant. The confidence interval approach thus appears to provide a more satisfactory answer.

One disadvantage of constructing conventional confidence intervals, which are symmetrical about the difference of the population means of the two formulations, is that when expressed in the form given earlier, $(\mu_1 - k_1 \leq \mu_2 \leq \mu_1 + k_2)$, they are *not* symmetrical; that is, k_1 and k_2 are not equal. Many clinical pharmacologists would, however, make their equivalence statements in a symmetrical form, for example: the second formulation is clinically equivalent to the first if its bioavailability is within 20% of that of the first formulation. This implies that k_1 and k_2 are equal in a confidence statement of the type just given. Westlake (45) suggested that the traditional method of constructing confidence intervals could be modified so as to ensure a symmetrical statement in which k_1 and k_2 are equal, and he gave a method for carrying out this procedure.

The author's opinion is that confidence intervals currently afford the best means of summarizing the results of a comparative bioavailability trial since, unlike hypothesis tests, they portray the information in a manner that can form the basis for an informed judgment by a physician or clinical pharmacologist. Unfortunately, this only seems to be true when one is discussing a simple quantity such as amount of drug absorbed or mean blood level over time. More complex concepts, such as the difference between the blood level patterns over time, would not seem to be amenable to a confidence interval treatment at present. Theoretically, there is no difficulty; the problem lies in the practical interpretation of confidence interval statements concerning more complex functions of blood levels.

SUMMARY

The discussion of the statistical methods used in the evaluation of *in vivo* performance of dosage forms has covered the characterization of dosage forms by biologically meaningful models and purely empirical means and the comparison of formulations. The subject seems to have resulted in very few novel developments in statistical theory and, by and large, existing statistical methods seem more than adequate to cope with the

problems that arise in this particular field. It seems questionable whether new statistical techniques could improve matters in the estimation of pharmacokinetic parameters in compartmental models. Rather, as Lanczos (16) so elegantly demonstrated, only an immense refinement in measuring techniques can improve the estimation of parameters in a function that is the sum of exponential terms. Even if this were possible, one would still face the fact that any compartmental model is a grossly simplified model of the biological reality. For the comparison of dosage forms or formulations, numerous methods of analysis are available but it has been suggested that a confidence interval approach, rather than the more usual hypothesis testing, may offer more hope for a realistic appraisal of comparative dosage form performance. Eventually, it seems that the major problem in the use of statistical methods is a nonstatistical one. That problem is the elucidation of the relationship of blood level characteristics to therapeutic use of the drug. It is only with improved knowledge in this area that we can hope to determine accurately those blood level characteristics which are of really critical importance. Statisticians can perform many types of analyses on blood level sequences; some of them are of considerable sophistication and complexity. But the value and significance of such analyses can hardly be assessed until a clear idea of the relation between blood levels and therapeutic use of the drug is available.

REFERENCES

(1) T. N. T. Olson and I. Lee, *J. Pharm. Sci.*, **55**, 1(1966).
(2) J. G. Wagner, "Biopharmaceutics and Relevant Pharmacokinetics," Drug Intelligence Publications, Hamilton, Ill., 1970.
(3) S. Riegelman, J. C. K. Loo, and M. Rowland, *J. Pharm. Sci.*, **57**, 117(1968).
(4) R. Nagashima, G. Levy, and R. A. O'Reilly, *ibid.*, **57**, 1888 (1968).
(5) S. D. Foss, *Biometrics*, **25**, 580(1969).
(6) H. O. Hartley, *Technometrics*, **3**, 269(1961).
(7) D. W. Marquardt, *J. Soc. Ind. Appl. Math.*, **2**, 431(1963).
(8) C. M. Metzler, "NONLIN. A Computer Program for Parameter Estimation in Non-Linear Situations," Technical Report 7292/69/7292/005, The Upjohn Co., Kalamazoo, Mich.
(9) D. W. Marquardt, DPE-NLIN Share General Library Program No. 7-1354.
(10) F. B. Smith and D. F. Shanno, *Technometrics*, **13**, 63(1971).
(11) E. Ackerman, E. H. Strickland, J. B. Hazelrig, and L. C. Gatewood, *Clin. Pharmacol. Ther.*, **8**, 170(1967).
(12) W. H. Lawton and E. A. Sylvestre, *Technometrics*, **13**, 461 (1971).
(13) R. G. Cornell, *Biometrics*, **18**, 104(1962).
(14) S. D. Foss, *ibid.*, **26**, 815(1970).
(15) J. W. Maas and M. Patel, *J. Pharmacol. Exp. Ther.*, **174**, 369(1970).
(16) C. Lanczos, "Applied Analysis," Prentice-Hall, Englewood Cliffs, N. J., 1956.
(17) D. G. Gardner, J. C. Gardner, G. Laush, and W. W. Meinke, *J. Chem. Phys.*, **31**, 978(1959).
(18) *Ann. N. Y. Acad. Sci.*, **108**, 1(1963).
(19) G. E. P. Box and H. L. Lucas, *Biometrika*, **46**, 77(1959).
(20) M. J. Box, *Technometrics*, **12**, 569(1970).
(21) J. G. Wagner, G. K. Aghajanian, and O. H. L. Bing, *Clin. Pharmacol. Ther.*, **9**, 635(1968).
(22) C. M. Metzler, *ibid.*, **10**, 737(1969).
(23) B. E. Rodda, C. B. Sampson, and D. W. Smith, Abstract No. 2057, *Biometrics*, **28**, 1180(1972).
(24) J. C. K. Loo and S. Riegelman, *J. Pharm. Sci.*, **57**, 918(1968).
(25) R. S. Julius, *Comput. Biomed. Res.*, **5**, 473(1972).

(26) W. J. Westlake, *J. Pharm. Sci.*, **60**, 882(1971).

(27) S. Wold, *Chem. Scr.*, **1**, 97(1971).

(28) C. M. Metzler, Abstract No. 2043, *Biometrics*, **28**, 1176 (1972).

(29) W. G. Cochran and G. M. Cox, "Experimental Designs," Wiley, New York, N. Y., 1957.

(30) W. J. Westlake, in "Current Concepts in the Pharmaceutical Sciences, Dosage Form Design and Bioavailability," J. Swarbrick, Ed., Lea & Febiger, Philadelphia, Pa., 1973, chap. V.

(31) J. G. Wagner, *Drug Inform. Bull.*, **January/June**, 45(1969).

(32) E. F. Lindquist, "Design and Analysis of Experiments in Psychology and Education," Houghton Mifflin, New York, N. Y., 1956.

(33) C. MacLeod, H. Rabin, J. Ruedy, M. Caron, D. Zarowry, and R. O. Davies, *CMA J.*, **107**, 203(1972).

(34) B. Bernard, E. J. Yin, and H. J. Simon, *J. Clin. Pharmacol.*, **11**, 332(1971).

(35) S. Siegel, "Non-parametric Statistics for the Behavioral Sciences," McGraw-Hill, New York, N. Y., 1956.

(36) B. J. Winer, "Statistical Principles in Experimental Design," McGraw-Hill, New York, N. Y., 1962.

(37) G. E. P. Box, *Biometrics*, **6**, 362(1950).

(38) W. J. Westlake, Abstract No. 2072, *ibid.*, **28**, 1184(1972).

(39) S. Greenhouse and S. Geisser, *Psychometrika*, **24**, 95(1959).

(40) J. W. C. Cole and J. E. Grizzle, *Biometrics*, **22**, 810(1966).

(41) A. Church, Jr., *Technometrics*, **8**, 229(1966).

(42) C. E. Redman, Abstract 1467, *Biometrics*, **24**, 228(1968).

(43) R. D. Snee, *Technometrics*, **14**, 47(1972).

(44) G. Levy, *CMA J.*, **107**, 722(1972).

(45) W. J. Westlake, *J. Pharm. Sci.*, **61**, 1340(1972).

RESEARCH ARTICLES

Conformation of Sodium Cromolyn in Aqueous Solution Using Light Scattering and Magnetic Birefringence

J. V. CHAMPION and G. H. MEETEN[▲]

Abstract ☐ The conformation of the sodium cromolyn (disodium cromoglycate) molecule is possibly relevant to its pharmacological activity in the treatment of bronchial asthma. Two optical techniques, depolarized light scattering and magnetic birefringence, were used to study the molecular conformation in dilute solutions (<1 g./100 ml.) of unassociated molecules. Both techniques confirm that the molecule in dilute solution has a planar conformation as found in the crystal.
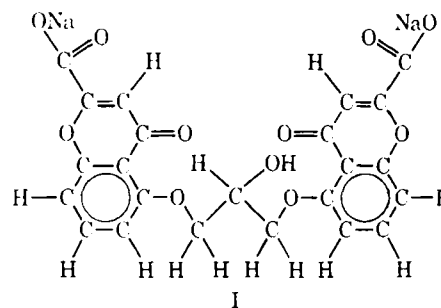
Keyphrases ☐ Sodium cromolyn (disodium cromoglycate)—molecular conformation in aqueous solution, light-scattering and magnetic birefringence techniques ☐ Molecular conformation, sodium cromolyn—determination in aqueous solution, light-scattering and magnetic birefringence techniques ☐ Light-scattering techniques—determination of sodium cromolyn conformation in aqueous solution ☐ Magnetic birefringence techniques—determination of sodium cromolyn conformation in aqueous solution

The conformation of the molecule of sodium cromolyn[1] (disodium cromoglycate) is of particular interest because the compound has been used as a new pharmacological approach to the treatment of bronchial asthma. The solid-state chemistry of sodium cromolyn was recently described (1), and the lattice parameters of the crystalline solid solution at high (90%) relative humidity indicate that the molecule is in a planar conformation. In the concentrated solution state, the formation of liquid crystals occurs.

The aim of this study was to determine the conformation of the molecule in aqueous solution. This may be important in assessing the activity and specificity of biological action at a membrane surface.

The chemical structure of sodium cromolyn in the planar conformation is shown here (I), and the molecule may be approximated to a pair of rigid chromones joined by a glycerol bridge [—O—CH$_2$—CH(OH)—CH$_2$—O—], the hydroxy, carbonyl, and ether groups giving many possibilities of hydrogen bonding, especially in the presence of water molecules. Simple molecular model building shows that numerous different conformations are possible by rotation about the bonds of the glycerol bridge, many of these having a small

I