FACTORIAL DESIGN AND COVARIANCE IN THE BIOLOGICAL ASSAY OF VITAMIN D

By C. I. Bliss
Sandusky, Ohio

The development by R. A. Fisher and his co-workers of statistical methods adapted to small numbers has provided the pharmacologist with an important tool. Recently Bliss and Marks (1) have shown some of the possibilities of these methods in the biological standardization of insulin. The technique is suitable for many other bioassays and in some cases it permits a fuller utilization of information inherent in past experiments that have complied with good pharmacological practice. A case in point is the standardization of vitamin D from the ash content of the femur of the rat. This assay has been described in detail by Coward (2), who has given in Table IX all essential data for computing the potency of a sample of cod liver oil by the newer procedures. The purpose of the present paper is to reanalyze this material by factorial analysis with covariance and to propose on this basis a more efficient criterion for measuring the potency of vitamin D.

Coward's model experiment was based upon 36 rats, comprising six litter mates from each of six litters. Soon after weaning, the rats were divided into six groups, each group containing one individual from each litter. Three groups were dosed with the International Standard of vitamin D at rates of 0.025, 0.05 and 0.1 units daily over an interval of six weeks and the other three groups with an unknown sample of cod liver oil at rates of 0.2, 0.4 and 0.8 mgms. daily for the same period. At the end of the test the femora were removed from each rat, fat-extracted with alcohol, dried to constant weight and ashed. The problem here under consideration is how best to determine the content of vitamin D in the sample of cod liver oil from these weights of dried, fat-extracted bone and of ash content.

Analysis in terms of the percentage ash. The experiment has been analyzed first in terms of the usual criterion of response, which is the ratio of these two weights or the percentage of ash in the femora. The determination of relative potency depends upon two factors, the difference in the response to paired dosages of the unknown and of the standard and the rate at which the response increases with dose as measured by the slope of the curve relating dosage and response. If the latter were invariably linear, two doses of each material would suffice, but in the absence of such prior knowledge the results from three or more doses provide a more satisfactory basis. In the present experiment the dosages of both standard and unknown were obtained by multiplying the smallest dose of each material by 2 and again by 2. As a result the difference between standard and unknown was tested at three corresponding levels and at the same time the slope and form of the curve relating dosage and response could be determined for both materials. The experiment was well adapted, therefore, to factorial analysis as described by Bliss and Marks (1).

The percentage of bone ash is given for each of the 36 rats in Table 1,

TABLE 1

PERCENTAGE OF BONE ASH IN AN ASSAY OF VITAMIN D (COWARD)

Litter no.	Percentage of ash following treatment						Total
	S_1	S_2	S_3	U_1	U_2	U_3	
1	20.6	37.6	35.5	24.9	35.2	43.0	196.8
2	25.1	30.5	41.3	32.2	36.4	46.0	211.5
3	24.0	32.9	41.9	33.7	36.1	47.3	215.9
4	23.0	33.3	42.8	31.4	37.9	46.3	214.7
5	24.7	29.5	42.2	24.5	35.7	48.8	205.4
6	22.8	32.4	41.1	34.2	37.7	43.8	212.0
Total	140.2	196.2	244.8	180.9	219.0	275.2	1256.3

where rows represent litters and columns treatments. Since one factor balanced the other, the total variation was divided by an analysis of variance into that attributed to litters, to treatments and to interaction or error. It is evident from Table 2 that the segregation of litter dif-

TABLE 2

ANALYSIS OF VARIANCE COMPUTED FROM THE ASSAY OF VITAMIN D IN TABLE 1

Variation due to	Degrees of freedom	Sum of squares	Mean square
Litter differences	5	42.712	8.5424
Unequal dosage of vitamin D	5	1908.348	
Experimental error	25	169.270	6.7708
Total	35	2120.330	

ferences contributed to the reliability of the experiment even though the improvement was small. For the isolation of the several treatment factors the factorial scheme proposed by Bliss and Marks (1) gave a satisfactory subdivision of the five degrees of freedom between the three treatments with the standard, S_1, S_2 and S_3, and the three treatments with the unknown, U_1, U_2 and U_3. The six subtotals (Y_p) from Table 1 have been entered in Table 3 beneath the corresponding columns of coefficients (x) and multiplied by the coefficients of each row

TABLE 3

FACTORIAL ANALYSIS OF VITAMIN D ASSAY IN TERMS OF PERCENTAGE ASH

Effect	Factorial coefficients (x) for dose						Divisor $NS(x^2)$	Sum of products $S(xY_p)$	Variance $\frac{S^2(xY_p)}{NS(x^2)}$	$\sqrt{[Y^2]}$
	S_1	S_2	S_3	U_1	U_2	U_3				
Samples	−1	−1	−1	+1	+1	+1	36	93.9	244.92 $=D^2$	15.65 $=D$
Slope	−1	0	+1	−1	0	+1	24	198.9	1648.38 $=B^2$	40.60 $=B$
Parallels	+1	0	−1	−1	0	+1	24	−10.3	4.42	
Curvature	+1	−2	+1	+1	−2	+1	72	10.7	1.59	
Opposed curvature	−1	+2	−1	+1	−2	+1	72	25.5	9.03	
Total Y_p	140.2	196.2	244.8	180.9	219.0	275.2	Error		6.77 $=s^2$	2.60 $=s$

to obtain the sum of products $S(xY_p)$ for each treatment effect, leading directly to the five variances, $[Y^2]$.

The treatment effects combined factorial and polynomial analysis. Biologically, the first measured the average difference in response between the three paired doses of standard and of unknown and the second the combined slope of the parallel straight lines relating percentage ash to the dose of vitamin D, where each dosage interval of 0.3010 logarithm is equal to 1. The remaining three terms showed that the relative potency and its error could be computed legitimately from these first two items. The third treatment effect or the interaction between samples and slope measured the divergence from parallel dosage-response curves for standard and unknown and was less than the mean square for error; the fourth row or the quadratic term in fitting a polynomial to both standard and unknown signified that the combined curvature from a linear relation was negligible; the fifth row or the interaction between the quadratic term and samples indicated a curvature in opposite directions that exceeded the experimental error by too small a margin to have any significance. Hence the last three variances passed the test as a valid assay.

Relative potency was determined from the variances for the difference between samples and for slope. As shown elsewhere (1), the log-ratio of potencies, M, can be computed from the equation

$$M = kID/B$$

and its standard error from

$$s_M = skI\sqrt{D^2 + B^2}/B^2,$$

where D^2 and B^2 are the treatment variances for sample and for slope, I is the interval between doses in logarithms or 0.30103 in the present experiment, s is the standard deviation from the mean square for error and k is based upon the ratio of $S(x^2)$ for slope to that for the difference between treatments and for three dosage levels is equal to $\sqrt{8/3}$ or 1.6330. Substituting the values from Table 3,

$$M = (1.6330)(0.30103)(15.650)/40.601 = 0.18949$$

and

$$s_M = (2.6021)(1.6330)(0.30103)\sqrt{244.92+1648.38}/1648.38 = 0.03377.$$

This log-ratio of potency of $M = 0.1895 \pm 0.0338$ had yet to be corrected for the potency of the unknown cod liver oil which was assumed in planning the experiment since U did not represent the same units as S. Eight mg. of cod liver oil were assumed to contain one unit of standard vitamin D, so that the logarithm of 8 was subtracted from the initial M to obtain $0.18949 - 0.90309 = -0.71360$ as the logarithm of the units of vitamin D in each milligram of oil. Transforming to original units, this particular sample of cod liver oil has been found to contain 193.4 ± 15.0 units of vitamin D per gram. Thus the factorial analysis in Table 3 has provided terms for measuring relative potency and a test of the hypothesis of parallel, linear dosage-response curves upon which the calculation depends.

Analysis in terms of the log-weight of ash. Although the percentage of ash in the fat-extracted bone has proved a valid index to the vitamin D in the diet, it involved an implicit assumption which may be questioned. It assumed that the percentage of ash was constant at any given dose of vitamin D, so that if the weight of ash be designated by A and the weight of fat-free organic matter by B, $A/(A+B) = K$. Then

$$A = KA + KB,$$

$$A(1 - K) = KB,$$

$$A = \frac{K}{1 - K} B$$

and

$$\log A = \log \frac{K}{1 - K} + \log B.$$

It is clear that the use of percentage ash involved the assumption that for individuals treated alike the logarithm of the ash content plotted against the logarithm of the organic content lost in ashing gives a straight line with a slope equal to 1. If the observed slope were to differ appreciably from unity or from linearity, the percentage ash would represent a less efficient criterion for measuring vitamin D than a more consistent value. In testing this postulate, however, it was essential to avoid the loss in precision due to differences between litters.

Covariance answered all of these requirements. For the dependent variate the percentage ash was replaced by the log-weight of the ash (Y), which was to be adjusted by covariance for differences in the log-weight of the organic matter lost by combustion (W). The results are given for each rat in Table 4, from which the sums of squares and

TABLE 4

ORGANIC AND ASH CONTENT OF DRY BONE OBSERVED IN ASSAY OF VITAMIN D

	Litter no.	Data for treatment						Total
		S_1	S_2	S_3	U_1	U_2	U_3	
Log-weight of organic content (W)	1	.792	.714	.893	.876	.801	.825	4.901
	2	.820	.818	.792	.790	.833	.829	4.882
	3	.877	.916	.893	.804	.925	.928	5.343
	4	.805	.753	.846	.787	.834	.885	4.910
	5	.802	.809	.811	.911	.719	.799	4.851
	6	.793	.774	.865	.783	.831	.886	4.932
	Total	4.889	4.784	5.100	4.951	4.943	5.152	29.819
Log-weight of ash (Y)	1	.207	.494	.633	.396	.537	.702	2.969
	2	.346	.459	.640	.467	.590	.760	3.262
	3	.377	.606	.751	.511	.678	.880	3.803
	4	.281	.450	.720	.449	.619	.820	3.339
	5	.318	.431	.674	.423	.464	.778	3.088
	6	.262	.453	.708	.498	.614	.778	3.313
	Total	1.791	2.893	4.126	2.744	3.502	4.718	19.774

products, $[W^2]$, $[WY]$ and $[Y^2]$, have been computed for differences between litters, for each of the treatment effects listed in Table 3 for and the error (Table 5), the calculation following the procedure described

TABLE 5

COVARIANCE IN THE ASSAY OF VITAMIN D FROM LOG-WEIGHT OF BONE ASH

Variation due to	Degrees of freedom	Sums of squares and products			Adjusted mean $[Y^2]$	$\sqrt{[Y^2]}$
		$[W^2]$	$[WY]$	$[Y^2]$		
Litter differences	5	.028473	.039365	.068422	.010393	
Unequal potencies	1	.002070	.016334	.128881	.121546 $=D^2$	.3486 $=D$
Combined slope	1	.007073	.073971	.773645	.740308 $=B^2$	.8604 $=B$
Lines not parallel	1	.000004	.000150	.005430	.005362	
Similar curvature	1	.005653	.005219	.004818	.002733	
Opposed curvature	1	.000578	−.000926	.001485	.001937	
Experimental error	25	.058279	.013277	.036914	.001412 $=s^2$	.0376 $=s$
Total	35	.102130	.147390	1.019595		
Coefficients for adjusting $[Y^2]$		.051901	−.455636	1	$b_c=\frac{.013277}{.058279}=.22782$	

by Bliss and Marks. From the row for error, $b_c=[WY]/[W^2]=0.228$, a value so much smaller than the 1 assumed in using the percentage ash as the criterion of response that it did not differ significantly from zero. After segregating differences due to the dosage of vitamin D and to litters, the amount of combustible matter in the bones had but little effect upon their ash content. Because the effect was so small, the weight of the organic matter in this assay could be ignored altogether and the potency of vitamin D determined from the log-weight of ash alone. In other cases, however, the weight of the organic matter may have a significant effect upon the weight of the ash, and even here the adjustment for W reduced the standard deviation (s) from 0.0384 to 0.0376, so that it has been retained. Since the experiment was designed to measure the potency of a sample of cod liver oil rather than to test the statistical significance of any given treatment effect, adjusted $[Y^2]$'s have been used and the coefficients b_c^2, $-2b_c$ and 1 have been entered in the appropriate columns to facilitate their computation.

The mean adjusted $[Y^2]$'s determined from the log-weight of ash (Table 5) may be compared with the corresponding terms based upon the percentage ash (Tables 2 and 3), noting first that the new criterion segregated a larger part of the total variation under differences between litters. Of the treatment effects the fifth term, that for opposed curvature, had about the same relation to the error in either unit, but with the new criterion both the term for combined curvature and especially that for departure from parallelism were relatively larger. However, the ratio of neither adjusted $[Y^2]$ to that for error exceeded the 5 per cent level and since the reduced $[Y^2]$'s, which are appropriate for tests of significance, would give no larger a ratio, the dosage-response curves for standard and unknown could be considered as parallel straight lines suitable for the determination of relative potency.

It may be contended, however, that the usual 5 per cent fiducial limit is not applicable in the present case. The curve relating log-dose and response probably flattens out as it approaches both an upper and a lower limit and within a middle dosage range a straight line is essentially a convenient approximation. To justify its use the experimental data should lie within the central part of the complete curve where the relation is so nearly linear that it cannot be distinguished experimentally from a straight line. Since this linear central portion has a limited range, any departure from parallelism or from linearity which exceeds the 20 per cent level of significance should be scrutinized. The present experiment is a case in point and the adjusted mean log-weight of ash for each dose has been plotted in Fig. 1 to see if the distribution of the observations indicated an approach to either a "floor" or a "ceiling." Instead the plotted points were found to differ in the opposite direction to that required by this hypothesis, the only alternative to a straight line which is biologically reasonable. In consequence, the customary 5 per cent fiducial limits applied here and the experiment could be accepted in terms of the log-weight of ash as a valid assay.

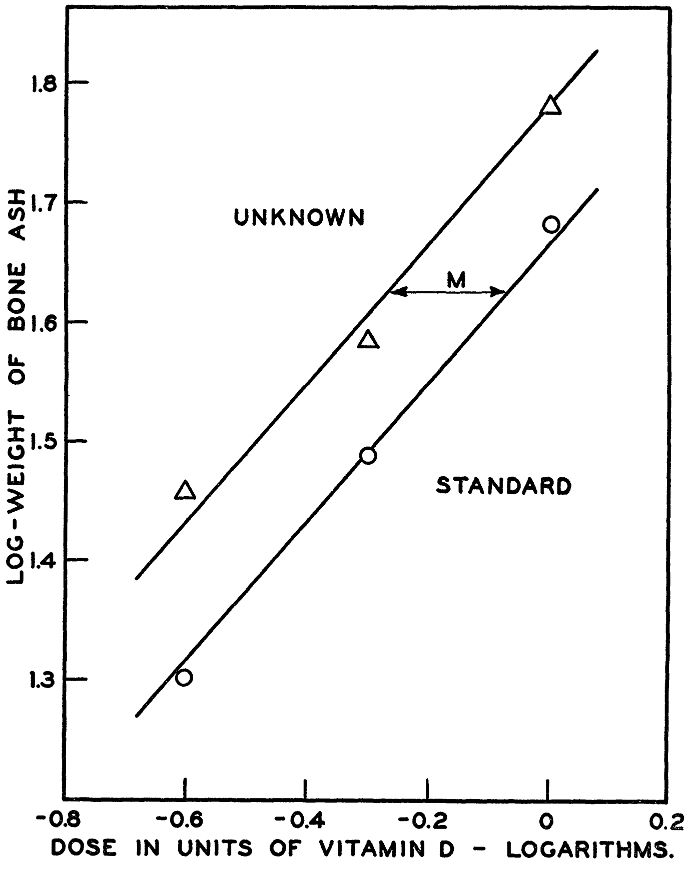

Fig. 1. Relation between the log-weight of bone ash and the log-dose of vitamin D as computed in Table 5, the means adjusted for the log-weight of organic content. Dosage of cod liver oil (the unknown) is given in assumed units of standard.

The computation of M and s_M from D, B and s of Table 5 paralleled that in the preceding section, giving $M = 0.1992 \pm 0.0232$ in arbitrary units. When corrected for the assumed potency of the unknown, 1 gram of cod liver oil assayed at 197.7 ± 10.6 units of vitamin D, differing but little from the 193.4 units computed from the percentage ash. By correcting the weight of the ash by its observed relation to the weight of the combustible matter (in logarithms) instead of by the artificial value assumed in using percentages, the standard error of the determination has been reduced from 15.0 to 10.6 units of vitamin D. To obtain a comparable improvement with the percentage ash as the criterion of response, the number of experimental animals would have to be doubled. If variations in the log-weight of the organic content of the bones had been ignored, 1 gram of cod liver oil would have assayed from the log-weight of ash alone as containing 198.4 ± 10.6 units of vitamin D, a value nearly identical with that obtained from the adjusted estimates.

It would be premature to conclude from this one assay that the weight of the organic component of the bone could always be neglected in the assay of vitamin D. It is probable, however, that any information in such measurements can be utilized effectively by covariance as illustrated above. Two possibilities which might complicate the correction have been examined. It is conceivable that the apparent unimportance of W was partly due to fitting a straight line instead of a curve. This has been tested by introducing W^2 as an additional term and fitting a polynomial but the variation in Y attributable to the quadratic term was negligible, being less than the mean square for error after adjustment by linear covariance. Another possibility is that the slope (b) of the straight line relating Y to W is itself a function of the dose of vitamin D. When checked by computing b for each dose separately, both with and without a correction for litter differences, no such relation could be detected. If the analysis of covariance as illustrated here does not decrease the standard deviation effectively, no other available correction would be expected to give a better result.

The rats in the assay included 16 males and 20 females, divided equally between standard and unknown. A possible sexual difference in response has been checked by adding a second concomitant measure to the analysis in Table 5, designating males by the numeral 1 and females by 0. In terms of $[Y^2]$ the variance attributable to differences between sexes was 0.001259 or less than the mean square for error where sex was disregarded (0.001412).

It has been assumed throughout that the variance was independent of treatment and, within the sampling error, constant throughout the experiment. In some assays the variance has proved to be a function of the mean, which, if true here, would require a more involved calculation. In this assay, however, the variance within doses showed no indication of being tied to the mean, whether computed from the observed log-weights of ash or from values that had been adjusted for differences both in W and between litters.

Summary. A reexamination of an experiment by Coward on the assay of vitamin D from the ash content of the femur of the rat shows that the estimation of relative potency and its error can be facilitated by factorial analysis. When the customary criterion of response, the percentage of ash, was replaced by the log-weight of ash, the increase in the precision of the assay was equivalent to doubling the number of observations.

References

(1) Bliss, C. I. and Marks, H. P. (1939) *Quart. J. Pharm. Pharmacol. 12*, 182.
(2) Coward, K. H. (1938) "The Biological Standardisation of the Vitamins." William Wood, Baltimore.