JOURNAL OF CLINICAL MICROBIOLOGY, Apr. 1980, p. 389-393
0095-1137/80/04-0389/05$02.00/0

Vol. 11, No. 4

Bacterial Abundance on Hands and Its Implications for Clinical Trials of Surgical Scrubs

C. THOMAS SPRADLIN

Lilly Research Laboratories, Indianapolis, Indiana 46285

The numbers of bacteria on the hands of 157 subjects volunteering for a clinical trial of a surgical scrub preparation were evaluated statistically. Differences among the volunteers with respect to day-to-day variability in bacterial counts were the most important source of variation in these counts. Generally, more bacteria were found on the left hand than on the right. The experimental plan, proposed by the U.S. Food and Drug Administration, contained criteria for acceptability of subjects which were found to exclude at least as many suitable volunteers as they admitted. The plan was also found to require more testing on more volunteers than was necessary to establish the efficacy of the surgical scrub.

The U.S. Food and Drug Administration has published a tentative final order providing guidelines for experiments designed to establish the efficacy of surgical scrub preparations (3). These guidelines require that volunteers be screened for base-line bacterial count on the hands by means of a glove juice test described below. A volunteer is eligible for entry into the experiment only if there are counts of 1.5×10^6 to 4.0×10^6 on each hand on a single day of screening. The guidelines further require that 30 such subjects participate in the experiment and that the bacterial counts of each be evaluated in triplicate on each of 2 additional base-line days before use of the test product begins.

This paper reports on the screening period of such an experiment, in which 157 volunteers were screened with the intention of finding 36 who could be admitted to the study. The bacterial counts on the hands were evaluated statistically, and the variation in the counts was partitioned into its component parts. These computed quantities were used to assess the probability that a randomly chosen volunteer would be deemed eligible to enter such an experiment, and to estimate the number of subjects and the amount of testing on each necessary to establish efficacy of the surgical scrub.

MATERIALS AND METHODS

A total of 157 adult volunteers of both sexes were recruited for the experiment. The volunteers had occupations which permitted them to come to the testing center on 6 days during a 2-week period, possibly to remain all day on some occasions. They were instructed not to use any products containing antimicrobial agents for 2 weeks before the base-line period, and they were issued rubber gloves to wear upon contact with detergents, acids, bases, or solvents.

On day 1 of the test, the volunteers arrived at the testing center without having cleaned the hands at home. Hands, including two-thirds of each forearm, were rinsed for 30 s with water at 35 to 40°C, and the nails were cleaned with a nail cleaner. The hands were then washed for 30 s with Camay soap and rinsed under running tap water. Excess water was shaken from the hands, and surgeon gloves previously washed with sterile distilled water were donned. A 50-ml amount of a solution of Triton X-100, 0.1% in 0.075 M phosphate buffer, was added to each glove, and the glove was massaged for 1 min. A 10-ml sample was then removed from each glove for microbial determinations.

Glove samples were diluted with 10 ml of sterile saline containing 2×10^{-3} M cysteine. Two 10-fold dilutions were made in sterile saline containing 2×10^{-3} M cysteine, and 1-ml samples of each dilution (1:2, 1:20, 1:200) were plated in triplicate on Trypticase soy agar containing 2×10^{-3} M cysteine. After 48 or 72 h of incubation at 37°C, the plates were evaluated for colony count.

For the volunteers who returned to continue the experiment, the same procedure was carried out on days 5 and 7.

RESULTS

The 157 volunteers who were screened can be categorized as follows. Group 1: 45 had each hand counted by means of triplicate plates on each of 3 days. Of these, 23 (group 1A) had bacterial counts in the specified range and were therefore eligible for entry into the experiment, and 22 (group 1B) did not. Group 2: 82 had each hand counted on triplicate plates on only 1 day. Group 3: 30 reported for a single day of screening but had counts that could not be used. Most of these (28 volunteers) had microorganisms too numerous to count on at least one hand. This group was not included in the detailed statistical analysis.

The common logarithms of the bacterial

counts were submitted to analysis of variance under the following model. $y_{ijkm} = \mu + S_i + D_j + H_k + SD_{ij} + SH_{ik} + DH_{jk} + SDH_{ijk} + \epsilon_{m(ijk)}$, where y_{ijkm} is the estimate from plate $m(m = 1, 2, 3)$ of the log number of bacteria on hand $k(k$ = right, left) of subject $i(i = 1, \ldots, 45$ for group 1, $i = 1, \ldots, 82$ for group 2) on day $j(j = 1, 5, 7$ for group 1, $j = 1$ for group 2). S_i is the random effect of subject i, D_j is the random effect of day j, H_k is the fixed effect of hand k, SD_{ij} is the random interaction of subject i with day j, SH_{ik} is the random interaction of subject i with hand k, DH_{jk} is the random interaction of day j with hand k, SDH_{ijk} is the random three-way interaction, and $\epsilon_{m(ijk)}$ is the random effect associated with the triplicate plate counts.

The random terms in the models are assumed independently distributed, with means 0 and with variances as follows. Variance $(S_i) = \sigma_S^2$; variance $(D_j) = \sigma_D^2$; variance $(SD_{ij}) = \sigma_{SD}^2$; variance $(SH_{ik}) = \sigma_{SH}^2$; variance $(DH_{jk}) = \sigma_{DH}^2$; variance $(SDH_{ijk}) = \sigma_{SDH}^2$; and variance $[\epsilon_{m(ijk)}] = \sigma_\epsilon^2$. The analog of the variance for the fixed hand effect is denoted by ϕ_H.

Table 1 provides means for the bacterial counts for the various groups and days of the experiment, and the analyses of variance, one each for groups 1 and 2, are presented in Table 2. The individual components of variance can be estimated by solving the equations resulting from equating the mean squares to their expected values, which were derived as previously described (4). However, Table 2 provides nine such equations for the estimation of eight variance components, and therefore unique estimates cannot be obtained in that manner. Instead, the equations were solved by the method of least squares (6).

The estimates of each of the variance components are listed in Table 3. The three negative variance components, those for days, subject-by-hand interaction, and day-by-hand interaction, indicate that those sources do not contribute significantly to the total variability. These findings can be interpreted as meaning that mean log bacterial counts, averaged over subjects, hands, and triplicate plates, remain constant from day to day. Also, the number of bacteria on the left hand, relative to the number on the right, is about the same for all subjects. Finally, when counts are averaged over subjects and plates, the difference between the hands remains essentially constant from day to day.

A total of 94% of total variability is accounted for by differences among subjects and the subject interactions. Variability among triplicate plates is small, as is the difference between the hands. The consistency of the difference between the hands is illustrated by the fact that 61.4% of the volunteers had more bacteria on the left hand than on the right on day 1 of screening.

Assuming that the log bacterial counts obey a normal probability law, it was possible to use the estimated quantities to assess the probability that a randomly chosen volunteer would have bacterial counts between 1.5×10^6 and 4.0×10^6 on each hand on any given day. Since both hands must simultaneously satisfy this condition, it was necessary to evaluate the probability under the bivariate normal surface by a method provided by Abramowitz and Stegun (1). Assuming that a volunteer is a "perfect" subject, i.e., has a true mean bacterial count of 2.45×10^6 on each hand, midway between 1.5×10^6 and 4.0×10^6 on the log scale, the probability that his mean (of triplicates) log bacterial count for both hands on any given day will fall within the specified range is estimated to be 0.30.

Table 4 uses the variance component estimates of Table 3 to compute the standard deviation of the mean log bacterial count of a single subject based upon counting single, duplicate, or triplicate plates on each of 1, 2, or 3 base-line days. It is clear that the counting of more than one plate decreases the standard deviation of the mean base-line count of a subject only negligibly, regardless of the number of days that subject is tested. The decrease in variability resulting from testing on multiple days is more

TABLE 1. *Geometric mean bacterial counts on the hands of normal volunteers*

Group	Day of count	Mean bacterial counts ($\times 10^6$)	
		Right hand	Left hand
1[a]	1	2.52	2.95
	5	2.46	2.98
	7	2.45	4.11
1A[b]	1	2.32	2.43
	5	2.99	2.90
	7	2.71	4.11
1B[c]	1	2.75	3.63
	5	2.00	3.08
	7	2.21	4.10
2[d]	1	1.77	2.34

[a] 45 subjects counted on days 1, 5, and 7.

[b] 23 subjects with counts between 1.5×10^6 and 4.0×10^6 on both hands on day 1.

[c] 22 subjects with counts outside the range of 1.5×10^6 to 4.0×10^6 on at least one hand on day 1.

[d] 82 subjects whose hands were sampled on only 1 day and whose counts on at least one hand were outside the range of 1.5×10^6 to 4.0×10^6.

TABLE 2. *Analyses of variance of the log bacterial counts on the hands of normal volunteers*

Group	Source of variation	Degrees of freedom	Mean square ($\times 10^{-3}$)	Expected mean square
1[a]	Subjects	44	749.11	$18\sigma^2_S + 6\sigma^2_{SD} + \sigma^2_\epsilon$
	Days	2	411.65	$270\sigma^2_D + 6\sigma^2_{SD} + \sigma^2_\epsilon$
	Hands	1	3,172.32	$405\phi_H + 9\sigma^2_{SH} + 135\sigma^2_{DH} + 3\sigma^2_{SDH} + \sigma^2_\epsilon$
	Subjects by days	88	331.92	$6\sigma^2_{SD} + \sigma^2_\epsilon$
	Subjects by hands	44	246.96	$9\sigma^2_{SH} + 3\sigma^2_{SDH} + \sigma^2_\epsilon$
	Days by hands	2	493.41	$135\sigma^2_{DH} + 3\sigma^2_{SDH} + \sigma^2_\epsilon$
	Subjects by days by hands	88	104.58	$3\sigma^2_{SDH} + \sigma^2_\epsilon$
	Triplicate plates	540	6.16	σ^2_ϵ
2[b]	Subjects	81	1,055.33	$6\sigma^2_S + 6\sigma^2_{SD} + \sigma^2_\epsilon$
	Hands	1	389.77	$246\phi_H + 3\sigma^2_{SH} + 246\sigma^2_{DH} + 3\sigma^2_{SDH} + \sigma^2_\epsilon$
	Subjects by hands	81	294.89	$3\sigma^2_{SH} + 3\sigma^2_{SDH} + \sigma^2_\epsilon$
	Triplicate plates	328	9.32	σ^2_ϵ

[a] 45 subjects were counted on each of 3 days.

[b] 82 subjects whose hands were sampled on only 1 day and whose counts on at least 1 hand were outside the range of 1.5×10^6 to 4.0×10^6.

TABLE 3. *Components of variance of log bacterial counts on the hands of normal volunteers*

Variance component	Estimate ($\times 10^{-3}$)	% of total
σ^2_S	16.22	6.4
σ^2_D	−0.63	0.0
ϕ_H[a]	7.57	3.0
σ^2_{SD}	95.77	37.6
σ^2_{SH}	−0.07	0.0
σ^2_{DH}	−4.98	0.0
σ^2_{SDH}	127.65	50.0
σ^2_ϵ	7.74	3.0
$(\sigma^2_S + \sigma^2_D + \sigma^2_{SD} + \sigma^2_{SH} + \sigma^2_{DH} + \sigma^2_{SDH} + {}^1/_3\sigma^2_\epsilon)$[b]	242.22	
$[\sigma^2_S + \sigma^2_{SH} + {}^1/_3(\sigma^2_D + \sigma^2_{SD} + \sigma^2_{DH} + \sigma^2_{SDH}) + {}^1/_9(\sigma^2_\epsilon)]$[c]	91.55	

[a] Although this is technically a fixed effect, it was treated as random for inclusion in this table.

[b] This is the variance of the mean of triplicate plates for one subject, one hand, 1 day.

[c] This is the variance of mean log bacterial count for one subject, three plates on each of 3 days.

TABLE 5. *Standard deviations of the difference between mean log base-line bacterial count and mean log bacterial count on a test day for a single subject for various numbers of days of base-line testing and various numbers of plates per day*

Plates/day	Standard deviation for no. of days of testing:		
	1	2	3
1	0.680	0.589	0.555
2	0.674	0.584	0.551
3	0.672	0.582	0.549

TABLE 4. *Standard deviations of mean log bacterial count for a single subject for various numbers of days of base-line testing and various numbers of plates per day*

Plates/day	Standard deviation for no. of days of testing:		
	1	2	3
1	0.497	0.363	0.305
2	0.493	0.360	0.303
3	0.492	0.359	0.303

substantial; counting bacteria on 3 days, rather than on 1 day, reduces the standard deviation of the mean base-line count by some 38%.

Since the reason for collecting base-line measurements is to use them as a yardstick for evaluating the efficacy of the surgical scrub preparation to be used subsequently, it is also of interest to estimate the variability of the difference between a mean base-line bacterial count and the mean count of replicate plates for the same subject on a subsequent day after use of the scrub. Table 5 gives estimates of the standard deviations of this difference for several values of days of base-line counting and numbers of plates used. It has been assumed that the same number of plates is used on the test day as on each day of the base line. Again, use of more than one plate per day is seen to be of little value in reducing variability.

Table 5 is of use in reaching a decision as to the precision with which the mean log bacterial counts must be estimated. Knowledge of the standard deviation of the difference between mean base-line count and the count on a test day for a single volunteer, coupled with knowledge of the degree of efficacy of the scrub which should be detected in the experiment, make it possible, using a standard text such as that of Cohen (2), to compute the number of volunteers to be enrolled. For example, if base-line meas-

urements are taken on triplicate plates on each of 3 days, and if it is desired to be able to detect with 80% probability a 1-log decrease in bacterial count after treatment in a one-tailed test in which $P < 0.05$ is to be accepted as statistically significant, only three subjects need be enrolled.

DISCUSSION

The principal conclusion of this analysis is that the experimentation required by U.S. Food and Drug Administration guidelines (3) is considerably more extensive than necessary to achieve the objectives of the study. First, the criterion that volunteers must have 1.5×10^6 to 4.0×10^6 bacteria on each hand on a single day of screening is overly restrictive, in that this criterion excludes at least as many suitable subjects as it admits to the study. Secondly, the guidelines require the use of more subjects and more testing per subject than necessary to establish efficacy of the surgical scrub preparation. The following discussion substantiates these conclusions in detail.

It was estimated above that a "perfect" subject had only a probability of 0.30 of having bacterial counts on the hands within the specified range on a single day of testing. It is clear from purely theoretical grounds that 0.30 must be an overestimate of this probability of the acceptability of the volunteer, as is demonstrated by the following arguments.

First, the exclusion of group 3 from the statistical treatment of the data necessarily resulted in an underestimate of the variability, since almost all subjects in that group had exceedingly high, but indeterminate, bacterial counts. Such an underestimation of the variation in the test system results in an overestimation of the probability of eligibility.

Also, 0.30 was computed under the assumption that the volunteer being considered was a hypothetical perfect subject. Any other volunteer, whose true mean log bacterial count may be well within the prescribed limits yet not exactly at the midpoint, would have an even smaller chance of being deemed eligible on the basis of a screen performed on a single day. Indeed, for a subject whose true means were the same as the averages observed on day 1 of this experiment, namely, 2.01×10^6 on the right hand and 2.54×10^6 on the left, this probability is estimated to be only 0.26.

Finally, the estimate of the probability of acceptability was computed under the assumption that the perfect subject had equal numbers of bacteria on both hands, whereas this experiment provides evidence that it is likely that more will be found on the left than the right. This finding decreases even further the probability that both hands will simultaneously satisfy the eligibility criterion.

It appears clear that 0.30 can be regarded as an upper limit for the probability that even a perfect subject would be deemed eligible for an experiment of this type. In this investigation, only 23 of 157 volunteers screened (14.6%) met the eligibility criterion. This 0.146 serves as another estimate that a randomly chosen volunteer would be deemed eligible, and might be more nearly accurate than 0.30. Even using the larger figure, it can be estimated that at least 100 volunteers would have to be screened, on the average, to include the prescribed 30 in the experiment. One test of a surgical scrub preparation conducted by these guidelines has been reported (5). Thirty volunteers were found with bacterial counts on the hands within the specified range, but no mention was made of the total number of subjects screened.

It is also of interest to compare those volunteers who were deemed eligible (group 1A) with those who were not (group 1B). No such statistical comparison is proper for day 1 of the base-line testing, since the groupings were made on the basis of the counts themselves. On days 5 and 7 of the base-line period, however, analysis of variance showed no significant differences ($P > 0.05$) between groups 1A and 1B with respect to either the mean or the variance of the number of bacteria on the hands. Furthermore, only 7 of the 23 volunteers in group 1A would have qualified for entry into the experiment on either of the other two days of the base-line period, whereas nine of the 22 in group 1B would have qualified had the decision been made on one of the other days. It appears that there is little, if any, difference between groups 1A and 1B with respect to the numbers of bacteria on their hands as sampled from day to day.

This study shows that the proposed range for acceptability of volunteers of 1.5×10^6 to 4.0×10^6 bacteria per hand on a single day is too narrow, in that it excludes at least as many suitable volunteers as it includes. The majority of the subjects excluded from this study had bacterial counts above that range, and yet their day-to-day counts did not differ significantly from those deemed acceptable. This finding implies that the range should be expanded to include numbers larger than 4.0×10^6. Since the surgical scrub was only used on those subjects with base-line counts within the prescribed range, there are no data to show that volunteers with large numbers of bacteria respond to the scrub preparation in the same way as those with smaller numbers. Nevertheless, it appears unlikely that those with base-line counts of 10.0×10^6 would respond demonstrably differently

from those with counts of 4.0×10^6. Consequently, the range for acceptability could be expanded, perhaps to between 1.0×10^6 and 10.0×10^6, without loss of precision.

The statistics given in Tables 4 and 5 indicate also that counting triplicate plates on each of 3 days for 30 subjects is in excess of the experimentation needed to detect a 1-log decrease in bacterial count. Conducting base-line counts on a single plate on a single day would reduce the cost of the experiment substantially, yet the increase in variability would be inconsequential. A 1-log reduction in this case is equivalent to a reduction of approximately 1.5 standard deviation units, and only about seven subjects in 100 would show such a reduction erroneously, i.e., if in fact the surgical scrub preparation is ineffective. Base-line counts on a single plate on a single day for four subjects would therefore be sufficient to achieve 80% probability of detecting a 1-log reduction in bacterial count in a one-tailed test, using $P < 0.05$ as a criterion of significance.

It is important to be able to quantify with precision the effectiveness of a surgical scrub preparation. Nevertheless, this study indicates that the efficacy of such a preparation could be demonstrated conclusively with only a fraction of the resources, both laboratory and human, which have been expended in the past.

ACKNOWLEDGMENTS

I acknowledge the able secretarial assistance of Marla A. Gutowski, who persevered through several revisions of this paper.

LITERATURE CITED

1. **Abramowitz, M., and I. A. Stegun.** 1964. Handbook of mathematical functions, Applied mathematics series 55, p. 936–939. National Bureau of Standards, Washington, D.C.
2. **Cohen, J.** 1977. Statistical power analysis for the behavioral sciences, p. 19–74. Academic Press Inc., New York.
3. **Food and Drug Administration.** 1978. Fed. Regist. **43:** 1210–1249.
4. **Hicks, C. R.** 1973. Fundamental concepts in the design of experiments, p. 177–179. Holt, Rinehart, and Winston, New York.
5. **Rosenberg, A., S. D. Alatory, and A. F. Peterson.** 1976. Safety and efficacy of the antiseptic chlorhexidine gluconate. Surg. Gynecol. Obstet. **143:**789–792.
6. **Searle, S. R.** 1971. Linear models, p. 448–451. John Wiley & Sons, Inc.