INTRODUCTION

Age-related changes in the eyelids include aponeurotic blepharoptosis and dermatochalasis. Blepharoptosis is diagnosed by decreased levator function, but dermatochalasis simultaneously presents in many cases, and a comprehensive evaluation is needed to improve the cosmetic aspect of age-related eyelid changes. In addition, the difference between the left and right side is also an important factor in cosmetic evaluation.

Cahill et al divided the indications for surgical treatment of blepharoptosis into objective symptoms and subjective symptoms, and, on the basis of a systematic review, cited margin reflex distance 1 (MRD1) in frontal vision and visual field disturbances in upward and downward vision as objective symptoms.1–4 Averbuch-Heller et al described surgical indications on the basis of levator contractile function, in which the distance of the eyelid margin was measured from the bottom to the top of vision.5 All of these indicators are based on the function of one eyelid.

In many patients, the appearance of the eyelids varies from side to side.6–8 As is known from Hering’s law, the direction of eyelid opening to the left and right levator palpebrae is controlled bilaterally.9–12 When laterality is present, the eyelid on the side with no or mild ptosis appears hyper-opened by bilateral signals. In daily practice, both eyelids are observed at the same time, and morphological differences between the left and right eyelids seem to be important information with which to determine the surgical indications.

Background: Although the functional and anatomical differences between the left and right eyelids are important in the evaluation of age-related changes in the eyelids, they have not been described clearly as indications for surgical treatments. This study aimed to investigate how laterality of the eyelids affects evaluation of age-related changes.

Methods: Photographs of either one or both eyelids of 100 people were evaluated in four stages by 10 plastic surgeons. To investigate the consistency of the results between evaluations, surgeons evaluated the single-eyelid photographs (group U) or two-eyelid photographs (group B). It was investigated whether the difference in margin reflex distance 1, height of the upper eyelid crease, height of eyebrow, and levator contractile function were associated with mismatched evaluations.

Results: The weighted kappa coefficient for groups B and U was 0.77 (substantial agreement). One-point difference in scores was observed in 23 cases. In the multiple logistic regression analysis, only the laterality the height of the eyelid crease was significantly different between patients whose evaluations were matched and those whose evaluations were mismatched (0.9 ± 0.1 mm versus 1.7 ± 0.2 mm; OR = 1.06, 95% CI: 1.01–1.10; P = 0.01).

Conclusions: Besides the structure and function of each eyelid, the laterality of the height of the eyelid crease was important in the evaluation of the age-related changes in the eyelids. This factor may be important in evaluating the aesthetic and visual impressions of age-related changes in the eyelids. (Plast Reconstr Surg Glob Open 2021;9:e3909; doi: 10.1097/GOX.0000000000003909; Published online 4 November 2021.)
We hypothesized that the evaluation of age-related changes in the eyelids might be aided by the integration of information about laterality with information about the functional and aesthetic indicators of each eyelid. Although various factors may be involved, we focused on four: the laterality difference in MRD1, the height of the upper eyelid crease, the height of eyebrow, and levator contractile function.

The primary endpoint of this study was how laterality in appearance of the eyelids affected the evaluation of the age-related changes in the eyelids. We investigated the consistency of the evaluation of severity of the age-related changes in the eyelids by experienced plastic surgeons between photographs of one and both eyelids. When the evaluation results were inconsistent, we verified which factors accounted for the inconsistency. Additionally, we compared the results of evaluations by plastic surgeons with limited clinical experience with those by experienced plastic surgeons. We believe that the results of this study will elucidate how well-experienced plastic surgeons recognize the age-related changes of the eyelid and will contribute to the development of theory and standardization of evaluation of the cosmetic aspect of the age-related changes of the eyelid.

### PATIENTS AND METHODS

This study was conducted at the Chiba University Hospital, Chiba, Japan, with approval from the relevant institutional review board and permission from the hospital ethics committee (No. 2910). This study is an observational study without intervention. Written informed consent to participate in this study was obtained from all selected study participants before their enrollment in the study.

Photographs of upward, frontal, and downward views of the eyelids were taken in 25 healthy volunteers and 75 patients with untreated aponeurotic blepharoptosis. Patients with myasthenia gravis, Horner syndrome, or neurologic disease were excluded. For each subject, photographs were taken of the two eyes together in the three views, including the eyebrows and the upper and lower eyelids (Fig. 1).

The photographs were taken by a digital camera with a flash, with a reference 10 mm² marker (CASMATCH; Bear Medic Corporation, Tokyo, Japan), which was used as an index to measure the absolute value of the length. We used Photoshop CS (version 5) software (Adobe Inc., San Jose, Calif.) to measure MRD 1, the height of eyelid crease, eyebrow height, and levator contractile function. The vertical distance from the imaginary line connecting the inner canthus on both sides to the upper edge of the eyebrow on the line through the center of the pupil was defined as the eyebrow height (Fig. 2). Height of the eyelid crease was also measured on the same line. These distances used in this study were the apparent heights in the standing frontal view, and they differed from the length.
in the supine position with eyelids closed. The distance along the upper eyelid margin from the upward to downward visual images was used as an index of levator contractile function.

Validation 1: Concordance between Evaluations of Ptosis in Two- and Single-eyelid Photographs by Plastic Surgeons with Extensive Clinical Experience

A diagram of the study outline is presented in Figure 3. Ten plastic surgeons who obtained specialist qualifications from the Japanese Society of Plastic Surgeons and had treated more than 100 cases of blepharoptosis were asked to evaluate ptosis in eyelid photographs of 100 patients.

Each plastic surgeon scored each eyelid on a four-point scale: 0 (no surgical indication for age-related change), 1 (mild age-related change as an indication for surgery), 2 (moderate age-related change), and 3 (severe age-related change) from the photographs without any patient information, including eye or age dominance. Each plastic surgeon evaluated both sides of the eyelid simultaneously in photographs of both eyelids of 50 patients (Fig. 1A), or...
each side of the left and right eyelid separately in photographs of a single eyelid of another 50 patients (Fig. 1B, C). The single-eyelid photographs of the same patient were evaluated not consecutively but at intervals to ensure that the judgments about one side did not affect those about the other. The surgeons were assigned to evaluate two-and single-eyelid photographs of different patients to reduce bias. Thus, for all 100 patients (200 eyelids), five judgments were obtained for each photograph. Results of evaluations of photographs of both eyelids were categorized as the “bilateral photograph evaluation group” (group B), and results of evaluations of photographs of one eyelid were categorized as the “unilateral photograph evaluation group” (group U). Within each group, the degree of agreement between the raters was determined.

To determine the final evaluation results of each group, the mode score of each group was adopted. If there were two mode scores, the higher score was used because it meant that a majority of five plastic surgeons judged that the eyelid was an indication for surgical treatment. The correlation between final evaluation score and MRD1, the height of the upper eyelid crease, the height of eyebrow, and levator contractile function were investigated in both groups B and U. The degree of agreement between the scores of groups B and U was then determined.

The cases with inconsistent scores between groups B and U were compared with those with consistent scores regarding lateral differences in MRD1, height of the upper eyelid crease, height of eyebrow, and levator contractile function (Fig. 3).

Validation 2: Comparison between Plastic Surgeons with Extensive Clinical Experience and Those with Limited Clinical Experience

With the same set of photographs as in validation 1, evaluations were conducted by 10 plastic surgeons who were certified by the Japan Society of Plastic Surgeons but had experience with fewer than 100 cases of blepharoptosis. The results of these evaluations were compared with those by plastic surgeons with extensive clinical experience. The degree of consistency between the final evaluation scores of groups U and B was compared between plastic surgeons with extensive experience and those with limited experience.

Statistical Analysis

JMP Pro (version 13.0.0) and SAS (version 9.4) (SAS Institute, Cary, N.C.) were used to conduct all statistical analyses. Weighted kappa coefficients were used to calculate the degree of agreement among the groups. To determine the rate of evaluation agreement within each group, weighted GLMM-based ordinal measure kappa coefficients with a confidence interval (CI) of 95% was calculated. The values of kappa coefficient were interpreted according to the criteria defined by Landis and Koch: −1.00, total disagreement; 0.00, no agreement; 0.01–0.20, slight agreement; 0.21–0.40, fair agreement; 0.41–0.60, moderate agreement; 0.61–0.80, substantial agreement; 0.81–0.99, almost perfect agreement; and 1.00, perfect agreement. Spearman’s correlation coefficient was used to determine the strength of the correlation among the four variables between patients with consistent scores in groups B and U and those with inconsistent scores in groups B and U. A multiple logistic regression analysis was performed to identify the variables associated with inconsistent scores, calculating the odds ratio (OR) with 95% CI and P values. The assessed variables were as follows: MRD1, height of eyelid crease, eyebrow height, and levator contractile function. After the univariate logistic regression analysis, statistically significant variables were substituted for the multiple logistic regression analysis.

A P value less than 0.05 was considered statistically significant. The data are presented as means ± SDs.

RESULTS

Basic characteristics of the 100 patients evaluated in this study are listed in Table 1. All the participants in this study were active clinicians working as plastic surgeons in Japan. The average number of cases that had been handled by the plastic surgeons with extensive clinical experience was 475.0 ± 519.3, whereas that handled by those with limited clinical experience was 55.5 ± 24.1.

Validation 1

The mean weighted GLMM-based ordinal measure kappa coefficients of the five evaluations by plastic surgeons with extensive clinical experience were 0.58 in group B and 0.59 in group U. The numbers of eyelids with final evaluation scores of 0, 1, 2, and 3 were 58, 71, 58, and 13, respectively, in group B and 59, 74, 50, and 17, respectively, in group U. The weighted kappa coefficient of the final score for groups B and U was 0.77, suggesting that the results of the evaluation based on bilateral and unilateral eyelid photographs were consistent, to some extent, when evaluated by plastic surgeons with extensive experience (Table 2). Nevertheless, a one-point deviation was observed in 26 (13%) of the
and 23 cases with inconsistent scores between groups B and U. The result of comparisons between cases with consistent scores and those with inconsistent scores is shown in Table 4. Between cases with consistent scores and those with inconsistent scores, univariate logistic regression revealed significant differences in the laterality of height of eyelid crease (0.9 ± 0.1 mm versus 1.9 ± 1.6 mm; OR = 1.06, 95% CI: 1.02–1.11; \( P = 0.003 \)) and in laterality of eyebrow height (1.3 ± 1.5 mm versus 2.3 ± 2.2 mm; OR = 1.03, 95% CI: 1.00–1.05; \( P = 0.03 \)), and no significant difference was observed in the laterality of MRD1 (0.9 ± 1.0 mm versus 1.1 ± 0.7 mm; OR = 1.02, 95% CI: 0.97–1.07; \( P = 0.40 \)) or the laterality of levator contractile function (1.1 ± 0.9 mm versus 1.5 ± 2.0 mm; OR = 0.99, 95% CI: 0.95–1.02; \( P = 0.33 \)). Multivariate logistic regression revealed a significant difference only in width of eyelid crease (OR = 1.06, 95% CI: 1.01–1.099; \( P = 0.01 \)).

**Validation 2**

In the results of evaluations by plastic surgeons with limited clinical experience, the mean weighted kappa coefficients were 0.62 for group B and 0.59 for group U. The numbers of eyelids with final evaluation scores of 0, 1, 2, and 3 were 78, 60, 42, and 20, respectively, in group B and 81, 58, 46, and 15, respectively, in group U. The weighted kappa coefficient for groups B and U was 0.51 (Table 5). In 63 eyelids of 52 cases, there were discrepancies between the results of groups B and U. The consistency of the results between groups B and U was higher among the plastic surgeons with extensive experience than among the plastic surgeons with limited experience (\( P < 0.0001 \), OR = 0.28, 95% CI: 0.15–0.51; Fig. 4). The weighted kappa coefficient for the final evaluation scores of group B by plastic surgeons with limited experience and those with extensive experience was 0.59.

**DISCUSSION**

Although the age-related changes of the eyelids can be evaluated based on structure and function of single eyelids, the actual control of eyelid opening is bilateral, and plastic surgeons evaluate the severity by observing both eyelids of each patient in daily clinical practice.\(^5\,6\,9\,12\,15\) In this study, we investigated the degree to which plastic surgeons evaluated age-related changes in the eyelids in this manner in single eyelids, and we examined what information about the contralateral eyelid was used to make a final evaluation.

In this study, even when evaluated by an extensively-experienced plastic surgeon, 13% of patients had differences in

---

**Table 2. Weighted Kappa Coefficient within each Group and between Groups Evaluated by Plastic Surgeons with Extensive Experience**

| Parameter                        | Right Side (95% CI) | Left Side (95% CI) | Mean       |
|----------------------------------|---------------------|--------------------|------------|
| Coefficient within group B\(^*\) | 0.57 (0.51–0.62)    | 0.50 (0.45–0.56)   | 0.58       |
| Coefficient within group U\(^†\)| 0.56 (0.51–0.62)    | 0.62 (0.58–0.67)   | 0.59       |
| Coefficient between groups      | 0.77 (0.67–0.87)    | 0.76 (0.66–0.86)   | 0.77       |

\(^*\)Photographs of two eyelids.  
\(^†\)Photographs of single eyelids.

---

**Table 3. Spearman’s Rank Correlation Coefficients between the Factors and the Score Evaluated by Plastic Surgeons with Extensive Experience**

| Factor                        | Group B\(^*\)       | Group U\(^†\)       |
|-------------------------------|---------------------|---------------------|
| Margin reflex distance 1      | −0.73 (\( P < 0.0001 \)) | −0.77 (\( P < 0.0001 \)) |
| Height of eyelid crease       | 0.24 (\( P = 0.0005 \))     | 0.25 (\( P = 0.0009 \))     |
| Eyebrow height                | 0.47 (\( P < 0.0001 \))     | 0.50 (\( P < 0.0001 \))     |
| Levator contractile function  | −0.62 (\( P < 0.0001 \))    | −0.56 (\( P < 0.0001 \))    |

\(^*\)Photographs of two eyelids.  
\(^†\)Photographs of single eyelids.

---

**Table 4. Multivariate Logistic Regression Analysis between the Patients with Consistent Scores (n = 23) and Those with Inconsistent Scores (n = 77)**

| Variables                      | Univariate Logistic Regression (n = 100) | P     | Multivariate Logistic Regression (n = 100) | P   |
|--------------------------------|----------------------------------------|-------|----------------------------------------|-----|
| Laterality in margin reflex distance 1 | OR (95% CI)                         | 0.40  | OR (95% CI)                             | 0.01|
| Laterality in height of eyelid crease       | 1.06 (1.02–1.11)                     | 0.003 | 1.06 (1.01–1.10)                       | 0.01|
| Laterality in eyebrow height                | 1.03 (1.00–1.05)                     | 0.03  | 1.02 (0.99–1.05)                       | 0.23|
| Laterality in levator contractile function  | 0.99 (0.95–1.02)                     | 0.33  |                                         |     |
evaluation based on bilateral and unilateral eyelid photographs. Although laterality in MRD1 and levator contractile function were strongly correlated with evaluation scores in both groups B and U, they were not related to whether the evaluations were consistent or inconsistent. Conversely, although the correlation coefficient with the score was not very high for the height of eyelid crease itself, the left-right difference of it was independently and significantly correlated with the consistency of scores between groups B and U. These results suggest that the evaluation of ptosis by experienced plastic surgeons is largely based on the structure and function of one eyelid along with information about laterality, such as the height of the eyelid crease.

As there is no criteria for correct evaluation of the severity of age-related changes in the eyelids, there was a similar variation among the subjects evaluated by both plastic surgeons with extended experience and plastic surgeons with limited experience. However, the evaluations of plastic surgeons with extended experience demonstrated a higher agreement rate between groups B and U than the plastic surgeons with limited experience. These results suggest that experienced plastic surgeons are less affected by the shape of the contralateral eyelid during evaluation.

MRD1 and levator contractile function have been proposed as indicators of visual field dysfunction in the treatment of blepharoptosis. Although the height of the eyelid crease also changes, setting criteria for surgical treatment at absolute value is difficult because this width varies greatly among people. A comparison of both eyelids reveals more clearly whether they are normal or abnormal. It has been widely recognized and reported that asymmetry in eyebrow position and height of eyelid crease reflects compensation for blepharoptosis and that correction of both is important in increasing patients’ satisfaction with treatment. In conjunction with the results of this study, recording changes before and after surgery is necessary because the laterality of the eyelid crease height is a factor that emphasizes age-related changes in the eyelids.

Is it possible to standardize the objective aspects of age-related changes in the eyelids for evaluation? It is difficult to conclude only from the results of this study. In clinical practice, evaluation has to be done on a patient and analysis of photographs cannot replace clinical evaluation. However, by analyzing the results of evaluation by experienced plastic surgeons, investigators in future studies may devise a highly reproducible algorithm for objective evaluation. Determining the morphological and functional abnormalities of individual eyelids and accounting for the left-right difference of height of eyelid crease may be important steps.

One of the limitations of this study is that the movement of the eyebrows is not restricted in measuring the levator contraction function. Thus, limiting the function of the frontal muscle to accurately measure the levator muscle function in the evaluation of aponeurotic ptosis.

### Table 5. Weighted Kappa Coefficient within and between Groups Evaluated by Plastic Surgeons with Limited Experience

| Parameter | Right Side (95% CI) | Left Side (95% CI) | Mean |
|-----------|---------------------|--------------------|------|
| Coefficient within group B | 0.60 (0.55–0.65) | 0.64 (0.59–0.69) | 0.62 |
| Coefficient within group U | 0.55 (0.50–0.60) | 0.63 (0.57–0.68) | 0.59 |
| Coefficient between groups | 0.49 (0.36–0.63) | 0.53 (0.40–0.66) | 0.51 |

*Photographs of two eyelids.
†Photographs of single eyelids.

---

**Fig. 4.** The rate of discrepancies in evaluations of groups B (photographs of both eyelids) and U (photographs of single eyelids) was significantly lower in evaluations by extensive experienced plastic surgeons than in those of plastic surgeons with limited experience ($P < 0.0001$, OR = 0.28, 95%CI: 0.15–0.51).
is necessary. This study evaluated the overall age-related changes in the eyelid and measured the distance of the movement of the eyelid margin, including the elevation of the eyebrow by the frontal muscle. However, the measured value was different from the accurate levator muscle function. Another limitation of this study is the classification of evaluators by the number of experienced cases. There was a large difference between the two groups, and as a result there was a difference between the two groups. However, it should be noted that the number of experienced cases does not always reflect diagnostic ability.

**CONCLUSIONS**

In evaluating age-related changes in the eyelids, scoring on a four-point scale, plastic surgeons with extensive experience showed fairly good consistency for photographs of both eyelids and of single eyelids, but one-point deviation was observed for 13% of the eyelids. Laterality in the height of eyelid crease was an important factor in the integration of left- and right-sided information. Height of eyelid crease may thus need to be included in an algorithm for evaluating age-related changes in the eyelids to include aesthetic aspects and visual impressions.

**Shinsuke Akita, MD, PhD**
Department of Plastic, Reconstructive, and Aesthetic Surgery
Chiba University Graduate School of Medicine
1-8-1, Inohana, Chuo-ku, Chiba City
Chiba 260-8670
Japan
E-mail: shinsukeakitaprs@gmail.com

**PATIENT CONSENT**
The patient provided written consent for the use of her image.

**ACKNOWLEDGMENT**
The authors thank Enago (www.enago.jp) for the English language review.

**REFERENCES**
1. Cahill KV, Bradley EA, Meyer DR, et al. Functional indications for upper eyelid ptosis and blepharoplasty: A report by the American Academy of Ophthalmology. *Ophthalmology*. 2011;118:2510–2517.
2. Putterman AM, Urist MJ. Müller muscle-conjunctiva resection. Technique for treatment of blepharoptosis. *Arch Ophthalmol*. 1975;93:619–623.
3. Small RG, Meyer DR. Eyelid metrics. *Ophthalmic Plast Reconstr Surg*. 2004;20:266–267.
4. Cahill KV, Burns JA, Weber PA. The effect of blepharoptosis on the field of vision. *Ophthalmic Plast Reconstr Surg*. 1987;3:121–125.
5. Averbuch-Heller L, Leigh RJ, Mermelstein V, et al. Ptosis in patients with hemispheric strokes. *Neurology*. 2002;58:620–624.
6. Macdonald KL, Mendez AI, Hart RD, et al. Eyelid and brow asymmetry in patients evaluated for upper lid blepharoplasty. *J Otolaryngol Head Neck Surg*. 2014;43:36.
7. Song WC, Kim SJ, Kim SH, et al. Asymmetry of the palpebral fissure and upper eyelid crease in Koreans. *J Plast Reconstr Aesthet Surg*. 2007;60:251–255.
8. Zhou Q, Zhang L, Wang PJ, et al. Preoperative asymmetry of upper eyelid thickness in young Chinese women undergoing double eyelid blepharoplasty. *Plast Reconstr Aesthet Surg*. 2012;55:1175–1180.
9. Parsa FD, Wolff DR, Parsa NN, et al. Upper eyelid ptosis repair after cataract extraction and the importance of Hering’s test. *Plast Reconstr Surg*. 2001;108:1527–36; discussion 1537.
10. Erb MH, Kersten RC, Yip CC, et al. Effect of unilateral blepharoptosis repair on contralateral eyelid position. *Ophthalmic Plast Reconstr Surg*. 2004;20:418–422.
11. Bodian M. Lip droop following contralateral ptosis repair. *Arch Ophthalmol*. 1982;100:1122–1124.
12. 1 Nemet AV. The effect of Hering’s law on different ptosis repair methods. *Aesthet Surg J*. 2015;35:774–81.
13. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33:159–174.
14. Ide K, Yamada H, Kitagawa M, et al. Methods for estimating causal relationships of adverse events with dietary supplements. *BMJ Open*. 2015;5:e009038.
15. Japan Society of Plastic and Reconstructive Surgery. Clinical Practice Guidelines for Plastic Surgery. Vol. 6: Blepharoptosis (in Japanese). Tokyo, Japan: Kanehara & Co., Ltd; 2015.
16. Shome D, Mittal ST, Kapoor R. Effect of eyelid crease formation on aesthetic outcomes post frontalis suspension for unilateral ptosis. *Plast Reconstr Surg Glob Open*. 2019;7:e2939.
17. Karlin JN, Rootman DB. Brow height asymmetry before and after eyelid ptosis surgery. *Plast Reconstr Aesthet Surg*. 2020;73:357–362.
18. Zhou X, Zhu M, Lv L, et al. Treatment strategy for severe blepharoptosis. *J Plast Reconstr Aesthet Surg*. 2020;75:149–155.