and retention time. Evidently, the adsorption takes place at the beginning of the column in order to meet the criteria of obtaining the same peak shape and retention time as the standard glycol solution.

We found that as long as an ethylene glycol standard is employed in the same concentration range as is found in the sample, the results are quantitative and no significant error is obtained even upon duplicate injections. Again, it is important to keep in mind that this phenomenon does exist, and it should be considered in the quantitative as well as the qualitative analysis of ethylene glycol.

(1) A. Davis, A. Roaldi, and L. E. Tufts, *J. Gas Chromatogr.*, **2,** 306(1964).

(2) R. G. Ackerman and R. D. Burghen, *Anal. Chem.*, **35,** 647 (1963).

(3) E. D. Smith and A. B. Gusnell, *ibid.*, **34,** 438(1962).

(4) T. J. Lemoine, R. H. Benson, and C. R. Herbeck, *J. Gas Chromatogr.*, **3,** 189(1965).

(5) D. A. M. Geddes and M. N. Gilmour, *J. Chromatogr. Sci.*, **8,** 394(1970).

(6) H. D. Spitz and J. Weinberger, *J. Pharm. Sci.*, **60,** 271(1971).

HARVEY D. SPITZ
Johnson & Johnson Research Center
New Brunswick, NJ 08903

Received April 5, 1972.
Accepted for publication May 15, 1972.

Use of Confidence Intervals in Analysis of Comparative Bioavailability Trials

Keyphrases ☐ Bioavailability trials, comparative—analysis, use of confidence intervals ☐ Confidence intervals—analysis of comparative bioavailability trials ☐ Drug formulations, comparison—use of confidence intervals in analysis of comparative bioavailability trials ☐ Clinical equivalence of drug formulations—use of confidence intervals in analysis of comparative bioavailability trials

Sir:

The current emphasis on the comparative bioavailability trial, in which a new formulation of a drug is compared against a standard formulation in human subjects, is attracting considerable attention. Today such a trial is a key element in the submission of an abbreviated New Drug Application, in which the object is to show that the *in vivo* characteristics of the new formulation are essentially identical to those of an already approved, standard formulation. Typically, a crossover trial is employed, and blood levels of the drug at various times after administration and/or amounts of drug excreted in the urine are determined. Results are often analyzed using the classical statistical theory of hypothesis testing, in which the null hypothesis of no difference between the formulations is tested with respect to some characteristics of the blood or urine levels of the drug. Presumably, if the difference between the formulations turns out to be significant at the 0.05 or 0.01 level, one rejects the null hypothesis and concludes that the new formulation does not match the standard.

The purposes of this communication are to suggest that classical hypothesis testing techniques may not be particularly relevant to this problem and to propose, as an alternative, an approach based on confidence intervals. To bring the problem into focus, it is assumed that the new and standard formulations are to be compared with respect to total amount of drug excreted in the urine for some specified period following administration. If the trial is then run under tightly controlled conditions (resulting in small error variance in the analysis) with a large number of subjects, it could happen that the formulation would give a urinary recovery only 1% different from the standard but that this difference would still be significant statistically. It should be kept in mind that however small the difference between the formulations, it will be detected as significant if the trial is sufficiently well controlled and the number of subjects employed is large enough. In such a situation, the reviewing clinical pharmacologist or physician might well feel that a 1% difference in absorption is of no clinical significance, but he might also be concerned that this difference is termed significant in the statistical sense. The dilemma is occasioned by use of an inappropriate tool—hypothesis testing. What he may need is not a test of whether the two formulations are identical but rather some degree of assurance that the mean amount of drug excreted using the new formulation is sufficiently close to the amount excreted using the standard. The usual hypothesis testing approach does not give this assurance; nevertheless, the analysis can easily be modified to provide it in a manner that will be meaningful and that can provide the basis for a rational decision by the clinical pharmacologist. The proposed approach is based on the use of confidence intervals and is described here.

Suppose, to fix ideas, that the total urinary excretion of the drug (or the mean drug blood level over a number of sampling times) is to be analyzed in a crossover trial in 12 subjects, with both formulations—new and standard—administered to each subject. With the usual analysis of variance, which is based on the presence of additive effects due to subject (11 degrees of freedom), day of administration (1 degree of freedom), and formulation (1 degree of freedom), one can easily verify that the error mean square, s^2, is based on the remaining 10 degrees of freedom. If μ_s and μ_n are the true population means of the mean total urinary excretion of drug for the standard and new formulations, respectively, and $\bar{x}_s$ and $\bar{x}_n$ are the corresponding sample means obtained from the trial, then with the usual normality assumptions, $\{(\bar{x}_s - \bar{x}_n) - (\mu_s - \mu_n)\}/(s/\sqrt{6})$ has the t-distribution with 10 degrees of freedom. Two constants, k_1 and k_2, can be chosen so that the integral of the t-distribution from k_2 to k_1 is 0.95. Then with 95% probability, the inequality:

$$k_2 s/\sqrt{6} \leq \{(\bar{x}_s - \bar{x}_n) - (\mu_s - \mu_n)\} \leq k_1 s/\sqrt{6} \quad \text{(Eq. 1)}$$

holds. This inequality can be rearranged to give:

$$\{\mu_s + k_2 s/\sqrt{6} - (\bar{x}_s - \bar{x}_n)\} \leq \mu_n \leq \{\mu_s + k_1 s/\sqrt{6} - (\bar{x}_s - \bar{x}_n)\} \quad \text{(Eq. 2)}$$

In normal statistical practice, k_1 is chosen to be equal to $-k_2$ to give a symmetrical confidence interval for $\mu_s - \mu_n$. However, in this instance, one is presumably interested in a symmetrical confidence interval for μ_n of the type:

$$\mu_s - \Delta \leq \mu_n \leq \mu_s + \Delta \qquad \text{(Eq. 3)}$$

where $\Delta = \{k_1 s/\sqrt{6} - (\bar{x}_s - \bar{x}_n)\} = -\{k_2 s/\sqrt{6} - (\bar{x}_s - \bar{x}_n)\}$. This relation implies that k_1 and k_2 must be chosen so that the second and third terms in this equality are equal. Thus, the confidence interval for the mean urinary excretion with the new formulation will be given as an interval which is symmetric about the corresponding mean for the standard formulation. To achieve this end, one must set:

$$2(\bar{x}_s - \bar{x}_n) = (k_1 + k_2)s/\sqrt{6} \qquad \text{(Eq. 4)}$$

and determine k_1 and k_2 accordingly.

For example, suppose in a particular crossover trial in 12 subjects that $\bar{x}_s = 11.5$, $\bar{x}_n = 10.75$, and $s = 0.75$. Then substituting in Eq. 4, one finds that $k_1 + k_2$ must equal 4.90. By examination of a tabulation of the probability integral for the t-distribution with 10 degrees of freedom[1], it can now be determined that k_1 and k_2 must be approximately equal to 6.72 and -1.82, respectively. The value of Δ can then be computed to be 1.31. With 95% confidence, it can be stated that the confidence interval $\mu_s \pm 1.31$ covers the mean μ_n or, alternatively, if μ_s is approximated by $\bar{x}_s$ (*i.e.*, 11.5), that with 95% confidence the mean total urinary excretion of the drug for the new formulation is within 11.4% (*i.e.*, 1.31/11.5) of the mean for the standard formulation.

The latter statement is now in a form that the clinical pharmacologist can use as the basis for a rational decision as to the clinical equivalence of the two formulations. For example, he may have decided that if the new formulation comes within 15% of the standard, it is essentially equivalent. In the numerical example given here, he would presumably decide that the new formulation is essentially equivalent to the standard even though (as one can easily verify from the numbers given) the difference between them is significant at the 0.05 level. To summarize: a final decision as to the equivalence of two formulations of a drug must involve some judgment on the part of the clinical pharmacologist or physician. To this end, it seems appropriate that the statistician's analysis should result in a statement that the former can use as the basis for his decision. If classical hypothesis testing techniques are used, however, and decisions are made on the basis of significance tests, then the decision is essentially made by the statistician and the clinical pharmacologist's judgment is bypassed. It is true that his judgment may have played an important part in designing the experiment, based on the available information concerning the inherent variability of the data. However, the judgment may be largely negated if the variability turns out, in the actual trial, to be either much less or much more than anticipated.

One interesting result of the proposed approach is that if a manufacturer produces a formulation that is almost identical to the standard, a well-controlled experiment of adequate size will give small confidence intervals and a high probability of demonstrating practical equivalence. A poorly controlled experiment of inadequate size, on the other hand, will give confidence intervals so large that practical equivalence cannot be demonstrated. In the latter case, with a hypothesis-testing approach to the analysis, exactly the reverse of this situation is true.

WILFRED J. WESTLAKE
Biostatistics Department
Smith Kline & French Laboratories
Philadelphia, PA 19101

Received December 22, 1971.
Accepted for publication May 17, 1972.

[1] See, for example, "Biometrika Tables for Statisticians," Vol. I, E. S. Pearson and H. O. Hartley, Eds., Cambridge University Press.

Determination of Meprobamate in Dissolution Studies: Shortcomings of Direct GLC and Development of a New Assay

Keyphrases ☐ Meprobamate—analysis in solution *via* hydrolysis and silylation, compared to direct GLC method ☐ GLC—analysis, meprobamate in solution *via* hydrolysis and silylation, compared to direct GLC method

Sir:

USP XVIII introduced dissolution tests for seven kinds of tablets and, in the case of meprobamate, a GLC procedure was specified for the determination of the drug in solution (1). The USP recommendation reads: "... the amount in solution being determined on filtered portions of the *Dissolution Medium* by means of gas chromatography, the internal standard consisting of a solution of dibutyl phthalate in anhydrous ether containing 0.4 mg. per ml." This approach was used in our laboratory and it was found that, under a variety of conditions, meprobamate cannot be accurately quantitated by direct GLC because of breakdown of the drug in the injection port. The first supplement to USP XVIII (2) revised this analytical procedure and replaced it by a colorimetric assay adapted from a method recently described by Poole *et al* (3). The purposes of this report are to point out the shortcomings of the USP's GLC method and to propose a specific and sensitive assay for the determination of meprobamate in solution.

The GLC method originally recommended by the USP is very similar to that proposed by Douglas *et al.* (4) since both utilize direct GLC of meprobamate as well as quantitation with dibutyl phthalate as an internal standard. Because the USP did not specify any GLC operating conditions, it appeared reasonable to