

Letter to the Editor: Masked Variability in Regression Analysis

RODNEY C. BRYANT

Good Samaritan Hospital, Woodbury, Tennessee, USA

Summary

Mackler *et al.* (1) found markedly elevated plasma concentrations of phenylalanine in rats made iron deficient; phenylalanine levels returned to control levels within one week after rats were treated with iron dextran. Plasma phenylalanine levels, the authors stated, "appeared to vary directly with the degree of iron deficiency."

Unfortunately, as evidence of this last point, the authors show, in their Figure 1, a linear regression line which is misleading. The ordinary least squares method is "blind" to the weight of points such as those shown in their Figure 1 (2). Examination of the figure and of their Table 1 shows that some points in the figure represent means for 5 animals, others 3 animals, and one point represents only 1 animal. The correlation coefficient for the points in Figure 1 is quite high, $r = 0.950$, seeming to allow 90% of the variability of percentage of increases in phenylalanine levels to be explained by variability in hemoglobin concentration in iron-deficient animals. However, for the actual data in their Table 1, the relationship appears much weaker. When the regression line is calculated correctly from the data in Table 1, $r = -0.564$, explaining only 32% of the variance instead of 90%. Mean hemoglobin concentration from Table 1 is 5.82 gm/dl; mean percentage of increase in phenylalanine is 39.6%. However, the mean of the X values (hemoglobin concentration) in Figure 1 is 5.75 gm/dl, and the mean of the Y values (% increase in phe) is 43.06%. The inappropriateness of the displayed regression line could be reme-

died by weighting the points according to the amount of evidence summarized by each, but then, why not just analyze the original data?

Needless to say, the increases in phenylalanine concentration observed in iron-deficient animals must be determined very indirectly as regards the manipulated variable of iron deficiency. Particularly surprising is that baseline phenylalanine levels are completely unrelated to phenylalanine levels in the iron-deficient state ($r = 0.029$). Initial phenylalanine levels were marginally related to hemoglobin levels in iron deficiency ($r = 0.537$), but the meaning of this is obscure unless one suspects that phenylalanine somehow conditions the development of iron deficiency! Also, the authors found that phenylalanine hydroxylase, measured as specific activity or as activity per g of liver, showed increased activity in iron-deficient rats instead of decreased activity, as one would postulate. This increase is of borderline significance ($0.05 < P < 0.1$) for both measures, but the probability is not >0.1 as the authors indicate in their Table 2.

Obviously, there are additional questions to be answered in this area. I look forward to further reports from the authors' laboratory on iron metabolism and its unexpected ramifications.

REFERENCES AND NOTES

1. Mackler, B., Person, R., Miller, L., and Finch, C.: Iron deficiency in the rat: effects on phenylalanine metabolism. *Pediatr. Res.*, 13: 1010 (1979).
2. Mosteller, F., and Tukey, J.: *Data Analysis and Regression*. (Addison Wesley, Reading, MA, 1977).