J. Soc. Cosmetic Chemists, **17**, 343–353(1966)

Use of Human Subjects for Product Evaluation: An Evaluation of Antibacterial Soap Bars

JOHN A. KOOISTRA, Ph.D., ELMER A. BANNAN, M.S., and R. OWEN CARTER, Ph.D.*

Presented September 21, 1965, Midwest Chapter

Synopsis—It is shown that *in vitro* microbiological testing of antibacterial-containing soaps is frequently unreliable and is influenced by the strain of the organisms used. On the other hand, *in vivo* testing under laboratory controlled or consumer testing conditions can be used to establish the degerming efficacy of sanitizing soap bars. Several techniques for conducting such tests are described.

For many years efforts have been made to improve the skin degerming potential of skin cleaning agents by enhancing the inherent antibacterial power of soap through the incorporation of specific antibacterial ingredients. Many antibacterial chemicals were tried. It was not, however, until the introduction of hexachlorophene in the mid-forties that a satisfactory agent was found to achieve this purpose. With the demonstration that skin degerming through the use of specific antibacterials incorporated into soap was feasible, the search for new and better antibacterial agents was intensified.

The apparent simplicity of the degerming process is likely to be misleading. Actually, the suppression of the cutaneous microflora through washing of the skin with surface active agents containing bacteriostatic compounds is a complex process. It involves a number of interactions,

* The Procter & Gamble Company, Ivorydale Technical Center, Cincinnati, Ohio.

most of which are understood incompletely at best, among the cutaneous microorganisms, the agent, and the environment. Individual microorganisms of the skin differ greatly among themselves in regard to sensitivity to a given antibacterial agent. Agents differ in their ability to deposit from the washing solution on the stratum corneum and to penetrate into the appendages of the skin. The effect of the matrix and the solubility of the antibacterial agent in the liquid phases of the skin surface influence the degree to which the bacterial population is affected. These factors play a role in determining the extent to which a practical objective is achieved; that is, to decrease the number of skin bacteria that can lead to infection and to axillary and other so-called body odors.

In view of such a multifaceted situation, methodology is obviously critical in evaluating new antibacterial agents and in comparing one material or product with another in regard to degerming efficacy. The purpose of this paper is to discuss some approaches currently used for evaluation of skin degerming efficacy and to highlight their advantages and limitations.

Methodology and Discussion

In Vitro Testing

In vitro procedures, because of their relative simplicity, are widely utilized. In order for an agent to warrant consideration at all as a candidate antibacterial ingredient, obviously it must first have antibacterial activity in the matrix in which it is to be used. *In vitro* procedures can be used for screening purposes to supply such information and thus exclude from further consideration compounds with inadequate activity. By their nature, however, these techniques rely on arbitrary selections of test organisms and exposure conditions. They can give reproducible results; however, extrapolation from *in vitro* data to more complex or practical situations has been found by experience to be unreliable.

Many *in vitro* procedures have been advocated, but the type that is most widely used at the present time involves determination of the minimum inhibitory concentration (MIC).

Part of the problem inherent in basing over-all conclusions regarding degerming efficacy on MIC data is illustrated by the information of Table I. This table presents MIC data comparing two antibacterial soaps against a group of 20 different *Staphylococcus aureus* strains isolated from pyogenic infections. Antibacterial Soap A was a milled bar

containing 2% of a mixture of the antibacterials: 3,5-di- and 3,4′,5-tribromosalicylanilides (BSA), 4,4′-dichloro-3-(trifluoromethyl) carbanilide (TFC), and 3,4,4′-trichlorocarbanilide (TCC). Antibacterial Soap B was a milled bar containing 0.75% 3,4,4′-trichlorocarbanilide (TCC) plus 0.75% hexachlorophene (G-11).

Considerable variation among the organisms in sensitivity to both soap bars was evident. Strain 22 required over four times as much of either soap for complete inhibition of growth as did Strain 32. Overall, the data show that the soaps were similar in regard to bacteriostatic

TABLE I

Minimum Inhibitory Concentration Against *Staphylococcus aureus* Strains

	mg./100 g. of product				
	Antibacterial Soap			Antibacterial Soap	
Strain	A[a]	B[b]	Strain	A	B
26	6	8	32	4	4
20	6	8	34	8	8
19	6	6	35	8	8
21	6	6	36	6	6
22	16	16	37	6	6
24	6	6	38	6	6
27	6	6	39	6	6
29	6	6	23	8	6
30	8	8	25	12	8
31	6	6	28	8	6

[a] Antibacterial additive: 2% mixture of BSA, TFC, and TCC.
[b] Antibacterial additive: 0.75% TCC + 0.75% G-11.

activity. If, however, either Strains 20 or 26 had been chosen as the only test organisms, the conclusion would have been that Soap A was superior to B. Had either Strains 25 or 28 been the only test organisms, Soap B would have been judged superior. With each of these strains, the recorded differences between the soap bars were reproducible, but, if the experiment were to be repeated with another group of strains of this same bacterial species, the results might very well be different.

The problem of interpretation of MIC data in terms of the relative skin degerming effectiveness of two products is further complicated when other species of bacteria are used as test organisms. Data in Table II illustrate this complication. MIC values were obtained for Soaps A and B with a group of 20 strains of *Corynebacterium minutissimum* isolated from toe webs of individuals with erythrasma. Overall, the data show Bar A to be superior to Bar B in control of these strains.

TABLE II

Minimum Inhibitory Concentration Against *Corynebacterium minutissimum* Strains

mg./100 g. of product

Strain	Antibacterial Soap A[a]	Antibacterial Soap B[b]	Strain	Antibacterial Soap A	Antibacterial Soap B
P17-1	8	14	P1-1	11	16
P16-1	8	14	P3-3	11	14
P6-1	8	14	P8-2	16	16
1-2	8	16	2A	14	16
10289	8	11	7A	12	16
3A	8	16	P22-1	16	16
12A	6	8	P7-3	16	16
4R	8	16	P1-4	16	16
8L	8	16	P20-1	16	16
P13-1	11	14	P26-3	14	12

[a] Antibacterial additive: 2% mixture of BSA, TFC, and TCC.
[b] Antibacterial additive: 0.75% TCC + 0.75% G-11.

TABLE III

Skin Degerming Effectiveness of an Antibacterial Bar Soap: Constancy of Results

Handwashing—Fifth Basin—Four Day Regimented Usage

Test Product	Test Date	Mean Bacterial Count Per Liter
Antibacterial soap A[a]	1-9-61	9,000
	4-17-61	17,800
	3-11-63	11,300
	1-27-64	13,100
	3-9-64	13,800
	10-23-64	8,900
	10-26-64	1,900
	11-9-64	11,000
	11-9-64	5,700
Control soap[b]		1,300,000

[a] Antibacterial additive: 2% mixture of BSA, TFC and TCC.
[b] No antibacterial additive.

However, if Strain P26-3 had been selected as the only test organism, Soap B would have been judged superior. Had any of the four strains just above it in the table been used, the bars would have been judged equally effective.

Obviously, decisions regarding relative efficacy of antibacterial products based on MIC data are capricious. Results are too dependent upon the bacterial strains that happen to be chosen for the test. For

this reason alone, MIC data do not offer a means of judiciously selecting any one antibacterial bar soap as being superior. Furthermore, the MIC test completely neglects to take into account the effect of the environment in which the product will be expected to perform its function.

Pillsbury (1), one of the pioneers in the field of skin degerming, stated that he was unable to demonstrate that *in vitro* studies were of value in predicting the antibacterial action of an agent on the skin. The way to determine the degerming effectiveness of a skin degerming product is to measure the reduction in the number of bacteria on the skin of panels of people using the product. Such *in vivo* methods are feasible and, if properly designed and conducted, can give reproducible and precise results.

In Vivo Testing

A reliable method for determining the size of the bacterial flora of the skin following various types of surgical scrub regimens was developed by Price (2). The procedure involved a sequence of standardized handwashings (with brush) and the estimation of the number of bacteria removed after each washing. Variations of the basic Price method have been reported (3–5). For the laboratory evaluation of the skin degerming effectiveness of antibacterial soap bars, a further modification of the basic handwashing method has been developed in this laboratory, which over a period of five years has proved to be a precise and practical investigative tool.

A pool of subjects is maintained by supplying individuals with soap bars containing no added antibacterial agents for their hygiene when they are not participating as members of a handwashing panel. Sufficient time is allowed between participation in different handwashing tests to permit the bacterial flora of the hands of each individual to return to its normal level.

From this pool, test panels of 10 individuals (5 males, 5 females) are drawn as required. During a test, each subject is given a bar of test soap to use at home for all handwashing and bathing and a second bar to keep with him for use during the day. In addition, the panelists wash their hands with the test soap under supervision in the laboratory three times daily according to a prescribed routine. In order to minimize possible extraneous factors, panelists are requested not to expose their hands during the test period to other soap solutions (which might remove the antibacterial compounds from the skin) such as in dishwashing, car

washing, shampooing, etc. Rubber gloves are provided for use by those participating in such activities.

At the beginning of the test on Monday morning, before using the test product, and on the following Friday morning, after exclusive use of the test product for four days, an estimate of the number of bacteria on the skin is made. Four successive standardized handwashings are performed in a stream of warm tap water, using a bar soap containing no antibacterial additives. The fifth handwashing (fifth basin) is performed using the same procedure, except that it is done in a basin containing one liter of sterile distilled water rather than in running tap water. The hands are carefully washed and rinsed into the water in this basin. Aliquot samples of the wash water are immediately taken for bacterial enumeration. Pour plates or filter membrane procedures can be used for enumeration. Raw count data are transformed to logarithms for statistical analysis.

Data presented in Table III show the constancy of counts determined over a three-year period when using this handwashing procedure as the analytical tool. The soap used by the panelists was the bar previously described as antibacterial Soap A containing the 2% mixture of BSA, TFC, and TCC. Nine different panels were subjected to the same four-day testing procedure. The number, 1.3×10^6, was included in the table as a reference figure. This bacterial count was obtained from nonantibacterial soap bar users and is an average based on more than 500 individual fifth basin handwashings.

Data showing the skin degerming effectiveness of various concentrations of the antibacterial ingredient, hexachlorophene, incorporated into a milled soap matrix are presented in Fig. 1. With an increase in concentration of the antibacterial ingredients, a progressive decrease in the mean number of bacteria recovered in the fifth basin was evident.

Figure 2 shows the mean number of bacteria recovered in the fifth basin after use of bar soaps containing as antibacterials the following agents:

2.0% Mercuric iodide
1.0% Hexachlorophene
0.5% 3,4,4′-trichlorocarbanilide plus 0.5% hexachlorophene
0.75% Brominated salicylanilides
1.0% 3,4,4′-trichlorocarbanilide

Considerable difference was evident as to their relative effectiveness in reducing the number of bacteria found on the skin.

The skin degerming effectiveness of Antibacterial Bar A (2% BSA, TFC, and TCC) and that of Antibacterial Bar B (0.75% TCC plus 0.75% G-11) are compared in Fig. 3. Each bar on the graph, the mean number of bacteria recovered in the fifth basin, is representative of a single 9–10 member panel. The data were collected over the same general time period. The mean number of bacteria found after four day's regimented use of Bar A ranged from 1,900 to 11,000 for the individual panels with an over-all mean of 5,700. For Bar B, the range

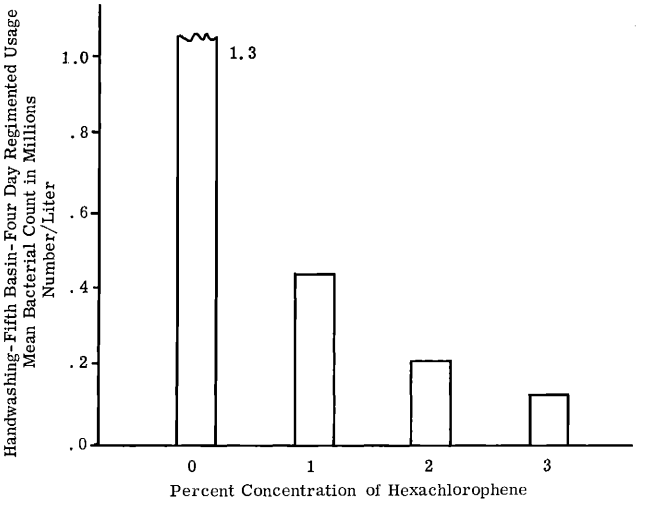

Figure 1. Skin degerming effectiveness of various levels of an antibacterial incorporated into a bar soap matrix

was 48,100 to 91,000 with a mean of 65,100. An analysis of variance showed that the difference between the two bars was statistically significant ($P < 0.01$).

The data presented in Table III and Figs. 1, 2, and 3 demonstrated the capabilities of a refined technique involving usage and handwashing sampling under controlled laboratory conditions. While these results are realistic and valid for the conditions of usage exposure prescribed for the panelists, the ultimate test is to determine degerming effects resulting from normal consumer usage.

For normal consumer usage studies, large panels of housewives are employed. Each subject is given unidentified bars (different shapes, colors, etc.) of the test soap to use at home for all handwashing and bathing. No other restrictions with regard to either soap or detergent exposure are imposed on the subjects. In order to make the test practical for use with large numbers of panelists, second basin counts are used

as a quantitative sample of the resident bacteria on the skin. Note should be made that these second basin values are influenced more by transient bacteria than the fifth basin values used in the laboratory

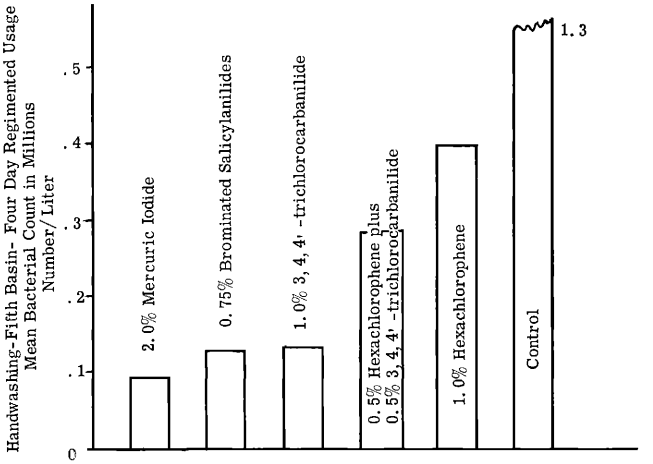

Figure 2. Skin degerming effectiveness of antibacterial bar soaps

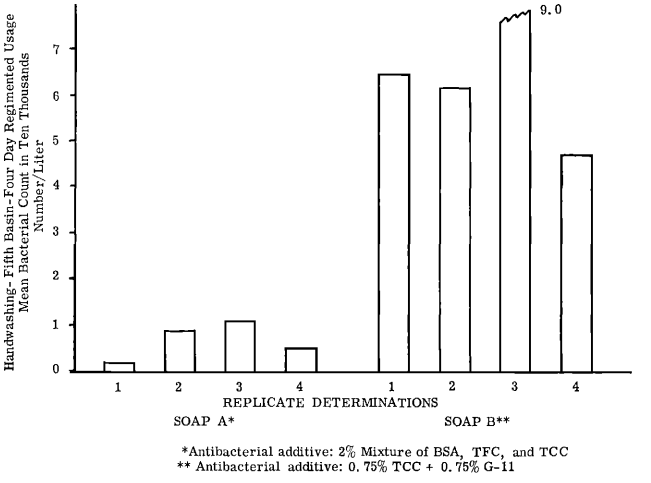

Figure 3. Skin degerming effectiveness of two different antibacterial systems

evaluations and are, therefore, somewhat less reproducible. The greater variability of the individual counts is compensated for by use of larger numbers of test subjects.

Three different consumer usage skin degerming evaluation tests will be described. The first was a test to determine effectiveness of Antibacterial Bar A relative to ordinary nonsanitizer soap. The study was run in the month of February in a northern city. Two groups of approximately 150 subjects each were sampled for the bacterial level on their hands and then were assigned on a random basis either the antibacterial soap or its control (without an antibacterial) to be used exclusively (*ad lib.*) in place of the customary bar soap product. After two weeks, the bacterial flora on the hands was redetermined. The data in Table IV show that the antibacterial soap users had a bacterial

TABLE IV

Skin Degerming: Effectiveness in a Temperate Environment under *Ad Libitum* Usage Conditions

Handwashing—Second Basin

Test Product	Mean Bacterial Count Per Liter	
	Initial Sampling	Two Week Sampling
Antibacterial soap A[a]	4,920,000	1,960,000
Control soap[b]	5,210,000	3,310,000

[a] Antibacterial additive: 2% mixture of BSA, TFC, and TCC.

[b] No antibacterial additive.

population on their hands lower than that of a comparable group using the control bar. An analysis of variance showed that this difference was significant at $P < 0.05$.

The second test was performed to determine the skin degerming efficacy of Antibacterial Bar A under hot, humid weather conditions. This study was conducted in a southern city during the month of July. Two groups of housewives were sampled for bacterial levels on the skin initially; one group was then assigned the antibacterial soap and the other the control soap without sanitizer additives. The test was conducted on a double-blind basis; neither the subjects nor the bacteriologists knew who was assigned the antibacterial product until the test was completed. The subjects were sampled for the levels of bacteria on the skin after two and four weeks of assigned product usage. The results of the handwashing tests (Table V) show again that there was a reduction in the bacterial flora on the skin of the subjects using the antibacterial soap. The difference between the bars was statistically significant ($P < 0.05$).

The results of this test, in general, corroborated the previous clinical test findings regarding the skin degerming efficacy of the antibacterial soap under *ad lib.* conditions. This test also showed that the antibacterial bar soap reduced the total skin bacterial level found on the hands as effectively in a sub-tropical climate as in a northern or temperate climate.

A third test was conducted to determine the capability of the handwashing test to detect differences in the relative skin degerming effectiveness of two antibacterial bar soaps containing different antibacterial systems. Degerming efficacy differences had already been detected under the above described laboratory handwashing conditions. The

TABLE V

Skin Degerming: Effectiveness in a Sub-tropical Environment under *Ad Libitum* Usage Conditions

Handwashing—Second Basin

Test Product	Two Weeks Usage	Four Weeks Usage
Antibacterial soap A[a]	665,000	441,000
Control soap[b]	1,381,000	1,126,000

[a] Antibacterial additive: 2% mixture of BSA, TFC, and TCC.
[b] No antibacterial additive.

TABLE VI

Skin Degerming Effectiveness of Two Different Antibacterial Systems

Handwashing—Second Basin—Ad Libitum Usage

Test Product	Mean Bacterial Count Per Liter	
	Two Weeks Usage	Four Weeks Usage
Antibacterial bar soap A[a]	649,000	657,000
Antibacterial bar soap B[b]	1,513,000	1,403,000
Control soap[c]	3,715,000	3,495,000

[a] Antibacterial additive: 2% mixture of BSA, TFC, and TCC.
[b] Antibacterial additive: 0.75% TCC + 0.75% G-11.
[c] No antibacterial additive.

test was conducted in a northern summer environment. Antibacterial Bar Soap A containing the 2% mixture of BSA, TFC, and TCC; Antibacterial Bar Soap B containing 0.75% TCC plus 0.75% G-11; and a control bar with no antibacterials were randomly distributed to approximately 500 housewives. Distribution was such as to obtain three equal-sized groups using each test bar. The test was conducted as a

double-blind study. Individual subjects were sampled for the levels of bacteria on their hands (second basin counts) after two and four weeks of assigned product usage. Over 160 subjects in each group completed the study. Results are presented in Table VI.

A statistically significant ($P < 0.05$) reduction in the bacterial flora of subjects using the antibacterial bar soaps when compared to the control group was evident at both the second and fourth week sampling. Moreover, the analysis of variance showed that the observed difference in skin degerming effectiveness between the two antibacterial bar soaps was significant at $P < 0.05$.

The ability to demonstrate statistically significant differences between two effective antibacterial soaps under normal usage conditions is evidence of the resolving power of which properly designed and executed testing techniques are capable.

Conclusions

Results of *in vitro* testing may be of value in screening for potential antibacterial ingredients. They are not reliable, however, for predicting actual skin degerming effectiveness of products. In the case of minimum inhibitory concentration tests, the results are too dependent upon the particular bacterial species or strains selected for the test to be meaningful in determining the relative effectiveness of various antibacterial bar soaps.

At the present time, the only reliable way to determine efficacy of antibacterial bar soaps is the measurement of a change in the number of cutaneous bacteria resulting from exposure of the skin of human subjects to the products. Procedures are available for conducting such studies. These are feasible not only for use under controlled conditions in the laboratory, but also for *ad lib.* consumer use studies in the field.

(Received December 2, 1965)

References

(1) Pillsbury, D. M., Livingood, C. S., and Nichols, Anna C., *Arch. Dermatol. Syphilol.*, **45,** 61 (1942).
(2) Price, P. B., *J. Infect. Diseases*, **63,** 301 (1938).
(3) Cade, A. R., *Soap Sanit. Chemicals*, **73,** 35 (1950).
(4) Quinn, H., Voss, J. G., and Whitehouse, H. S., *Appl. Microbiol.*, **2,** 202 (1954).
(5) Travers, J. J., Rohloff, A. C., Ambye, E. L., and Vinson, L. J., *J. Soc. Cosmetic Chemists*, **10,** 410 (1959).