This work was supported by a grant to T.N.C. from the Peel Medical Research Trust.

Department of Pharmacology and Therapeutics,
University of Liverpool,
Liverpool L69 3BX, U.K.

T. N. CALVEY
D. J. BACK

March 19, 1971

REFERENCES

GAUDETTE, L. E. & BRODIE, B. B. (1959). *Biochem. Pharmac.*, **2**, 89–96.
MAZEL, P. & HENDERSON, J. F. (1965). *Ibid.*, **14**, 92–94.
RANDALL, L. O. (1950). *J. Pharmac. exp. Ther.*, **100**, 83–93.
SOMANI, S. M., WRIGHT, A. & CALVEY, T. N. (1970). *Europ. J. Pharmac.*, **12**, 109–113.

A common error in assessing the significance of percentage change in neuropharmacology

It has become common practice to express values for drug-treated animals as a percentage of (or as a percentage change from) values observed in saline controls. This procedure would be straightforward and easily understood if one could measure the effect of drugs in each experimental animal both after a saline control treatment and after a drug treatment, or if there existed some rational justification for pairing individual saline and drug-treated animals and if the numbers of animals employed in each treatment were the same. However, the calculation and interpretation of percentages is not so simple when these requirements are not met.

We have noted that the standard errors of mean percentages published in many experiments of this kind are erroneously small. In these instances, the standard error is unjustifiably small because it reflects only the variation within the drug sample and does not reflect the variation inherent within the control sample.

The error arises in one of two ways. It is made in one way when the standard error of the mean percentage change is calculated from a series of percentages, each derived from an observation made on an individual drug-treated animal, by dividing the mean percentage change by the mean of the controls. It is made in another way when the standard error of the mean percentage change is obtained by dividing the standard error of the drug mean by the mean of the controls.

Let us consider a hypothetical experiment involving a certain drug and brain 5-hydroxytryptamine concentrations. We assume that this drug has no effect on the concentration of 5-HT and hence that the experimental animals should have the same average value for 5-HT as do the saline controls. Let us assume that the available animals are from a normal population having a 5-HT concentration with a mean and standard deviation of 1000 and 150 ng/g brain. Thus, if one randomly takes 25 animals for each of the two groups the standard deviation of the sample mean of the control as well as the treatment group would be $150/\sqrt{25} = 30$ ng/g. Thus, it would be quite reasonable to obtain sample means of 1045 and 965 ng/g 5-HT for the saline and drug group, respectively. If one ignores the variability of the control group we can estimate that the figure for the drug group is 92·3% of that of the control group with a standard error of $(30 \times 100)/1045$ or 2·87%. We then see that the drug appears to lower brain 5-HT by 7·7% which is 2·68 standard errors lower than the control and normally the conclusion would be made that the drug significantly lowers brain 5-HT. However, if the variability in the control group is considered, the standard error for the ratio of the two means, say, $\bar{X}/\bar{Y}$, must be

obtained. An estimate (Duncan, 1965) of this standard error is given by:

$$\sigma_{\bar{X}/\bar{Y}} = ((\bar{Y}^2\sigma^2_{\bar{X}} + \bar{X}^2\sigma^2_{\bar{Y}})/\bar{Y}^4)^{1/2} \quad .. \quad .. \quad .. \quad .. \quad (1)$$
$$= [(1045^2 \times 900 + 965^2 \times 900)/1045^4]^{1/2}$$
$$= 0{\cdot}039$$

or on a percentage basis this would be 3·9%. Thus, the decrease of 7·7 would only be 1·93 standard errors and hence would not be as readily accepted as statistically significant.

When nine animals are used for the control group and 25 for the drug group the results are even more startling. The standard error of the saline group would be 50 ng/g, and a sample mean of 1060 would not be unreasonable. The former method of calculation would produce a percentage decrease of 8·96% compared to the same standard error of 2·87% which is 3·12 standard errors (a highly significant value). The standard error of the ratio $\bar{X}/\bar{Y}$, however, is:

$$\sigma_{\bar{X}/\bar{Y}} = [(1060^2 \times 900 + 965^2 \times 2500)/1060^4]^{1/2}$$
$$= 0{\cdot}048$$

or 4·8%. Thus, the decrease of 8·96% is only 1·87 standard errors compared to the 3·12 standard errors obtained above.

The consequence of this kind of error is also important in the comparison of the relative effect of a drug upon substances in the brains of animals arising from two different prior treatments. The example presented here will exhibit the situation where the saline control groups for the two treatment conditions are significantly different. It should be evident in this circumstance that absolute drug-induced changes could differ significantly although the proportional changes might not be significantly different, or *vice versa*. Thus, the formulation of the hypothesis before experimentation becomes important.

In such cases, one popular erroneous method of testing for significant differences in the effect of a drug for different prior treatments involves: (i) expressing the

Table 1. *Three methods for testing for significant differences of drug-induced changes.*

	Prior treatment 1		Prior treatment 2	
	Saline (X)	Drug (Y)	Saline (X)	Drug (Y)
Number of samples	30	30	30	30
5-HT (ng/g), Mean and standard error	702 ± 19·4	794 ± 22·0	913 ± 25·3	935 ± 25·9
Drug-induced changes (Δ)				
Method				
I Standard error of Δ obtained by $\left(\sigma^2_{\bar{X}} + \sigma^2_{\bar{Y}}\right)^{1/2}$.	92 ± 29·3		22 ± 36·2	
Test of difference in Δ's		70 ± 46·6 t = 1·50		
II Standard error of Δ percentage obtained by equation (1)	13·10 ± 4·49		2·40 ± 4·01	
Test on difference in Δ's		10·70 ± 6·02 t = 1·78		
III Standard error of Δ percentage obtained by $\left(\sigma_{\bar{X}}\right)/\bar{Y}$	13·10 ± 3·13		2·40 ± 2·83	
Test on difference in Δ's		10·70 ± 4·22 t = 2·54		

Table 2. *Analysis of variance table for 5-HT measurements in ng/g.*

Source of variation	Degrees of freedom	Sums of squares	Mean square	F	$\sqrt{F}$
Prior treatments (T) ..	1	929 808	929 808		
Control vs. drug (d) ..	1	97 527	97 527		
T D interaction ..	1	36 855	36 855	2·27	1·51
Error	116	1 885 476	16 254		
	119	2 949 666			

Table 3. *Analysis of variance table for the logarithm of the measurement.*

Source of variation	Degrees of freedom	Sums of squares	Mean square	F	$\sqrt{F}$
Prior treatments (T) ..	1	0·259 777	0·259 777		
Control vs. drug (D) ..	1	0·031 360	3·031 360		
T D interaction ..	1	0·014 531	0·014 531	3·22	1·80
Error	116	0·522 801	0·004 507		
	119	0·828 469			

standard error of drug-treated animals for each prior treatment as a percentage of the mean saline control for that prior treatment, and (ii) using the "standard errors" thus derived in a *t*-test to compare the percentage drug-induced change for the different prior treatments. It is the standard error of these differences that is in error. Table 1 summarizes a set of figures analysed for significance, first by the actual measurements in ng/g, second by the use of equation (1), and third by using actual measurements and testing for percentage differences by ignoring the uncertainty in the saline control groups. Tables 2 and 3, respectively, were constructed by making an analysis of variance on the absolute measurement data and on the logarithm of each measurement. The first test in Table 1 is equivalent to the test on the interaction term in Table 2. It should be recalled that the F distribution with 1 and n degrees of freedom is equal to the square of the *t* distribution with n degrees of freedom. By comparing the results of the second test in Table 1 with the interaction test in Table 3 one sees that the analysis of variance using logarithms is equivalent to the test for significant percentage changes. The third test in Table 1 ignores the variability in the saline control samples and therefore gives the erroneous impression that animals experiencing the two prior treatments respond differently to the drug.

Department of Statistics,
The University of Tennessee,
Knoxville, Tennessee 37916. ROBERT A. MCLEAN

Maryland Psychiatric Research Center and The Johns Hopkins University School of Medicine,
Baltimore, Maryland 21228, *U.S.A.* BRUCE L. WELCH

March 22, 1971

REFERENCE

DUNCAN, A. J. (1965). *Quality Control and Industrial Statistics*, 3rd edn, p. 94. Homewood, Illinois: Richard D. Irwin, Inc.