*Letter to the Editor*

# Evaluation of the Proposed FDA Pilot Dose-response Methodology for Topical Corticosteroid Bioequivalence Testing

The Food and Drug Administration's Office of Generic Drugs issued a guidance in 1995 for documentation of *in vivo* bioequivalence of topical dermatologic corticosteroids (1). The guidance recommends a pilot dose-response study and a pivotal bioequivalence study to be conducted. The pilot study is carried out to explore the dose-response relationship of reference listed drug (RLD) in question. The pivotal bioequivalence study is performed to determine bioequivalence of a multisource dermatologic corticosteroid. The dose at which the generic product is compared with the RLD is approximately equal to the RLD $ED_{50}$ as determined in the pilot study. The combined pilot-pivotal study design was based on the methodology endorsed by the September 12–13, 1994, meeting of the Generic Drugs Advisory Committee with representation of Dermatologic Drugs Advisory Committee. The guidance was issued after consultation with experts from outside the Agency.

An evaluation of the proposed pilot study recently appeared in *Pharmaceutical Research* (2). We are delighted to note the interest shown by scientists at Rhodes University School of Pharmaceutical Sciences (South Africa) and acknowledge their enthusiasm in evaluating it. However, we believe that the above publication did not provide proper evaluation of the pilot study recommended in the Agency guidance.

The main objective of the pilot study is to determine the population $ED_{50}$ of RLD's. The Agency Guidance recommended the following two approaches for data analysis to determine $ED_{50}$: (i) nonlinear mixed effect modeling which accounts for both intra-and inter-subject variability, and (ii) the "naive pool" method where observations from various subjects are pooled before analysis. Although Demana et. al. did not state which of the above two methods was used to determine population $ED_{50}$, their use of the computer software PCNONLIN (V 4.2, SCI Software, Lexington, KY) indicates that the data analysis did not employ mixed effect modeling. PCNONLIN is a robust program for analyses of individual subjects data. However it does not perform mixed effect modeling. It can only provide approximate estimates of population parameters based on the "naive pool" method or by fitting model to the population mean data. In either case it does not account for both intra- and inter-subject variability, nor does it take into consideration the nature of distribution (normal or lognormal) of population parameters. Determination of population parameters should preferably be based on mixed effect modeling. Based on our experience with analyses of pilot study data for dermatologic corticosteroids, mixed effect modeling and "naive pool" analysis provide distinctly different $ED_{50}$ values. Differences in $ED_{50}$ values, based on these two methods of modeling, are observed for both chromameter and visual assessment of vasoconstriction, regardless of corticosteroid potency. Model predictions based on the "naive pool" analysis poorly correlate with the observed data, and population parameters do not represent the study population. However, predictions based on mixed effect modeling represent the study population and provide acceptable posterior Bayesian estimates for all subjects' data. Therefore, an evaluation of the proposed pilot dose-response study should include comparison of results based on both methods of modeling listed in the guidance.

The Agency guidance not only provided an outline of methodology for dermatologic corticosteroids, but it also included an example of analyses of the pilot and pivotal study data. An evaluation of this methodology should be based on data analysis performed in the manner described in the guidance. Any deviation from the proposed analysis may provide different results, and influence the outcome of evaluation. The Agency guidance recommends measurement of baseline values for each designated treated and untreated skin site because, based on chromameter assessment of skin blanching, there is notable inter-site variation in baseline values in the same individual (3). Changes in chromameter values due to corticosteroid skin blanching represent only a small fraction of the baseline values. Demana et. al. did not correct data for baseline, as recommended in the guidance. The implications of not correcting data for baseline value may be two fold: (i) a portion of the pharmacodynamic metric (area under the effect curve, AUEC) may not be related to the drug effect, since changes in chromameter values as a result of skin blanching represent only a small fraction of the respective baseline values, and (ii) AUEC for vasoconstrictor response in the absence of drug is not equal to zero, if the data are not corrected for the baseline. Analysis of such data using the simple $E_{max}$ model [$E = E_{max} * Dose/(Dose + ED_{50})$] is not appropriate because, based on this model, the value of pharmacodynamic effect in the absence of drug reduces to zero. $E_{max}$ models suitable for data not corrected for baseline are described by the equations, $E = E_0 - (E_{max} * Dose/(Dose + ED_{50}))$ or $E = E_0 + (E_{max} * Dose/(Dose + ED_{50}))$, where $E_0$ is the baseline value predicted by the model and it is estimated as an additional parameter (4). Demana et. al. did not use the latter model even though the AUEC data were not corrected for baseline values for each treated spot. Based on our experience the values of $ED_{50}$ for data not corrected for baseline may vary considerably depending upon which of the above models were used for pharmacodynamic analysis. It is important to use the appropriate model depending upon the nature of data correction.

In their analysis of vasoconstrictor data, these authors considered a sigmoidal $E_{max}$ model [$E = E_{max} * Dose^{\gamma}/(Dose^{\gamma} + ED_{50}^{\gamma})$, where $\gamma$ is the sigmoidicity constant that influences slope in the region of $ED_{50}$] to be more appropriate for the analysis of visual scores data. Sigmoidicity in the observed dose response may partly be due to the limited ability of human eye to detect subtle changes in skin blanching induced by short dose durations. In fitting the sigmoidal model the authors reported gamma values of 7.23 and 7.19. These values of gamma are indicative of very steep dose response which may not be consistent with the action of medium potency corticosteroids. In addition, such high values of gamma are rare. A survey of

0724-8741/98/0100-0004$15.00/0 

literature indicates that the average value of the sigmoidicity constant may be approximately 2 (5). Demana et. al.'s paper did not contain any figures showing observed and fitted data. Therefore, it is difficult to comment on validity of the reported high values of gamma. Furthermore, since the modeling did not use a population model representing intra- and inter-individual variation, parameter values may not represent the study population.

Demana et. al. also asserted superiority of visual assessment of skin blanching over chromameter even though visual scoring has been considered to be highly subjective by other investigators (3, 6–8). However, suitability of visual scores data for product evaluation is limited not only by its subjectivity, but also due to the inherent problems with pharmacodynamic modeling and statistical analysis of nonparametric data. Unlike the chromameter data, visual scores are not continuous *i.e.*, a score of 2 may not be 1 + 1. Furthermore, visual scoring may not allow determination of precision of the method as recommended in the guidance. On the other hand, chromameter assessment allows determination of reproducibility of method by measuring precision of multiple intra- and inter-site readings. Nonetheless, the Agency guidance did not exclude documentation of bioequivalence based on visual assessment of skin blanching, if a correlation between the visual and chromameter assessments is demonstrated.

Consistent with Good Guidance Practices the Agency periodically assesses guidances and updates them as required. Thus, the June 2, 1995, guidance for dermatologic corticosteroids was a final version of the interim guidance issued previously (9), based on an assessment of the interim guidance after its release. Similarly, the Agency has now conducted an evaluation of the June 2, 1995, guidance based on several pilot and pivotal studies submitted by the industry. These studies were conducted on corticosteroid preparations over a wide range of potencies. Results of these retrospective analyses will be presented at the upcoming annual meeting of American Association of Pharmaceutical Scientists to be held November 2–6, 1997, in Boston, Massachusetts.

Gur Jai Pal Singh,
Nicholas Fleischer,
Lawrence Lesko
and
Roger Williams
*Office of Pharmaceutical Science,*
*CDER, US FDA,*
*Rockville, Maryland*

## REFERENCES

1. Guidance: Topical dermatologic corticosteroids *in vivo* bioequivalence. Division of Bioequivalence, Office of Generic Drugs, Food and Drug Administration, Rockville, MD. June 2, 1995.
2. P. H. Demana, E. W. Smith, R. B. Walker, J. M. Haigh, and I. Kanfer. *Pharm. Res.* **14**:303–307 (1997).
3. L. K. Pershing. *Drug Info. J.* **29**:923–34 (1995).
4. N. H. G. Holford. *Clin. Pharmacokinet.* **6**:429–53 (1981).
5. S. Dutta, Y. Matsumoto, and W. F. Ebling. *J Pharm Sci.* **85**:232–239 (1996).
6. V. P. Shah. C. C. Peck, and J. Skelly. *Arch Dermatol.* **125**:1558–1561 (1989).
7. L. K. Pershing, L. D. Lambert, V. P. Shah, and S. Y. Lam. *Int. J. Pharm Sci.* **86**:201–210 (1992).
8. J. Ashworth. *Acta Derm. Venereol.* **151**(Suppl.):20–25 (1989).
9. Interim guidance for topical corticosteroids: *In vivo* bioequivalence and *in vitro* release methods. Office of Generic Drugs, Food and Drug Administration, Rockville, MD. July 1, 1992.

### The authors reply:

We appreciate the opportunity given to us to respond to the comments of Singh et al. presented in their letter to The Editor regarding our evaluation (1) of the pilot study methodology recommended by the Food and Drug Administration (FDA) Guidance. Policies adopted by the FDA tend to have ramifications, especially regulatory implications, in countries other than the U.S.A. Our evaluation (1) was stimulated, in part, by our concerns based upon the long association and experience (2,3) this laboratory has had with the development and application of the visually-assessed human skin blanching assay in basic research and contractual bioequivalence evaluations. We were privileged to be invited by the FDA to comment on the draft version of the Guidance and, hence, we deemed it appropriate to maintain our research interest in the implementation of the adopted methodology. We fully understand and wholeheartedly support the need for an objectively-assessed protocol to replace the subjective visual grading system that we have documented at length. However, we feel that any instrumental methodology that is adopted should, at least, be as accurate, precise and robust as visual assessment at monitoring the induced vasoconstriction phenomenon.

While the data modeling is an essential component of the Guidance methodology, the overall intention of our discussion was to address the numerous variables that produce so many permutations of the methodology that no two laboratories could be expected to obtain the same results. This is especially apparent when examining the different software programs available for data modelling. In their submitted letter, Singh et al. concede that "distinctly different" $ED_{50}$ values are obtained when using the nonlinear mixed effect and the naïve pool methods for data analysis. This is inconsistent with the Guidance recommendation in that the mixed effect model is advocated in preference to the naïve pool analysis.

The main intention of our study was to compare the visual blanching assessment with data obtained by using a chromameter in accordance with the FDA Guidance. In this respect, we draw attention to the relationships between effect and dose duration as depicted in the figures of our evaluation paper and the summary diagrams presented here. These plots clearly indicate that the relationships obtained by visual assessment. (Figure 1) show the typical pattern expected for dose-response

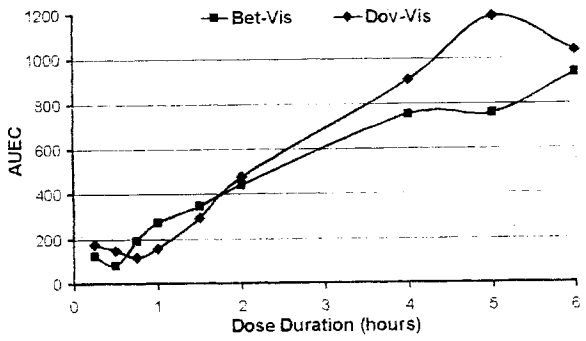


**Fig. 1.** Visual response.

curves. The relationships based upon measurements by the chromameter (Figure 2) are, in contrast, less apparent. Consequently, visual blanching can readily be seen to yield much more precise and reproducible results than those from the chromameter. This was the essential statement of our paper.

The comments of Singh et al. do not address these important considerations at all. On the other hand, their letter is largely preoccupied with various aspects of modelling and computational methodologies, which we address as follows:

1) Singh *et al.* allude to the fact that fitting corticosteroid data to a sigmoid model is inappropriate. It is reasonable to apply the sigmoid model relating drug effect (E) to dose (D) to the clearer, visually-assessed blanching curves, rather than the simple model recommended by the Guidance. The steepness of the curve is not accommodated by the simple model that implies an almost 100-fold dose range between responses of 10% to 90% of the maximum. There is also a noticeable initial sigmoidicity in these curves which the more simple model is not able to accommodate.

2) Since the objective of estimating an $ED_{50}$ is to establish the median of the steepest part of the dose-response curve, with sufficiently "clear" data almost any reasonable modelling procedure can be expected to yield similar $ED_{50}$ results. It should make little difference whether the PCNONLIN or NONMEM procedures are used, or whether the naïve pool or mixed-effect model is applied. In fact, the estimated $ED_{50}$ of 3.2 hours is very plausible when one assesses the AUEC *versus* Dose Duration curves for this data. In the analysis of "clean" data sets, as obtained from the visual assessment procedure, the choice of computational method has little influence on the final result. In contrast, when the data are "noisy" resulting in a less obvious trend, such as the chromameter data obtained in our evaluation, different estimation procedures could yield markedly different results and associated interpretations.

3) The estimated gamma value of 7.2 is quite realistic. A gamma value of 2 discussed by Singh et al. would require an almost 10-fold dose range for evoking responses between 10% and 90% of the maximum. The data do not substantiate this possibility. Furthermore, Singh et al. make reference (4) to published sigmoidicity data. This cited reference concerns mathematical Monte Carlo simulation of theoretical clinical drug concentrations following dosing via routes where blood concentrations can be measured. There is no reference to any named drugs or any implication that these simulations refer to topical drug delivery. The validity of comparing the values from this paper to the topical corticosteroid data that we collected in a clinical trial is, at best, questionable. However, these authors (4) also report highly variable sigmoidicity parameters obtained from their simulations, confirming our experience, and re-iterating the problems associated with attempting to model "noisy" data sets.

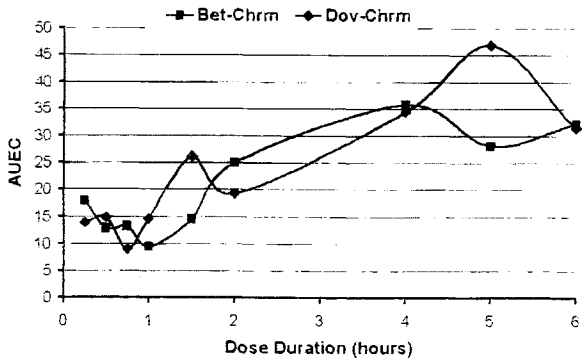


**Fig. 2.** Chromameter response.

4) With regard to the correction of the recorded chromameter data for baseline values, we, and other researchers (5), have not observed the notable inter-site variation in baseline values quoted by Singh et al. Since we both quote the same reference (6) to validate opposite viewpoints, there is a difference in the interpretation of chromameter data by the two groups! Even the FDA sample data quoted in the Guidance (page 26) have a relative standard deviation (RSD) of only 7.2% for the 16 baseline values quoted for one subject. This variance is relatively minor when assessed in relation to the quoted AUEC data for the 12 subjects which at the 0.5 hour dose duration, for example, have a RSD of 250% of the mean. It is interesting to note that the sample AUEC data are depicted in the Guidance as mean value ± the SEM and not ± the SD which would more dramatically portray the variation obtained about the mean data points. Even if this variation in baseline values does in fact exist, then correction for both baseline and unmedicated site values (as the Guidance advocates) is unnecessary (1). The net arithmetic effect of this manipulation is the subtraction of the *uncorrected* unmedicated site value from the *uncorrected* medicated site value. Data documenting this aspect of the handling of chromameter values has been submitted to *Pharmaceutical Research* (7) which shows that there is no significant difference in the response profile regardless of the method of correction (baseline, unmedicated or baseline and unmedicated) of the medicated values. This corroborates the data handling procedures suggested by other researchers (5). Hence, since all the data that we modelled for the publication were corrected for unmedicated site values, the majority of the concerns of Singh et al. relating to baseline correction are unfounded.

5) When the terminal regions of the sigmoid curve cannot be measured precisely or the extent of the asymptote is not tested experimentally, ED50 estimations are unreliable (4,8). In this assay the chromameter is unable to distinguish significantly between the weak skin blanching induced by very short formulation-skin contact times, as seen by the greater imprecision of those data sets. In addition, dose-duration times greater than 6 hours were not advocated in the Guidance which further prevents full characterisation of the upper portion of the sigmoid.

6) The guidance demands inclusion of the data from all the subjects of the pilot trial in the modelling procedure. If a subject responds poorly (or negligibly) in the pilot study then the results from this and similar subjects will skew the modelling results obtained (8). "Inappropriate" subject data are excluded from the pivotal study data pool (by reference to the results from the pilot study) but no assessment of the quality of the pilot study data is made. Moreover, it has been suggested (8) that it is not always necessary to determine the $ED_{50}$ of the dose-response relationship in a pilot study in order to mount a pivotal assessment.

Thus, in addition to emphatically reinforcing our results, we have reviewed the comments of Singh et al. on various technical details and clarified their questions and concerns. In

summary, therefore, the corticosteroid-induced skin blanching response was determined far more precisely and reproducibly by visual assessments than by chromameter measurements.

E. W. Smith, R. B. Walker, J. M. Haigh, and I. Kanfer
*School of Pharmaceutical Sciences*
*Rhodes University*
*Grahamstown 6140*
*South Africa*

## REFERENCES

1. P. H. Demana, E. W. Smith, R. B. Walker, J. M. Haigh, and I. Kanfer. Evaluation of the proposed FDA pilot dose-response methodology for topical corticosteroid bioequivalence testing. *Pharm. Res.*, **14**:303–307 (1997).
2. G. L. Coleman, I. Kanfer, and J. M. Haigh. Comparative blanching activities of proprietary diflucortolone valerate topical preparations. *Dermatologica,* **156**:224–230 (1978).
3. J. M. Haigh, E. Meyer, E. W. Smith, and I. Kanfer. The human skin blanching assay for in vivo topical corticosteroid assessment. I. Reproducibility of the assay. *Int. J. Pharm.* **152**:179–184 (1997).
4. S. Dutta, Y. Matsumoto, and W. F. Ebling. Is it possible to estimate the parameters of the sigmoid $E_{MAX}$ model with truncated data typical of clinical studies? *J. Pharm. Sci.*, **85**:232–239 (1996).
5. M. J. Waring, L. Monger, D. A. Hollingsbee, G. P. Martin, and C. Marriott. Assessment of corticosteroid-induced skin blanching—evaluation of the Minolta chromameter CR200. *Int. J. Pharm.*, **94**:211–222 (1993).
6. L. K. Pershing. Assessment of topical corticosteroid-induced skin blanching response using the visual McKenzie-Stoughton and Calorimetric methods. *Drug. Inf. J.*, **29**:923–934 (1995).
7. E. W. Smith, R. B. Walker, and J. M. Haigh. Analysis of chromameter results obtained from corticosteroid-induced skin blanching. I. Manipulation of data. *Pharm. Res.*, (in press).
8. L. Endrenyi, M. Gao, and Y. Hao. Exploratory and population pharmacodynamic analysis of pilot dose-response data in the bioequivalence testing of topical corticosteroid preparations. *Drug Inf. J.*, **31**: (in press).