J. Soc. Cosmet. Chem., 43, 187–193 (July/August, 1992)

Comparison of exaggerated and normal use techniques for assessing the mildness of personal cleansers

B. H. KESWICK, K. D. ERTEL, and M. O. VISSCHER,
The Procter & Gamble Company, 11520 Reed Hartman Highway, Cincinnati, Ohio 45241.

Received December 31, 1991.

Synopsis

A variety of exaggerated-exposure wash techniques have been used to evaluate the mildness of personal cleansing products in lieu of clinical evaluation during normal use. This study compares two exaggerated methods, a forearm wash method and a flex wash method, to home-use studies to determine how well the exaggerated methods approximate *ad lib* usage. The results indicate that the forearm wash method is a better predictor of product mildness under home-use conditions when soap bars are used. Both exaggerated methods yield similar mildness pictures when syndet bars are used, although the forearm wash method is more discriminating. The flex wash implement (sponge) induces greater damage to the stratum corneum than does the forearm wash implement (towel). However, this does not account for the total damage observed when product is applied in the flex wash method. A significant amount of damage is apparently due to product reaching and interacting with lower layers of the skin once the barrier's integrity has been compromised by the sponge. The results indicate that the forearm and flex wash methods are based on different exposure models, and that the model on which the forearm wash is based is more predictive of actual consumer use conditions than is the model on which the flex wash is based.

INTRODUCTION

An important characteristic of bar soap products is their mildness, as this is often a major determinant of consumer acceptance. Methods originally developed to assess personal cleanser mildness, such as patch and chamber tests, have been criticized because they do not reflect consumer usage patterns (1–3). More recently, exaggerated protocols (2–7) have been developed to mimic consumer-use conditions and allow mildness predictions to be made in a relatively short time. The following work examines two five-day exaggerated wash protocols based on published procedures—a flex wash method (7) and a forearm wash method (6). Mildness results obtained by each method are compared to results obtained in home-use studies to see how well the exaggerated wash protocols predict results generated under actual use conditions.

EXPERIMENTAL

STUDY POPULATION

Healthy male and female volunteers were recruited as test subjects. All prospective

subjects were screened prior to enrollment to assure that they had no history of sensitivity or allergy to soap or detergent products, and that they had not used prescribed anti-inflammatory or antibiotic drugs for at least three weeks prior to study start-up. Female subjects were also screened to assure that they were not pregnant or lactating. Informed consent was obtained from all subjects prior to their being issued product or beginning treatments.

MATERIALS

Test products were either commercially available cleansing bars or cleansing bar prototypes. The major formula components of the test products are reported in Table I. Products provided to the test laboratories were coded such that subjects and test site personnel, including trained graders, were unaware of product identities. Masslinn® towels (Chicopee Mills, New Brunswick, NJ) were used as wash implements in the forearm wash method; JAECE Identi-Plug® (size D, JAECE Industries, Inc., Tonawanda, NY) or Cerafoam (1.75″ diameter, Wilfred Heath Ltd., Stoke-On-Trent, England) sponges were used as wash implements in the flex wash method.

TEST METHODS

All studies were conducted at independent testing laboratories. The home-use, forearm wash, and flex wash studies were run during the period from winter to early spring. The flex wash implement study was run in late summer; however, this method is reported to yield results that are free from seasonal variation (7). The numbers of subjects enrolled in the exaggerated method studies were consistent with those specified in the literature; the number of subjects enrolled in each of the home-use studies was based on prior testing experience with similar products.

Subjects in the home-use studies were randomly assigned a single product to take home and use for four (soap bars) or twelve (syndet bars) week periods. Subjects periodically returned to the test facility for a visual evaluation of the redness and dryness induced by product usage. The exaggerated studies were run as paired comparisons, with a single product randomly assigned for use on each arm. Key features of the exaggerated wash methods are summarized in Table II.

Table I
Major Formula Components of the Test Products Used in These Studies

Test code	Major formula components
A	Sodium tallowate, sodium cocoyl isethionate, sodium cocoate, stearic acid, sodium iesthionate, coconut fatty acid
B	Soap (sodium tallowate and sodium cocoate or palm kernalate types), sodium cocoglyceryl ether sulfonate, glycerin, coconut or palm kernel acid, polyquaternium-7, guar hydroxypropyltrimonium chloride
C	Sodium alkylglyceryl ether sulfonate, sodium lauroyl sarcosinate, sodium soap, stearic acid, lauric acid, Polyquarternium-7, Polyquarternium-10
D	Sodium cocoyl isethionate, stearic acid, sodium tallowate, sodium isethionate, coconut acid, sodium stearate, sodium dodecylbenzenesulfonate, sodium cocoate or sodium palm kernelate

Table II
Key Features of the Flex and Forearm Wash Methods Used in This Study

	Flex wash method	Forearm wash method
Implement used	Identi-Plug® or Cerafoam sponge	Masslinn® towel
Bar/implement lathering	10 Seconds—rub sponge on bar	6 Seconds—rub towel on bar
Wash area	Inner elbow crease	Inner forearm
Wash time	60 Seconds	10 Seconds
Residence time	None	90 Seconds
Rinse time	10–15 Seconds	15 Seconds
Wash visits/day (final day)	3	2 (1)

A separate flex wash study was conducted to evaluate the effect of the implement on the subjects' skin and on the study outcome. A procedure identical to the normal flex wash method was used, except that a single product was applied with either a sponge or a Masslinn® towel. Treatment (implement) assignments were made randomly so that the implements were used for an approximately equal number of times on the left and right arms. An evaporimeter (model EP-1C, ServoMed, Uppsalla, Sweden) was used to record transepidermal water loss (TEWL) values for treated sites to provide an indication of stratum corneum barrier integrity.

STATISTICAL ANALYSIS

Visual attribute scores at each time point were subjected to ANOVA to account for subject, side (left vs right), and product differences. Least squares attribute means for each test product at each evaluation time point were compared using t-tests.

RESULTS AND DISCUSSION

Ideally, an exaggerated wash protocol should have the ability to quickly and accurately predict mildness trends while maintaining its relevance to conditions of actual consumer use. The exaggerated protocols examined here are of five days' duration, significantly shorter than either of the home-use studies. Both use an area of the forearm as the wash site, and both use an implement to apply product. The methods differ in the type of implement used: the flex wash method uses a sponge to apply product while the forearm wash method uses a Masslinn® towel.

Results obtained from exaggerated studies performed using two bar soap products coded A and B are reported in Table III. Only erythema (redness) results are reported for the flex wash method, since this method is not useful for measuring dryness (7). Both exaggerated wash methods differentiate between the products; however, the flex and forearm methods yield different mildness pictures. In the flex method, product A induces significantly less erythema than product B, indicating that the former is the milder of the two products. The opposite was found in the case of the forearm wash method; the data for both erythema and dryness indicate that product B is significantly milder than product A.

Table III
Comparison of Visual (Mildness) Results Obtained When Two Personal Cleansing (Soap) Bars Were Tested in 5-Day Exaggerated-Use and 4-Week Home-Use Studies

Test product	Flex test	Forearm wash test		Home-use test			
				Forearms		Legs	
	Erythema	Erythema	Dryness	Erythema	Dryness	Erythema	Dryness
A	0.77	1.61	1.76	0.96	1.17	0.94	2.43
B	2.45	0.76	1.09	0.85	1.07	0.82	2.47
p-value	0.0001	0.0001	0.0001	0.24	0.37	0.11	0.05

Values reported are mean endpoint attribute scores; differences are considered significant if $p \leq 0.05$.

While these findings are useful from developmental and marketing standpoints, they do not answer the question of how mild the products will be under conditions of actual consumer use. Insight into the question is provided by data generated in a home-use study (Table III). Because this is an *ad lib* study, the differences observed are not as great as those in the exaggerated studies, and in only one instance is a significant difference found between the products. However, there is a clear trend in the data indicating that product B is milder than product A. This ranking is consistent with the ranking predicted by the forearm wash method, opposite that predicted by the flex wash method.

To determine whether the contradictory mildness picture yielded by the flex and forearm wash methods was due to product composition (bars A and B both contain a high percentage of soap), the exaggerated wash procedures were repeated using two synthetic detergent (syndet) bars, coded C and D. The results of these studies, along with results generated in a 12 week home-use study, are reported in Table IV. The home-use results show product C to be milder than product D. Both exaggerated methods show a similar mildness trend, although the flex test fails to find a significant mildness difference between the products.

During the performance of the flex studies, subjects developed a noticeable abrasion at the wash sites. This abrasion is atypical of normal use, and was not observed in any of the subjects participating in the forearm wash studies. As noted earlier, one of the

Table IV
Comparison of Visual (Mildness) Results Obtained When Two Personal Cleansing (Syndet) Bars Were Tested in 5-Day Exaggerated-Use and 12-Week Home-Use Studies

Test product	Flex test	Forearm wash test		Home-use test			
				Forearms		Legs	
	Erythema	Erythema	Dryness	Erythema	Dryness	Erythema	Dryness
C	0.52	1.27	1.11	0.28	0.76	0.45	2.16
D	0.56	1.82	1.61	0.31	0.98	0.70	2.70
p-value	-NS-	0.0001	0.0001	-NS-	0.031	0.016	0.003

Values reported are mean endpoint attribute scores; differences are considered significant if $p \leq 0.05$. P-values ≥ 0.5 are indicated by -NS-.

differences between these exaggerated methods is the implement used to apply product. The sponge is, by nature, rough, while a towel presents a smooth surface; hence, the implements differ in their abrasive potential.

To determine what impact, if any, implement abrasiveness had on the study outcome, a modified flex study was run in which a single product was applied in the usual fashion, or with a Masslinn® towel. Visual grades assigned during the course of this experiment are summarized in Figure 1. The implement effect is clearly visible in these data; both the rate and magnitude of erythema development are greater when the product is applied with a sponge.

In addition to visual grades, TEWL values were recorded throughout the implement study to provide a measure of the damage induced in the stratum corneum by the treatments. These data are summarized in Figure 2. To provide a more quantitative evaluation, the data were analyzed by approximating the change in mean TEWL value, with time for each treatment as a linear function and calculating least-squares regression coefficients for the fitted lines. The calculated intercepts provide an indication of the stratum corneum integrity at the start of the study, i.e., the baseline TEWL value. The intercepts calculated for the sponge and towel treatments are not significantly different ($p \leq 0.05$), showing that both treatment groups began the study with equal barrier function. The calculated slopes provide an indication of the rate at which each treatment induces stratum corneum damage, a higher slope indicating a greater rate of damage.

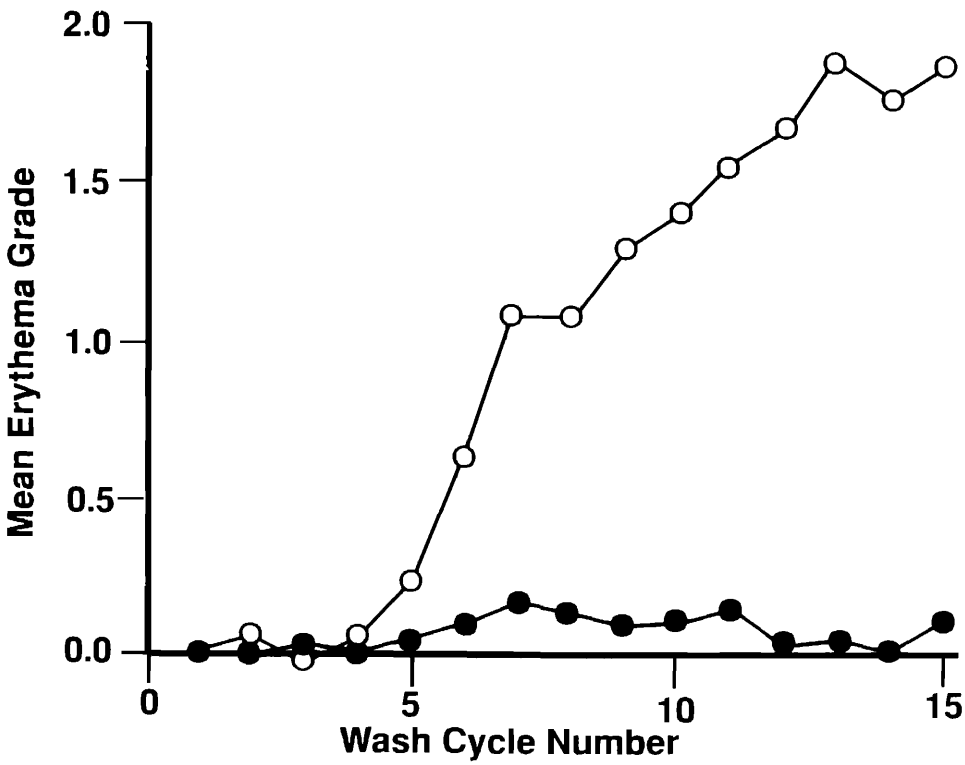

Figure 1. Mean erythema scores of subjects' arms washed with a sponge or towel in the flex test plotted as a function of wash cycle number. ○ = sponge; ● = towel.

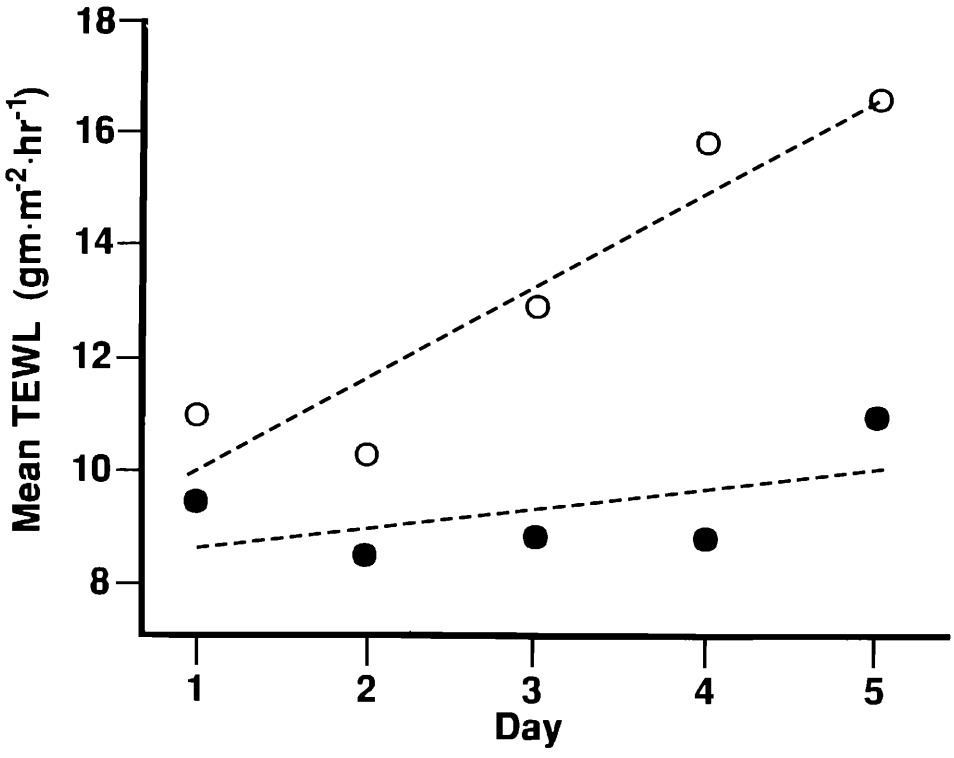

Figure 2. Mean daily transepidermal water loss values obtained from subjects' arms washed with a sponge or towel plotted as a function of time. The fitted lines were obtained by least-squares regression. ○ = sponge, y = 8.29 + 1.70X; ● = towel, y = 8.33 + 0.33X.

Although both treatments induce damage, the rate is greater when a sponge (slope = 1.70) is used to apply the product rather than a towel (slope = 0.33).

The greater rate of stratum corneum damage induced in the flex wash method is not due solely to the abrasiveness of the implement. To demonstrate this, a test leg was included in the modified flex study in which no product was applied, i.e., the subjects' arms were "washed" with a moistened sponge to provide an indication of how much barrier damage was due to the sponge alone. Treating the TEWL data generated in this test leg as described above yields a line having a slope value of 0.39, which is only slightly greater than that observed when the test product is applied with a towel. Clearly, both the sponge and the product contribute to the greater rate of stratum corneum damage. The small amount of damage induced when the product was applied with a towel, however, indicates that the product contribution is not due to skin surface effects; the contribution must be due to product interacting with lower skin layers exposed after the stratum corneum is damaged by the sponge. This suggests that a two-stage mechanism is operating in the flex method: an initial stage in which the stratum corneum is damaged by the sponge, followed by a stage in which product comes in contact with lower layers of the skin, inducing an inflammatory response that contributes to further barrier breakdown.

These findings point out a key difference between the forearm and flex wash methods. The forearm wash method provides an indication of product effects on the surface of the

skin, mimicking most normal-use conditions. Results from the flex wash method, however, are dependent on product interaction with the surface and lower layers of the skin. While there are instances where this type of exposure could occur during home use (for example, wound cleansing or product use on sun-damaged skin), this certainly cannot be considered the norm. Thus, while the forearm wash and flex wash methods both provide measures of product mildness, these measures are based on different exposure models. Given the consistency of the forearm wash and home-use data generated with both soap and syndet bars, the forearm wash exposure model appears to be more relevant to actual consumer use conditions than is the flex wash exposure model.

SUMMARY

The mildness of several personal cleansing products was measured using two exaggerated wash protocols—a forearm wash method and a flex wash method. The forearm wash method more accurately predicted home-use results when soap bars were tested. Both exaggerated methods yielded similar mildness pictures when syndet bars were tested; however, the forearm wash method was more discriminating. Experiments designed to measure implement effects show that the towel used in the forearm wash method induces minimal damage to the stratum corneum, while damage induced in the flex wash method is due to both implement and product effects. The results indicate that while both exaggerated methods provide a measure of product mildness, the methods are based on different models of consumer use and exposure.

REFERENCES

(1) P. J. Frosch and A. M. Kligman, The soap chamber method for assessing the irritancy of soaps, *J. Am. Acad. Dermatol.*, **1**, 35–41 (1979).
(2) P. J. Frosch, "Irritancy of Soaps and Detergent Bars," in *Principles of Cosmetics for the Dermatologist*, P. H. Frost and S. T. Horwith, Eds. (C. V. Mosby, St. Louis, 1982), pp. 5–12.
(3) G. Sauermann, A. Doerschner, U. Hoppe, and P. Wittern, Comparative study of skin care efficacy and in-use properties of soap and surfactant bars, *J. Soc. Cosmet. Chem.*, **37**, 309–327 (1986).
(4) B. Komp, Skin compatibility tests—Importance in skin cleansing product development, *Cosmet. Toiletr.*, **102**, 89–94 (1987).
(5) B. Komp and A. K. Reng, Developing ether sulfate-free surfactant formulations, *Cosmet. Toiletr.*, **104**, 41–45 (1989).
(6) M. F. Lukacovic, F. E. Dunlap, S. E. Michaels, M. O. Visscher, and D. D. Watson, Forearm wash test to evaluate the clinical mildness of cleansing products, *J. Soc. Cosmet. Chem.*, **39**, 355–366 (1988).
(7) D. D. Strube, S. W. Koontz, R. I. Murhata, and R. F. Theiler, The flex wash test: A method for evaluating the mildness of personal washing products, *J. Soc. Cosmet. Chem.*, **40**, 297–306 (1989).