CXXXV. THE DETERMINATION OF VITAMIN D IN THE LINE TEST BY MEASUREMENT.

By RONALD SYDNEY MORGAN.

*From the Food Research Laboratory, Lever Brothers Limited,
Port Sunlight, Cheshire.*

(Received June 18th, 1932.)

INTRODUCTION.

SHIPLEY *et al.* [1921] showed that when cod-liver oil is administered to rats in which a rachitic condition has been induced by a deficient diet, calcium is deposited in the proliferative cartilage of the cartilage-shaft junctions of the long bones, forming a line across the bone at right angles to the longitudinal axis, the amount of the deposit being an indication of the extent to which healing has taken place. On the basis of this observation, McCollum *et al.* [1922] developed a technique for the determination of the ability of any agency to heal rickets, to which they gave the name "line test." In recording the results of tests carried out by this method, McCollum and his associates distinguished four grades of healing from + to + + + +, no healing being represented by a minus sign. Bills and McDonald [1926] later published coloured illustrations of stained bone sections typical of each healing grade. No special quantitative significance was attached to the differences in grade: they were adopted in the first place as a convenience for recording data. However, as data accumulated, Bills was able to establish a relationship between healing and dosage and he published [Bills *et al.*, 1931] a description of an improved technique of the "line" test by which the potency of a substance is estimated as the dose required to produce an average healing of + + in a large number of rats. From this is calculated a cod-liver oil coefficient, from the known dose of average cod-liver oil required to produce + + healing.

Coward [1928] had earlier introduced the use of a standard preparation of vitamin D with which the potencies of other substances could be compared, but this standard had not been adopted in the U.S.A. at the time of Bills's publication.

In the technique of the "line" test described by Dyer [1931] seven degrees of healing (from 0 to 6) are distinguished and the variation between average healing and dosage of the standard preparation has been determined. In the technique for the estimation of vitamin D by radiography described by Bourdillon *et al.* [1931] a similar procedure is followed, but the degree of healing is estimated from X-ray photographs instead of from stained sections, and 13 degrees of healing (0 to 12) are distinguished.

An essential feature of the recently described improved techniques is the estimation of the degree of healing by comparison with an ideal series, but in the method here described, the ideal series is dispensed with, and the degree of healing is expressed as the area of new calcification measured on magnified camera-lucida drawings of the bone sections. This has the advantages that a much finer gradation of healing is distinguished than is possible by the method of eye judgment, and that uncertainty in allotting the healing figure is eliminated. The error in matching the picture given by the test rat is considered by Bourdillon *et al.* to be responsible for one-third of the total probable error of the estimation carried out by their technique. In the technique of Bills *et al.* the difficulty involved in matching the picture against the ideal series is not so important, since only four grades of healing are distinguished.

EXPERIMENTAL.

The breeding stock are fed on a cooked diet consisting mainly of cereal offals with 10 % Trufood-brand skimmed milk powder, 5 % of Gromax and 2 % of red palm oil with 2.5 units of added vitamin D per g. In addition they receive about 10 cc. of milk daily and 10 g. of cabbage 3 times a week. On this diet, the bucks attain an average weight of 200 g. and the does an average weight of 150 g. in 95 days.

When rats are required for test, the mother of a litter is given the stock diet without milk, cabbage or the oil containing vitamin D from 1 week after the birth of the litter. The litters are not reduced in size, and we have not observed any objection to using litters of 10 to 13 rats for the test. The litter is weaned at 23 days, and if the rats are under 40 g. in weight, they are continued on the same diet as the mother had for a week. They are then transferred to the rickets-producing diet (Steenbock 2965), and maintained on it for 3 weeks to a month. An experimental period of 10 days is used. Doses of standard and other oil solutions are administered directly into the mouth from calibrated dropping tubes. Doses of margarine and butter are measured as cut lengths of ribbons obtained from calibrated glass syringes.

The day after the tenth dose has been given, the rats are killed by chloroform vapour, the distal ends of the radii and ulnae are dissected out, and left to harden in formalin overnight. A pair of bones is then sectioned and stained with silver nitrate. The stained sections are covered with water in a small porcelain dish which is placed on a platform attached to the upright of a Giltch drawing stand. They are viewed through a Zeiss 8× aplanatic magnifier, and the drawings are made with the aid of an Abbé camera lucida. With the distance between the instrument and the drawing paper which has been selected, an overall magnification of 11 diameters is obtained.

The operations of sectioning, staining and drawing, require about 10 minutes per rat.

Measurement of the healing obtained.

When the initial rickets is very severe, the healing produced by a small dose of standard is usually shown by a dense black "line" with well defined edges, as shown in Fig. 1. The area of this "line" can easily be measured by a planimeter.

If the initial rickets is not very severe, or if the dose of vitamin D is large, the newly-calcified area is not completely separated from the shaft of the bone, but it can be distinguished by the pattern of the calcification, usually combining a dense black area with a number of radial lines, as shown in Fig. 2. The actual area of new calcification cannot then be measured so easily,

Fig. 1.

Fig. 2.

Fig. 3.

Fig. 1. Appearances of healing produced by small doses of vitamin D in rats with severe initial rickets. $\times 5.5$.

Fig. 2. Appearance of healing produced by a moderately large dose of vitamin D. $\times 5.5$.

Fig. 3. Litter-mates which received the same dose of vitamin D ($\frac{1}{2}$ unit). Though the general rachitic condition is more severe in rat 999 than in rat 1001, they show practically the same measured healing. $\times 5.5$.

and it is our practice to measure the total densely calcified area *plus* half the area covered by scattered calcification. In the example given the area of dense calcification (on the original $11\times$ magnified drawing) is 127 mm.^2 , and the area of scattered calcification is 121 mm.^2 The figure recorded is

$$127 + \frac{121}{2} = 187 \text{ mm.}^2$$

An illustration of the elimination of uncertainty in estimating healing produced is given in Fig. 3. The pictures reproduced are from litter-mates each receiving 0.5 unit of standard. On the basis of inspection one might place Rat 1001 in a higher grade than Rat 999, since the rachitic condition is less severe; but the measured areas of new calcification are practically equal. Though litter-mates, the two rats showed different degrees of severity of initial rickets.

It sometimes happens that the negative controls show some calcification similar in appearance to the new calcification of healing rickets. This may represent actual healing during the experimental period due to an error in feeding or due to loss of weight by the animals, or it may merely indicate that the initial rickets was not very severe because of a too generous supply of vitamin D in the mother's diet. In the last case the calculation of the assay is not affected (as will be shown below), but in calculating the general average values of the healing produced by doses of the standard preparation (plotted in Fig. 4) a correction has been made by simple subtraction when negative controls have shown such calcification. When the calcification represents healing produced by loss in weight, it is obviously exceptional and does not affect the remainder of the litter.

Fig. 4.

The relation between dosage and healing.

The values of the healing produced by different doses of standard in tests over a period of several months have been collected, and the average values are given in Table I. The value for 1 unit is based on the average of 6 rats only. For the three other dosage levels, the standard deviations were in fairly close agreement (29 to 34.5), and the probable errors have been calculated from the general value of the standard deviation determined for the whole series, which equals 33.

Table I.

No. of rats	Dose of standard preparation units	Average healing produced (mm. ² on 11 × magnified drawing)	Probable error
44	0.125	72	3.4
58	0.25	125	2.9
43	0.50	185	3.4
6	1.00	247	12.2

In Fig. 4 the average values given in Table I are plotted against $\log_{10}$ dose of standard.

The unshaded area on either side of each point represents three times the probable error of the value. It is seen that the linear relationship fits the figures obtained within the limits of error.

The logarithmic relationship implies that doubling or halving the dose, whatever its value, will increase or reduce the measured healing by a constant

amount. This value has been determined by calculating the average value of the difference in healing between litter-mates receiving doses of vitamin D (whether of known value or unknown) in the ratio 2/1. The value obtained is 59.0 mm.², and the line given in Fig. 4 has the slope determined by this. The equation to the curve is

$$y = 196 \log x + 245$$

where y is the healing measured in mm.² on the 11× magnified drawing of radius + ulna, and x is the dose of vitamin D in units. In interpreting the results of assays, the value of the constant 245 does not enter into the calculation, since the litters are always divided between the unknown substance and the standard preparation.

Another implication of the logarithmic relationship is the existence of a threshold dose which will just fail to produce any healing. This has frequently been confirmed in tests in which the lower of two doses in the ratio 2/1 fails to produce any healing or produces only a few mm.², while the higher dose produces healing of the order 60 mm.² The average value of the threshold dose is given by the equation as 0.056 unit.

Interpretation of assays.

Assays may be interpreted by a general method for litters in which the doses are distributed in any way by reducing all the standards to one level and all doses of unknown substance to one level by calculation. The average values are then compared and an assay made for each litter. A weighted average of the separate litter assays is then taken as the final assay. This average must be calculated geometrically (*i.e. via* the logarithms of the values) since variations in doses corresponding to the normally variable healing will not be normally distributed.

Table II. *Assay of a sample of vitamin D concentrate.*

Litter	Rat	Dose	Area of new calcifi- cation mm. ²	Litter	Rat	Area of new calcifi- cation mm. ²
146	1103	Negative control	0	148	1118	0
	1110	"	0		1119	0
	1102	$\frac{1}{2}$ unit standard	92		1115	115
	1112	"	88		—	—
	1113	"	103		—	—
	1105	$\frac{1}{2}$ unit standard	198		1116	160
	1108	1 unit standard	243		—	—
	1101	0.125×10^{-5} g. conc.	119		1117	131
	1107	"	99		—	—
	1111	"	127		—	—
	1104	0.25×10^{-5} g. conc.	159		1114	147
	1109	"	167		—	—
	1106	0.5×10^{-5} g. conc.	167			

When the litters contain equal numbers of rats on comparable doses of the standard and the unknown, the animals may be considered in pairs and the differences between the individuals of each pair averaged. From this

difference the ratio of the doses may then be calculated by reference to the curve given in Fig. 4 or from the formula

$$(y_1 - y_2) = 196 \log \frac{x_1}{x_2}$$

where $(y_1 - y_2)$ is the difference in healing and x_1/x_2 is the ratio of the doses.

The calculations are illustrated by the assay in Table II.

Method I. (General method for any distribution of doses.)

The average healings produced by $\frac{1}{2}$ unit standard and by 0.25×10^{-5} g. of concentrate are calculated as follows.

Litter	Rat	Healing produced by $\frac{1}{2}$ unit standard	Rat	Healing produced by 0.25×10^{-5} g. concentrate
146	1102	92 + 59 = 151	1101	119 + 59 = 178
	1112	88 + 59 = 147	1107	99 + 59 = 158
	1113	103 + 59 = 162	1111	127 + 59 = 186
			1104	= 159
	1105	= 198	1109	= 167
	1108	243 - 59 = 184	1106	167 - 59 = 108
		Average = 168.4		Average = 159.3
148	1115	115 + 59 = 174	1117	131 + 59 = 190
	1116	= 160	1114	= 147
		Average = 167		Average = 168.5

From litter 146 the value of 0.25×10^{-5} g. of concentrate in terms of standard is given by

$$168.4 - 159.3 = 196 \log \frac{0.5}{x},$$

$$x = 0.5/1.113$$

$$= 0.45 \text{ unit.}$$

From litter 148 the value of 0.25×10^{-5} g. of concentrate in terms of standard is given by

$$168.5 - 167 = 196 \log \frac{x}{0.5},$$

$$x = 0.5 \times 1.017 = 0.508 \text{ unit.}$$

The geometrical average of these two values, weighted for the number of rats in each litter (excluding controls) is 0.466 corresponding to

$$0.466/(0.25 \times 10^{-5}) = 186,000 \text{ units per g.}$$

in the original concentrate.

Method II. (Special method when litters are divided into pairs of rats.)

If one of the rats on 0.25×10^{-5} g. of concentrate had been omitted, the above two litters could have been considered as seven pairs on comparable doses of standard and unknown. Omitting rat 1104, the differences between the pairs are:

$$\begin{array}{rcl} 92 - 119 & = & -27 \\ 88 - 99 & = & -11 \\ 103 - 127 & = & -24 \\ 198 - 167 & = & 31 \\ 243 - 167 & = & 76 \\ 174 - 190 & = & -16 \\ 160 - 147 & = & 13 \\ \hline \text{Average} & = & 6.0 \end{array}$$

This average difference corresponds to a dosage ratio of standard to unknown, in any of the pairs, of 1.073, *i.e.* 0.5×10^{-5} g. concentrate contains $1/1.073 = 0.932$ unit, or 1 g. contains 186,000 units.

Calculation of errors.

The errors to be expected in practice have been calculated by two of the methods described by Bourdillon *et al.* [1931]. The total number of rats available for consideration, however, is only about 200, and the experiments extend over a period of 7 months. The figures are, therefore, somewhat provisional.

The mean value for the increment obtained on doubling the dose is, as given above, 59.0 mm.² The standard deviation of this value is 30.8 mm.², *i.e.* the probable error of any result calculated from n pairs of rats is $\frac{2}{3} \cdot \frac{30.8}{\sqrt{n}}$.

In many litters several rats have received uniform doses of standard or unknown substance, or have been divided between two levels of dosage. When at least 4 rats have received the same dose, or can be reduced to the same level by calculation, the deviations from the mean for the litter have been calculated. The standard deviation calculated from these figures is 20.6. The standard deviation between two means is thus $20.6 \times \sqrt{2} = 29.2$, a value agreeing with the standard deviation for pairs of rats receiving doses in known ratio. It is of interest to note that the standard deviation from the mean for animals receiving the same dose is 20.6 if the animals considered are litter-mates, and 33 if the animals are not litter-mates (calculated from values for rats receiving the standard preparation over the whole period). Thus confining the comparisons to litter-mates reduces the probable error of the result by about one-third.

Taking 30 as the value of the standard deviation of the difference between two means, the errors to be expected in an experiment with n pairs of rats have been calculated.

Table III.

No. of pairs of rats	Percentage error corresponding to			
	$\frac{2}{3} \frac{\sigma}{\sqrt{n}}$		$\frac{2\sigma}{\sqrt{n}}$	
	"Probable error"		3 times "probable error"	
1	+27	-22	+102	-51
2	18	16	65	40
3	15	13	50	33
5	11	10	37	27
10	7.7	7.0	25	20
20	5.4	5.1	17	15
30	4.4	4.2	14	12
50	3.4	3.4	11	9.5
100	2.4	2.4	7.3	6.8

The values of percentage error corresponding to $\frac{2\sigma}{\sqrt{n}}$, *i.e.* the error likely to be reached or exceeded once in 22 groups, are plotted in Fig. 5, which shows

also the values calculated similarly by Bourdillon *et al.* [1931] for the estimation of vitamin D by radiography. It is observed that the errors now estimated are about two-thirds those given by Bourdillon *et al.*, which accords with their finding that about one-third of the probable error is due to error in the diagnosis of healing. With this reduction in the probable error, the number of rats required to produce an assay of a given degree of accuracy is approximately halved.

Sources of error.

Errors in measurement. The error in measuring the area of cross section of the new calcification deposit varies according to the type of picture presented. Small areas such as are shown in Fig. 1 can be measured to within a few mm.², but in measuring larger areas the error is more considerable. It is, however, not likely to be of importance when the healings produced in the animals compared are of approximately the same order. Assays based on a comparison of considerable healing produced by one substance with slight healing produced by the other are necessarily unsatisfactory, and can only be considered as approximate.

Slight errors in measurement may also be caused by the variation in the deposit at different planes of section and by slightly oblique viewing of the bone when making the drawing. These again will only affect to an appreciable extent the measurement of large areas.

Since the relation between healing and dosage is logarithmic, the errors in the assay from these causes are determined by the absolute errors in measurement and not the percentage errors. It therefore appears to be more satisfactory to select doses that produce slight healing ($\frac{1}{8}$ to $\frac{1}{4}$ unit). The variation in response in this neighbourhood does not appear to be greater than at higher levels of dosage.

The error in measurement could be reduced by using a higher magnification in making the drawing and by using both legs of each animal. The extra labour involved, however, does not appear to be justified by the slightly increased accuracy that would be obtained. It might, however, be possible to develop other more accurate methods of measuring the deposit, such as a photometric determination of the average blackness in the area of the line, or a micro-calcium estimation on the complete disc of newly calcified tissue dissected out from the remainder of the bone.

Fig. 5. Relation between error and number of pairs of rats in an assay. The curves show the percentage error corresponding to three times the probable error. The thin line is plotted from the values given for the estimation of vitamin D by radiography [Bourdillon *et al.*, 1931].

It may be mentioned here that in one routine assay the tibia was used as well as the radius + ulna. The calcium deposits were much less dense than in the radius + ulna, and were less suitable for accurate measurement.

Error due to variation in response. This error is responsible for nearly the whole of the error observed in practice. As is shown above, the error is much reduced by confining comparisons to litter-mates.

Other sources of error. The individual records in a number of vitamin D tests have been examined to determine the effect of the following variables: sex, weight of the animal at the beginning of the experiment, fore-period on the deficient diet, age and growth during the experimental period. None of these appears to exert any considerable influence. Growth during the experimental period appears to be inversely related to healing in many litters, but the effect is not uniform. The results are summarised below.

Influence of sex. The values of the healing produced by the standard preparation in 16 litters were considered. $\frac{1}{2}$ -unit figures were converted to $\frac{1}{4}$ -unit figures by adding 59, and $\frac{1}{2}$ -unit figures were converted by subtracting 59. The healings for bucks and does in each litter were averaged separately and the litter averages were then averaged as follows: bucks 128 mm.², does 128 mm.²

Influence of weight at the beginning of the experimental period. Considering the same figures, the weights of the rats and the healings were averaged irrespective of sex, and then the deviations from the mean healing for each litter were tabulated separately according to whether the individual exceeded or fell short of the litter average weight. For 30 rats exceeding the average weight, the average of the deviations was -1 mm.² For 29 rats of less than the average weight, the average of the deviations was 0 mm.²

Influence of the length of the fore-period. The values from 22 litters expressed as $\frac{1}{4}$ -unit standard were averaged for each litter.

4 litters had a fore-period of 25/27 days, mean healing	122 mm. ²
8 " " less than 25 days, mean healing	126 "
8 " " more than 27 days, mean healing	128 "

Influence of age at beginning of experimental period. 23 litters were considered. All were weaned at 23 or 24 days.

9 litters were less than 29 days from weaning, mean healing	127 mm. ²
5 " 31 to 33 days from weaning, mean healing	116 "
9 " more than 33 days from weaning, mean healing	129 "

Influence of growth during the experimental period. The deviations from the mean healings for each litter were tabulated separately according to whether the individual's growth was more or less than the mean for the litter. Of 27 animals showing less than average growth, 16 showed more than average healing and 11 showed less. The average deviations were -2.25 g. and $+2.7$ mm.² Of 27 animals showing more than average growth, 15 showed less than average healing and 12 showed more. The average deviations were $+2.0$ g. and -4.9 mm.²

It is seen that, so far as the figures considered show, none of the variables exerts an appreciable effect, certainly nothing for which an allowance can be made. When a litter is being taken for test, pairs can be selected according to sex and weight, but it is impossible to pair animals according to the growth that they will make during the experimental period. Age and length of fore-period are the same for all the animals of one litter.

*Variation in the healing produced by the standard preparation
at different periods of the year.*

The healings produced by different doses of the standard preparation in routine tests since September 1931, when the method was first applied, have been averaged throughout as $\frac{1}{2}$ unit. The results are summarised in Table IV.

Table IV.

Period covered	No. of litters represented	No. of rats	Mean healing as $\frac{1}{2}$ unit standard
18. ix. 31 to 23. x. 31	6	11	173
26. x. 31 to 18. xi. 31	7	19	186
21. xi. 31 to 21. xii. 31	6	26	187
21. xii. 31 to 6. i. 32	7	15	200
10. ii. 32 to 8. iii. 32	11	28	182
19. iii. 32 to 5. iv. 32	7	21	194
6. iv. 32 to 4. v. 32	10	23	167

During the corresponding period of an earlier year, Bourdillon *et al.* [1931] obtained a variation in healing produced by $\frac{2}{3}$ unit of standard, from about 6 scale divisions in September to 4.7 in February and 8.2 in April, corresponding to a spread of about 100 mm.² in healing measured by the present method. It appears likely that the greater variation obtained is to be explained by a difference in the criterion of healing adopted. The present method does not take into account the general rachitic condition, but only the calcification in the line. As was shown in Fig. 3, two rats may show the same healing while showing general rachitic conditions of different degrees. So long as a line is obtained, the values are not likely to be affected by differences in the degree of severity of rickets developed in the animals before they are taken for test.

SUMMARY.

A method is described for the assessment of the degree of healing in the "line" test for vitamin D by measuring the area of new calcification shown on magnified camera-lucida drawings of the bone sections. In this way a much finer gradation of healing is distinguishable than is possible by the usual method of eye judgment, and uncertainty in estimating the degree of healing is eliminated.

Within the range studied, the healing measured thus is found to be proportional to the logarithm of the dose of vitamin D.

From an examination of the results so far available it appears that the expected probable error in an assay based on 10 pairs of rats is + 7·7 or - 7 %, and the error expected once in 22 tests is + 25 or - 20 %.

The chief source of error is the variable response of litter-mates to a given dose of vitamin. The influence of a number of variables such as sex and weight has been studied, and appears to be negligible.

REFERENCES.

- Bills, Honeywill, Wirick and Nussmeier (1931). *J. Biol. Chem.* **90**, 619.
— and McDonald (1926). *J. Biol. Chem.* **68**, 821.
Bourdillon, Bruce, Fischmann and Webster (1931). *Med. Res. Council Special Report Series*, No. 158. (London, H.M. Stationery Office.)
Coward (1928). *Quart. J. Pharm. Pharmacol.* **1**, 27.
Dyer (1931). *Quart. J. Pharm. Pharmacol.* **4**, 503.
McCollum, Simmonds, Shipley and Park (1922). *J. Biol. Chem.* **51**, 41.
Shipley, Park, McCollum, Simmonds and Parsons (1921). *J. Biol. Chem.* **45**, 343.