ACCURACY OF HEMOGLOBIN DETERMINATIONS ON FINGER-TIP BLOOD¹

DOROTHY G. WIEHL

N analysis of the sources of errors and their magnitude for hemoglobin determinations made by the photoelectric method is presented in this report. The hemoglobin estimates were part of serial examinations on a group of employed women. In order to evaluate fluctuations in the hemoglobin levels at different times, it is necessary to have a measure of the variation in estimated amounts of hemoglobin in the blood that is to be expected as a result of the experimental error for the method used in determining the hemoglobin values. Since many studies on hemoglobin are made which involve exact comparisons of values obtained at different times or under changed conditions the accuracy shown for the photoelectric method by this analysis is of interest.

The blood sample used for determining the amount of hemoglobin was taken by pricking the finger-tip. The finger-tip sample is easily and quickly obtained and, as a rule, the subject is less likely to object to having the finger pricked than to having a venipuncture sample taken. For surveys and in the routine practice of most physicians, a sample taken from the finger or ear-lobe is usually the preferred method of taking the blood specimen.

Studies Comparing Venous and Cutaneous Blood

The comparability of cutaneous or peripheral blood with venous blood in respect to hemoglobin content and the number of red blood cells has been investigated in a number of studies. These studies have been reviewed and if the data were published in sufficient detail, the findings are summarized in Table 1, and computations have been made to estimate the significance of differences shown for venous and cutaneous blood specimens. It was found that in the

¹ From the Milbank Memorial Fund.

1

ß

5

15

oť

ŀ

Ķ

:w he

'n

ĩ

.:

15

ö

ċ

X

by or

ņ

d 12

re

1

Д,

ĉ,

Table 1. Summary of results of various studies comparing the hemogroun content and the red blood cell counts for venous blood and cutaneous blood with new estimates¹ of the statistical significance of differences in reported values.

Source of Blood and Values Compared	Num- ber of Cases	Mean Differ- ence	St'd. Error of Mean	Proba- bility of Chance Occur.	Basis For Estimated Standard Error of Mean			
		ADULTS						
<i>Hemoglobin;</i> Vein—Finger Andresen and								
Mugrage, 1938 (2)	30	— .06 gm.	±.061	>.30	Reported C. V. of 2.2% for 10 de- terminations on same sample			
Rud, 1922 (6)	18	+ .33%	±.396	>.40	Differences between paired values			
Vein—Ear Price-Jones, 1935 (3)	90	-1.21%	±.168	<.001	Reported C. V. of 1.2% for values on 10 days for same person			
Rud, 1922 (6)	27	+ .26%	±.554	>.60	Differences between paired values			
Vein—Artery² Gibbs, 1942 (4)	50	054 gm.	±.0107	<.001	Differences between paired values			
Red Blood Cells;		Millions						
Vein—Finger Andresen, 1938 (2)	30	037	±.012	<.01	Reported C. V. of 1.3% for 10 de- terminations on same blood sample			
Rud, 1922 (6)	18	+ .056	±.038	>.10	Differences between paired values			
Bogendorfer, 1921 (7)	46	+ .518	±.067	<.001	Differences between paired values			
Vein—Ear Price-Jones, 1935 (3)	96	078	±.0056	<.001	Reported C. V. of 0.7% for values on 10 days for same person			
Rud, 1922 (6)	41	075	±.034	.0205	Differences between paired values			
Duke, 1922 (8)	9	+ .082	±.047	>.10	Differences between paired values			
Bing, 1919 (9)	10	320	±.208	/ /.10	Differences between panet values			
	CHILDREN AGED 2-14 YEARS							
Hemoglobin; Vein—Finger			Lorr	20	Perperted C. V. of 2.2%			
Andresen (2)	30	+ .03 gm	±.033	2.30				
Vein—Ear Rud (6)	4	o						
Red Blood Cells; Vein—Finger		Millions			Perented C. V. of a off			
Andresen (2)	30	012	±.011	>.20	Reported C. v. of 1.3%			
Vein—Ear Rud (6)	11	023	±.087	>.70	Differences between paired values			

¹ See text, pp. 7 and 8. ² Blood from internal jugular vein and from the femoral, radial or brachial artery. Original data given in oxygen combining capacity.

•

ú

=

U

i nà

т<u>è</u>

7.2

nh:

inia

i Tân

101

÷

indes

l tala

173

12

1

2

Ţ.

original articles either no statistical test of significance for differences was applied to the data or the test used was not the most appropriate one for the problem.

If the results from samples of venous and of cutaneous blood were published for individuals, differences between the two values for each person were computed.² A frequency distribution of differences is obtained for which a standard deviation of the distribution and standard error of the mean of differences are computed in the usual way. Since it is desired to know whether there is any constant or systematic difference between venous and cutaneous blood, differences are computed with signs. Thus, for Table 1, the value for cutaneous blood was always subtracted from the value for venous blood and a minus difference indicates that cutaneous blood had the higher value. The mean of the differences computed in this way is the same as the difference between means for any set of original values. The reliability of this mean difference is tested only by the variation associated with the two types of blood samples for the same person and the variation in blood values for different individuals is eliminated since this variation is not related to the problem.

In two extensive statistical studies of venous and cutaneous blood, Andresen and Mugrage (2) and Price-Jones, Vaughan and Goddard (3), data for individuals were not published and the statistical significance of mean differences was measured in terms of the standard errors of the means derived from the distribution of values for different individuals and not, as described above, in terms of the variation between paired values for the same individual. Since hemoglobin values vary widely among individuals, the standard error of the difference in means for the observed values is considerably larger than the standard error of the mean of differences between paired values and means were found to be not significantly

^a For examples of the use of differences between paired values and a discussion of their use in place of a comparison of the two distributions of values, *see* Snedecor (1).

different by this method which undoubtedly would be significantly different if judged by the variation between paired values. In both of these studies, some data are given on the accuracy of the hemoglobin method used and of the red cell counts and from this information a standard error of the mean difference has been computed which measures the probability that the observed mean difference would occur if the experimental error of the method were the only source of variation between the paired values. Thus, if the mean of the differences is significantly greater than is expected on the basis of error variation, it may be concluded that some other factor was operating, and in this case the factor presumably would be the source of the blood samples. If the standard deviation for experimental error of the method is based on a sufficient number of observations which include all sources of error which may affect the determinations on venous and cutaneous blood, the estimated standard error for a mean difference affords a valid measure for judging the observed difference. In these two studies, the coefficient of variation for the error of the method is given and is based on ten determinations which are too few observations to give a highly reliable estimate of the error. Furthermore, Andresen and Mugrage used ten determinations on the same blood sample and consequently their experimental error does not include any variation that might be associated with independent samples from the same person. Therefore, the estimated standard errors of mean differences' which would be expected from experimental error in the

* The standard error of the mean difference was computed from the coefficient of variation for experimental error as follows:

1. Coefficient of variation = $\frac{S.D. x}{Mean}$ x 100; therefore, the standard deviation = $\frac{C.V. x}{100}$ Mean

This standard deviation is the experimental error for a single determination. And resen and Mugrage made duplicate determinations for each value and used the average; and the experimental error of this average value would be $\sqrt{\text{SD}^a}$.

2. For a series of differences between two values, each of which has a given experimental error (S.D.), the standard deviation = $\sqrt{2 (S.D.)^2}$.

3. The standard error of the mean of the differences is the standard deviation for the differences divided by \sqrt{N} and N is the number of pairs or number of differences.

Accuracy of Hemoglobin Determinations

i

5

ľ

2

11

10

ŀ

Ľ

ż

2

ند: ند:

þ

Ŀ

method are somewhat too low for the Andresen and Mugrage data. From the study by Price-Jones, *et al.*, the coefficient of variation for determination on venous specimens taken from the same person on ten different days has been used for the experimental error.

From Table 1, it is evident that the difference between hemoglobin determinations and between red cell counts on venous and cutaneous blood is very small both for adults and for children aged 2 to 14 years. For hemoglobin, only the difference reported by Price-Iones between determinations on venous blood and blood from the ear-lobe is found to be significant,⁴ and the probability that the difference would occur as the result of accidental variation is less than one in a thousand. Although the hemoglobin values on venous blood determined by Price-Jones are significantly lower than those for blood from an ear-lobe, the difference was only 1.21 per cent hemoglobin, or about .17 gm., and for most clinical purposes so small an amount would not be important. Data from Gibbs, et al. (4), which compares hemoglobin in venous and arterial blood have been included in Table 1, and indicate a significant tendency for venous blood also to be very slightly lower in hemoglobin than blood from the femoral, radial or brachial artery. The difference is only .054 gm. and is of interest chiefly because it suggests that slightly higher hemoglobin values for blood from the ear-lobe and finger-tip may not be entirely due to irritation and congestion from the incision, as suggested by Drucker (5), or to other artificial causes.

Comparisons of red cell counts in venous and cutaneous blood are available from a larger number of studies and, in general, show results similar to those for hemoglobin. In the majority of studies, there is a tendency for the cell count to be somewhat lower for venous blood than for blood from a finger-tip or an ear-lobe, but the difference is not always statistically significant. For the Price-Jones data, the difference is very significant⁴ and the data from Andresen and

⁴ In the original report, using the standard error of the difference in means for the distributions of values for individuals studied, the authors conclude that the difference is (Continued on page 10)

Mugrage show a significant difference for adults but not for children.⁵ Values from Rud (6) also show a significantly lower cell count for venous blood for adults as compared with blood from the ear-lobe, but the counts on venous blood did not differ significantly from those on finger-tip blood. From one study, that of Bogendorfer and Nonnenbruch (7), a significantly higher cell count on venous blood samples was reported. Since Bogendorfer and Nonnenbruch found that this difference could be eliminated by putting the finger in hot water before taking the blood, it is possible that the puncture had not been sufficiently deep to establish a free flow of blood or that other artificial factors affected the findings. From the available data, it seems that the red cell count on venous blood may be very slightly lower than on that from the finger or ear-lobe. Inconsistent results and lack of significance in some studies could be explained as the result of the effect of such factors as the relatively large experimental error for red cell counts, the small number of cases compared, the technique of taking blood samples, and artificial and other factors affecting circulation of blood in the capillaries and arterioles.

Hemoglobin values and red blood cell counts seem to show a greater difference for venous and cutaneous blood in infants than in adults and older children. For thirty infants aged 1 to 19 months, Andresen and Mugrage obtained hemoglobin values on venous blood that averaged 0.55 gm. less than values on blood from the heel, and red cell counts were .051 million lower. For thirty infants one-half hour to nineteen days old, venous blood values averaged 0.77 gm. of hemoglobin and 0.264 million red cells lower. These

not significant. If the experimental error used for the standard error of the mean of the differences in Table 1 were increased about three times for hemoglobin determinations and five times for red cell counts, the mean difference would still be significant.

^b Andresen and Mugrage also compared the volume of packed cells. For adults, the mean volume for venous blood was the same as the mean for finger blood; and for children, the mean volume for venous blood was slightly higher. Thus, the results for hemoglobin, number of cells, and volume of cells are not consistent. Furthermore, the validity of the significantly lower cell count on venous blood for adults is questionable since the reported experimental error by which the difference was tested probably is too low for error variation for independent samples.

Accuracy of Hemoglobin Determinations

ŀ

1

28

μ

İtr

ΰij

œ

le T

tere

άc

<u>j</u>hla

m

YC:

Ľà

x

<u>CD</u>

 \mathbf{r}

Ð

0¥.

Ē

203

212

ico Qui

::: 0::

ipe*

differences would almost certainly be very significant statistically if tested by the method of differences between paired values. It is apparent that the difference was much larger for the very young infants. For six infants five to twenty-four days old, Haden and Neff (10) reported red cell counts for blood from the longitudinal sinus much lower than counts for blood from the heel in five cases (differences were 0.68 to 2.71 million), and higher by 0.11 million for one case. On the other hand, Lucas, et al. (11) reported higher hemoglobin values and red cell counts for blood from the longitudinal sinus than for unspecified peripheral blood for infants one to eight days old. Average values for sixty to 100 infants are given but comparisons are not made between averages for the same infants. Thus, the findings on infants have not been consistent, but the more controlled, recent study by Andresen and Mugrage strongly supports the view that in infants, especially the new-born, hemoglobin values and red cell counts are considerably lower for blood taken from the vein than for that from the heel.

In general, it may be concluded that, except for infants, blood from the finger or ear-lobe is not appreciably different from blood taken from a vein, although there is some evidence that both hemoglobin values and red blood cell counts tend to be slightly lower for venous blood. For most purposes,⁶ cutaneous blood gives satisfactory results which are comparable with determinations on venous blood. However, since there may be a small, systematic difference, for special investigations in which small changes or differences are studied and tested statistically, values for venous blood and cutaneous blood should not be used interchangeably or considered to be identical. This precaution in making comparisons of values for blood from the two sources is necessary whether the difference is in part a true difference in venous and capillary or arteriolar blood

⁶ Comparisons of venous and cutaneous blood have been for so-called healthy or normal persons. Duke and Stofer (8) found much higher red cell counts for blood from an ear-lobe than for venous blood in cases of pernicious anemia, but not in cases of secondary anemia. For hematological study of abnormal bloods, venous specimens seem preferable.

or is entirely due to the effects of the capillary technique, such as irritation, constriction, and congestion.

METHODS AND DATA COLLECTED

The hemoglobin values for finger-tip blood used in the following analysis were obtained by the Evelyn method (12) for photoelectric determination of oxyhemoglobin. A single Evelyn Photoelectric Colorimeter was used for all determinations. The instrument had been purchased in 1941, approximately a year before it was first used for the serial examinations of this study. The manufacturer's calibration and K_2 value were accepted. Any error in calibration would be constant for all values and would have no effect on this analysis of experimental error.

The procedure was to pierce the tip of a finger with a spring lancet and draw 20 cu. mm. of blood into a calibrated capillary pipet from which it was discharged into 10 cc. of distilled water in a colorimeter tube. The pipet was thoroughly rinsed by drawing the distilled water into the pipet about three times. After about ten minutes one drop of 28 per cent ammonia water was added and the tube was thoroughly shaken. The outside of the tube was carefully wiped before it was inserted in the colorimeter. The galvanometer was read to the nearest one-quarter unit of scale.

Two technicians made all the hemoglobin determinations on finger-tip blood samples. A single technician carried out the entire procedure for the hemoglobin determinations used in this analysis. One technician made all the determinations in a given examination period except for a few cases which have not been used in this report. Technician B made the determinations in two periods and Technician C in four periods. Both technicians had had considerable previous experience with this method of hemoglobin determination.

One technician (B), used the same capillary pipet for all blood specimens at both examinations and variation in determinations due to differences in accuracy of pipets is eliminated. The other technician (C), used a different pipet for the right and left-hand blood specimens and had new pipets for examinations in each period. In a given period, only two pipets were used except for an occasional sample taken with a third extra pipet. The two pipets were used at random for right and left specimens. Between specimens the pipet was thoroughly cleansed with distilled water and alcohol, then with acetone.

Hemoglobin determinations were made on a group of employed women who were healthy enough to be at work. Six examinations were made at about six-month intervals but many did not have all examinations. Except at the first examination, determinations were made for the right and left hand. These two independent hemoglobin values furnish the data for an estimate of the experimental error for the determinations on finger-tip blood.

DIFFERENCE BETWEEN TWO SAMPLES OF FINGER BLOOD

The mean difference between right and left-hand hemoglobin values for determinations made in each of five periods is shown in Table 2. These means are for differences with signs and the lefthand value was always subtracted from the right-hand value. A plus value, therefore, indicates that the right-hand value was higher.

Technician and Examination	Number of Cases	Mean Difference Gms. of Hb. (Rt.—Lt.)	Standard Error of Mean Diff.	Standard Deviation	
Total	391	+.062	.013	.250	
Technician C	261	+.037	.015	.244	
Examination 2 Examination 4 Examination 6	112 95 54	+.016 +.053 +.051	.023 .027 .028	.241 .264 .208	
Technician B Examination 3 Examination 5	130 85 45	+.115 +.159 +.030	.022 .030 .029	.256 .273 .197	

Table 2	. Difference	between	hemoglobin	values	for	blood	taken	from	a	finger
tip of the	right and le	ft hand.	-							-

ÿ

æ

In four of the five periods, the mean difference between right and left-hand values is very small, ranging from +.016 gm. to +.053gm. These mean differences are not significant in a statistical sense except the highest value of .053 gm. which is just at the conventional line of significance and has a probability of occurring from chance variations of five in one hundred times. A much higher mean difference (+.159 gm.) is found for the two blood samples taken by Technician B in the third examination period and this mean is very significant. It will be noted that at each examination the sign of the mean difference is plus, and for the total cases for each technician the mean difference is significant. Apparently some factor operated to produce values that were slightly higher for the right hand. The cause of this bias is unknown. The most likely cause would seem to be some defect in technique. From the data available, the bias may be due either to a tendency for right-hand values to be too high or left-hand values to be too low.

As previously stated, Technician C used different pipets for the right and left-hand blood specimens and new pipets at each examination. The standard deviations for differences between the two specimens are about the same for both technicians and, therefore, there is no evidence that differences in pipets affected the variation between specimens taken by Technician C. Although the small mean difference between the right and left-hand values could have resulted from pipet differences, if the pipets were not used at random, as the technician believed, it is very unlikely that inaccuracies in the pipets would account for consistently higher right-hand values in several different examination periods.

For the relatively large mean difference for Technician B in the third examination period, there is reason to think that values for the left hand were too low." Since the same pipet was used for both hands and the right specimen was always taken first, it seems pos-

⁷ The mean of hemoglobin values for all women was slightly lower in the third period than at any other examination. If the left values were too low, the effect would be to lower the mean for the group (*See* Table 4).

Accuracy of Hemoglobin Determinations

1

d

3

Ż

ų

12

.

5

1

1

-

Db

سیم آغذ

Πũ

ſŻ

12

;阆

jr.

1

<u>____</u>

hm

. سند ا

100

1213

12

3Ĩũ

100

; p06

pool

lond

sible that at times enough moisture remained in the pipet after cleansing to have a slight effect on the volume of the left-hand blood specimen.^{*} If this dilution factor did affect the left-hand values obtained by Technician B in the third examination, it may have been in addition to the undetermined factor which caused a slight bias in the values obtained by Technician C.

The very small but significant bias disclosed for these duplicate determinations is an example of systematic error which is apt to escape notice.

Accidental Error of One Hemoglobin Determination

The observed difference between the two hemoglobin values for the same person at a given time is the result of the systematic bias plus the accidental error⁶ for each of the determinations. The bias error is small and its effect on the variation between two samples may be eliminated by subtracting the mean difference for all examinations in a specific period from each difference between paired values for that period. This adjusted difference is equal to the deviation of each observed difference from the mean difference, and the standard deviation for these deviations is shown in Table 2. Thus, the standard deviation in Table 2 represents the variation between the two samples which resulted from accidental error for the two determinations.

On the assumption that each determination contributed one-half of the variation between samples which resulted from accidental error, the standard deviation for error variation in one determination is the standard deviation for accidental variation between samples divided by the $\sqrt{2}$. The standard deviation for error in

⁹ The term accidental error is used here to include all random or non-systematic variation associated with taking the blood samples and the technique of making a hemoglobin determination.

⁸ In the re-examination of the group in December, 1945, Technician B took the left-hand blood specimen first and, as before, used the same pipet. For fifty-one paired values, the mean difference is $-.055 \pm .023$, and the right-hand value is significantly lower than the left-hand value. This apparently confirms the hypothesis that, in spite of routine use of alcohol and acetone, there was a tendency for some moisture to adhere to the interior of the pipet.

one hemoglobin determination is shown in Table 3 for each examination period, and the average standard error is given for each technician and for all determinations.²⁰

From the total experience the average standard error of a hemoglobin determination is 0.174 gm., that is, a hemoglobin value on fingerprick blood in this study was accurate within plus or minus 0.17 gm. in two out of three times and within 0.35 gm. in ninety-five out of one

Technician and Examination	Number of De- terminations	Standard Error Gms. of Hb.					
Total	782	.174					
Technician C	522	.172					
Examination 2	224	.171					
Examination 4	190	.187					
Examination 6	108	.147					
Technician B	260	.176					
Examination 3	170	.193					
Examination 5	90	.139					

Table 3. Standard error for accidental variation of a single hemoglobin determination.

hundred times so far as error from accidental sources in the technique and from blood sampling are concerned.

There was considerable variation in the errors estimated for the five different series of determinations. The standard errors for different periods ranged from .14 to .19 gm., as shown in Table 3. Technician B had both the highest and the lowest accidental error, and these standard deviations are significantly different; but the error for his total determinations was about equal to the standard error for all determinations by the other technician. The highest standard error for Technician C also differed significantly from his lowest error. Obviously, the error variation was not constant, even though estimated from a rather large number of determinations in each period. The combined experience probably affords the best estimate of the accidental error in this type of hemoglobin determination.

Nearly all hemoglobin values for this group of women were

¹⁰ The average standard errors for each technician and for all five examination periods shown in Table 3, were computed as weighted averages of the variance for error in each period and are not exactly one-half the square of the standard deviations given in Table 2.

between 11.50 gms. and 14.50 gms. and the mean was about 13 gms. There was no evidence of a consistent difference in the magnitude of the standard error for hemoglobin values at different levels within this range. The accidental error of .174 gm. is about 1.3 per cent of the mean hemoglobin level for this group, that is, the coefficient of variation for accidental error is 1.3 per cent. Since the estimated hemoglobin value may be either higher or lower than the true value, the true value represented by a determination will be within a range of four times the error in ninety-five out of a hundred times, and a 1.3 per cent error means that a hemoglobin estimate is unlikely to differ by more than 5.2 per cent from the true value.

SOURCES OF ACCIDENTAL ERRORS

The error in these duplicate, right and left-hand, determinations is the composite or net result of errors from several sources. In addition to variation due to blood sampling, the technician may make errors in measuring the sample or reading the galvanometer, undetected turbidity or imperfectly cleansed colorimeter tubes may cause error, and the performance of the photoelectric colorimeter may be a source of slight variation. The contribution of sampling variation for independent finger-tip blood specimens to the total accidental error has been estimated by comparing this error with that found when duplicate hemoglobin determinations were made by the same photoelectric method on two subsamples of a venous specimen.

The accidental error of a hemoglobin determination on a subsample of a blood specimen taken by venipuncture was found to be .145 gm. per 100 ml. of blood. This was estimated from duplicates on two hundred persons examined in a survey of nutritional status of high school students in New York City (13). Estimates of the error were made from data on two groups of one hundred persons examined at different times, and the two estimates were almost identical although different technicians made the determinations and the work of several technicians was included in each series of one hundred. This accidental error of .145 gm. for the procedure of making a determination¹¹ is somewhat lower than the error of .174 gm. found for finger-prick samples. Since hemoglobin values were determined by identical methods in the two studies, it is reasonable to conclude that the higher error in the values on finger blood was due to differences in the blood samples which are not present when two subsamples of blood are measured from a single venous specimen.

The difference between the variance¹² for error of hemoglobin values on the finger-tip blood and that for venous blood gives a measure of the variance for the error associated with finger-prick blood sampling. This difference is as follows:

Variance for error of Hb. values on finger-tip blood	.030184
Variance for error of Hb. values on venous blood	.021077
Difference	.009107

The standard deviation for variation in finger-prick blood samples, as measured by these data, is $\sqrt{.009107}$ or 0.095 gm. Thus, if there were no procedural error in determining hemoglobin, a value derived from finger-prick blood could be expected to be accurate within plus or minus 0.19 gm. in ninety-five out of one hundred determinations. The sampling error for finger blood was less than the procedural error³³ and the finger-prick samples apparently were

¹³ The small variation in these blood samples taken by puncturing the finger-tip is in agreement with the findings of Berkson, Magath, and Hurn (14) who studied errors in (Continued on page 19)

¹¹ It is of interest to compare the procedural error estimated for this large series of determinations made by the Evelyn photolometric method with the error estimated from data reported by Evelyn (12) as typical of the accuracy of the method. From duplicate values for ten blood samples reported, we have calculated a standard error for one determination of .087 gm. Although this error is appreciably smaller than the error of .145 gm. for our routine, survey duplicates, the method has proven highly accurate for routine work.

¹⁹ The variance for the error is the square of the standard deviation of error. When variation arises from the effects of several factors, the total variance is the sum of the variances for the variation contributed by the separate factors. Thus, if total variance is known, and variance for one factor also is known, the difference between these variances may be taken as a measure of the variance for other factors.

affected very little by either natural or artificial causes.¹⁴ At an average hemoglobin level of 13.0 gms., this sampling error is about 0.7 per cent.

Variation in samples of venous blood was studied by Walters (15) who withdrew ten samples of blood from the same venipuncture and varied the position of the needle between samples. A series of samples was taken from ten young men. For each sample, six readings were made by the Newcomer method in a colorimeter and averaged. The standard deviation for the variation among the ten average values for the same person ranged from .143 gm. to .419 gm.; the mean variation was .282 gm. and the coefficient of variation was 1.75 per cent. This variation for venous samples may have been significantly affected by the error of the average values used in spite of having made six determinations for each sample. Variation in the volume of packed cells also was reported for the same venous samples, and the mean coefficient of variation was 2.05 per cent. Since the procedural error for volume of cells is small, most of this 2 per cent variation would be caused by sampling variation.

In a study of diurnal variations of hemoglobin McCarthy and Van Slyke (16) estimated hemoglobin by the CO capacity method for venous specimens taken six times during the day from twenty-three subjects. The average change in values from 2: to 5: p.m. was $-.051 \pm .105$ volume per cent of CO capacity and therefore not significant.¹⁵ If the variation between these twenty-three paired

¹⁴ Although the sampling variation is relatively small, it should not be neglected as a source of variation. For example, sampling variation as well as procedural error affects differences between venous and cutaneous blood. The reliability of any observed variation in hemoglobin at different times or under different conditions is affected by sampling error.

¹⁵ The hours 2: and 5: p.m. were selected because the difference was less than that between any other two periods and the variation between samples is least affected by any real diurnal variability. A careful analysis of the data from McCarthy and Van Slyke has been published by Mole (17) and indicates a significant diurnal fluctuation from 9: to 11: p.m. but not from 9: to 5: p.m. From the six diurnal observations from 9: to 11: p.m., by (Continued on page 20)

erythrocyte counts and found that variation in counts for different finger-tip specimens from the same person was not significantly greater than the variation for subsamples of the same specimens. Since the procedural error for cell counts is large, a small but real sampling variation would be easily masked.

values is assumed to be the result of procedural error in each determination plus sampling variation, and the standard deviation for this total experimental error of one hemoglobin value is estimated in the same way as described for the paired values on the right and left hand, we obtain a standard error for one hemoglobin estimate of .36 volume per cent of CO or an error of 1.8 per cent. The CO capacity method has a very low procedural error³⁶ and most of the variation of 1.8 per cent can be attributed to the blood samples. It is almost the same as that noted above for sampling variations in the study by Walters (15). In both studies, the variation indicated for venous blood specimens is greater than that estimated from our data for blood from a finger-tip.

Since it is common practice to make two determinations of hemoglobin and average them in order to obtain more accurate values, the amount of reduction in the experimental error obtained by averaging two determinations may be considered. If the determinations are made on two independent finger-tip blood specimens, both the error of sampling and of the method are reduced. The variance (standard error squared) for the average is one-half the variance for sampling error plus one-half the variance for procedural error or .021007 + .009107, and the standard error of the average is

 $\sqrt{.015092}$ or .123 gm. as compared with .174 gm. for one determination on a finger-prick sample. But if two readings are made from the same specimen, variance for the average is one-half the variance for procedural error $\left(\frac{.021077}{2}\right)$ plus variance for the

sampling error (.009107) and the standard error of the average is

analysis of variance Mole obtains a standard deviation of .46 volume CO for "uncontrolled error" after eliminating variance for average diurnal change and for individual averages. This standard deviation for "uncontrolled error" is affected to some extent by irregular diurnal variation as well as technical error and sampling variation.

¹⁰ McCarthy and Van Slyke (16) reported that the average difference between duplicates was .06 volume per cent of CO capacity but give no data on standard deviation or range for differences.

 $\sqrt{.019646}$ or .140 gm. For most purposes, the slight reduction in error attained by puncturing two fingers to obtain independent samples does not seem worth while. Although the error is not greatly reduced by duplicate determinations, two estimates are so easily made on a photoelectric colorimeter that the greater accuracy is obtained with little effort and may be desirable at least for studies.

.

PHOTOELECTRIC METHOD COMPARED WITH OTHER METHODS

The foregoing estimates for variation in determinations of hemoglobin by the photoelectric method indicate a high degree of reproducibility. The method has been found highly accurate by others. Karr and Clark (18) compared hemoglobin values obtained by oxygen capacity method and by the photoelectric method using a Sheard-Sanford photelometer and concluded that "after an electric photometer is calibrated it gives more consistently accurate results than the oxygen capacity." However, under careful laboratory conditions, the Van Slyke oxygen capacity method has a somewhat lower error for duplicate determinations³⁷ than the photelometric method but it is less suitable for surveys and routine work.

From a series of ten consecutive determinations on the same blood specimen made by the Hellige hemoglominometer, Andresen and Mugrage (2) reported a coefficient of variation of 2.2 per cent³⁸ which may be compared with the procedural error for our routine determinations by the photelometric method of .145 gm. or about 1 per cent.

Most statistical data on accuracy of methods is found in studies comparing one method with another. In comparisons of hemoglobin estimates on subsamples of the same blood specimens by

¹⁷ The standard error of a Van Slyke oxygen determination estimated from duplicates on ten blood samples reported by Evelyn (12) was .07 gm. and the error from duplicates on seven samples made by an experienced laboratory technician for calibration of the photoelectric colorimeter used for the nutrition survey (13) determinations was .05 gm.

¹⁹ The error of a hemoglobin determination is commonly reported as a percentage or coefficient of variation. Our data and data from other studies which we have examined indicate that error variation is more constant in absolute amounts of hemoglobin than in percentage.

different methods, variation is the result of procedural error plus any bias or systematic difference in the measurement of hemoglobin. Therefore, the differences between estimates by different methods are not comparable with the differences for repeated determinations by the same method which measure only procedural error, but with adjustment of differences between two methods for any average difference, the variability due to a systematic factor, such as calibration of instruments, is eliminated and the remaining variation is largely that associated with the method. Karr and Clark (18) studied the accuracy of a number of methods in general use and had laboratory technicians and physicians make hemoglobin determinations on subsamples of the same blood specimen by various methods. They reported the per cent of determinations which differed by not more than .5 gm. and not more than 1.0 gm. from the value obtained by the photelometric method. After corrections for systematic bias due to calibration of instruments or personal bias in matching color, the best results were obtained by one technician using a Hellige-wedge type colorimeter who had 83 per cent of his determinations within .5 gm. and 93 per cent within 1.0 gm. of the standard value, and by one technician using a Haden-Hausser colorimeter who had 80 per cent of his determinations within .5 gm. and 98 per cent within 1.0 gm. of the standard. On the basis of our standard error of .145 gm. for each value, differences between two photelometric determinations would have a standard deviation of .205 gm. and 98 per cent would be expected to differ by not more than .5 gm. and none would differ by more than 1 gm. Each technician made hemoglobin estimations on about forty bloods, and allowing for the statistical reliability of their percentages, these two technicians obtained results only a little less accurate than might occur for photelometric determinations. But for four other technicians using the Haden-Hausser colorimeter, only 45 to 55 per cent of their determinations differed by not more than .5 gm. Two technicians using Sahli instruments had 46 and 54 per cent of the values within .5 gm. of the standard; one technician using a Bausch and Lomb Newcomer instrument had 49 per cent of his determinations within .5 gm. of the standard. It is apparent that all of these methods furnished very unreliable estimates of the amount of hemoglobin, and if no correction in results is made for systematic bias the magnitude of the differences is much increased for most of the technicians. Results for the physicians' readings were less accurate than for technicians.

Periodic Variation in Determinations

The accidental error measurable from duplicate values does not reveal inaccuracies in the determinations which may result from some constant source of error. Inaccurate calibration of an instrument will give results which are consistently too high or too low by a constant amount, and an inaccurately calibrated pipet will have the same effect. These are systematic errors which must be guarded against and can be eliminated. But other systematic errors which may affect the determinations are more difficult to detect and unless a careful study is made the presence of bias or systematic error remains unknown. The personal equation is an important factor in methods which require the technician to match colors, and Karr and Clark (18) have shown that many persons tend to consistently underestimate or overestimate hemoglobin when matching color standards. The photoelectric method eliminates this personal factor. Evidence of other systematic bias was found in the determinations by the two technicians who examined the same group of women several times.

Comparison of the mean hemoglobin levels at six different periods for this group of women which are given in Table 4 shows that relatively high values were obtained at two examination periods. The higher values were for the second and sixth examinations, both by Technician C. In the first period, only one determination was made for each person and for the other five periods an average of the right and left-hand determinations was used. Not

Examination Period	TECHNICIAN OF WOMEN ¹		Mean Gms. Per 100 Ml.	Number of Women ²	Mean Gms. Per 100 Ml.		
I. Oct., 1942 II. July, 1943 III. Nov., 1943 IV. May, 1944 V. Nov., 1944 VI. May, 1945	C C B C B C	63 63 63 63 38 43	12.94 13.44 12.85 13.03 12.92 13.58	36 36 36 36 36	12.90 13.38 12.83 13.20 13.57		

¹ Same 63 women for periods I.-IV... ² Same 36 women for all five periods.

- Same 30 women for all five periods.

Table 4. Mean hemoglobin values for a group of women examined at approximately six-month intervals.

all women were examined at every period, and in the fifth and sixth periods both technicians made some hemoglobin determinations. Furthermore, women are excluded for whom the hemoglobin value was less than 11.0 gms. at any examination. Therefore, in Table 4, average hemoglobin values are shown for sixty-three women examined at each of the first four periods and for all those examined by a given technician in the fifth and sixth periods. It is apparent that the group mean values for the two periods in which Technician B made the determinations are very similar, and that the group mean values for two of the four periods in which Technician C made the determinations closely approximate the means for Technician B's values. But in two periods Technician C obtained values which were, on the average, higher than those in the other four periods. There was no consistent trend in the means and the two high periods were about two years apart³⁰ with the lower means prevailing for three examinations in between them. The conclusion seems justified, though not definitely proved, that some systematic overestimate of hemoglobin occurred in the second and sixth periods.

¹⁹ The second examination (high mean level) was in July 1943 and the sixth examination (high mean level) in May 1945. Thus, one was a summer period and one a late spring period. The fourth examination also was in the spring, and the other three were in October, November, or December. It is possible that there was some seasonal variation, but it does not seem likely that an average decrease in hemoglobin of .59 gm. occurred between July and November.

Source of Variation	Degrees of Freedom	VARIANCE	Ratio	Probability	
Individual Differences Period Differences Remainder—Uncontrolled	35 4 140	1.7786 3.5370 0.3950	8.95	<.001	

Table 5. Analysis of variance of hemoglobin values for thirty-six women at five different periods.

If the fifth examination period is excluded, hemoglobin values are available for thirty-six women for the other five periods. Average hemoglobin values for these thirty-six women are given in Table 4, and an analysis of the statistical significance of the variation among average values for the five periods is given in Table 5. The variation among mean hemoglobin values for the different periods is very significant, that is, it is much greater than would be expected as the result of random variability due to accidental error and other influences on hemoglobin fluctuation.

For the thirty-six women, the average hemoglobin value in the fourth examination period is significantly higher^{∞} than that for the first and for the third periods, and not significantly lower than the average for the second period. Thus, for this smaller group of women, the determinations by Technician C are relatively high in three of the four periods in which he did the examinations. The mean value for the thirty-six women in the fourth period is somewhat higher than the mean for the larger group of sixty-three women and it seems probable that the tendency to obtain higher estimates of hemoglobin was operating to some extent but not as consistently as in the second and sixth periods.

Evidence that the performance of the photoelectric colorimeter was not a factor in obtaining higher hemoglobin values in some periods is obtained from ten determinations made by Technician C

²⁰ The standard error of the difference between any two means is \pm .148 gm. This is obtained from the uncontrolled variance in Table 5 and is $\sqrt{2 \times 0.3950}$.

CRAMMAN CONTRACTOR

in the fifth examination period when Technician B made all other determinations. For the ten women examined by Technician C, the averages of hemoglobin values in the fourth and fifth periods were 12.43 gms. and 13.16 gms., respectively, indicating a significant increase in hemoglobin at the later period. But determinations by Technician B show no similar tendency to be higher in the fifth period. This finding points strongly to some shift in the technique of Technician C. The most probable explanation seems to be that he had a tendency to measure the sample somewhat generously and possibly also to read the colorimeter scale with a slight bias.

It is of interest and important that the personal factor in the technique of Technician C was not constant throughout his four examination periods. He had both high and low periods and these were not associated with any change in the accuracy of his determination as revealed by variation between right and left-hand values. Apparently any shift in his attitude or standards for measurements was fairly constant throughout a specific examination period.

Two troublesome problems in laboratory data are illustrated by these findings for hemoglobin values at six different examination periods. These are (1) using the identical technique and equipment, technicians may obtain significantly different results due to certain personal bias or attitudes; and (2) the same technician may change with respect to these personal factors over a period of time. Only the most careful and frequent evaluation of the work of technicians will reveal the presence of systematic bias due to the personal equation or bias from other sources. Regardless of the high degree of accuracy of the method used, such variations in results may occur. If undiscovered, a systematic bias in data easily may lead to erroneous conclusions.

SUMMARY

Sources of variation in estimates of hemoglobin content of blood made by the Evelyn photoelectric method are discussed.

From duplicate determinations on venous blood specimens from

200 subjects examined in a survey, the standard error of one determination is estimated as \pm .145 gm. of hemoglobin per 100 ml. This is a procedural error only.

For 391 examinations in which two independent determinations were made using blood from a finger of the right and of the left hand, the standard error of one determination is estimated as \pm .174 gm. This error includes the procedural error and blood sampling variation. On the assumption that the procedural error was equal in the two studies, the standard error for sampling variation is \pm .095 gm. for finger-tip blood specimens.

An analysis of differences between hemoglobin values obtained on the right and left-hand blood samples indicated that values for the right-hand specimens, which were always taken first, were on the average slightly but significantly higher than those for the left hand. This difference could not be explained.

The determinations on finger-tip blood were from serial examinations made at approximately six-month intervals on the same group of women. One technician made the determinations in four periods and a second technician made them in two periods. Comparison of the average hemoglobin levels for the six examination periods indicates a significant variation in the average levels which seemed to be the result of some technical variation. A possible explanation is that one technician had a tendency at times to measure the blood samples slightly generously.

References

1. Snedecor, George W.: STATISTICAL METHODS. Iowa State College Press, Ames, Iowa, 1940, chapter 2, p. 74.

2. Andresen, Marjorie I. and Mugrage, Edward R.: Venous and Peripheral Red Blood Cell Values. American Journal of Clinical Pathology, 1938, 8, pp. 46-51.

3. Price-Jones, C.; Vaughan, Janet M.; and Goddard, Helen: Hematological Standards of Healthy Persons. *Journal of Pathology and Bacteriology*, May 1935, 40, pp. 502-519.

4. Gibbs, E. L.; Lennox, G. W.; Nims, L. F.; and Gibbs, F. A.: Arterial and Cerebral Venous Blood. *Journal of Biological Chemistry*, 1942, 144, pp. 325-332.

5. Drucker, Paul: Investigations on the Normal Values for the Hemoglobin and Cell Volume in the Small Child. *Acta Paediatrica*, 1923, iii. pp. 1-39.

6. Rud, Einer J.: Le Nombre des Globules Range chez les Sujets Normaux et leur Variations dans les Diverses Conditions Physiologique. *Acta Medica Scandinavica*, 1922-1923, 57, p. 168.

7. Bogendorfer and Nonnenbruch: Vergleichende Bestimmung dir Blutkörperschenzahl, des Serumeiweisses und Serumkochsalzes in Venem- und Kapillarblut. *Deutsches Archiv* für klinische Medizin, September 1920, 133, pp. 389-396, Leipzig.

8. Duke, W. W. and Stofer, D. D.: A Comparison of Capillary and Venous Blood in Pernicious Anemia. Archives of Internal Medicine, 1922, 30, p. 94.

9. Bing, H. I.: Antallet af rode Blodlegemer i de forskellige Aldre og under forskellige Forhold. Ugeskrift for Laeger, September 18, 1919, 812, No. 38, pp. 1483-1489.

10. Haden, R. L. and Neff, F. C.: The Volume Index and Color Index of the Red Blood Corpuscles in Newborn Infants. *American Journal of Diseases of Children*, 1924, 28, pp. 458-463.

11. Lucas, W. P.; Dearing, B. F.; Hoobler, H. R.; Cox, Anita; Jones, M. R.; and Smyth, F. S.: Blood Studies in the New-Born. *American Journal of Diseases of Children*, December 1921, 22, pp. 525-559.

12. Evelyn, K. A.: A Stabilized Photoelectric Colorimeter with Light Filters. Journal of Biological Chemistry, August 1936, 115, p. 63.

13. Kruse, H. D.; Palmer, C. E.; Schmidt, W.; and Wiehl, Dorothy G.: Medical Evaluation of Nutritional Status. I. Methods Used in a Survey of High School Students. The Milbank Memorial Fund *Quarterly*, July 1940, xviii, No. 3, pp. 257-298.

14. Berkson, Joseph; Magath, Thomas B.; and Hurn, Margaret: Laboratory Standards in Relation to Chance Fluctuations of the Erythrocyte Count as Estimated with the Hemocytometer. Journal of the American Statistical Association, June 1935, XXX, No. 190, pp. 414-426.

15. Walters, Orville S.: Variation of Erythrocytes, Hemoglobin and Packed Cell Volume in Immediately Consecutive Samples of Venous Blood. *American Journal of Physiology*, 1934-1935, 110, pp. 37-41.

16. McCarthy, Earl F. and Van Slyke, Donald D.: Diurnal Variations of Hemoglobin in the Blood of Normal Men. Journal of Biological Chemistry, May 1939, 128, pp. 567-572.

17. Mole, R. H.: Diurnal and Sampling Variations in the Determination of Haemoglobin. Journal of Physiology, June 1945, 104, No. 1, pp. 1-5, London.

18. Karr, W. G. and Clark, J. H.: Comparison of Various Hemoglobin Methods as Performed in Hospitals and Physicians' Laboratories. *American Journal of Clinical Pathology* (Technical Section), September 1941, 5, pp. 127-147.