

## ORIGINAL ARTICLE

# Establishment of grading criteria for acne severity

Nobukazu HAYASHI,<sup>1</sup> Hirohiko AKAMATSU,<sup>2</sup> Makoto KAWASHIMA,<sup>1</sup>  
Acne Study Group\*

<sup>1</sup>Department of Dermatology, Tokyo Women's Medical University, Tokyo, and <sup>2</sup>Department of Dermatology, Fujita Health University School of Medicine, Toyoake, Japan

## ABSTRACT

For the epidemiological surveys and evaluations of therapy, it is essential to evaluate the severity of diseases. There are several reported methods of assessment for acne severity including lesion counting, comparison of the patient's to a photographic standard and comparison of the patient's to a text description. But all of these are based on opinions of specialists. In this study, we attempted to make an evidence-based grading criteria for acne severity, which was expected to yield consents from most dermatologists. The dermatologists consulted classified the global severity of acne patients without any standard and then counted the numbers of eruptions. Three independent expert dermatologists graded the photographs of these patients. We compared the verdicts of the consulted dermatologist and three experienced dermatologists, and analyzed the relationships between these classifications and numbers of eruptions. Our results showed that most of the dermatologists have similar latent recognitions of acne severity. We selected representative photographs as standards, which would contribute to making adjustments for judgments. Global classifications of dermatologists correlated with numbers of inflammatory eruptions (papules plus pustules), but did not with numbers of comedones. The appropriate divisions of inflammatory eruptions of half of the face to decide classifications were: 0–5, “mild”; 6–20, “moderate”; 21–50, “severe”; and more than 50, “very severe”.

**Key words:** acne, classification, counting, grading, photographs, severity.

## INTRODUCTION

Identification of disease severity is very important to characterize and classify patients for epidemiological surveys and evaluation of therapy. There are various methods to assess acne severity, including lesion counting on all or part of the face, comparison of the patient's to a photographic standard and comparison of the patient's to a text description.<sup>1</sup> Many of them use the terminology, “mild”, “moderate” or “severe.” Others use numerous scores.<sup>1</sup> However, these methods have been established independently and there is no relationship between these terminologies and scores. These classifications are usually based

on expert opinions, and no evaluation study has been conducted to confirm these classifications. Recently, evidence-based medicine is widely accepted and many guidelines are published. To establish guidelines, evidence-based criteria to classify acne severity is required. The purpose of this study was to establish scientifically evidence-based classification that is easy for clinical dermatologists to accept and use.

## METHODS

### Subjects

Two hundred and forty-four Japanese acne patients who visited the dermatologists listed in Table 1

---

Correspondence: Nobukazu Hayashi, M.D., Ph.D., Department of Dermatology, Tokyo Women's Medical University, 8-1 Kawada-cho, Shinjuku-ku, Tokyo 162-8666, Japan. Email: hayashi@derm.twmu.ac.jp

\*Acne Study Group: M. Ito, Niigata University; M. Otsuki, Jichi Medical University; M. Kawashima and N. Hayashi, Tokyo Women's Medical University; R. Tsuboi, Tokyo Medical University; H. Nakagawa, Tokyo Jikei-kai Medical University; S. Watanabe, Teikyo Medical University; K. Matsunaga and H. Akamatsu, Fujita Health University; Y. Miyachi, Kyoto University; F. Furukawa, Wakayama Medical University; K. Iwatsuki, Okayama University; Y. Kubota, Kagawa University; Y. Tokura, University of Occupational and Environmental Health; and M. Furue, Kyushu University.

Received 10 July 2007; accepted 18 January 2008.

**Table 1.** Members of Acne Study Group and three independent expert dermatologists

Members of Acne Study Group		
Institute	Core member	Collaborators
Niigata University	M. Ito	R. Watanabe, H. Shimura, S. Sawada
Jichi Medical University	M. Otsuki	S. Murata, S. Toda, H. Abe, M. Fujimoto
Tokyo Women's Medical University	M. Kawashima, N. Hayashi	T. Kusunose, M.I. Maroto, E. Kan, C. Narita, K. Yanagisawa
Tokyo Medical University	R. Tsuboi	E. Nakazaki, T. Miyakura, Y. Okubo
Tokyo Jikei-kai Medical University	H. Nakagawa	Y. Nobeyama, N. Tsukahara
Teikyo University	S. Watanabe	H. Mori
Fujita Health University	K. Matsunaga, H. Akamatsu	Y. Washimi, M. Nakagawa
Kyoto University	Y. Miyachi	M. Nakamura
Wakayama Medical University	F. Furukawa	A. Kishioka, Y. Yamamoto
Okayama University	K. Iwatsuki	T. Oono
Kagawa University	Y. Kubota	Y. Matsuoka, J. Katsuura, I. Takai, M. Masada, C. Sadahira, T. Morikami
University of Occupational and Environmental Health	Y. Tokura	M. Kobayashi, T. Mori
Kyushu University	M. Furue	M. Toyota, A. Toshitani
Independent expert dermatologists		
Aoyama Beauty Research Clinic	K. Abe	
Kanto Medical Center NTT EC	A. Igarashi	
Aizawa Dermatology Clinic	H. Aizawa	

participated. Informed consent was obtained before the study commenced.

### Comparison of global severity classifications by consulted dermatologists with photograph-based ones by three expert dermatologists

The dermatologists consulted classified acne global severity grade of half of the face of each patient as "mild", "moderate", "severe" or "very severe" without any standard. Three expert dermatologists, who had board certifications issued by the Japanese Dermatological Association, also graded the global severity based on a photograph of half of the face independently without any consultation or additional information. The photographs of half faces were taken at an approximately 70 degree angle from the front of patients and all photographs were taken in a similar way. Grading of consulted dermatologists and expert dermatologists were compared, and correspondence rates were calculated.

### Lesion counting

Consulted dermatologists counted numbers of open and closed comedones, papules, pustules, cysts and nodules on half of each patient's face. We divided the eruptions into three categories: (i) comedones;

(ii) inflammatory eruptions that included papules and pustules; and (iii) severe eruptions that included cysts and nodules. We compared the numbers of these categories with the classification of dermatologists consulted. We drew scatter blots and calculated the coefficients of simple and multiple regressions.

Next, we selected the cases which had been classified identically by at least two of the three dermatologists consulted and named these as the consensus cases. We compared the number of eruptions with the classifications of consensus cases, and decided the best divisions of numbers for classifications.

### Establishing standard photographs

We selected the cases that matched the counting-based classifications established by consulted dermatologists from consensus cases. The representative photographs were selected as standard.

## RESULTS

### Patient distribution and adopted cases

Two hundred and forty-four acne patients, consisting of 43 males and 199 females with average ages of  $21.4 \pm 5.4$  and  $24.1 \pm 5.8$  years, respectively, were

**Table 2.** Difference between classifications by consulted dermatologists and photograph-based ones: complete accordance in classifications by four dermatologists was observed in 176 faces (45.1%). Classifications by consulted dermatologists were identical by at least two of three expert dermatologists on 268 faces (68.7%)

No. of identical judgments	Faces	Percentage
All three	176	45.1%
Two of three	92	23.6%
One of three	58	14.9%
None	64	16.4%

enrolled. The sex of two cases escaped record. Four hundred and eighty-six half faces of these patients were used for this survey. Two half faces were excluded at this point because of lack of grading by consulted dermatologists. Of these faces, 271 were classified as “mild”, 195 as “moderate”, 16 as “severe” and four as “very severe” when classified by the consulted dermatologists. Photographs of 390 half faces were successfully taken and used for photograph-based classifications.

#### Comparison of classifications by consulted dermatologists with photograph-based ones

The summary of classifications by consulted dermatologists and photograph-based ones are listed in Table 2. Complete accordance in classifications by four dermatologists was observed in 176 faces (45.1%). Classifications by consulted dermatologists were identical by at least two of three expert dermatologists in 268 faces (68.7%).

A total of 1170 photograph-based judgments (390 faces) were also compared with the classifications of three consulted dermatologists (Table 3). Seven hundred and seventy decisions (65.8%) showed conformity between consulted dermatologists. One

**Table 3.** Number of photo-based judgments that were the same as one other consulted dermatologist. Comparison of 1170 photograph-based judgments (390 faces) with classification of consulted dermatologists revealed that 65.8% of classifications were identical, and 99.6% of them were acceptable

Photo-based classification	No. of faces	Percentage
One grade lower than CD	87	7.4%
Identical with CD	770	65.8%
One grade higher than CD	308	26.3%
Two grade higher than CD	5	0.4%

CD, classification by consulted dermatologists.

thousand one hundred and sixty-five decisions (99.6%) stayed within one grade of difference in classification of consulted dermatologists.

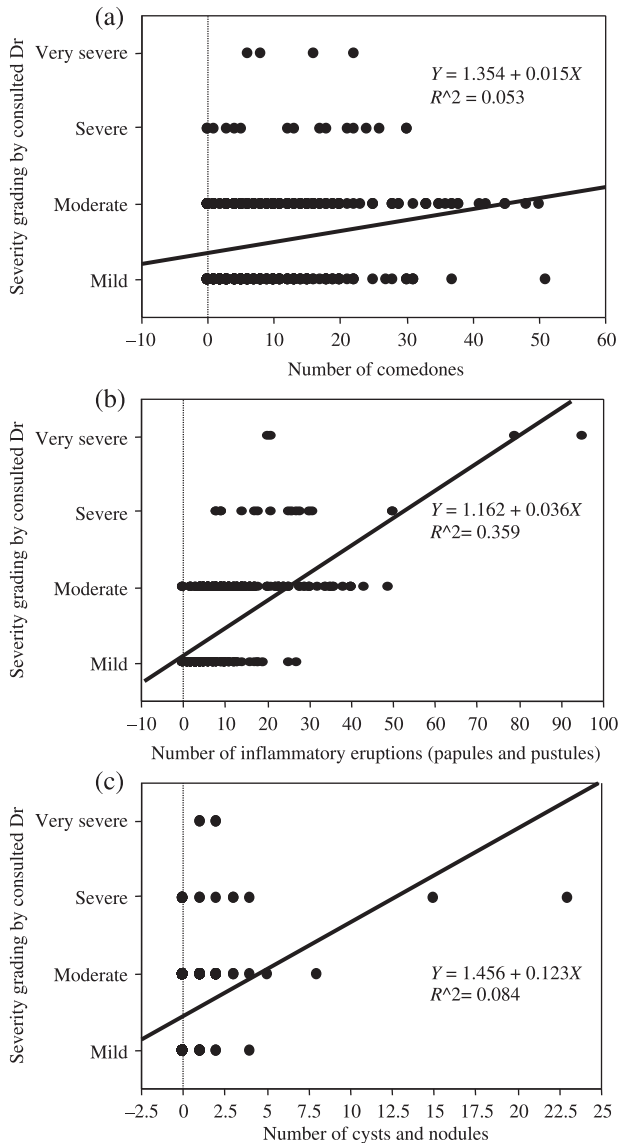
#### Relationship between lesion counting and classification of consulted dermatologists

The results of calculated coefficients of multiple regressions between lesion counting and classification of consulted dermatologists are summarized in Table 4. Scatter blots (Fig. 1) show the relationship between the numbers of inflammatory eruptions and classifications of consulted dermatologists ( $r = 0.68$ ). Comedones (Fig. 1a) did not correlate with global classifications ( $r = 0.28$ ). Cysts and nodules (Fig. 1c) showed high values of multiple regression coefficients (0.11) and seemed to be important, but 95% confidence intervals (CI, 0.082–0.138) were larger than other eruptions, because severe eruptions that consisted of cysts and nodules were rarely observed and not enough information was obtained from our experiments. Therefore, we decided to limit our target to inflammatory eruptions (Fig. 1b), which showed high values (0.03) and narrower 95% CI.

**Table 4.** Coefficients of multiple regressions between lesion counting and classification of consulted dermatologists

Eruption type	Multiple regressions*		
	Coefficients of regressions	95% confidence interval	Significance of probability
Comedones	0.011	0.007–0.015	<0.0001
Papules + pustules	0.034	0.030–0.038	<0.0001
Nodules + cysts	0.110	0.082–0.138	<0.0001
Intercept	1.047	0.982–1.112	<0.0001

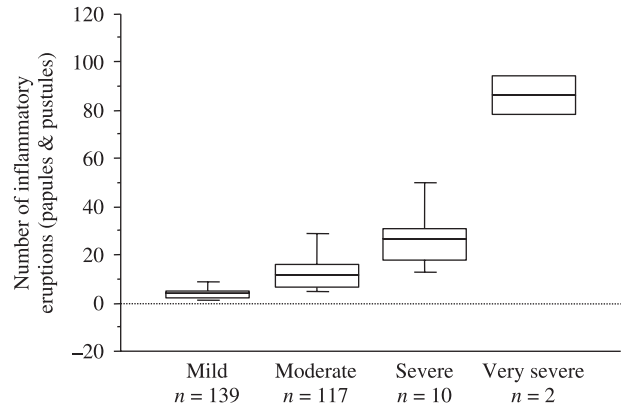
\*Acne severity grade was converted to: 1, “mild”; 2, “moderate”; 3, “severe”; 4, very severe.



**Figure 1.** Scatter blots show the relationship between eruptions and classification of consulted dermatologists. (a) Comedones; (b) inflammatory eruptions; (c) cysts and nodules.

**Classification of acne severity on lesion counting**

We classified the consensus cases with the numbers of inflammatory eruptions, and analyzed their distributions (Fig. 2). We selected several numbers which satisfied 75–90% or 10–25 % of total distributions. For example, distribution of the number of inflammatory eruptions (Fig. 2) suggested that the best division between mild and moderate should exist



**Figure 2.** Marginal blot shows distribution of numbers of inflammatory acne eruptions classified with consensus grade. The end of lines presented 10th to 90th percentiles and interquartile range exhibited 25th and 75th percentiles. Median is the line of interquartile range.

**Table 5.** Classification of acne severity on lesion counting of papules and pustules per half face

Mild	0–5
Moderate	6–20
Severe	21–50
Very severe	>50

between four and six. Then, the conformity rates were calculated to establish the best division. Consequently, the maximum conformity rate was 75.1% and the best divisions were 0–5 for mild, 6–20 for moderate, 21–50 for severe and more than 50 for very severe in the half faces (Table 5).

**Standard photograph for the classification of acne severity**

We selected the representative photographs from the cases, which satisfied our lesion counting-based classifications from consensus cases and decided to use them as standard photographs (Fig. 3).

**DISCUSSION**

Standardization of acne severity classification is important to assess clinical trials, because acne severity is the most important patient characteristic especially in discussing individual treatment profiles. There are more than 25 methods of assessing acne

### Standard photographs for acne severity grading



**Figure 3.** Standard photographs for the classification of acne severity.

severity, but they have differed between studies. Many trials have not reported their method of assessment.<sup>1</sup> Lehmann *et al.* presented similar combined acne severity classification for the convenience of methodological review.<sup>1</sup> They evaluated grading system reporting in various methods, analyzed the mean and standard deviation of each trial reporting lesion count, and then converted these with their standard. However, they declared that their criteria should be used to classify studies into severity groups and was not intended for specific research studies.<sup>1</sup> Cunliffe *et al.* also showed combined criteria, but for the purpose of a specific clinical trial and there was no evidence for the definition of classification.<sup>2</sup> Our experiments showed that global classification by consulted dermatologists without a standard related to the global grading by experts using photograph and counting of inflammatory eruptions. These results support the adequacy of photograph-based global assessment and the conversion of these three methods for assessment.

The presentation of standard photographs is expected to adjust classifications and make them

more accurate. In reported acne grading,<sup>3,4</sup> experts defined severity grading and selected standard photographs. One grading used approximately 1000 photographs to be categorized by several experts, but the process of definition was unclear.<sup>4</sup> In our study, nomination of standard photographs was based on scientific procedure, therefore our standard photographs should be more reliable.

The number of comedones did not correlate with global estimation by dermatologists. The numbers of cysts and nodules showed correlation, but we could not make appropriate divisions because the numbers were small and many cases did not have these severe inflammatory eruptions. In consequence, the application of our criteria is limited to the cases that mainly have papules and pustules. Comedones should be estimated with numbers or other criteria, and cysts or nodules should be treated differently.

From our comparison between global classifications without any standard by consulted dermatologists and expert dermatologists, we believe that most of dermatologists have similar latent recognitions of

acne severity, which crystallized into our criteria. The appropriate divisions of inflammatory eruptions to decide classifications were 0–5 for mild, 6–20 for moderate, 21–50 for severe, and more than 50 for very severe in half of the face. These results are similar to expert opinion<sup>5</sup> and demonstrate the propriety of expert opinions. However, our results are more accurate and would be easy to accept for most dermatologists because they are based on the global estimations of consulted dermatologists without any standard.

Our criteria of acne severity are the only methods based on scientific procedure and should be more reliable and useful for the estimation of acne severity. Global estimation is easy to use and standard photographs would contribute to making adjustments for difference of judgments between the estimators, and expected to make the grading more precise. Counting of the inflammatory eruptions may be useful to analyze precise change, and our criteria enable the counting to convert into global estimation. Our grading is very useful in comparing one experiment to another. We hope our criteria will be some help for the studies of acne.

## ACKNOWLEDGMENTS

We thank Mr Kazuo Matsumoto at Access Laboratory for his statistical advice, and Eisai, Otsuka Pharmaceutical and Sanofi-Aventis for their financial support of this project.

## REFERENCES

- 1 Lehmann HP, Robinson KA, Andrews JS, Holloway V, Goodman SN. Acne therapy: A methodologic review. *J Am Acad Dermatol* 2002; **47**: 231–40.
- 2 Cunliffe WJ, Meynadier J, Alirezai M *et al.* Is combined oral and topical therapy better than oral therapy alone in patients with moderate to moderately severe acne vulgaris? A comparison the efficacy and safety of lymecycline plus adapalene gel 0.1%, versus lymecycline plus gel vehicle. *J Am Acad Dermatol* 2003; **49**: S218–S226.
- 3 Cook CH, Centner RL, Michaels SE. An acne grading method using photographic standards. *Arch Dermatol* 1979; **115**: 571–575.
- 4 O'Brien SC, Lewis JB, Cunliffe WJ. The Leeds revised acne grading system. *J Dermatol Treat* 1998; **9**: 215–220.
- 5 Plewig G, Klingman AM. Classification of acne vulgaris. In: *Acne, Morphogenesis Treatment*, Springer-Verlag, Berlin, 1974; 162–163.