

TABLE II  
R.B.C. Li—plasma Li ratio in mixed psychosis

Patient	R.B.C. lithium		Clinical response
	Plasma lithium		
R.L.	0.29 ± 0.01*		+†
M.M.	0.36 ± 0.02		—
C.G.	0.42 ± 0.01		+
L.T.	0.44 ± 0.01		+
G.A.	0.52 ± 0.07		—
A.R.	0.72 ± 0.05		—

\* $\bar{x} \pm$  S.E.M.

† + responders; — non responders.

#### REFERENCES

- CRAWFORD, J. & MCPHAIL, N. I. (1973) Measure of depressive mood at monthly intervals. *British Journal of Psychiatry*, **122**, 447.
- FRAZER, A., MENDELS, J., SEGUNDA, S. K., COCHRANE, C. M. & BIANCHI, C. (1973) The prediction of brain lithium concentrations from plasma or erythrocyte measures. *Journal of Psychiatric Research*, **10**, 1.
- SEGUNDA, S. K. & MENDELS, J. (1972) A method for the determination of sodium, potassium, magnesium and lithium concentration in erythrocytes. *Clinica Chimica Acta*, **36**, 499.
- GLENN, A. I. M. & READING, H. W. (1973) Regulatory action of lithium in manic-depressive illness. *Lancet*, **ii**, 1239.
- MENDELS, J. & FRAZER, A. (1973) Intracellular lithium concentration and clinical response: towards a membrane theory of depression. *Journal of Psychiatric Research*, **10**, 9.
- PERRIS, C. (1973) The genetics of affective disorders. In *Biological Psychiatry* (ed. Mendels). New York: Wiley & Sons.

#### INTER-RATER RELIABILITY OF WARD RATING SCALES

DEAR SIR,

The enthusiasm with which Dr. Hall encourages the wider use of the weighted kappa statistic in estimating inter-rater reliability of ward rating scales (*Journal*, September 1974, **125**, 248–55) tends to obscure two features common to many statistical computations—that they are prone to errors of arithmetic (especially when computed by hand) and that they must be interpreted with care.

Evidence of the former is taken from Dr. Hall's illustrative examples, in which at first reading seven

mistakes were noted in the calculation. It is simplest to list them.

From the first example (on pp. 251–2), the correctly computed values are:

- (i) weighted kappa = +0.6719;
- (ii) quantity A = 0.025534 (has some of step 7 been mislaid?);
- (iii) quantity B = 0.020621;
- (iv) quantity C = 0.95367;
- (v) variance of weighted kappa = 0.004685.

In the practical example of observations on 12 patients:

- (vi) a marginal frequency was incorrectly summed to 27 (instead of 28), and was the probable cause of
- (vii) an inaccurate weighted kappa value, which should have been +0.68557.

Secondly, the relative interpretations to be attached to the three reliability measures given in the practical example were not obvious. The data arose from the pooled observations of 12 patients for 12 items on a five-point scale. An implied comparison was made between the product-moment correlation coefficient (computed from the *total* data to have a value of +0.98) and weighted kappa (calculated from the *pooled* data as 0.68557). Such a comparison may be misleading, unless the weighted kappa is corrected to take account of the effects caused by pooling the data. An approximate correction may be made by application of the Spearman-Brown formula (1) which estimates the reliability,  $r_n$ , for a test which is  $n$  times the length of a test of known reliability  $r$ , by

$$r_n = \frac{n \cdot r}{1 + (n-1) \cdot r}$$

Since in the case presented  $n = 12$  and  $r = 0.68557$ , the appropriate estimate of the total reliability,  $r_n$ , is 0.97, which is a more realistic value for comparison with the other measures of reliability of total scores quoted by Hall.

IAN M. LEITCH.

Department of Psychological Medicine,  
Royal Victoria Infirmary,  
Newcastle upon Tyne, NE1 4LP.

#### REFERENCE

1. GUILFORD, J. P. (1956) *Fundamental Statistics in Psychology and Education*. New York: McGraw-Hill.