J. Soc. Cosmet. Chem., 43, 101–112 (March/April 1992)

Quantitative assessment of cyanoacrylate follicular biopsies by image analysis

DAVID G. GROH, OTTO H. MILLS, JR., and
ALBERT M. KLIGMAN, *Skin Care Laboratory, Personal Care Products Research and Development, Amway Corporation, Ada, MI 49355 (D.G.G.), Hill Top Research, Inc., East Brunswick, NJ 08816 (O.H.M.), and University of Pennsylvania, Philadelphia, PA 19104 (A.M.K.).*

Received November 27, 1991. Presented at the Annual Scientific Meeting of the Society of Cosmetic Chemists, New York, December 5–6, 1991.

Synopsis

A method has been developed for quantifying the comedogenic effect of topically applied products using image analysis. Cyanoacrylate follicular biopsies are evaluated under polarized light, which brightly illuminates horny casts and microcomedones, enabling rapid measurement of density and size distribution. A number of substances were tested on the backs of human subjects. There was a high correlation between stereomicroscope scores for comedogenicity and the readout by digital image analysis ($r > 0.9$).

INTRODUCTION

The rabbit ear model was the first widely used procedure for determining whether a topically applied substance could induce comedones. In the early eighties Mills and Kligman developed a human model for assessing comedogenic substances (1). The "non-animal" aspect of this model has greater meaning today than when it was first developed.

In that model, test substances were applied under occlusion for one month on the upper part of the backs of young adult black men having large follicles. The suitability of each subject was checked by the noninvasive "follicular biopsy" technique (2). The follicular biopsy (FB) samples the contents of sebaceous follicles. To obtain a follicular biopsy, the skin surface was coated with a thin layer of methyl cyanoacrylate glue and a glass slide was applied firmly. After one minute, the slide was carefully peeled off, bringing with it a thin layer of stratum corneum and follicular horny extensions. Hyperkeratotic follicles on the follicular biopsy appeared as cylinders of horny material surrounding extracted vellus hairs (Figure 1). Follicular biopsies obtained at the end of the four weeks were then evaluated using a standard four-point scale: 0, noncomedogenic; 1, smallish horny cylinders, involving at least half of the follicles; 2, moderately sized horny masses

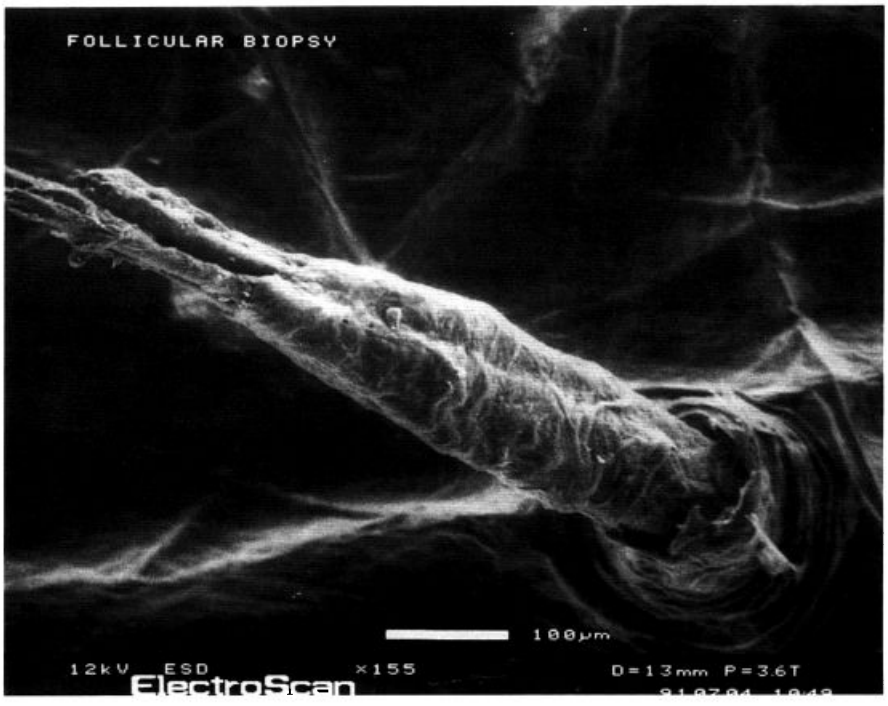


Figure 1. An ESEM (environmental scanning electron micrograph) of a follicular biopsy.

over most of the field; and 3, larger, globoid microcomedones over the entire field. This study showed that, although the human model was less sensitive, there was reasonably good correlation between the human and rabbit ear models. Recently, Mills *et al.* (3) compared black males and white females and found a good correlation between microcomedone formation on the backs of these two groups. Image analysis (IA) has been used to evaluate skin surface topography (5–10), histology (11–14), and more recently, facial comedones by porphyrin fluorescence (15). We evaluated follicular biopsies by image analysis under polarized light. Polarized light has been effectively used for the photography of the skin (16) and consists of wavetrains whose planes of vibration are oriented in a parallel manner (17).

METHODS AND MATERIALS

CLINICAL STUDY

The current study used a variation of the Mills and Kligman model. In the study the product was occluded on the upper backs of twelve white females with a history of acne. Each panelist was screened using the "follicular biopsy" technique, and a score of 1 or greater was required for inclusion in the test. The test was concluded in October of 1990.

The test samples were: (A) a cleansing masque; (B) a moisturizer; (C) a positive control, acetylated lanolin alcohol (Acetulan®); and (D) a negative (no product) control. Test samples A and B were chosen because they were originally tested using the modified

Kligman rabbit ear assay (4). In the rabbit ear model, sample A received a score of 0 and sample B received a score of 1+.

The test agents were blind coded and applied in a randomized manner on the upper back in five-centimeter squares using Dermapore® tape. The agents were liberally applied to each square, with reapplication on Mondays, Wednesdays, and Fridays for four consecutive weeks, providing 28 days of occlusion.

At the end of four weeks, FB slides were obtained and evaluated under a stereo microscope by an expert grader. These interpretive evaluations were conducted in a blind manner using the standard four-point scale mentioned previously. In this paper the "interpretive score" refers to the expert grader's evaluation of a follicular biopsy slide.

DIGITAL IMAGE ANALYSIS

Images of follicular biopsy slides are captured through an Olympus SZH light microscope with a 1× objective, overhead fiberoptic ring light illumination (Highlight® 3000, Olympus Corporation, Lake Success, NY), two polarizing filters, and a high-resolution Nuvicon® video camera (Series 68, Dage-MTI Inc.). An Intel® PC-AT system with a Joyce-Loebl® vision card set (Mini Magiscan, Compex Inc., Mars, PA) was used to digitize and analyze the captured video image.

In our system a polarizing filter is placed over the fiberoptic light source and another polarizing filter is attached to the microscope lens. Figures 2 and 3 show a follicular biopsy slide under regular and polarized light, respectively. The horny impactions are barely visible under regular light (Figure 2), while under polarized light they appear

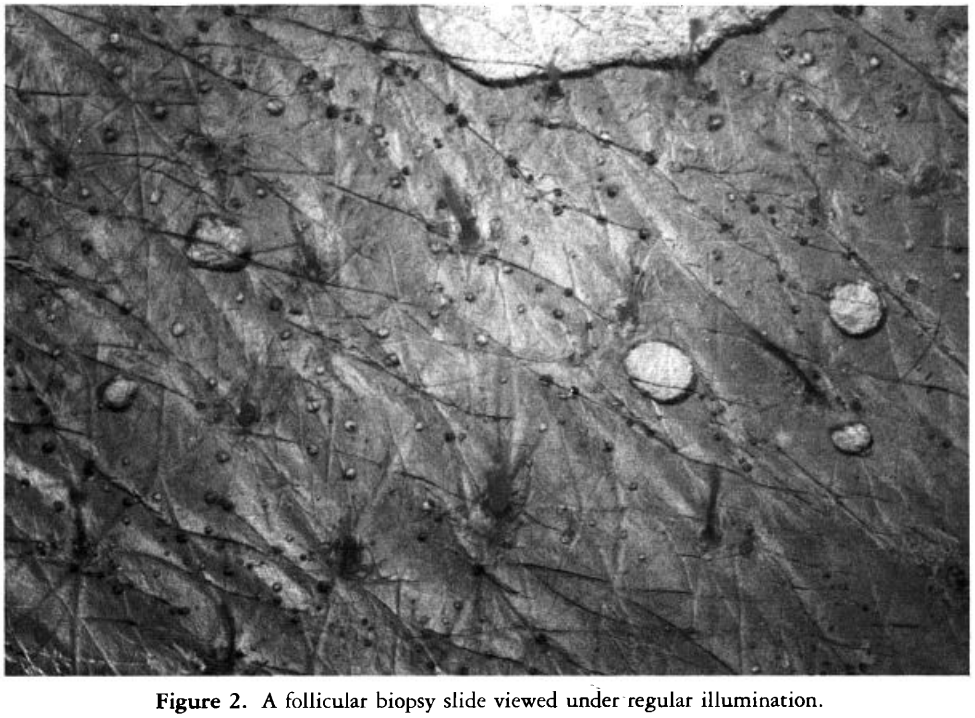


Figure 2. A follicular biopsy slide viewed under regular illumination.

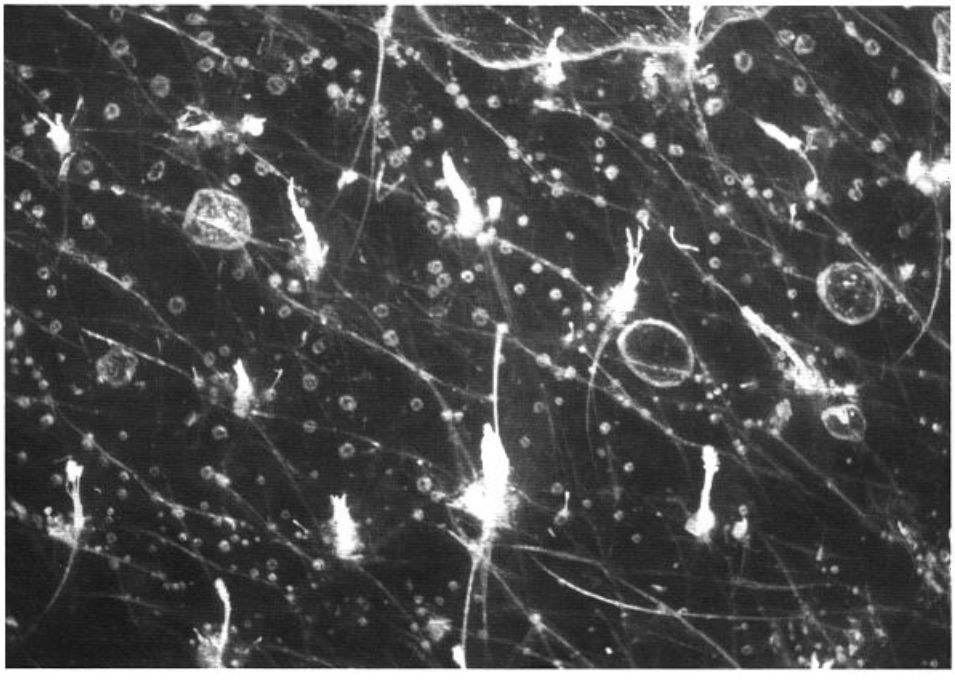


Figure 3. The same follicular biopsy as in Figure 2 viewed under polarized illumination. Note the contrast between the horny impactions and the background.

more intense, standing out from the background (Figure 3). This contrast between the impactions and the background allows the samples to be evaluated by image analysis for total number and size distribution of impactions.

A task (program) for evaluating FB slides under polarized illumination was developed that automatically counts the number, area, and size distribution of horny impactions. An IA task consists of the following basic steps:

1. Image capture (digitizing)
2. Segmentation
3. Object detection
4. Measurement/data storage
5. Analysis

The task automatically controls what objects to evaluate (based on intensity range), what type of editing to perform, and what measurements to make and record. An example of a task used to evaluate follicular biopsy slides is illustrated in Figure 4. Image analysis data was obtained from the FB slides of the 12 panelists for each test sample.

The follicular biopsies were evaluated at 15×, giving a field of view 9.4 mm × 10.0 mm. The measurements are in units of pixels for comparative purposes.

RESULTS AND DISCUSSION

When looking at the individual data from each panelist for sample A (cleansing masque), we see a relationship between the interpretive scores and the number of objects

```
{ Original task file COMEDO7.TSK }
{ Task list for GENIAS25 2.1 }
MAIN MENU
Capture
CAPTURE
Correct
Change
COMEDO
CAPTURE
Quit
MAIN MENU
Grey
GREY OPS
Linear
LINEAR FILTERS
Mean
LINEAR FILTERS
Quit
GREY OPS
Quit
MAIN MENU
Threshold
THRESHOLD
Manual
50              {SLICE : LOW }
63              {SLICE : HIGH }
THRESHOLD
Quit
MAIN MENU
Binary
BINARY OPS
Object-based
OBJECT-BASED OPS
Delete
DELETE OBJECTS
Area
 0.00000E+00 { Minimum?   }
 1.00000E+01 { Maximum?   }
Inclusive
DELETE OBJECTS
Quit
OBJECT-BASED OPS
Quit
BINARY OPS
Quit
MAIN MENU
Measure
Objects
FIELD MEASUREMENTS
Quit
0               { Nesting levels? [0 = no nesting]   }
Yes
MEASUREMENTS 1
Select-all
MEASUREMENTS 1
Quit
MAIN MENU
Quit
Yes
END
```

Figure 4. An image analysis program (task list) used to evaluate follicular biopsy slides.

from IA (Figure 5). The general trend for this data indicates that when the interpretive score increases, so does the number of objects from IA.

The relationship between the individual interpretive scores and the number of objects from IA for sample B is not as strong as in sample A, but there is still a trend (Figure 6). This is due to the low density of microcomedones in the field of view at the 15× magnification. The interpretive evaluations are conducted at a lower magnification, which gives a larger field of view. For sample B, one half of the panelists received an interpretive score of 0.5 or less.

Samples A and C did not receive any scores of 0.5 or less, and sample D had four panelists with a score of 0.5 or less. This magnification difference with sample B is not a factor with the other samples due to the higher numbers of microcomedones overall.

The graph of sample C (positive control) shows a large difference between the interpretive scores and the number of objects from IA (Figure 7). A closer review of the FB slides from sample C reveals an increased number of microcomedones and a large degree of scales and flaking of the stratum corneum surface. The IA system picked up the flaking as background noise, thereby giving higher results for object number and area. This surface disorder is most likely an irritation response caused by the four-week occlusion of acetylated lanolin alcohol. Follicular biopsy slides were not taken on three of the 12 panelists due to irritation produced by the positive control (Figure 7).

A graph of the data from each panelist for sample D (negative control) again shows a

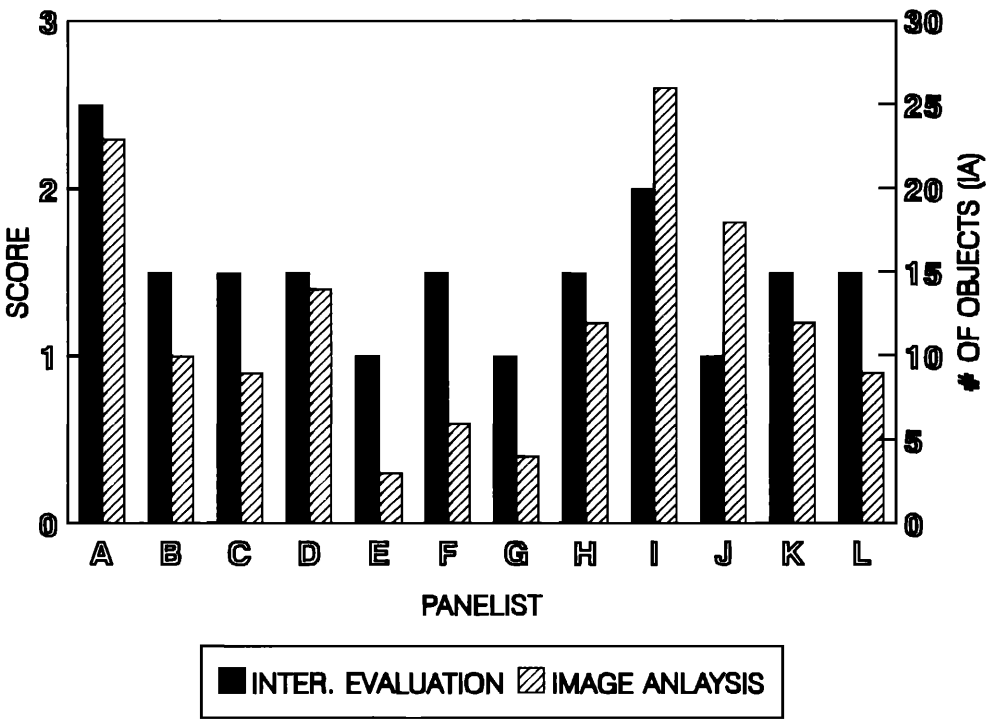


Figure 5. A graph of the individual data from sample A (cleansing masque), showing the interpretive score and the IA data for all 12 panelists.

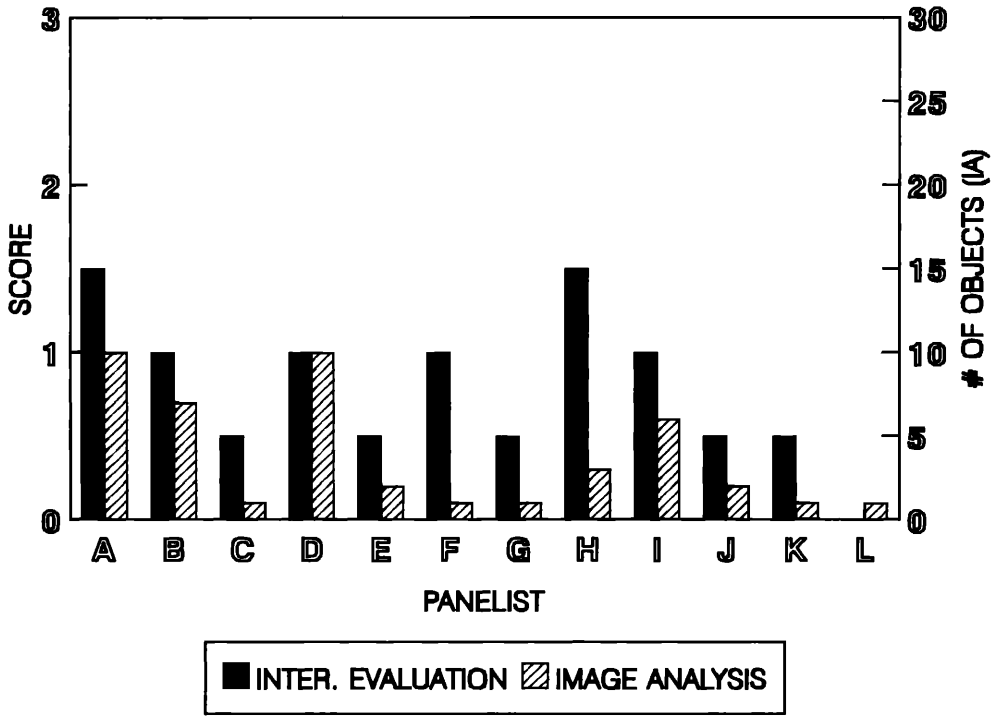
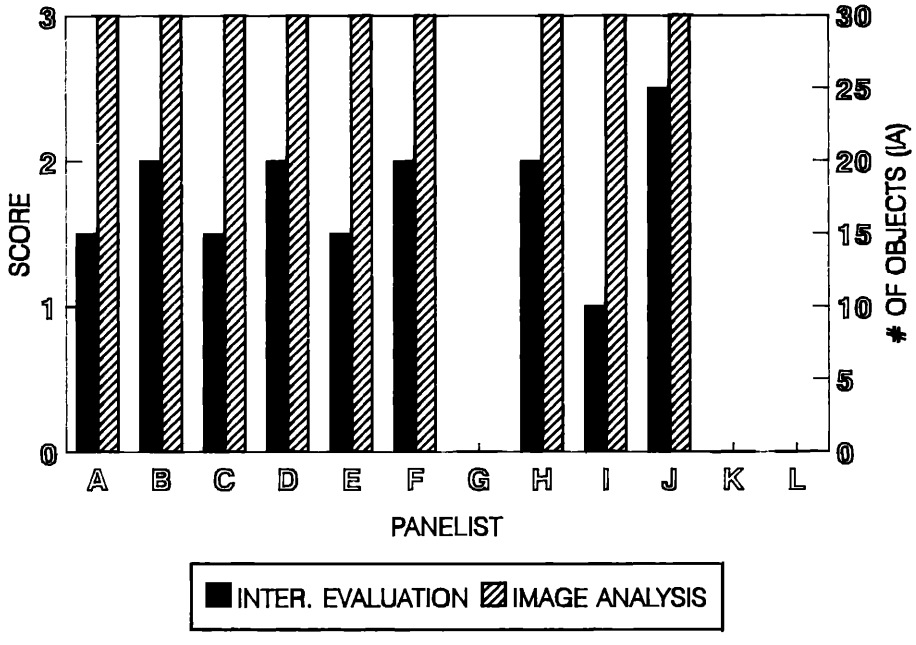


Figure 6. A graph of the individual data from sample B (moisturizer), showing the interpretive score and the IA data for all 12 panelists.



PANELISTS G, K & L WERE NOT SAMPLED DUE TO IRRITATION

Figure 7. A graph of the individual data from sample C (positive control), showing the interpretive score and the IA data for all 12 panelists.

relationship between the interpretive scores and the number of objects from IA as seen with sample A (Figure 8).

The differences between the individual interpretive scores and the number of objects from IA for the test agents indicates the nature of the evaluations. The interpretive score is based on a subjective human evaluation of a follicular biopsy slide, which takes into consideration the number and size of the comedones. The image analysis evaluation of a follicular biopsy represents an objective numeric count of the total number of comedones. If we look at the average of the individual data for each sample, we see a good correlation between the interpretive evaluations and image analysis numbers.

A graph of the average number of objects for each test sample from IA versus the average interpretive score for each sample gives a straight line (Figure 9). The linear correlation of this line is greater than 0.9, which indicates a direct relationship between image analysis data and the interpretive evaluation scores. It should be noted that data from the positive control was not included in Figure 9 due to the large increase in the number of objects.

A graph of the average object area for each test sample from IA versus the average interpretive score for each sample yields a flat, almost horizontal line (Figure 10). This data indicates that the average object area is very similar for all test samples being evaluated (i.e., the computer is evaluating similar objects for all three samples). It should be noted that data from the positive control was not included in Figure 10 due to the large increase in the object area (Figure 11).

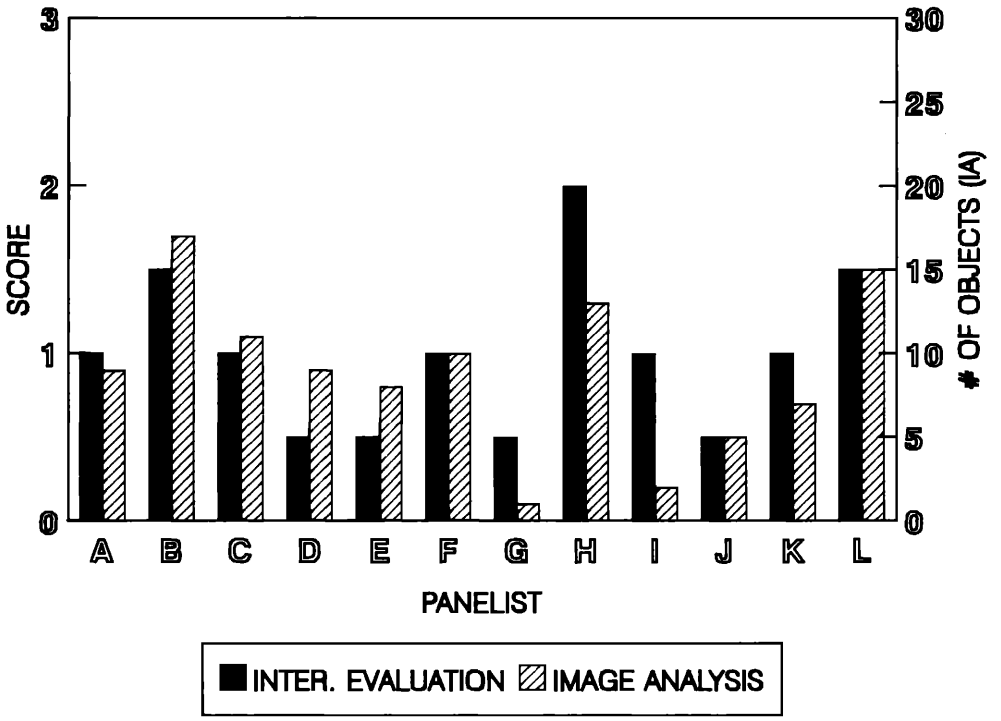
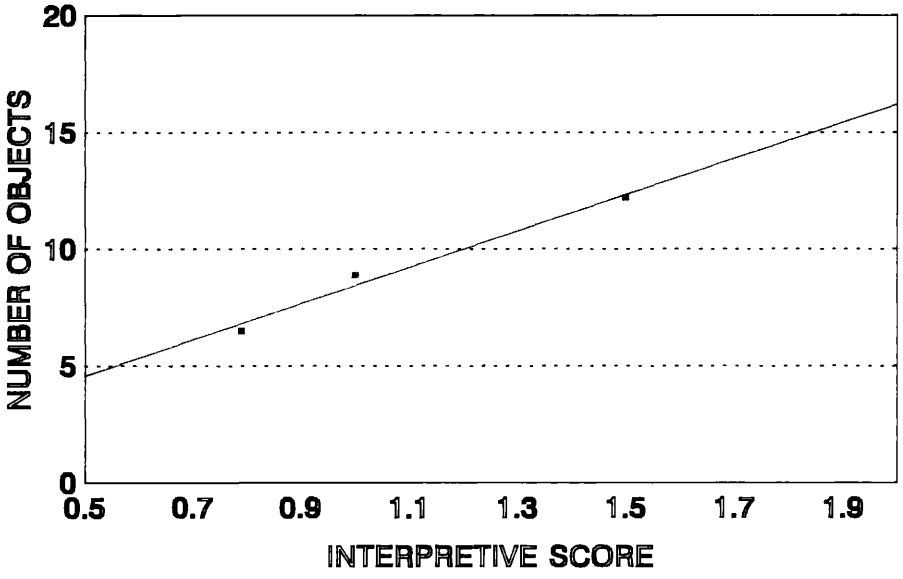
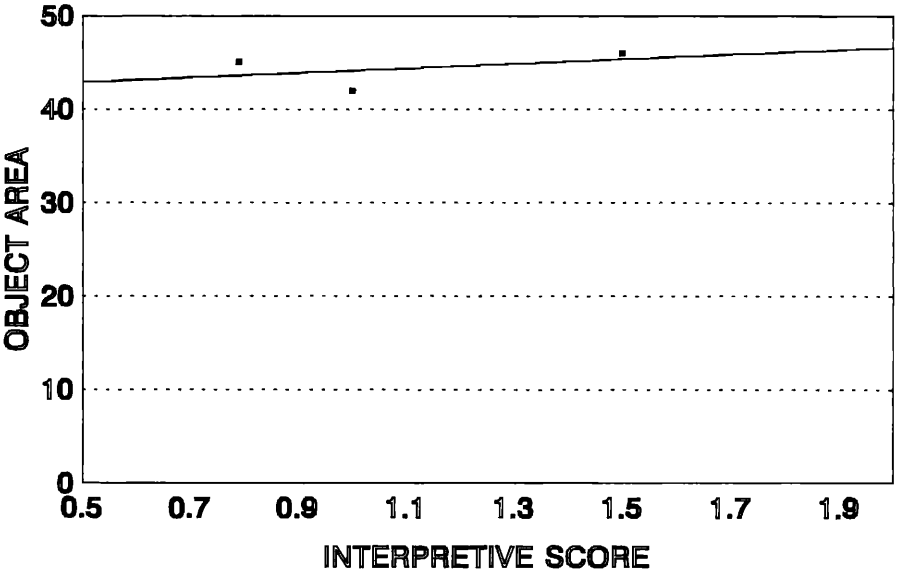


Figure 8. A graph of the individual data from sample D (negative control), showing the interpretive score versus the IA data for all 12 panelists.



$r > 0.9$

Figure 9. A graph of the average number of objects for each test sample from IA versus the average interpretive score for each sample.



$r > 0.9$

Figure 10. A graph of the average object area for each test sample from IA versus the average interpretive score for each sample.

Statistical analysis (ANOVA and Newman-Keuls) of the IA results indicates a significant difference at the 99% confidence level between test sample C (positive control) and the other three test samples (A, B, and D). No significant differences were noted

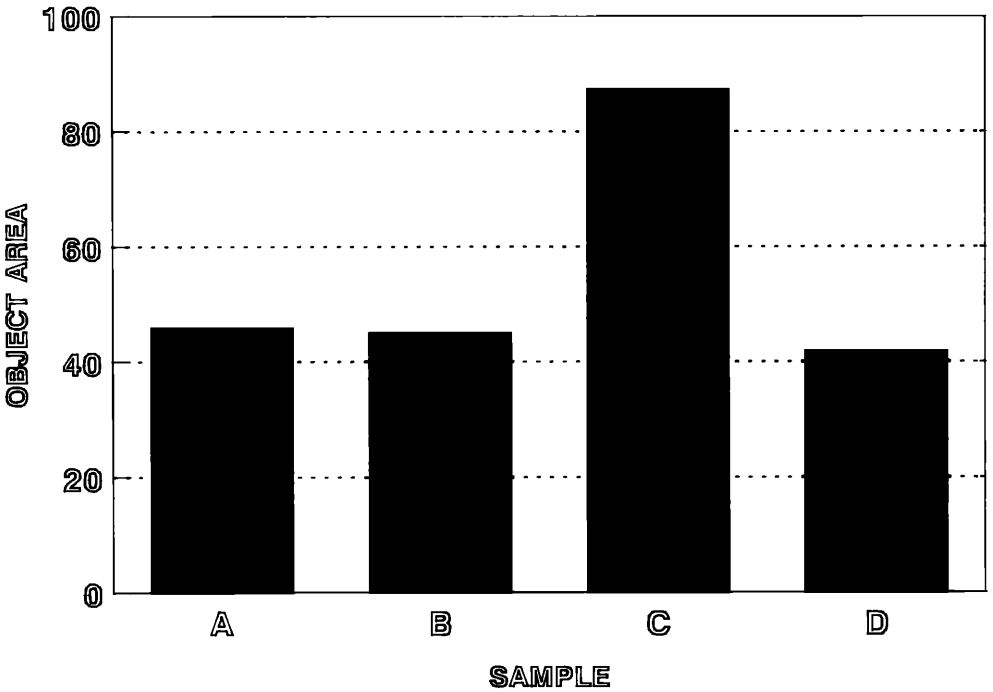


Figure 11. A graph of the average object area for each test sample.

between samples A, B, and D. These results correlate well with the results from the interpretive evaluations, where sample C was considered comedogenic.

Image analysis numbers obtained from evaluations of the same slides on separate days consistently give a correlation of 0.9. Repeated measurements on the same slide, on the same day, show a variation of less than 0.10% in the data. These results indicate that IA data is reproducible over time. The only change in instrumental set-up that has an effect on IA results is the intensity of light and degree of polarization. These can be controlled by adjusting to a predetermined threshold prior to performing IA evaluations.

When evaluating follicular biopsy slides by image analysis there are a number of factors that can influence the results. The most obvious is image processing. Follicular biopsy slides from this study were evaluated using five different IA tasks or programs, and although the numbers were different with each task (slightly larger or smaller), the data consistently gave a linear correlation greater than 0.9 when compared to the interpretive scores. This indicates that image processing is less of a factor, providing that all of the test samples are evaluated using the same task under to same conditions.

Artifacts in the FB slides present another problem. One must be sure to use slides with a confluent adhesive coating. Occasionally the glue does not adhere well to the slide and part of the "sample" is lost, leaving an open area. Follicular biopsy slides cannot have any missing sections or the IA system will give artificially low numbers. Large flakes and clumps of material should be avoided, as they will cause the IA system to give artificially high numbers. In one instance, a slide from one panelist was getting very high numbers with IA, while the interpretive score was very low. On review, this may have been

caused by the presence of tiniea versicolor on the subject's back. The clinical expert grader was able to discern this condition and interpret accordingly. The IA system, however, confused the dermatological problem with horny impactions. The above problems can be addressed by a brief inspection of the slides under a microscope prior to IA processing.

Finally, irritation of the stratum corneum may cause problems when evaluating FB slides. In our study, sample C (acetylated lanolin alcohol) caused an increase in uplifted scaly material on the stratum corneum surface. When evaluated by IA, this sample gave higher than expected numbers for number of objects and object area, approximately double those of the other samples. The IA task we used is more sensitive to this type of background noise and may also be helpful in detecting irritation during a human comedogenicity study.

CONCLUSION

Image analysis has been used to quantify follicular biopsy slides obtained during a human comedogenicity study. Results indicate that there is a linear correlation between IA data and the interpretive score. Image analysis data is reproducible over time, in that the system is non-biased. This method of evaluation is more sensitive to irritation on samples than previous methods. IA can reduce errors during the evaluation because it is semiautomated. Central data interpretation of studies conducted at multiple investigator sites can also be achieved through the use of digital image analysis. The computerized nature of this method allows for mathematical analysis of results and easy data storage.

REFERENCES

(1) O. H. Mills and A. M. Kligman, A human model for assessing comedogenic substances, *Arch. Dermatol.*, **118**, 903–905 (1982).
(2) O. H. Mills and A. M. Kligman, Follicular biopsy, *Dermatologica*, **167**, 57–63 (1980).
(3) O. H. Mills *et al.*, "Assessing Comedogenicity: Current and Future Trends," in *Clinical Safety and Efficacy Testing of Cosmetics*, W. C. Wagginer, Ed. (Marcel Dekker, New York, 1990), pp. 83–91.
(4) A. M. Kligman and T. Kuong, An improved rabbit ear model for assessing comedogenic substances, *Br. J. Dermatol.*, **100**, 699–702 (1979).
(5) A. Awajan, D. Rondot, and J. Mignot, Quick method of measuring the furrows distribution on skin surface replicas, *Med. Biol. Eng. Comput.*, 379–389 (July 1989).
(6) G. L. Grove, M. J. Grove, and J. J. Reydon, Optical profilemetry: An objective method for quantification of facial wrinkles, *J. Am. Acad. Dermatol.*, **213**, 631–637 (1989).
(7) R. Marks, Methods for the assessment of the effects of topical retinoic acid in photo-aging and actinic keratosis, *J. Int. Med. Res.*, **18**, 29C–34C (1990).
(8) M. Monto and R. Caputo, Chemical efficacy and patient tolerance of topical trentinoin therapy in photo-aging, *J. Int. Med. Res.*, **18**, 35C–40C (1990).
(9) J. Bhawan *et al.*, Effects of trentinoin on photodamaged skin, *Arch. Dermatol.*, **127**, 666–672 (1991).
(10) G. L. Grove *et al.*, Skin replica analysis of photodamaged skin after therapy with trentinoin emollient cream, *J. Am. Acad. Dermatol.*, **25**, 231–237 (1991).
(11) T. J. Flotto *et al.*, A computerized image analysis method for measuring elastic tissue, *J. Invest. Dermatol.*, **93**, 358–362 (1989).

(12) D. T. Woodley *et al.*, Treatment of photoaged skin with topical tretinoin increases epidermal-dermal anchoring fibrils, *J.A.M.A.*, **263**, 3057–3059 (1990).
(13) C. R. Roquet and M. Kermici, A method for measuring the various constituents of the human hair follicle, *J. Microsc.*, **156**, 115–123 (1981).
(14) V. A. Moss, D. M. Jenkinson, and H. Y. Elde, Automated image segmentation and serial section reconstruction in microscopy, *J. Microsc.*, **158**, 187–196 (1990).
(15) G. Sauermann, B. Ebens, and U. Hoppe, Analysis of facial comedones by porphyrin fluorescence and image analysis, *J. Toxicol.–Cut. Ocular Toxicol.*, **8**, 369–385 (1989/1990).
(16) R. R. Anderson, Polarized light examination and photography of the skin, *Arch. Dermatol.*, 127 (July 1991).
(17) D. Halliday and R. Resnick, *Physics* (John Wiley & Sons, New York, 1978), pp. 1069–1073.