

THE USE OF r VALUE

Dear Sir,

It was interesting to read the article 'Transcutaneous Bilirubinometry in Malay, Chinese and Indian Term Neonate' in the *Med J Mal* 1984; 1: 35-37.

However, I would like to point out that though there is positive correlation between serum bilirubin (μ mol/l) in all the three races, the use of r value to indicate the degree of correlation between the two variables tend to give the impression of a stronger relationship than is actually present. For example the r value of +.66 in the Chinese babies seem to represent a high degree of association, but since the coefficient of determination r^2 is only equal to 0.43, less than half of the total variance in Y has been explained by the regression equation. (The results would have been better if the outliers had been excluded from the study.)

Similarly the coefficient of determination in the Indian group is 0.39 and the Malay babies have a coefficient of determination of 0.68. This seems to indicate that the Malay babies have a higher coefficient of determination than the Chinese and the Indian babies. This high coefficient of determination amongst the Malay babies means that the regression equation in this sample population explains the 'total

variance' better than the Chinese and the Indian babies. The 'unexplained variance' is high in the Indian and Chinese babies in this sample. This means that there are other factors in these populations that will reduce the 'unexplained variance'.

The discussion of the bilirubin indices can also be misleading. (Table II). The values constructed using the respective equations have not taken into account the standard error of estimate. The standard error of estimate is simply the scatter of observed values around the computed Y_c values around the regression line. For example the bilirubin indices 12.5 for the 85 serum bilirubin levels have a range of values (prediction interval) which have to be taken into consideration. The prediction interval in this case would be $Y_c \pm t \times S_i$, where Y_c is the individual forecast as determined from the sample regression line, S_i is the standard error of forecast, and t value for $n-2$ degrees of freedom for the desired confidence interval.

By calculating these prediction intervals a range of values are obtained. These range of values may be a better way indicating the action levels rather than individual values.

Dr Hematram Yadav, MBBS, MPH
Public Health Institute, Kuala Lumpur

THE USE OF r VALUE: REPLY

Dear Sir,

From the statistical point of view, I agree with the comment made by Dr. Hematram Yadav regarding the use of coefficient of determination and the standard error of estimate in our study.

However, from the practical point of view, in a population with wide variation of skin pigmentation, the transcutaneous bilirubinometer should not be relied on as the sole instrument to

determine which jaundiced neonate needs treatment. A reading which is considered "outlying value" in the study may in real life situation, when ignored, result in kernicterus in a neonate.

Dr N. Y. Boo
Department of Paediatrics
Faculty of Medicine
Universiti Kebangsaan Malaysia
Kuala Lumpur, Malaysia