crystallized pure from water by maintaining a solution slightly below saturation at 5° C. for two weeks. Crystallization from ether-alcohol mixtures was unsuccessful.

SUMMARY.

The following new quinine salts were prepared with the weak acids added to the weak nitrogen: quinine.HCl.acetate, quinine diacetate, quinine.HCl.propionate, quinine.HCl.valerate and quinine.HCl.lactate. The latter is soluble 115 Gm. per 100 cc. of solution with a p_H of 4.13.

REFERENCES.

(1) Johnson, F. F., refer to Part I.
(2) Guigues, M. P., *J. pharm. chim.* (6), 24, 301 (1906).
(3) Merck and Co., Brit. 241,109; through *C. A.*, 20, 2564 (1926).

THE RANDOM SAMPLING ERROR AS A POSSIBLE ANSWER TO THE APPARENT VARIATIONS IN ANTISEPTIC TEST DATA.[1]

BY ARTHUR R. CADE.[2]

Lack of uniformity in the results of duplicate antiseptic tests, as obtained either by the same worker on different days using supposedly identical cultures and techniques, or by different workers using the same procedures, has brought about recently a somewhat extended discussion as to the cause for these differences in findings. Variation in the day-to-day resistance of the test organism has been suggested as the most probable cause, in a recent series of papers published by members of the AMERICAN PHARMACEUTICAL ASSOCIATION Committee who have investigated antiseptic testing procedures. As a result of this work by Gathercoal and his co-workers (1) there has been established and incorporated into the latest National Formulary a standard of resistance for the test organism *Staphylococcus aureus*, which appears to be slightly inferior to the standard set by the Food and Drug Administration of the United States Department of Agriculture (2). The specifications of the latter state that the test organism must live in 1–80 phenol for 5 minutes, and must be killed by the same concentration in 10 minutes, at 37° C. At the same time, the organism should live in 1–90 phenol solution for 10 minutes at 37° C. The National Formulary standard states that the organism must be killed in the 1–80 phenol solution in 10 minutes, but live in the 1–90 strength for 10 minutes at 37° C. Thus, the difference is that the National Formulary does not require that the organism remain alive in 1–80 phenol for 5 minutes, at 37° C., which requirement the F. D. A. insists upon. The National Formulary specifications have been so drawn up, it is claimed, because experience has taught that it is difficult to get an organism which will retain this resistance with any degree of

---

[1] The experiments reported in this paper form the basis for a thesis presented by Arthur R. Cade in partial fulfilment of the requirements for the degree of Doctor of Philosophy, at the University of Minnesota, December 1933.

[2] Consulting Chemist and Bacteriologist, Minneapolis, Minn.

practical consistency, even though it does show exactly this resistance at intervals. The newly adopted requirement, on the other hand, can be consistently obtained by workers experienced in the field of antiseptic testing. Opposing this viewpoint, Reddish (3), who himself assisted in drawing up the F. D. A. specifications, says: "It is unfortunate that the National Formulary VI standard of resistance should be different from the F. D. A. standard" because "this variation will cause considerable confusion and will even lead to the use of weak cultures by unscrupulous manufacturers of antiseptics. . . . . .The use of such weak cultures should not be allowed in any test." Reddish also states "Meyer and Gathercoal are the only bacteriologists, as far as I know, who have experienced any difficulty in getting the resistance to phenol specified in circular 198, when the F. D. A. strain (No. 209) of *Staphylococcus aureus* is used. . . . .It has been my experience, as well as that of others working in this field, that the F. D. A. strain of *Staphylococcus aureus* (No. 209) very seldom varies from the maximum standard defined in circular 198, and almost never falls below the minimum standard. When the culture does fall below the minimum standard of resistance, it means that the organism has become weakened for some reason."

In an attempt to assist in solving this interesting controversial point, the writer has made a study of the data collected by him over a period of several years of antiseptic testing. The results seem to indicate that the two points of view are not as far apart as they may appear. In fact, the two apparently contradictory sets of data may actually be in rather close agreement. The difficulty is that past workers have failed to interpret their data mathematically, and have ignored what the writer chooses to call the error of random sampling. When this important factor is taken into consideration, it may be seen that it, as well as the day-to-day variation in test organism resistance, plays a most important rôle in explaining why certain control test data do not check, especially with regard to getting an organism to live for 5 minutes in 1–80 phenol at 37° C. in every test. After consideration of this factor, it would seem that to insist that the test organism must always survive 5-minute contact with 1–80 phenol, in the F. D. A. technique, is a bit too rigid for practical purposes, and that failure to meet this requirement does not necessarily mean that the test organism is weaker than one which does meet it. The writer agrees that a culture of lowered resistance should not be used in a standard antiseptic test, but he attempts to show by these data that negative results may sometimes be obtained in the five-minute test with 1–80 phenol even with an organism of standard resistance.

A test procedure is also suggested here which gives data which are quantitatively more accurate than those obtained with the standard procedure of the F. D. A. and which enables the worker, better to evaluate the resistance of his test organism when the 1–80 phenol control test is negative rather than positive. A brief discussion of the random sampling error and its application to these test data follows.

It is obvious that when a person removes a given sample from a suspension which contains relatively few bacteria, the sample removed may or may not contain organisms. The chances of getting a sterile sample depend upon the original population in the suspension; therefore, if the sample removed contains no organ-

isms, it is not safe to conclude that the original suspension was sterile. A somewhat similar situation exists when one considers the likelihood of removing a sample containing a given number of organisms. The possibility of obtaining this given number is naturally dependent upon the original population. When a sample is removed, a number greater or less than that looked for may be obtained. For example, suppose that the suspension contains 25 bacteria per cc., and that 1-cc. samples are removed. Some of these samples will contain 25 bacteria, others less, perhaps only 15, and still others more than 25, perhaps as many as 40. These variations are controlled by the laws of probability, and to calculate the error, it is necessary to resort to the rules of statistics.

When using a technique such as the Rideal-Walker or that of the F. D. A. test, one should remember that the removal of a loopful of solution to broth subculture tubes with no resultant growth in the latter does not necessarily indicate that the seed culture was sterile. The loop may simply have removed a sterile sample from a suspension containing viable organisms. In determining the limiting concentration of a chemical that will kill a broth suspension of organisms, one should bring about a condition wherein almost but not quite complete sterilization is obtained. When working with such borderline concentrations of germicides, it will be found that duplicate samples will not produce identical results. If one obtains no growth in a single test, or even in several tests, it certainly will not be safe to conclude that the original suspension was sterile. Another investigator, using the same materials and the same technique, may get growth in single or in duplicate tests, but this would not necessarily mean poorer germicidal efficiency.

With the plating technique, an occasional plate may show a count considerably above or below the average. This may be due to errors in technique, but it may equally well be due to the error of random sampling. The mathematical principle here involved has been carefully studied and reported in a series of papers by Halvorson and Ziegler (4) who show, for example, that if a 1-cc. sample is removed from a suspension containing 10 organisms per cc., there exists a reasonable chance that the sample removed will be sterile. On the other hand, if a 1-cc. sample is removed from a suspension containing 1 organism in 10 cc., there is still a reasonable chance that a positive sample will be obtained. This error of random sampling is always present, even though all other variables are completely eliminated. In actual practice, of course, other errors such as varying resistance of the organism or variation in the size of loopful, add to the random sampling error and make the range of counts even wider. On the basis of the above, it is obvious that very little confidence can be placed in a single test.

However, one can determine the most probable number of organisms present by applying the formula: $q/n = e^{-ax}$, wherein q is the number of tubes showing no growth, n is the total number of tubes inoculated, and e^{-ax} is the percentage of sterile tubes. Halvorson and Ziegler have calculated for various values of e^{-ax}, or percentage of sterile tubes, the corresponding value of x (where $a = 1$), which is the most probable number of organisms per cc. which could be present and produce the given result. Reversing these calculations, the percentage of times one could expect to get no growth from loopfuls removed from suspensions containing a given number of organisms per cc. would be as follows:

If There Are in the Tube:	Percentage of Times Loopful Would Be Sterile (e^{-ax}).
10 organisms per cc.	90.4 %
25 organisms per cc.	72.8 %
45 organisms per cc.	63.7 %
50 organisms per cc.	60.6 %
69 organisms per cc.	50.1 %
100 organisms per cc.	36.7 %
130 organisms per cc.	27.2 %
150 organisms per cc.	22.3 %
200 organisms per cc.	13.5 %
250 organisms per cc.	8.2 %
300 organisms per cc.	4.9 %
500 organisms per cc.	0.6 %
1000 organisms per cc.	0.005%

Thus, for example, if the seed culture contains 69 bacteria per cc., the subculture tubes would be negative 50.1% of the time. Also, with 25 bacteria per cc. in the seed tube, the subculture tubes would be negative 72.8% of the time. Ten organisms per cc. would show no growth 90.4% of the time. Therefore, for a single test, several subcultures must be made before one can conclude positively that the seed-culture tube was completely sterile.

In Gathercoal's most recent paper (5) he presented tables which showed variations in duplicate test results on control tests with the standard organism in contact with 1–80 and 1–90 phenol solutions at 37° C. for 5 minutes. Sometimes results were negative, sometimes positive. On the basis of these data, the author states that one cannot obtain positive results consistently in this test. He explains the contradictory results as due to a day-to-day variation in resistance of the test organism. Considering his data in the light of the random sampling error we may conclude that such results are to be expected and are not to be considered contradictory. Quantitative data from plate counts, however, rather than merely positive and negative readings, are needed to prove this point conclusively.

For example, in that part of Chart III which gives the results with the F. D. A. standard organism strain No. 209, tested with 1–80 phenol for 5 minutes, it will be noted that there are three negative findings out of 11 tests, or 27.2% negative results. Likewise, observing the other borderline data to which the random sampling error applies, namely, the 1–90 phenol, 15-minute tests, we find the author has obtained 7 negative results out of 11 tests, or 63.7%. Referring to the above table, we find that the most probable population which would give 27.2% negative results is 130 bacteria per cc. For 63.7% the corresponding population is 45. Thus the random sampling error, rather than variation in organism resistance, may be the explanation of why Gathercoal found organisms growing in some subculture tubes and not in others, when in reality probably none of the original seed culture tubes was sterile, but contained approximately the same number of viable organisms.

A more quantitative study of this problem is indicated by the following tables of data taken from the writer's findings. In all this work, the standard F. D. A. procedure was followed, and strain No. 209 was used. Plate counts, however, were taken to determine the numbers of organisms present at the end of the contact

period, instead of the ordinary procedure of transferring loopfuls of broth and recording only absence or presence of growth.

In Table I are presented data taken from control tests made over a period of several years, showing the number of viable organisms per loopful left in the seed tubes after the standard organism had contacted the different phenol concentrations for 5 minutes.

TABLE I.—COLONY COUNTS FOR LOOPFUL OF CULTURE TO AGAR AFTER BEING PLACED IN PHENOL OF DILUTIONS AS NOTED FOR 5 MINUTES AT 37° C.

Test Number.	Plate Counts. 1–80 Phenol.	1–90 Phenol.	1–100 Phenol.	Test Number.	Plate Counts. 1–80 Phenol.	1–90 Phenol.	1–100 Phenol.
1	3	53	850	41	2	125	350
2	0	43	400	42	4	150	400
3	0	9	40	43	0	69	200
4	2	55	125	44	19	57	110
5	0	40	200	45	26	100	200
6	1	30	150	46	3	53	850
7	4	95	350	47	0	43	400
8	3	28	850	48	1	55	125
9	4	200	550	49	0	40	200
10	0	49	440	50	1	30	150
11	29	250	850	51	4	95	350
12	0	53	200	52	3	28	850
13	6	100	225	53	0	49	450
14	1	19	150	54	0	53	200
15	29	150	200	55	1	19	150
16	0	37	150	56	1	37	160
17	1	19	150	57	1	19	170
18	0	24	160	58	0	24	150
19	0	31	225	59	0	31	225
20	1	88	400	60	3	150	375
21	1	250	650	61	2	250	200
22	0	17	150	62	0	300	500
23	14	85	350	63	2	450	600
24	0	112	300	64	1	70	450
25	1	125	500	65	20	40	100
26	5	150	375	66	5	55	550
27	1	15	450	67	1	450	450
28	2	49	100	68	0	25	250
29	3	40	175	69	5	14	400
30	4	85	300	70	3	350	450
31	7	72	150	71	1	150	320
32	5	150	200	72	19	67	300
33	7	18	450	73	0	75	250
34	4	175	500	74	0	90	220
35	1	19	100	75	0	43	400
36	0	10	100	76	1	59	210
37	23	40	175	77	0	40	200
38	0	37	250	78	1	30	150
39	1	25	150	79	4	95	350
40	6	35	800	80	3	28	550

These data show that the average number of viable bacteria remaining after the standard organism has contacted 1–80 phenol for 5 min. at 37° C. is between 1

and 2 per loopful, or 150 ± 50 per cc. Theoretically, therefore, one would expect to get negative results between 13.5% and 36.7% of the time. Actually, these tests show negative results 31% of the time, which is well within the theoretical range. There are eight counts for the 1–80 dilution marked with an asterisk because they are abnormally high. These figures alone would lead one to believe that the organism was of extra high resistance, but when considered in the light of the corresponding data for the 1–90 and 1–100 dilutions, it appears that the organism is about normal, and that the figures marked are merely abnormal variations from the average, as might be expected in biological data. These 8 figures, of course, are not included in calculating the average, as it would be mathematically inaccurate to do so.

Table II gives average data obtained from several hundred other tests. These data are included to show the average counts obtained, from which the writer determined the maximum and minimum counts allowable to decide whether the organism used in each series of germicidal tests was of the standard resistance. These data show that the results obtained for one dilution may not be at all typical, and that it is necessary to have data from at least two, namely, 1–80 and 1–90 dilutions, and preferably from three, including 1–100, before one can be sure that the organism really is of average resistance. Thus, if the 1–80 phenol data are not normal, as may frequently be the case, the 1–90 and 1–100 dilution data will reveal the facts, and thereby prevent the making of false deductions.

The data in Table II were compiled several years ago from hundreds of tests carried out previous to that time. It has been used as a standard for determining whether a given culture to be used in tests carried out by the writer was of standard resistance.

TABLE II.—PLATE COUNTS OF LOOPFULS TRANSFERRED FROM SEED CULTURES OF STRENGTHS AND FOR CONTACT TIMES AS NOTED.

Phenol Dilution.	Period of Contact. 5 Min.	10 Min.	15 Min.
1–80	2 5 8	0 0 0	0 0 0
1–80	0 1 1	0 0 0	0 0 0
1–80	2 0 0	0 0 0	0 0 0
1–80	1 5 0	0 0 0	0 0 0
1–90	250 55	8 0 12	0 0 1
1–90	300 450	0 2 22	0 2 0
1–90	450 25	20 8 50	0 0 0
1–90	70 14	85 2 10	1 0 0
1–100	200 550	85 110	150 5
1–100	500 450	50 400	200 4
1–100	600 250	25 450	3 3
1–100	450 400	50 200	100 15
Averages:			
1–80	2	0	0
1–90	202	18	0.3
1–100	420	170	60
Counts considered as showing test organism to have standard resistance:			
1–80	0– 15	0	0
1–90	50–500	1–100	0– 10
1–100	100–800	50–450	5–200

It is interesting to note that these data also substantiate the findings as in Table I above, and the results of Gathercoal, namely, that contact of the standard test organism with 1–80 phenol for 5 minutes at 37° C. produces 33% negative results. The results in Table I, taken from tests carried out several years later, showed 31% negative, and Gathercoal's data show 27.2%. These percentages check within mathematical error. Likewise, the 1–90 15-minute phenol data in Table II show 75% negative, while Gathercoal found 63.7% again showing agreement, and indicating that the + and − variations in findings are due to the random sampling error and not to a change in the resistance of the test organism.

The writer has made a further study of other variables which affect the results of standard antiseptic tests and the degree to which they must be controlled in order to be within the probable error of the testing technique as a whole. These findings will be reported later in a separate paper. However, there is included here one of these other variables, namely, the effect of variations in the purity and strength of phenol solutions, because it applies more or less directly to the specific problem discussed in this paper. An attempt has been made to find how wide a variation in these phenol factors is permissible before they affect the test results sufficiently to be of greater importance than the random sampling error, which, it is assumed, sets the limit of accuracy for the test itself.

The data as presented in Table III show the variations in findings when using three different grades of pure phenol and weighing these in two different precision manners. The three phenols used included Mallinckrodt's Reagent Grade (A), Merck's Blue Label Grade (B), and a phenol solution standardized by an outside consulting chemical laboratory (C). The first two grades were weighed on an analytical balance accurate to the third decimal point, with a duplicate series weighed on an ordinary laboratory trip balance, accurate only to the first decimal point.

The results show that all samples were reasonably equivalent in their effect upon the test organism, and that the two unstandardized grades were equal in effect to the standard solution. In addition, it is noted that a 1% error in weighing the phenol did not produce any material disagreement in the test results.

TABLE III.—VARIATIONS IN THE STANDARD PHENOL SOLUTION.

Exp. No.	Solution.	Age of 5% Phenol Solution When Dilution Was Made.	Age of Dilution.	Counts after 5 Min. at 37° C. 1–80.	1–90.	1–100.
1	A	1 month	1 day	2	100	...
2	A	1 day	1 day	4	75	...
3	A	1 day	1 day	5	125	...
4	A	1 day	1 week	0	10	60
5	A	2 months	1 day	1	14	28
6	B	1 day	1 day	0	36	...
7	B	1 day	1 day	1	19	100
8	B	1 month	1 day	0	110	300
9	B	1 month	1 day	3	40	175
10	B	1 month	1 week	2	65	240
11	AB	1 month	1 week	0	21	126
12	AB	1 month	2 weeks	5	18	320
13	AB	2 months	1 day	1	26	450
14	AB	2 months	2 weeks	3	150	300
15	AB	2 months	2 weeks	0	29	...

TABLE III.—(*Continued from page 1239.*)

16	C	1 day	1 day	2	100	...
17	C	1 day	1 day	6	29	...
18	C	1 week	1 day	1	50	220
19	C	1 week	1 day	0	17	120
20	C	1 month	1 day	2	61	...
21	C	1 week	1 week	2	49	100
22	C	1 week	1 week	5	54	210
23	C	1 week	2 weeks	1	150	400
24	C	1 week	2 weeks	2	100	350
25	C	1 week	3 weeks	0	150	375
26	C	1 week	3 weeks	2	100	450
27	C	1 month	1 day	5	88	420
28	C	1 month	1 day	0	84	350
29	C	2 months	1 day	2	100	210
30	C	2 months	1 day	4	110	150

A, Mallinckrodt's Reagent Grade Phenol.
B, Merck's Blue Label Grade.
AB, Mixture of equal parts of A and B.
C, Standardized phenol solution.

The data in Table III show that the stock solution of phenol, 5% keeps uniform in strength for several months, if kept away from the action of light in a glass-stoppered bottle, and that the dilutions 1–80, 1–90 and 1–100 keep satisfactorily for a few weeks.

CONCLUSIONS.

Evidence is presented in this paper to show that the random sampling error may account for the apparent discrepancy in results obtained by both Gathercoal and Reddish, as presented in their recent papers, and also that one may expect negative results with 1–80 phenol at 37° C., using the standard F. D. A. test organism, No. 209, up to approximately 35% of the time, even with an organism of standard resistance and an equivalent actual population in each of the seed tubes.

A new procedure is presented which may be used to determine quantitatively whether the test organism is of normal resistance, and whether the test data obtained with it are reliable and comparable to other data.

Experimental evidence presented indicates that a 5% stock solution of phenol will keep several months; and that the dilution used for the control tests will keep several weeks without changing sufficiently to affect materially the test results.

REFERENCES.

(1) Meyer, Esther, and Gathercoal, E. N., "The F. D. A. Test for Antiseptic Value of Liquid Antiseptics," JOUR. A. PH. A., 25, No. 3, 212-215 (March 1936).

(2) U. S. Food and Drug Administration Methods for Testing Antiseptics and Disinfectants, Circular 198 (1931).

(3) Reddish, Geo. F., "Antiseptic Testing," *Soap*, 109 (June 1936).

(4) Halvorson, H. O., and Ziegler, N. R., "Application of Statistics to Problems in Bacteriology," *J. Bact.*, 25, No. 2, 101–121 (February 1933).

(5) Vicher, E. E., Meyer, Esther, and Gathercoal, E. N., "Phenol Resistance of *Staphylococcus aureus*," JOUR. A. PH. A., 25, No. 7, 590–593 (July 1937).