REPRINTED FROM:

ACTA PHYSIOLOGICA
ET
PHARMACOLOGICA NEERLANDICA

SUBSIDIZED BY

the
HOLLANDSCHE MAATSCHAPPIJ
DER WETENSCHAPPEN
(Holland Society of Sciences)

Volume III, Nr. 3. 1954
Date of printing October 1954

NORTH-HOLLAND PUBLISHING COMPANY
AMSTERDAM

Acta Physiol. Pharmacol. Neerl. 3; 299—464 (1954)

The Acta Physiologica et Pharmacologica Neerlandica are the continuation of the two Dutch physiological journals Archives néerlandaises de Physiologie de l'Homme et des Animaux and Acta Brevia Neerlandica de Physiologia.
This quarterly contains, besides original papers on physiological, pharmacological and biochemical subjects from Dutch laboratories, the proceedings of the meetings of the Dutch Physiological and Pharmacological Society and of the Dutch Endocrinological Society, as well as the abstracts of communications of interest to physiologists and pharmacologists, made in the meetings of the Medical and Biological Sections of the Association for the Advancement of the Physical and Medical Sciences, the Netherlands Ophthalmological Society, the Netherlands Oto-rhino-laryngological Society, the Netherlands Electro-encephalographic Society and the Amsterdam Neurological Society.
A volume consists of 4 issues of 160 pages each.
Executing Editor: Dr D. K. de Jongh, Nedchem Institute for Pharmaco-Therapeutic Research, P.O.B. 761, Amsterdam, to whom all correspondence and manuscripts should be sent.
Publisher: North-Holland Publishing Company, P. O. Box 103, Amsterdam, to whom applications for subscriptions may be sent.
Subscriptions will also be accepted by any bookseller.
Subscription-price Guilders 35.— ($ 9.25; £ 3.15.—).

Instructions to authors can be obtained through the publishers.

CONTENTS OF VOLUME III NUMBER 3:

(Continued on p. 3 of Cover)

Acta Physiol. Pharmacol. Neerl. 3; 355—377 (1954)

Department of Preventive Medicine of the Royal Netherlands Army Medical Corps,
Department of Pharmacology, University of Groningen,
Statistical Department of the Mathematical Centre, Amsterdam.[1])

ERRORS IN COUNTING THE NUMBER OF EOSINOPHILS IN BLOOD

BY

CHR. L. RÜMKE

Received: July 1, 1954. Accepted for publication July 14, 1954.

INTRODUCTION

In an attempt to study the factors which control the level of circulating eosinophils in recruits during the first three months of training (vide RÜMKE, 1953) a great number of eosinophil counts have been performed. Three technicians (A, B and C) were involved in this experiment. Two of them (B and C) counted the eosinophils in the two counting chambers of FUCHS-ROSENTHAL hemocytometers. The calibration of the leucocyte-pipets has been controlled with the aim to correct deviations of the diluting factors. In order to get an impression of the accuracy of our method of counting control counts have been performed. The statistical analysis of this material as far as it had any bearing on the method of counting will be reported in this paper.

THE COUNTING OF THE EOSINOPHILS

After shaking a pipet in a shaker for 5 minutes the two counting chambers of a hemocytometer were filled. A period of about 15 minutes was allowed for sedimentation. Most of the counts have been performed by technician C; about $^1/_6$ of the counts were performed first by B and thereafter by C, the latter not knowing the results of the former. Both technicians first

[1]) This paper is Report S141 of this department of the Mathematical Centre.

counted the number of eosinophils in the upper counting chambre (P) and thereafter that in the lower one (Q).

In 34 out of 186 cases the difference between B's and C's eosinophil counts in counting chamber P (B_P–C_P, if B_P represents B's eosinophil count in counting chamber P, etc.) appears to be positive. This indicates that the counting level of technician B is lower than that of technician C ($p_t < 10^{-5}$). [1])

As was shown by application of TERPSTRA's test against trend (1952) to the differences B_p–C_p, this difference diminished in the last 10 weeks of the experiment ($p < 0{,}0002$).

It is to be expected that a technician will find about the same number of eosinophils in the two counting chambers of a hemocytometer. The number of eosinophils in counting chamber P will approximately be as often larger as smaller than that in Q.

In order to examine whether this was so we have, for the major part of the counts by C, determined the signs of the differences between the counts in counting chamber P and those in counting chamber Q (C_P–C_Q). If $C_P > C_Q$ the sign is positive; if $C_P < C_Q$ it is negative. As we had given a number to each of the eight hemocytometers in use, the number of positive and negative signs of the above differences could be determined for each hemocytometer separately. As by application of the sign test a significantly larger number of negative signs was found, the number of eosinophils in counting chamber P of every hemocytometer proved to be smaller than that in counting chamber Q.

With regard to technician B, who performed less counts than C, the same effect could be found in 7 out of 8 hemocytometers (table I).

This effect is likely to be caused by either (*a*) the blood-samples, (*b*) the counting chambers, or (*c*) the technicians.

Ad (a). A difference between the number of eosinophils in the first sample and that in the second taken out of the pipet can

---

[1]) p indicates the double tail probability. The index of p indicates the test used in the statistical analysis. r = Rank Correlation Method (Kendall's τ, KENDALL, 1948). s = Test for Symmetry (HEMELRIJK, 1950, R_2, page 953). t = Sign test (DIXON and MOOD, 1946). w = Wilcoxon's Two Sample Test (WILCOXON, 1945; MANN and WHITNEY, 1950; VAN DER VAART, 1950). 0,05 is taken as level of significance.

TABLE I

hemocytometer	C_P–C_Q			B_P–B_Q		
	+	of total	p_t	+	of total	p_t
1	66	185	$< 0{,}01$	5	22	$< 0{,}05$
2	64	188	$< 0{,}01$	5	23	$< 0{,}05$
3	64	201	$< 0{,}01$	8	27	$= 0{,}052$
4	70	201	$< 0{,}01$	6	27	$< 0{,}01$
5	71	196	$< 0{,}01$	4	22	$< 0{,}01$
6	62	193	$< 0{,}01$	6	34	$< 0{,}01$
7	59	172	$< 0{,}01$	7	28	$< 0{,}05$
8	48	167	$< 0{,}01$	4	21	$< 0{,}01$
total . . .	504	1503		45	204	

be excluded, because counting chamber P and counting chamber Q were alternatively filled first.[1])

Ad (b) and (c). Unfortunately the hemocytometers were not at our disposal anymore at the moment we had detected that a comparison of the volumes of the two chambers was desirable. We were able, however, to investigate this in another way. In some cases the method of counting was different from the one described earlier. In half of these cases C counted first the number of eosinophils in counting chamber P, subsequently B counted the number of eosinophils in counting chamber Q; in the other half B counted first the number of eosinophils in counting chamber P and afterwards C counted that in counting chamber Q. The technicians did not know each other's results.

If the technicians find a different number of cells in the two counting chambers, this difference may be caused by deviations in the hemocytometers or by a difference in counting level of the technicians. The effect of the latter difference, however, is neutralized if we compare B_P–C_Q with B_Q–C_P. Applying the method of the 2 × 2 table (FISHER 1950 p. 96) we found in 2 out of the 8 hemocytometers the chances of a positive (or

[1]) We have examined whether a significant difference exists between the number of eosinophils in the counting chambers which were filled first and second. Such difference could not be proved to exist, which indicates that the pipets had been shaken sufficiently.

negative) sign for both differences to be different. Moreover these chances appeared to differ also significantly for all hemocytometers together. This may be an indication for differences between the volumes of the two counting-chambers in some hemocytometers, that is to say that in most cases counting chamber Q is larger (table II).

TABLE II

hemocytometer	B_P–C_Q +	B_P–C_Q –	B_Q–C_P +	B_Q–C_P –	p
1	10	12	11	12	1
2	12	14	9	12	1
3	13	15	20	5	*0,05*
4	10	14	14	11	0,5
5	8	16	18	6	*0,01*
6	12	16	14	10	0,4
7	10	15	15	11	0,3
8	12	11	10	11	1
total	87	113	111	78	*0,004*

This might be an explanation for the fact that the two technicians systematically found less cells in counting chamber P than in counting chamber Q, provided that the ratio of the number of positive and negative differences is in accordance with that found earlier.

We have compared the ratio of the number of positive and negative differences of B_P–C_Q (vide table II) with that of C_P–C_Q (vide table I). In this comparison we used B_P–C_Q and not C_P–B_Q, because by using B_P–C_Q the counting level between the results of the two technicians increases the number of negative differences, which means that only a part of the negative differences may be attributed to the difference between the volumes of the counting chambres. If nevertheless we should find that the number of negative differences would be larger with C_P–C_Q, this would certainly not be the result of the difference in counting level. Actually we found in a 2 × 2 table that the chance of negative differences with C_P–C_Q is significantly larger (p = 0,03). This means that the difference between the numbers of eosinophils, which was found by one technician in the two counting-chambers

cannot be ascribed to a difference between the volumes of the two chambers only. The difference mentioned can also be explained by systematical errors made by the technicians. For example they might have taken into account, wittingly or unwittingly, the number of eosinophils just found in counting chamber P, when counting the number in Q.

From this we may draw the conclusion that the results of the counts performed immediately one after the other in two counting chambers cannot be considered as independent observations. Therefore the counts of counting chamber Q were not considered if the counts in both chambers were performed by the same technician.

Further we investigated whether there existed differences between the volumes of the counting chambers P of our hemocytometers. About 1900 counts belonging to the 8 hemocytometers and all performed by technician C were compared. After testing the normality of the distribution of the counts in each of the 8 hemocytometers by a graphical method an analysis of variance was applied in which the 8 classes represent the 8 hemocytometers. No significant difference proved to exist between these classes ($p = 0{,}8$). If there are differences between the volumes of the P-chambers of the hemocytometers used, such differences need not be taken into account.

THE CALIBRATION OF LEUCOCYTE PIPETS

It is a wellknown fact that the dilution factors of commercial leucocyte pipets frequently deviate from 10. Therefore blood was taken from each person as much as possible with the same pipet. In this way the results of the same testperson became very well comparable. Some weeks after the beginning of the experiment it appeared that the dilution factor of one of the pipets in use was larger than 30. Therefore it was thought desirable to verify the calibration of all pipets used. Usually this kind of pipet is calibrated with the aid of mercury. The pipet is filled with mercury up to mark 1 and mark 11. The two quantities of mercury used are weighed. The dilution factor of the pipet can be found by subtracting 1 from the quotient of the two weights.

The reason of this subtraction is that the solution which is drawn last into the stem of the pipet, does not take part in diluting the blood. As this method was not suitable in the prevailing circumstances, another method was devised, which is as follows:

Some liquid (we used Randolph's solution, RANDOLPH, 1944) is poured upon a plate and from this the pipet is filled up to mark 1 (Fig. 1, A). Next the pipet is removed from the liquid for a moment and an airbubble is drawn up (B). Then liquid is drawn

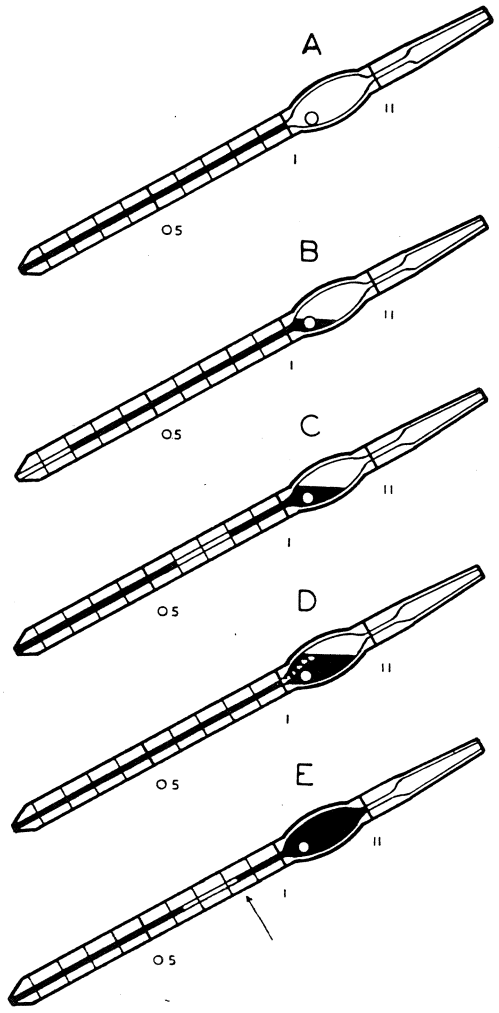

Fig. 1.
Calibration of leucocyte pipets according to the airbubble-method.

up again until the upper level of the new column has reached mark 1. The airbubble escapes in the meantime from the bulb (C and D). Again an airbubble is drawn up. One continues in this way until the bulb of the pipet is filled with liquid to mark 11 (E). When this is the case another airbubble will generally be found somewhere in the stem. The place of its upper boundary can be read from the calibration on the stem. In this way it can be determined with great accuracy how often the stem must be filled to fill the pipet up to mark 11. The dilution factor is equal to this number, diminished with 1.

This "airbubblemethod" can be executed without using much laboratory equipment. It was used for the calibration which took place shortly after the beginning of the investigations. After some time (on 5, 6 and 7 June 1951) 120 pipets were calibrated once more with the same method (calibration I).[1]) In most cases the results obtained were in accordance with the earlier results, some however showing some differences.

For this reason a number of pipets has been calibrated once more on 12 and 13 June 1951 (calibration II). The dilution factors of the pipets regularly used appeared to be generally lower than in the first calibration, whereas the results of the two calibrations of the pipets not used since 7 June differed only slightly. With regard to this fact the question was raised how the use of a pipet could affect the dilution factor.

In the course of the investigation the pipets obtained a blue-violet colour. This was probably caused by a layer of dirt which contained methylene blue. The dilution factor of a pipet may be increased by this layer, for the volume of the stem of a pipet is surrounded by a surface which is relatively larger than the surface surrounding the bulb.

During the first months of the investigations the pipets were cleaned by drawing water through them. Thereafter the pipets were dried with acetone. As the pipets got blue it was evident that the cleaning was not sufficient. Therefore (following a short advice from the "Tijdschrift voor Medische Analysten" *5*, 96,

[1]) In numbering the calibrations the calibration at the beginning of our investigations is not taken into account, because from these results no conclusions have been drawn.

1950) we now and then filled the pipets with a mixture of 96 % alcohol and sodiumnitrite and left them in this state for some time. After this treatment the pipets were clear again. It is possible, however, that only a discoloration and not a cleaning had occurred. Afterwards we cleaned our pipets several times with a mixture of bichromate-sulfuric acid, but it appeared that through this treatment the colouring of the eosinophils often failed, in spite of the fact that the pipets had been thoroughly washed with water.

To investigate whether the dirt in the pipets resulted in an increase of the dilution factors we dropped the treatment with bichromate-sulfuric acid or with alcohol-sodiumnitrite after 11 June 1951. One half of the pipets was used intensively, the other half was not used at all. On 27 and 28 June 1951 we calibrated a number of pipets (calibration III) in order to see if the influence of dirt could already be observed.

At the end of the experiments all pipets were calibrated on 10 and 11 July 1951 (calibration IV). Thereafter they were treated with bichromate-sulfuric acid and calibrated once more on 12 and 13 July 1951 (calibration V).

In general at the end of the investigation the dilution factors of the pipets used, appeared to have increased; this was not the case with the unused pipets. We did not find, however, a difference between the results of the calibrations II and III and between those of the calibrations III and IV. A change of the dilution factors after a fortnight's use only could not be proved to exist. After treatment with bichromate-sulfuric acid the dilution factors of the pipets used appeared to have decreased. The dilution factors of the unused pipets had hardly changed.

From the results of calibration V we have made up a list of correction factors for the results of the countings of the number of eosinophils. To these correction factors another correction has been applied with regard to the dirt by adding 0,01 or 0,02 to each correction factor when the pipets had been used five or ten times respectively after the last cleaning with bichromate-sulfuric acid. For the accurate description of the calculation of these last corrections we may refer to the section on the statistical analysis (p. 368).

To compare the reliability of the airbubble-method with a wellknown and current method 35 pipets were calibrated in the way described earlier with the aid of mercury. Immediately following these calibrations the dilution factor found with the airbubble-method was determined in the same pipets. The variances of the results that are obtained with both methods do not differ very much. The difference in counting level is, however, considerable. Besides there exists no good correlation between the results obtained by both methods in the various pipets (Fig. 2).

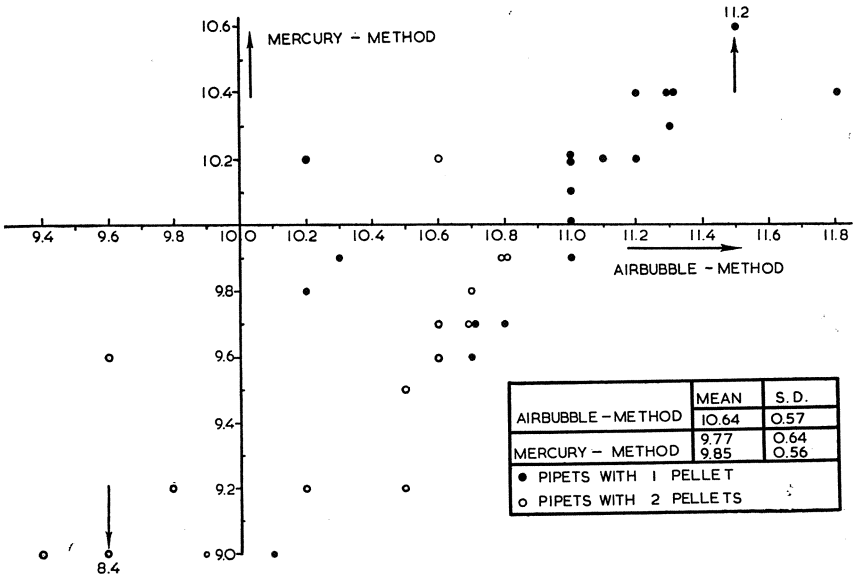

Fig. 2.

The average dilution factors of 35 leucocyte pipets according to the airbubble- and mercury-method.

Calibrating the pipets using the mercury-method it was observed that after drawing the mercury into the pipet the space between the wall of the bulb and the pellet(s) was not filled properly in some cases, the mercury not penetrating into this space on account of its surface tension (Fig. 3). When there were two pellets in the bulb, a still larger space was not filled up. Using the airbubble-method the pellet was entirely surrounded by

liquid. This may explain why lower dilution factors were found with the mercury-method than with the airbubble-method.

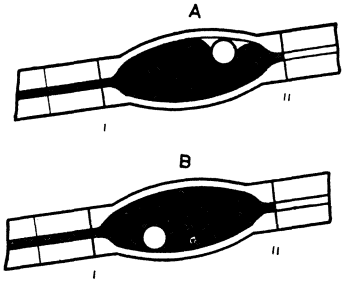

Fig. 3.
The pellet in the bulb. A. If calibrated using the mercury-method. B. If calibrated using the airbubble-method.

Besides it appeared that in the results of one of the two investigators who applied the mercury-method, this difference was more pronounced in pipets with one pellet than in those with two pellets. This explains partly that no good correlation exists between the results of the calibrations by both methods.

Moreover it was observed that for the 35 pipets involved in this investigation, and for both methods used the dilution factors of the pipets with one pellet were significantly larger than those of the pipets with two pellets.

STATISTICAL ANALYSIS OF THE ABOVE OBSERVATIONS

A. Influence of Dirt and Treatment with Bichromate-Sulfuric Acid

The group of sixty pipets which was used in the experiment (group 1) has been compared with a second group of sixty which was not used (group 2). Five series of calibrations have been performed:

I: 5–7 June 1951, both groups, after group 1 had been used for three weeks without treatment with bichromate-sulfuric acid.
II: 12–13 June 1951, both groups, immediately after treatment with bichromate-sulfuric acid.
III: 27–28 June 1951, 26 pipets out of group 1, after a further fortnight's use.
IV: 10–11 July 1951, both groups, after continued use of group 1.
V: 12–13 July 1951, both groups, immediately after treatment with bichromate-sulfuric acid.

In every series the technicians A and C each performed two determinations with each pipet in such a way that a pipet was never calibrated twice on the same day by the same technician.

First of all we compared the accuracy of the two technicians by considering whether the difference between the two determinations of technician A was systematically smaller or larger than the difference between the determinations of technician C. For both technicians the absolute values of the differences between the results of the two determinations performed with each pipet were calculated. The sign test was applied to the signs of the differences between these values (i.e. to $| A_1-A_2 | - | C_1-C_2 |$, if the two results of technician A in a certain pipet were indicated by A_1 and A_2 and those of technician C by C_1 and C_2).

In calibrations I and II no difference in accuracy could be shown ($p_t = 0{,}6$ and $= 0{,}4$ respectively). In III, IV and V it appeared that C had worked more accurately than A (for III $p_s = 0{,}04$; for IV and V $p_t = 0{,}05$ and $0{,}02$).

In the same way we investigated for both technicians whether the accuracy of the calibrations increased during the course of the experiment. The absolute values of the differences between the two results for each pipet, which were obtained by the same technician in two series of calibrations, were determined and the sign test was applied to the signs of the differences between these absolute values (e.g. to $| A_1^{\mathrm{I}}-A_2^{\mathrm{I}} | - | A_1^{\mathrm{IV}}-A_2^{\mathrm{IV}} |$, if the two results of calibration I, obtained by A, are indicated by A_1^{I} and A_2^{I} and of calibration IV by A_1^{IV} and A_2^{IV}). It appeared that calibration IV was performed more accurately than calibration I by A as well as by C ($p_t < 5.10^{-6}$ for both of them). Comparing the accuracy of the calibrations II and V no difference could be shown ($p_t = 0{,}6$ and $0{,}2$ respectively). We also compared the accuracy of the calibrations which were performed, shortly after each other, before and after the cleaning of the pipets. This gives an indication whether the accuracy would decrease, due, for instance, to fatigue, when the calibrations were performed during some days in succession. We found that A had carried out calibration II more accurately than calibration I ($p_t = 0{,}03$); with regard to C no difference could be shown ($p_t = 0{,}5$). Neither of the two technicians showed a difference in accuracy of the calibrations IV and V ($p_t = 0{,}4$ and $0{,}2$ respectively).

Consequently the accuracy of the calibrations increased in the

course of time. It did not diminish when the calibration was performed some days in succession.

In group 2, consisting of unused pipets, higher dilution factors were found in the determinations which were performed later than in the earlier calibrations. We suppose that this is due to (unknown) small changes in the way the method was applied. We shall indicate this phenomenon as shifting. This shifting was found by applicating the sign test to the signs of the difference between the sums of the results, which were obtained for each pipet by one technician in two series of calibrations [so: $(A_1^{\mathrm{I}} + A_2^{\mathrm{I}})-(A_1^{\mathrm{IV}} + A_2^{\mathrm{IV}})$ etc.]. A as well as C found larger dilution factors in calibration IV than in calibration I ($p_t < 3.10^{-4}$ and $< 0{,}02$ respectively). Also in calibration V both found larger dilution factors than in calibration I ($p_t < 0{,}05$ and $< 6.10^{-3}$ respectively). Comparing the dilution factors found in calibration II and IV it appeared also that on the whole an increase of the two results had occurred ($p_t < 4.10^{-3}$ and $< 0{,}03$ respectively).

Considering the results of the calibrations I and IV of the pipets of group 1 it appeared that for both technicians the dilution factors had increased ($p_t < 6.10^{-5}$ and $< 9.10^{-4}$ respectively). The two calibrations were performed three and four weeks after cleaning respectively. There may be two reasons for this increase, the above mentioned shifting and the effect of the difference in usage of the pipets after the last cleaning. We have arranged the frequencies with which differences of a certain magnitude between the results of the calibrations I and IV of technician C occurred in the pipets of the groups 1 and 2 in a 2×3-table. No difference in frequency could be shown between the differences in both groups ($p = 0{,}95$). From this it appears that the difference between the results of calibration I and IV for the pipets of group 1 is not due to a difference between the layers of dirt after a three or four weeks' use. We made use of this result in our subsequent determination of the correction factors. The remaining comparisons appear to confirm the expectation that dirt causes an increase of the dilution factor and that the latter decreases after treatment of the pipets with bichromate-sulfuric acid.

It was expected that the dilution factors of the pipets of group

2 would remain the same in all calibrations before and after cleaning, because this group had hardly been used. We did not expect that they remained equal in the time between the calibrations II and IV because of the shifting found. No difference could be proved to exist between the results of the calibrations I and II (p_t = 0,7 and = 0,3 respectively for the two technicians) and neither between the calibrations IV and V (p_t = 0,7 and = 1 respectively for the two technicians).

With respect to the pipets of group 1, both technicians found a significant decrease between the calibrations I and II (p_t = 0,02 and p_s = 0,03 respectively) and between IV and V ($p_t < 7.10^{-4}$ and $< 5.10^{-6}$ respectively). Between the two cleanings, i.e. between the calibrations II and IV, we found an increase, which may either be the result of dirt or of the shifting already observed in the pipets of group 2 ($p_t < 3.10^{-4}$ and $< 5.10^{-4}$ respectively).

The difference between the results of the calibrations I and V will be influenced by the cleaning just before calibration V, which made us expect a decrease, as well as by the shifting, which would have caused an increase. No difference was observed (p_t = 0,2 and = 0,4 respectively), so the two factors were likely to be balanced.

We have extended the analysis of the effect of dirtiness on the calibration factors by comparing the differences between the results of the calibrations I and II with those between the calibrations IV and V by means of the sign test [applied to $(A^{I}-A^{II})-(A^{IV}-A^{V})$, if A^{I} indicates the average of the values found by technician A of one pipet in calibration I, etc.]. We found p_t = 0,7 and = 1 respectively for the two technicians. From this may be concluded that the effects of 3 and 4 weeks of use without cleaning do not differ significantly.

The result of calibration III has been used to investigate whether the dirt reached a maximum while the pipets were in use. Neither between the results of the calibrations II and III nor between those of the calibrations III and IV, however, a difference could be shown (p_s = 1 in both cases) after applying a correction in view of the shifting described earlier.

Comparing the results of the calibrations I and IV it was

evident that between a layer of dirt after a three weeks' use or one after a four weeks' use no difference could be established. The pipets were used about four times a week. For these reasons we have assumed for the calculation of the correction with regard to the dirtiness, that the latter reached a maximum after the pipets had been used ten times without cleaning them in the meantime. In addition it was assumed that after using the pipets five times half of this dirtiness had occurred. To calculate the maximal effect of the dirt our starting-point has been the difference between the calibrations IV and V, which was found by technician C (who worked most accurately). The difference was 0,17. This same technician found the average difference between the calibrations I and II to be 0,21. These values did not differ significantly. In view of these facts we have added 0,2 to the dilution factor of each pipet used ten times after a cleaning with bichromate-sulfuric acid, and 0,1 when the pipets had been used five times.

B. Comparison of the Results with the Airbubble- and the Mercury-Method

35 pipets have been calibrated twice by two investigators, D and E, using the mercury-method. With the test for symmetry, applied to the differences found for each of the pipets between the absolute values of the differences of the results of the two investigators (i.e. to $| D_1-D_2 | - | E_1-E_2 |$), no difference in accuracy of the two results could be shown ($p_s > 0{,}1$). Between the sums of the two determinations there does exist a significant difference [i.e. between $(D_1 + D_2)-(E_1 + E_2)$] ($p_s = 0{,}01$); E's level proved to be higher.

The average and the standard deviation of the dilution factors of these 35 pipets determined by the mercury-method amounted to 9,77 and 0,64 for D, and to 9,85 and 0,56 for E.

After calibration by the mercury-method the same 35 pipets were calibrated in duplicate by the airbubble-method. This was carried out by technician C. The accuracy of his results was compared with that of the results of D and E by applying the test for symmetry to the differences between the absolute values

of the differences between the results of C and D and of C and E in each of the pipets (i.e. to $| C_1-C_2 | - | D_1-D_2 |$, etc.). No difference could be shown in the accuracy of the results determined by both methods (p_s in both cases $> 0{,}1$). However, a considerable difference in level was found by applying the sign test to the signs of the difference of the sums of the duplicates of C and D, or C and E respectively [i.e. to $(C_1 + C_2)-(D_1 + D_2)$, etc.] ($p_t < 10^{-6}$). The average and the standard deviation of the dilution factors of these 35 pipets, as determined by C using the airbubble-method, amounted to 10,64 and 0,57.

To investigate whether the dilution factors, determined by the mercury-and airbubble-method, are interchangeable if the difference of the levels is eliminated, the 70 results of technician C were divided by the average of these 70 observations, and the results of D and E were divided by their averages. In this way relative values of the observations of the three groups were obtained.

The frequency function of the number of possible inversions between the relative values under the hypothesis that both methods give comparable relative results, is known (Wilcoxon's "U", subsequently defined).

The two relative values of C's observations obtained with the airbubble-method may each be represented by c, and those of D obtained with the mercury method by d. U represents the number of pairs (c, d), for which $c < d$. Ranking the four values obtained for each of the pipets separately by C and D (or by C and E) in order of magnitude, the following rankings are possible, each of them giving a special value to U.

$$\begin{aligned} ccdd;\ U &= 4 \\ cdcd;\ U &= 3 \\ cddc;\ U &= 2 \\ dccd;\ U &= 2 \\ dcdc;\ U &= 1 \\ ddcc;\ U &= 0 \end{aligned}$$

The values $U = 1$ and $U = 3$ belong to those permutations in which the observations of the investigators are separated from each other; $U = 2$ belongs to those permutations, in which the observations of one investigator are separated, and those of the other one are not. With $U = 0$ and $U = 4$ the groups are entirely apart from each other. If every arrangement has the same probability, the probability that $U = 0$, 1, 3 or 4 is for each $\frac{1}{6}$; the probability that $U = 2$: $\frac{2}{6}$.

If for each pipet the values of U with the relative results found by means of the airbubble-method and the mercury-method by D and E respectively are determined, the following distribution was found:

$U =$	0	1	2	3	4
C and D	14	3	7	3	8
C and E	11	4	9	1	10
Expectation . . .	$5\frac{5}{6}$	$5\frac{5}{6}$	$11\frac{2}{3}$	$5\frac{5}{6}$	$5\frac{5}{6}$

Comparing these results with the frequencies expected from the distribution of Wilcoxon's U we found by means of a χ^2-test applied to the groups of values $U = 0$, 1, 3 or 4 and $U = 2$, double tail probabilities 0,002 and 0,01. So we see that the values, in which both pairs of observations are apart from each other, do not appear often enough. From this it may be concluded that even after elimination of the difference in level the two methods of calibration are not equivalent.

To investigate whether this phenomenon is influenced by the presence of one or two pellets in the bulb of the pipet, the frequencies of the various values of U for the pipets with one pellet (group I) and for the pipets with two pellets (group II) were investigated. The result is summarized in the following tables:

C and D $U =$	0	1	2	3	4
Group I	5	1	6	2	6
Group II	9	2	1	1	2

C and E $U =$	0	1	2	3	4
Group I	6	2	4	0	8
Group II	5	2	5	1	2

A difference between the pipets with one and two pellets would show in relatively too low dilution factors of the pipets with two pellets if determined by the mercury-method in comparison with those obtained by the airbubble-method. This would mean that the rankings, which result in $U = 0$ (*ddcc*) or $U = 1$ (*dcdc*) would occur more often in group II than in group I. The method of the 2×2-table was applied to both tables to investigate whether the values $U = 0$ and $U = 1$, taken together, occur more often in group II than in group I. This actually was the case for the first table ($p = 0{,}04$), but not for the second ($p = 0{,}8$).

From this we may conclude that D, applying the mercury-method found relatively lower dilution factors for pipets with

two pellets than C found with the airbubble-method, whereas E did not.

Investigations were made whether the above phenomenon can explain the difference of the level between the results of D and E mentioned earlier. Therefore the analysis of the difference in level (by applying the test for symmetry to the differences, for each pipet, between the sums of the results of the two determinations for both investigators) was performed for groups I and II separately. It appeared that with respect to the pipets of group I no significant difference in level could be found between the results of the two investigators ($p_s > 0{,}1$), but with regard to the pipets of group II a difference did occur ($p_s = 0{,}001$). From this it follows that the difference in level between the results of the two investigators could indeed be explained by the fact that D found relatively lower dilution factors than E in the pipets containing two pellets.

As yet it has been shown only that in the presence of two pellets the relative values of the dilution factors determined by D differed more from those of C than those of E did. To investigate whether the presence of one or two pellets influenced the results of all three investigators, the dilution factors of the pipets of group I and those of group II were compared by means of Wilcoxon's-test The dilution factors of the pipets with one pellet proved to be higher (C: $p_w < 0{,}0001$; D: $p_w < 0{,}0007$; E: $p_w < 0{,}0001$).

Thus for both calibration methods it was found that in these 35 pipets the dilution factors of the pipets with two pellets were smaller than those of the pipets with only one pellet.

THE EFFECT OF THE CORRECTION FOR CALIBRATION

The calibrations have been performed to enable us to correct the results of the counts. To investigate the effect of the correction the following experiment was carried out:

On the 7th of July 1951 blood was taken intravenously from three test persons. Coagulation was prevented by adding sodium citrate to the blood. From each of the three bloodsamples a series of 15 leucocyte pipets was filled in the usual way. We shall indicate these series as series I, II and III. Technician C counted

the number of eosinophils of each pipet. The results of the counts in counting chamber Q were left out of consideration for reasons described earlier.

The ratio of mean and standard deviation was calculated for the eosinophils counted in counting chamber P, for the dilution factors of the pipets used and for the number of eosinophils, which were found by applying the correction factor to the original counts. The correction factor was made on the basis of the results of the calibration on July 10 and 11, 1951 (calibration IV, see page 362), which took place only 4 days after this experiment. The results of this calculation for the three series are recorded in the following table.

	Ratio of mean and standard deviation		
	of the number of eosinophils counted in counting chamber P	of the dilution factors	of the corrected number of eosinophils
Series I	0,09	0,03	0,09
Series II	0,09	0,04	0,09
Series III	0,07	0,04	0,07

As was to be expected the ratio of mean and standard deviation of the number of eosinophils counted in the hemocytometers is larger than that of the dilution factors of the pipets. We had hoped to enlarge the accuracy of the countings by applying a correction for the deviations of the pipets. This proved to be unsuccesful. In view of this the effect of the corrections was analyzed more precisely.

When in a hemocytometer the number of eosinophils from one bloodsample is counted by means of various pipets, one may expect a larger number of eosinophils if the dilution factor is smaller, and vice versa. Therefore we investigated whether a negative correlation existed between the number of eosinophils and the dilution factors of the pipets. No such correlation was found [Series I: $p_r = 0{,}3$ (–) [1]); Series II: $p_r = 0{,}8$ (–); Series III: $p_r = 1$ (–)].

[1]) The sign (+) of (—) behind the double tail probability indicates that the value of the rank correlation coefficient found is positive or negative.

Combining the probabilities of the three series (FISHER, 1950 p. 99) $p_r = 0{,}7$ (–) is found. All correlation coefficients are negative; however, even after combination the correlation is not significant.

An existing negative correlation should disappear if a correction for the deviations of the dilution factors is applied to the results of the countings. No significant correlation was found between the dilution factors and the corrected numbers of eosinophils from these pipets [I: $p_r = 0{,}6$ (+); II: $p_r = 0{,}2$ (+); III: $p_r = 0{,}1$ (+); combined: $p_r = 0{,}08$ (+)]. It is striking, however, that an indication of a positive correlation exists in all the three series, the probability after combination being rather small (0.08). An "overcorrection" may have taken place.

Between the number of eosinophils counted and the numbers that are found after applying the correction a manifest positive correlation appears to exist [I: $p_r = 0{,}0001$ (+); II: $p_r = 0{,}0005$ (+); III: $p_r = 0{,}0004$ (+)]. This may mean that the differences between the number of eosinophils, which occur on account of diluting the blood and filling the hemocytometers, are considerably larger than the differences to which the correction is related.

From the results of this statistical analysis it is evident that the effect of the applied correction is not very satisfactory.

It was also investigated whether a correction, applied on the basis of the results of the calibration after cleaning the pipets with bichromate-sulfuric acid (calibration V, see p. 362), corrected with regard to the dirtiness, had a better result. This proved not to be the case. The results of this statistical analysis are in close agreement with the results described above. No negative correlation was found between the number of eosinophils counted in counting chamber P and the dilution factors of the pipets calculated in this way [I: $p_r = 0{,}6$ (+); II: $p_r = 0{,}6$ (–); III: $p_r = 0{,}1$ (–); combined $p_r = 0{,}4$ (–)].

After application of the correction to the number of eosinophils found, again there appeared to exist a nearly significant positive correlation between these numbers and the dilution factors of the pipets calculated in the way as described above [I: $p_r = 0{,}08$ (+); II: $p_r = 0{,}2$ (+); III: $p_r = 0{,}5$ (+); combined $p_r = 0{,}07$ (+)].

We found a positive correlation between the numbers of eosinophils counted and the numbers that can be calculated on the basis of the dilution factors [I: $p_r = 0{,}0001$ (+); II: $p_r = 0{,}0005$ (+); III: $p_r = 0{,}0023$ (+)].

So a correction with the aid of the dilution factors calculated in this way has no other effect than the one described earlier.

The indications of overcorrection suggest that as regards the dilution the eosinophils in the pipets behaved in another way than the liquid that was used for the calibration. Unfortunately too few pipets used in this experiment had been calibrated by the mercury-method (see p. 359) to allow to investigate whether the effect of correction is more favourable when it is applied on the basis of the dilution factors determined in this way. We have already pointed out, however, that the values found by the mercury-method are probably too low. The comparison of the results of counting the number of eosinophils with various pipets from one blood sample may be considered as a calibration too. Such a calibration should even be preferred because a criterium is chosen that corresponds with the normal use of the pipet. On the basis of the available data it may be considered possible that the results obtained with the mercury-method are better comparable with those according to the "blood-calibration method" than with those obtained with the airbubble method. It would be worth while to investigate this more closely by calibrating a greater number of pipets of which the dilution factors differ considerably by these three methods.

With regard to the experiment which yielded these eosinophil counts (vide RÜMKE, 1953) we may ask ourselves if the validity of our results would diminish if the corrections on the basis of the calibrations are omitted. This is probably not the case. After a correction which is worthy of this name no correlation should exist between the results of the counts and the dilution factors. We have observed, however, that *after* the correction there are *more* indications for such a correlation than before, so that the correction causes rather a more unfavourable situation. The strong positive correlation between the number of eosinophils before and after the correction is a further indication for the fact that the fluctuations, which already existed before the corrections, predominate. So this means that the influence of pipet errors is relatively small in proportion to these fluctuations.

CONCLUSIONS

It is evident that a difference in counting level exists between the results of the technicians B and C. This difference appears to decrease in the course of the experiment, which made it impracticable to apply a correction.

Whether the difference is caused by a change of only one or of both counting levels remains unknown. In future experiments of this kind a control of the course of the counting level seems desirable. This might perhaps be done by taking photographs of the contents of the counting chambers with regular intervals and comparing the numbers counted on these photographs with those recorded by the counters.

A second result of the statistical analysis is the observation that the results of the counts in two counting chambers performed by the same technician are not independent of each other. In other experiments this dependence might be avoided by having the contents of the two counting chambers counted by different technicians. If the same investigator should count both it must be avoided that he knows the result of the first count. This is possible, for instance, if an assistant records the number, or if the investigator himself uses a counting apparatus.

The result of the calibrations confirms the fact already known, that the dilution factors of the leucocyte pipets supplied by the trade rather often deviate from 10. Comparing the results obtained with the newly devised airbubble-method with those of the current mercury-method, it appeared that these did not agree very well. The airbubble-method yielded greater factors, which can be explained by the fact that the mercury does not entirely surround the pellets. However, the standard deviations of the results which are obtained with the methods differed only slightly. The results of the counts obtained with a number of pipets from one blood sample have been corrected with the dilution factors which were determined according to the airbubble-method. With these dilution factors, however, no satisfactory correction-formula could be found.

SUMMARY

In order to increase the accuracy of an investigation on the course of the number of eosinophils in recruits the leucocyte pipets used were calibrated. The calibration was executed on a large number of pipets in three ways, each of them with another result. No satisfactory correction for the deviations of the dilution factors was possible.

The accuracy of the results of the counts was not increased by counting the number of cells in two counting chambers instead of in one, because it could not be excluded that the result of the second count is dependent of that of the first.

There existed a difference in level between the counts of the different investigators. This difference in counting level appeared to decrease in the course of the investigation.

It is discussed in which way the accuracy of an investigation on the course of the number of eosinophils during a certain period could be increased.

RÉSUMÉ

Pour augmenter l'exactitude d'un expériment sur les changements des nombres d'éosinophiles chez des récrues pendant l'entrâinement les facteurs de dilution des pipettes sont vérifiés. La vérification fut exécutée sur un grand nombre de pipettes en trois façons, chaque façon produisant un autre résultat. Une correction satisfaisante n'était pas possible pour les déviations des facteurs de dilution.

L'exactitude des résultats des numérations n'était pas augmentée en comptant le nombre d'éosinophiles dans deux hématimètres au lieu d'une, parce que la dépendance des résultats de la seconde numération du première ne pouvait pas être exclue.

Il existait une différence de niveau entre les résultats des compteurs. Cette différence diminuait pendant l'expériment.

Une méthode pour augmenter l'exactitude d'un experiment sur les changements des nombres d'éosinophiles pendant une certaine période est discutée.

ZUSAMMENFASSUNG

Um die Genauigkeit einer Untersuchung bei Rekruten über dem Verlauf der Eosinophilen während des Trainings zu vergrössern wurden die Leukozytenpipetten geeicht. Die Eichung fand in drei Weisen bei einer grossen Anzahl von Pipetten statt; aber jede Weise zeichte ein anderes Resultat. Eine befriedigende Korrektion der Verdünnungsfaktoren konnte nicht erreicht werden.

Die Genauigkeit der Zählungsresultaten wurde nicht erhöht durch die Ausführung der Zählungen in zwei Zählkammern anstatt in einer, weil

es nicht ausgeschlossen werden konnte, dass das Resultat der zweiten Zählung von dem der Ersten abhängig war.

Es bestand eine Niveaudifferenz zwischen die Ergebnisse verschiedener Zähler. Diese Differenz wurde im Laufe des Experimentes kleiner.

Es wird besprochen in welcher Weise die Genauigkeit einer Untersuchung über dem Verlauf der Eosinophilen während einer bestimmten Zeit erhöht werden kann.

REFERENCES

DIXON, W. J. and A. M. MOOD, J. Am. Stat. Ass. **41**; 556 (1946).
FISHER, R. A., Statistical Methods for Research Workers, London (1950).
HEMELRIJK, J., Proc. Kon. Ned. Ak. v. Wet. **53**; 945 (1950).
KENDALL, M. G., Rank Correlation Methods, London (1948).
MANN, H. B. and D. R. WHITNEY, Ann. Math. Stat. **18**; 50 (1950).
RANDOLPH, TH. G., J. of Allergy **15**; 89 (1944).
RÜMKE, CHR. L., Thesis, Groningen (1953).
TERPSTRA, T. J., Memorandum S73 (M28) of the Statistical Department of the Mathematical Centre, Amsterdam (1952).
VAART, H. R. VAN DER, Report S32 (M4) of the Mathematical Centre, Amsterdam (1950).
WILCOXON, F., Biometrics **1**; 80 (1945).

ACKNOWLEDGEMENTS

Thanks are due to Prof. Dr. D. VAN DANTZIG, director of the Statistical Department of the Mathematical Centre at Amsterdam, and to Prof. Dr. J. HEMELRIJK, Chief of Statistical Consultation in this Department. Through their hospitality and advice the author was enabled to perform the statistical analysis of the data of this experiment as a guest-fellow-worker at their department.

Jkvr. H. D. SANDBERG performed the larger part of the many computations.