Effects of Topical Retinoid Therapy on Acne Lesions: A Psychometric Assessment

James J. Leyden, MD; Norman Preston, PhD; Lori A. Johnson, PhD; Ronald Gottschalk, MD, FRCPC

Topical retinoids are believed to increase inflammatory lesions within the first few weeks of treatment. We evaluated data from several clinical trials for evidence of a signal for retinoid aggravation of inflammatory lesions using a psychometric method and the proportion of participants who demonstrated varying degrees of increased lesion counts. We first determined the validity of a psychometric method based on Stevens' power law called the visual logarithmic scale (VLS) used to evaluate the perceived changes in inflammatory lesions. There was concurrence between the VLS model and the dermatologists' visual assessment of a flare in 80.0% (32/40) of participants (P =.0258). A subsequent analysis was performed using data from clinical trials to assess the occurrence of flares using the VLS model or percentage-based definitions (5%, 10%, or 20% increase) following the first week of treatment with various adapalene gel formulations. In this analysis, no evidence of worsening or a flare was seen by either the VLS model or percentage-based definitions.

The VLS model is valid for assessing the changes in acne severity. Topical retinoid treatment was not associated with a flare as measured by either the VLS model or the proportion of participants who showed an increase in inflammatory lesions.

Cutis. 2012;90:46-50.

Topical retinoids are a mainstay therapy for patients with acne vulgaris based on their effectiveness in reducing both comedones and inflammatory lesions.[1] However, it is common belief that the number of inflammatory lesions may increase within the first few weeks following the initiation of therapy and the perceived worsening of lesions commonly is referred to as a flare in clinical practice.[2] Reports of flaring date back to the late 1960s[3]; therefore, many clinicians have had concerns about using topical retinoids in the early stages of treating patients with inflammatory lesions. However, the available evidence that characterizes flares is relatively limited and findings have questioned if this phenomenon does occur.[4] Clinical trial data represent average changes across the entire study population in inflammatory lesions with topical retinoids at early time points; however, these data do not reflect the changes that occur in study participants. Therefore, clinical trial data do not exclude the potential for increased inflammatory lesions in individual participants. Furthermore, percentage-based definitions for a flare (ie, 10% or 20% increase in inflammatory lesions) may not be an optimal approach in clinical practice, particularly when there are a limited number of lesions involved. For example, a patient with 20 lesions prior to treatment would be designated as having experienced a flare (20% increase) with the addition of only 4 more lesions following the first week of treatment. However, this patient likely would not be designated as having experienced a flare based on visual assessment.

Dr. Leyden is from the University of Pennsylvania Medical School, Philadelphia. Drs. Preston, Johnson, and Gottschalk are from Galderma Laboratories, LP, Fort Worth, Texas.
Funding for these studies and for manuscript preparation was provided by Galderma Laboratories, LP. Dr. Leyden is an advisory board member and consultant for Allergan, Inc; Galderma Laboratories, LP; and Medicis Pharmaceutical Corporation. Drs. Preston, Johnson, and Gottschalk are employees of Galderma Laboratories, LP.
Correspondence: Ronald Gottschalk, MD, FRCPC, Medical Affairs, Galderma Laboratories, LP, 14501 North Freeway, Fort Worth, TX 76177 (Ron.Gottschalk@Galderma.com).

For our study, a model called the visual logarithmic scale (VLS) was used to assess perceived changes in acne to determine if topical retinoid therapy aggravated the erythema of acne either by intensifying erythema of existing lesions or inducing formation of new inflammatory lesions. The VLS model was based on the Weber-Fechner law and Stevens' power law, which state that the perception of a stimulus is proportional to the logarithm of the stimulus intensity.[5,6] The purpose of our study was to test the hypothesis that the erythema associated with inflammatory lesions is noticeable and that it follows the VLS model. Subsequently, analyses were performed using data from clinical trials with various formulations of a topical retinoid (adapalene) alone and in fixed combination with benzoyl peroxide to determine the proportion of patients who exhibited evidence of increased inflammation according to the VLS model. In addition, as has been done in prior studies, clinical trial data were analyzed for the proportion of participants who showed varying degrees of increased inflammatory lesions within the first 2 weeks of treatment.

Methods

Stevens' Power Law and the VLS Model—Stevens' power law states that the $\log_{10}$ of a response is equal to $a \times \log_{10}(S-S_o) + \log_{10}(k)$ where a and k are constants derived experimentally, S is the magnitude of the postbaseline stimulus, and S_o is the magnitude of the stimulus at baseline.

Applying Stevens' power law to the assessment of changes in inflammatory lesions and erythema due to the lesions only, the VLS model of $1.7 \times \log_{10}$ (difference in the number of inflammatory lesions from baseline) $+ \log_{10}(0.05)$ was used to determine the occurrence of a flare.[7] The constants 1.7 and 0.05 have been determined by prior experimentation and perception of redness conduced by Stevens.[7] Based on the VLS model, a flare occurs if the results of the model provide a value equal to or greater than 1, whereas values less than 1 indicate no flare. To generate a result equal to or greater than 1 with the model, the number of lesions would need to increase from baseline by at least 20 to 25.

Validation of the VLS Model—To determine if the VLS model is suitable for assessing the occurrence of flares at the outset of topical retinoid therapy, a panel of 6 dermatologists considered to be acne experts were convened to evaluate photographs of 40 participants in a vehicle-controlled clinical trial with adapalene 0.1%–benzoyl peroxide 2.5% gel. There were 2 photographs per participant (80 photographs in total); one of the photographs was obtained prior to treatment and the other was obtained following the initiation of treatment (with either active treatment or vehicle). The testing for visual assessment of flare was conducted in 2 separate phases and each phase involved a brief session to acclimate the dermatologists to the training tools and presentation style. In phase 1, the panel of dermatologists viewed the 80 photographs in a random order. For each photograph viewed, the dermatologists were asked to assign a number from 1 (clear) to 9 (very severe) that represented the severity of the acne lesions. For phase 2 of testing for visual assessment of flare, following a short interval of 15 minutes, the same dermatologists were shown the same 80 photographs, but the photographs for each participant were displayed together. Therefore, 40 sets of photographs (1 photograph taken prior to treatment and the other photograph taken following the initiation of treatment) were presented together. The 40 sets of photographs were presented in a random order. The dermatologists were asked if the 2 photographs (before and after initiating therapy) indicated that the participant experienced a flare with a categorical yes or no response. Lesion counting was not permitted during the testing for visual assessment of flare.

The dermatologists' responses were collected using an electronic data capture system. Participants were considered to have experienced a flare if any of the 6 dermatologists rated the change in lesions as a yes for flare. Complete consensus among the dermatologists was needed to categorize a change in lesions as a no designation for a flare. The designation of a flare during phase 2 was compared with the ratings of severity during phase 1 to confirm the occurrence of a flare.

The next step in the validation process involved applying the VLS model to the lesion counts on the participants' faces in the photographs that were evaluated by the panel of dermatologists (testing for visual assessment of flare). Results obtained using the VLS model were compared to the dermatologists' responses using the Fisher exact test. To assess the bias due to the use of photographs, the Fisher exact test was repeated with the VLS model using the dermatologists' severity ratings of the photographs from phase 1. In addition, 2 participants who experienced an extreme flare and 2 participants who experienced no flare were included as control participants to validate the dermatologists' responses. The control participants were identified by a separate panel of 10 raters, including 2 dermatologists, who consistently and unanimously designated the control participants as having experienced a flare or no flare following 2 separate testing sessions.

Concurrence between the dermatologists' assessment of a flare and the VLS model occurred when

the VLS model produced a value equal to or greater than 1 and at least 1 dermatologist rated the participant's acne as having flared, or when the VLS model produced a value less than 1 and none of the dermatologists rated the participant's acne as having flared.

Application of the VLS Model—A retrospective analysis was performed using the VLS model and the data from phase 2 and phase 3 clinical trials to determine if acne flares occurred following the initiation of the adapalene gel formulations. The formulations that were included in the clinical trials were adapalene 0.1%–benzoyl peroxide 2.5% gel and adapalene gel 0.1% or adapalene gel 0.3%. The analysis compared the proportion of participants who experienced a flare according to various percentage-based definitions (ie, 5%, 10%, 20% increase from baseline in inflammatory lesions) and the proportion of those participants who experienced a flare according to the VLS model. A value equal to or greater than 1 from the VLS model indicated a flare. The proportion of participants designated as having experienced a flare based on the VLS model or on percentage-based definitions was compared between treatment groups (ie, those receiving active treatment vs those receiving control or vehicle). *P* values were generated using the normal approximation for proportions.

Results

Validation of the VLS Model—Of the participants in the photographs, most were men (23/40 [57.5%]) and white (29/40 [72.5%]). Participants with marked erythema or notable lack of erythema also were included in this analysis (14/40 [35.0%]). The 2 control participants who experienced no flare and the 2 control participants who experienced extreme flare were designated accordingly by the panel of dermatologists.

Overall, there was concurrence between the dermatologists' visual assessment and the VLS model in 32 (80.0%) participants in the validation phase. The percentage agreement that was observed suggested that the VLS model accurately predicted both flares and improvements significantly better than random chance ($P=.0258$). Of the 40 sets of photographs, there were 30 instances (75.0%) in which the VLS model provided a value equal to or greater than 1 suggesting a flare and 10 instances (25.0%) in which the model provided a value less than 1 suggesting no flare. Among the 30 instances in which the VLS model suggested a flare, there were 28 instances (93.3%) in which there was concurrence with the dermatologists' assessment of a flare (eg, at least 1 of 6 raters designated that a flare had occurred). For 24 participants, more than 50% of the panel (at least 3 of 6 raters) designated that a flare had occurred, with unanimous agreement in 14 instances.

Among the 10 instances in which the VLS model suggested no flare, there were 4 instances in which there was concurrence with the dermatologists' visual assessment of a flare (ie, unanimous agreement from the panel that a flare did not occur). Among the 6 instances in which the VLS model suggested no flare, but the dermatologists' visual assessment indicated flare, there was unanimous agreement from the panel (all 6 raters) that a flare had occurred in 3 participants. In the other 3 instances, the dermatologists' designation of a flare or no flare was mixed (ie, 1 of 6, 2 of 6, and 3 of 6 raters designated flares in each, respectively).

When we compared the results that were obtained using the VLS model with the dermatologists' responses, we observed that a doubling in the number of inflammatory lesions was needed for dermatologists to designate that a flare had occurred.

Application of the VLS Model—Data from 14 treatment groups were included in this analysis, comprising a total of 1798 participants assigned to treatment with an active adapalene formulation or control (vehicle or benzoyl peroxide). The Table summarizes the range of participants across the 14 treatment groups who were designated as having experienced a flare following 1 week of treatment according to the various percentage-based definitions or the VLS model.

The results from our analysis of the 14 treatment groups indicated that flares did not occur following 1 week of treatment according to the percentage-based definitions (ie, 5%, 10%, 20% increase) or the VLS model ($P>.05$ for all). One observation from our analysis showed that during all of the postbaseline visits and in all of the study groups, the proportion of participants designated as having experienced a flare was higher when the 20% definition for worsening was used versus the VLS model. For example, when the 20% definition for worsening was applied for one of the studies included in our analysis, more than 12% of participants were designated as having experienced a flare compared with only 3% when the VLS model was applied. When the number of inflammatory lesions was charted at baseline and following 1 week of treatment (data not shown), there was no apparent increase in inflammatory lesions.

Comment

In our analyses, we demonstrated that the VLS model, a psychometric analysis based on Stevens' power law, is a valid method for assessing the changes in inflammatory lesions during acne therapy. We also demonstrated that topical retinoid therapy is not associated with an increase in inflammatory lesions following the first week of treatment,

Range of Participants Designated as Having Experienced a Flare Following 1 Week of Treatment in Phase 2 and Phase 3 Clinical Trials With Adapalene (N=1798)

	Participants With Flare		
Method for Defining a Flare	Topical Retinoid[a]	Vehicle	*P* Value[b]
5% increase in inflammatory lesions	7.55%–21.17%	14.81%–21.80%	NS
10% increase in inflammatory lesions	7.55%–17.51%	11.11%–16.54%	NS
20% increase in inflammatory lesions	3.77%–12.31%	7.25%–9.02%	NS
VLS model	1.87%–5.84%	3.70%–7.03%	NS

Abbreviations: NS, not statistically significant; VLS, visual logarithmic scale.

[a]Adapalene gel 0.1%, adapalene gel 0.3%, or adapalene 0.1%–benzoyl peroxide 2.5% gel.

[b]*P* values were associated with the comparison of proportions using average approximations. All *P* values for comparisons between topical retinoid and vehicle were not statistically significant.

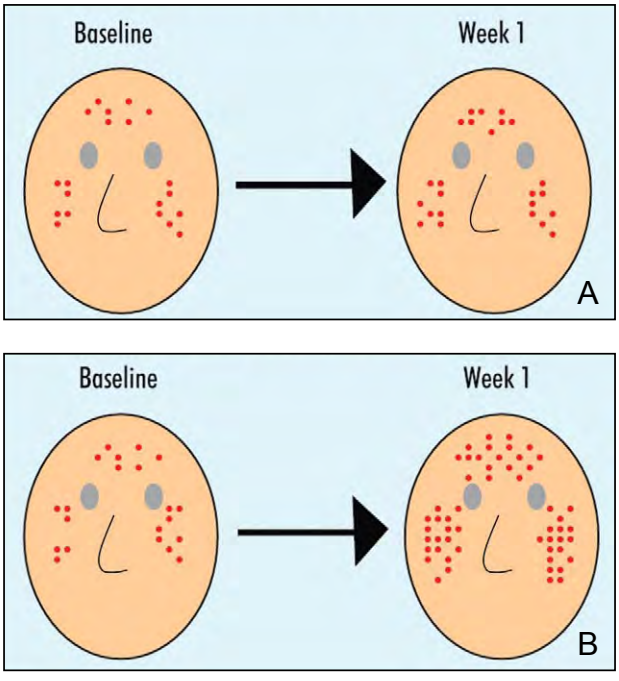
regardless of using the VLS model or percentage-based definitions.

The dermatologists' assessments that designated a lack of a flare during the validation phase were observed, despite stringent criteria that made it more likely that a participant was designated as a yes for a flare (designated as a yes by 1 or more of the dermatologists) versus a no for a flare (complete consensus). Although there generally was concordance between the VLS model and the clinical assessment of changes by a panel of dermatologists (80.0%; *P*=.0258), discrepancies between the 2 methods were observed. Overall, there were 8 (20%) instances in which there was no concurrence between the VLS model and the dermatologists' visual assessment of a flare. Of these instances, 2 participants were designated as not having flared by all 6 dermatologists but were designated as having flared according to the VLS model. Discrepancies between the 2 methods (VLS or percentage based) may be attributable to the lesion characteristics or to differences in photographic or presentation methods. In particular, the presence of large or more inflamed or erythematous clusters of inflammatory lesions appeared to influence the dermatologists' judgment of change, which suggested that these factors should be considered when using percentage-based definitions to assess the changes in inflammatory lesions.

The results from our analyses were consistent with reports on the phenomenon of a flare following the initiation of topical retinoid therapy for acne,[4] which expressed that topical retinoids increased inflammatory lesions during this period. In a study by Leyden and Wortzman,[2] data from a large clinical trial were analyzed for evidence of an increase in inflammatory lesions in participants by using arbitrary thresholds of change of a 10% or 20% increase in the number of inflammatory lesions as indication of possible aggravation of inflammatory lesions. Based on this method, the authors showed no significant differences between topical tretinoin, alone or in combination with clindamycin, and topical clindamycin or vehicle in the proportion of participants who experienced an increase in inflammatory lesions following 2 weeks of treatment.[2] In a literature review by Yentzer et al,[4] data from 17 clinical trials that involved various retinoid compounds were evaluated, and there was no evidence of an increase in total lesion count or inflammatory lesion count within the first or second week of treatment. In 1 study that used percentage-based definitions, the occurrence of a flare was higher in the vehicle-controlled group, though the difference with the active treatment group was not statistically significant.[8]

Overall, potential sources of bias in the validation phase included photographic lighting (eg, shadowing of the affected area), participant positioning in the photographs (eg, making lesions appear larger or smaller), variability of color calibration on the viewing monitors, and image size on the viewing monitor. The data analysis did not control for these potential sources of bias. A limitation of our

analyses showed that the clinical trials were retrospectively performed using prospective clinical trial data with only 1 topical retinoid agent (adapalene). It is unknown if acne flares would occur with other formulations because of excipients or other factors. Additionally, a 20% increase in lesions may not be an appropriate threshold for detecting a clinically relevant acne flare, especially when the number of lesions is limited (Figure). However, in our analyses, when the VLS model was applied to clinical trial data, 20% definition designated a greater number of participants who experienced a flare compared with the results with the VLS model.



Hypothetical example of percentage-based definitions and the visual logarithmic scale (VLS) model compared with visual perception. The first panel (A) depicts a patient with 20 lesions at baseline and 24 lesions at week 1, which would be considered a flare as defined by a 20% increase in lesions. However, without counting, the visual perception of a flare is not apparent in this example. The second panel (B) depicts a patient with 21 lesions at baseline and 54 lesions at week 1, which would be considered a flare as defined by the VLS model (the model produces a result of ≥1). Without counting, the visual perception of a flare is apparent and is agreeable with the VLS model.

Conclusion

The VLS model is a valid approach to assess the changes in acne severity over time and may be preferable to the use of percentage-based definitions, which tend to overestimate the occurrence of a flare. Treatments with various formulations of adapalene were not associated with a flare, regardless of using the VLS model or percentage-based definitions.

Acknowledgments—The authors wish to acknowledge Hilary Baldwin, MD; James Q. Del Rosso, DO; Jerry Tan, MD; Diane Thiboutot, MD; and Guy Webster, MD, PhD (for participating as panelists for the validation phase). The authors also acknowledge the role of Kevin Jarvis, PharmD, from BioCentric, Inc, for his help with preparing the manuscript (drafting an outline and manuscript first draft), compiling author comments, and styling and submitting the manuscript to the journal.

REFERENCES

1. Strauss JS, Krowchuk DP, Leyden JJ, et al. Guidelines of care for acne vulgaris management [published online ahead of print February 5, 2007]. *J Am Acad Dermatol*. 2007;56:651-663.
2. Leyden JJ, Wortzman M. A novel gel formulation of clindamycin phosphate-tretinoin is not associated with acne flaring. *Cutis*. 2008;82:151-156.
3. Kligman AM, Fulton JE Jr, Plewig G. Topical vitamin A acid in acne vulgaris. *Arch Dermatol*. 1969;99:469-476.
4. Yentzer BA, McClain RW, Feldman SR. Do topical retinoids cause acne to "flare"? *J Drugs Dermatol*. 2009;8:799-801.
5. Laming D. *The Measurement of Sensation*. Oxford, England: Oxford University Press; 1997.
6. Stevens SS. On the psychophysical law. *Psychol Rev*. 1957;64:153-181.
7. Atkinson R. Herrnstein R, Lindzey G, et al. *Stevens' Handbook of Experimental Psychology*. New York, NY: Wiley; 1988.
8. Schlessinger J, Menter A, Gold M, et al. Clinical safety and efficacy studies of a novel formulation combining 1.2% clindamycin phosphate and 0.025% tretinoin for the treatment of acne vulgaris. *J Drugs Dermatol*. 2007;6:607-615.