EFFECT OF DATA TRANSFORMATION ON *t*-TEST SENSITIVITY

In the study of biological and pharmacological variables it is common practice to identify significance by rejecting the 'null hypothesis' thus demonstrating the probability of distinct populations. Statistical procedures incorporating a variance function, such as Student's *t*-test, require that the observations satisfy three assumptions: (1) the data are normally distributed, (2) homoscedasticity (equality of variance), (3) the experimental effects are additive. It has been said that the *t*-test is remarkably insensitive to small or moderate departures from normality (Bulmer, 1967). In addition conclusions reflecting parent populations are generally drawn from samples of those populations and the central limit theorem states that the means of samples from a given population tend more toward a more normal distribution than is found with the parent data. As a result researchers have generally accepted that a normal distribution exists for their data and that departures from normality do not affect the sensitivity of the *t*-test or other statistical procedures which are dependent on the normality assumption. The qualifying statement (Bulmer, 1967), indicating that departures from normality should not be extreme or the number of observations too small, in many experimental situations, appears to have been ignored. Furthermore, statements indicating that the homoscedasticity and additivity assumptions are satisfied are singularly lacking.

If the data do not satisfy the first two of these assumptions in particular, the effect is generally to improperly estimate the error variance and hence make it more difficult to detect real differences when they do occur (Horsnell, 1953). In fact, the sensitivity of the test is reduced. By applying an appropriate

transformation, however, the first two assumptions may be satisfied and decreases in sensitivity of the particular statistical test reduced (Box, 1953). One might predict, therefore, that fewer experimental observations would be required in order to evaluate biochemical or pharmacological effects and the loss of valuable significant results may be prevented. Further, the relevance of this problem should be increased when dealing with small sample numbers with a moderate to large variance.

It is known that mathematical transformation of the data prior to a variance-related statistical test may more completely satisfy the assumptions described, thus restoring the power of the test employed. Two experiments were performed to demonstrate the practical significance of such a procedure. In both experiments the excretion of urinary D-glucaric acid was selected as an example of a biological variable with a moderate variance. D-glucaric acid is a product of glucuronic acid metabolism and its synthesis is enhanced after the administration of anticonvulsant drugs to man (Latham, Millbank, Richens & Rowe, 1973) and phenobarbitone to guinea pigs (Hunter, Maxwell, Stewart & Williams, 1973).

In the first experiment complete 24 h urine collections were provided by 47 normal drug-free adult subjects and 126 adult epileptic patients who were receiving chronic anticonvulsant drug therapy. Aliquots of these 24 h urine collections were stored at -15°C until required for the measurement of urinary D-glucaric acid (Latham, 1974). The frequency of D-glucaric acid results falling into particular pre-determined ranges was plotted as a histogram (Figure 1). Following logarithmic transformation of

Figure 1 Histograms showing the positively skewed distribution of D-glucaric acid excretion in (a) 47 normal subjects and (b) 126 epileptic patients.

Figure 2 Histograms showing the distribution of D-glucaric acid excretion in (a) 47 normal subjects and (b) 126 epileptic patients following logarithmic transformation of the data.

Figure 3 Plot of t and % probability (17 d.f.) as a function of time for the excretion of D-glucaric acid in guinea pigs treated with phenobarbitone. ■ raw data; ● logarithmically transformed data; ▲ $\sqrt{}$ transformation; * $P \leq 0.05$.

the data, the histograms were reconstructed (Figure 2). The transformation apparently improved the normality of the distributions obtained.

The second experiment was designed to demonstrate the effect of suitable transformations on the variance ratio and the t value obtained from a controlled study measuring urinary D-glucaric acid excretion. Small numbers of randomly bred guinea pigs (with a predictably large error variance) were therefore used. Thus 13 guinea pigs were treated with 40–60 mg/kg phenobarbitone i.p. twice daily for 24 days and 6 control guinea pigs were treated with isotonic saline i.p. twice daily. Twenty-four hour urine collections were obtained twice weekly from each animal and the urinary D-glucaric acid of these specimens determined (Latham, 1974). While Bartlett's F -test showed a significant difference

between the variances with the raw and logarithmically transformed data ($P < 0.001$), the square root transformation obviated this difference ($P > 0.3$). The mean D-glucaric acid excretion of the treated group was compared with that of the control animals by means of the t -test for each set of observations (twice weekly for 24 days). The t -tests were then repeated after applying both a square root and a logarithmic transformation to the data. The results (Figure 3) clearly demonstrate a sustained increase in the value of t after these transformations. On days 15, 18 and 22, of the experiment, the increase in the t value calculated from logarithmically transformed data was sufficient to define a significant drug effect ($P < 0.05$) which was not apparent from analysis of the original data.

These studies suggest a need for careful evaluation of data prior to interpretation of results. The laboratory studies, which demonstrated a loss of significant experimental findings when the assumptions were not satisfied, emphasize the importance of the assumptions upon which the validity of the statistical procedure depends.

The assistance and advice of Dr C.H. Goldsmith and Mr P.A. Johnson is gratefully acknowledged. Thanks are expressed to the human volunteers who were patients at the National Hospital, Chalfont Centre for Epilepsy.

A.N. LATHAM*

McMaster University Medical Centre, Hamilton, Ontario, Canada L8S 4J9

Received April 17, 1978

* Present address: Wyeth Laboratories, Huntercombe Lane South, Taplow, Maidenhead, Berks SL6 0PH.

References

- BOX, G.E.P. (1953). Non-normality and tests on variance. *Biometrika*, **40**, 318–335.
- BULMER, M.G. (1967). Tests of significance. In *Principles of Statistics*, p. 152. (2nd Ed.) Edinburgh: Oliver and Boyd.
- HUNTER, J., MAXWELL, J.D., STEWART, D.A. & WILLIAMS, R. (1973). Urinary D-glucuric acid excretion and total liver content of cytochrome P-450 in guinea pigs—Relationship during enzyme induction and following inhibition of protein synthesis. *Biochem. Pharmac.*, **22**, 743–747.
- HORSNELL, G. (1953). The effect of unequal group variances on the F test for the homogeneity of group means. *Biometrika*, **40**, 128–136.
- LATHAM, A.N. (1974). D-glucuric acid as an index of hepatic enzyme induction by anticonvulsant drugs. *J. Pharm. Pharmac.*, **26**, 284–286.
- LATHAM, A.N., MILLBANK, L., RICHENS, M.B. & ROWE, D.J.F. (1973). Liver enzyme induction by anticonvulsant drugs, and its relationship to disturbed calcium and folic acid metabolism. *J. clin. Pharmac.*, **13**, 337–342.

BIOAVAILABILITY OF TWO FRUSEMIDE PREPARATIONS

It has been shown that the bioavailability of the diuretic frusemide as Lasix tablets is similar to that of the drug given as a water solution (Kelly, Cutler, Forrey & Kimpel, 1974).

There are at least two marketed brands of frusemide, Impugan and Lasix. No paper has been published which compares the bioavailability of those two preparations.

Five healthy male subjects volunteered for the investigation. Their age ranged between 25 and 29 years. They were healthy according to history, physical examination and laboratory tests (Hb, WBC, thrombocytes, ESR, S-ALAT, S-ASAT, creatinine clearance and S-electrolytes i.e. Na⁺, K⁺, Ca⁺⁺ and Mg⁺⁺). The subjects were asked to refrain from the use of alcohol and heavy physical strain during 3 days before every test day.

Each subject took part in three experiments at least 1 week apart in a randomized order. They were given at 08.00 h (time 0) (a) placebo, (b) 40 mg frusemide as Impugan (A/S Dumex, Copenhagen, Denmark) or (c) 40 mg frusemide as Lasix (Hoechst AG, Frankfurt (M), FRG), after fasting overnight.

The subjects were given tap water (150 ml) each hour from 2 to 10 h except at 4, 5 and 6 h, when 75 ml of the tap water was replaced by Trilgar (carbohydrates 0.53 g ml⁻¹, Na⁺ 30 µmol ml⁻¹, K⁺ 0.8 µmol ml⁻¹, Ca⁺⁺ 1, 5 µmol ml⁻¹, Mg⁺⁺ 2.5 µmol ml⁻¹, Cl⁻ 8.5 µmol ml⁻¹ and PO₄³⁻ 6.3 µmol ml⁻¹). After 10 h the subjects were permitted to drink, eat and smoke.

Venous blood samples were collected at 15, 30, 45, 60, 90, 120 and 150 min and thereafter at each h to 12 h. After centrifugation, plasma was stored at -20°C until analysis. The subjects emptied their bladders immediately before the start of the experiment. The urine was collected every 0.5 h until 4 h, thereafter hourly until 10 h.

The concentrations of frusemide in plasma and

urine were determined with GLC (Lindström & Molander, 1974) and HPLC (Lindström, 1974) respectively. The GLC assay was used with a modification: 1% Se 30 column (1 m) operated at 220°C, injector temp 230°C, detector temp 300°C. Each series of plasma and urine specimens was analyzed on the same occasion.

The urine specimen was analyzed for its Na⁺ and K⁺ content by flame photometry (Flamephotometer 3 L 143, Instrumentation Inc.) and those of Ca⁺⁺ and Mg⁺⁺ by atomic absorption spectrophotometry (Perkin-Elmer 306) at the Department of Clinical Chemistry at Serafimerlasarettet.

The area under the plasma concentration time curve was calculated by the trapezoidal rule and extrapolated to infinity (AUC_{0-∞}) from the last measurable plasma concentration (> 10 ng ml⁻¹).

Statistical comparisons were made with Student's paired *t*-test. Averages are presented as mean ± 1 s.d.

After administration of 40 mg Impugan, maximal plasma levels were seen at 0.80 ± 0.33 h on average and they amounted to 2079 ± 540 ng ml⁻¹. The corresponding figures after 40 mg Lasix were 1.75 ± 1.22 h and 1836 ± 800 ng ml⁻¹ respectively. Although the peaks after Lasix tended to be less high and to appear later than after Impugan, the differences were not significant. AUC_{0-∞} after Impugan and Lasix averaged 3044 ± 763 and 3193 ± 760 ng ml⁻¹ h respectively (NS).

The urinary recovery of frusemide during 10 h after 40 mg Impugan amounted to 16.5 ± 4.4 mg in mean and that after Lasix to 18.2 ± 6.0 mg (NS).

The diuresis during the first 6 h after 40 mg Impugan was on average 1084 ± 395 ml higher than after placebo. The corresponding figure after Lasix was 987 ± 127 ml (NS). Likewise there was no significant difference between the increases in urinary excretion of Na⁺ (135 ± 51, 133 ± 25 mmol) K⁺ (20 ± 14, 22 ± 11 mmol) Ca⁺⁺ (1.02 ± 0.87,