Development of Definitive and Reliable Grading Scales for Meibomian Gland Dysfunction

REIKO ARITA, IZUMI MINOURA, NAOYUKI MORISHIGE, RIKA SHIRAKAWA, SHIMA FUKUOKA, KEI ASAI, TATEKI GOTO, TAKAHIRO IMANAKA, AND MASATSUGU NAKAMURA

• PURPOSE: To develop and validate grading scales for meibomian gland dysfunction (MGD) that allow consistent diagnosis of MGD and are suitable for clinical studies.

• DESIGN: Development and validation study of grading scales.

• METHODS: Lid margin and meibomian gland photographs were taken in the multicenter, prospective cross-sectional study for MGD and control subjects. New grading scales for MGD signs (abnormal lid margin findings of vascularity, plugging of gland orifices, lid margin irregularity, lid margin thickening, partial glands, and gland dropout) in both upper and lower eyelids were developed. Three MGD experts, 3 general ophthalmologists, and 3 non-physicians independently tested the scales by evaluating photographs. The levels of interrater and intrarater agreement for each grading scale were estimated with the use of kappa statistics.

• RESULTS: Thirty-eight patients with MGD and 20 control subjects were enrolled and photographed. New grading scales were developed using a total of 226 photographs. The interrater kappa values for MGD experts and for general ophthalmologists and non-physicians with reference to an MGD expert ranged from 0.36 to 0.87 (median of 0.66), 0.41 to 0.73 (0.60), and 0.30 to 0.77 (0.59), respectively. Those for intrarater reliability for 2 MGD experts ranged from 0.49 to 0.93 (0.82).

• CONCLUSIONS: New grading scales for MGD signs were developed and found to have appropriate inter- and intrarater reliabilities for grading MGD. These grading scales are suitable for MGD diagnosis and application to multicenter trials. (Am J Ophthalmol 2016;169:125–137. © 2016 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).)

AJO.com Supplemental Material available at AJO.com.

Accepted for publication Jun 13, 2016.

From the Department of Ophthalmology, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan (R.A., R.S.); Department of Ophthalmology, Itoh Clinic, Saitama, Japan (R.A.); Research and Development Division, Santen Pharmaceutical Co Ltd, Osaka, Japan (I.M., K.A., T.G., T.I., M.N.); Department of Ophthalmology, Graduate School of Medicine, Yamaguchi University, Yamaguchi, Japan (N.M.); and Tokyo Kyosai Hospital, Tokyo, Japan (S.F.).

Inquiries to Reiko Arita, Department of Ophthalmology, Graduate School of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan; e-mail: ritoh@za2.so-net.ne.jp

THE TEAR FILM THAT COVERS THE OCULAR SURFACE is protected from evaporation by a thin layer of lipid secreted by the meibomian glands.[1] Meibomian gland dysfunction (MGD) is one of the most common disorders encountered in ophthalmic clinics and is now recognized as a major cause of dry eye syndrome.[2–5] It can result in tear film instability, damage to the ocular surface epithelium, chronic blepharitis, and contact lens intolerance.[6–10] MGD is commonly characterized by a chronic, diffuse abnormality of meibomian glands, terminal duct obstruction, and qualitative or quantitative changes in the glandular secretion.[11] The key signs of MGD are meibomian gland dropout, altered meibomian gland secretion, and changes in morphology of the lid margin. The lid margin abnormalities may become detectable with a slit-lamp microscope as the disease progresses.[12] Lid margin findings and meibomian gland morphology are therefore important for diagnosis of MGD. Diagnosis and quantification of MGD thus require assessment of symptoms, altered meibomian gland secretion, changes in lid morphology, and meibomian gland dropout.[12] Evaluation of meibomian gland expressibility as a dynamic process is also important. Assessment of the efficacy of treatment for MGD requires precise evaluation of changes in lid morphology, meibomian gland dropout, and meibomian gland expressibility. The ability to perform an objective evaluation of MGD based on photographs would be useful as a standardized procedure for multicenter clinical trials.

Grading scales for diagnosis of MGD have been proposed and adopted in clinical practice.[13–18] These scales are based on assessment of lid margin findings[13–15] or of meibomian glands, with the latter scales being based on the proportion of meibomian glands showing dropout in only the lower tarsal plate[16,17] or on the number of whole glands and proportion of partial glands.[18] The application of these existing scales to the clinic is difficult, however, because there are many grading subdivisions for each sign or their targets are limited to the lower eyelids. Moreover, information on grading reliability has been available for only a few scales. There is thus still an unmet need for reliable and widely adoptable grading scales based on

0002-9394/$36.00
http://dx.doi.org/10.1016/j.ajo.2016.06.025

TABLE 1. Background and Clinical Characteristics of Subjects Evaluated for Establishment of Grading Scales for Meibomian Gland Dysfunction

Characteristic	Controls (N = 20)	Initial MGD Patients (N = 38)	Additional MGD Patients (N = 18)
Sex (male/female)	8/12	13/25	8/10
Age (y)	64.5 ± 6.7	66.9 ± 15.0	69.1 ± 13.2
BUT (s)	6.5 ± 1.6	3.4 ± 2.1	2.7 ± 1.1
Schirmer test value (mm)	16.4 ± 9.9	10.6 ± 7.2	7.9 ± 3.7
Corneal staining score[a]	0.0 ± 0.0	0.5 ± 0.6	0.8 ± 0.4
Conjunctival staining score[b]	0.0 ± 0.0	0.3 ± 0.6	0.4 ± 0.5
Telangiectasia[c]			
Upper	0.1 ± 0.3	1.7 ± 0.9	1.8 ± 0.9
Lower	0.1 ± 0.2	1.3 ± 0.8	1.3 ± 0.6
Plugging[d]			
Upper	0.1 ± 0.2	1.3 ± 0.7	1.5 ± 0.5
Lower	0.0 ± 0.0	1.0 ± 0.7	1.1 ± 0.4
Meiboscore[e]			
Upper	0.4 ± 0.5	1.9 ± 0.8	2.0 ± 0.8
Lower	0.2 ± 0.4	1.9 ± 0.9	1.6 ± 0.8
Meibum grade[f]			
Upper	0.1 ± 0.4	1.5 ± 0.9	1.6 ± 0.5
Lower	0.2 ± 0.5	1.6 ± 0.7	1.4 ± 0.6
Gland dropout[g]			
Upper	0.1 ± 0.3	1.1 ± 0.8	2.0 ± 0.8
Lower	0.0 ± 0.0	1.1 ± 0.8	1.7 ± 0.8

MGD = meibomian gland dysfunction.
Data are means ± SD.
[a]Corneal staining was scored 0-3 for the entire cornea.
[b]Conjunctival staining was scored 0-3 for each of the nasal and temporal conjunctiva and then summed.
[c]Telangiectasia was assessed on a scale of 0-3.
[d]Plugging was assessed on a scale of 0-2.
[e]Partial or complete loss of meibomian glands was scored 0-3 (meiboscore).
[f]Meibum grade was scored 0-3.
[g]Complete loss of meibomian glands was scored 0-2.

evaluation of both upper and lower eyelids for the consistent diagnosis of MGD.

We have now developed new grading scales for MGD that can be used by ophthalmologists without special experience, and we have performed a validation study to confirm the robustness of these scales. We propose that these grading scales are suitable for the diagnosis of MGD or for its evaluation in clinical studies.

METHODS

• **STUDY DESIGN AND TARGET POPULATION:** This study was conducted at The University of Tokyo Hospital, Itoh Clinic (Saitama City, Saitama, Japan), and Maeda Ophthalmic Clinic (Aizuwakamatsu City, Fukushima, Japan). MGD patients and control subjects were randomly enrolled from outpatients who visited the 3 medical facilities from December 4, 2012 to December 7, 2013. The study adhered to the tenets of the Declaration of Helsinki and was prospectively approved by the Institutional Review Board of Tokyo University School of Medicine. All subjects provided written informed consent before entry into the study.

The MGD patients were aged ≥20 years and were diagnosed on the basis of previously described criteria[19,20]: (1) at least 1 symptom, such as ocular fatigue, discharge, foreign body sensation, dryness, uncomfortable sensation, sticky sensation, pain, epiphora, itching, redness, heavy sensation, glare, excessive blinking, burning sensation, and ocular discomfort on arising; (2) at least 1 abnormal lid margin finding, such as vascular engorgement, anterior or posterior replacement of the mucocutaneous junction, and irregular lid margin; and (3) plugged meibomian gland orifices and poor meibum expressibility in the target eye. The control subjects had never been diagnosed with blepharitis or MGD, were aged ≥20 years, and had no history of contact lens wear or eye surgery. Individuals with severe systemic illness or with squamous

cell debris (collarette) around the base of the eyelashes were excluded.

• **CLINICAL ASSESSMENT AND IMAGE COLLECTION:** One eye was selected as the target eye in each subject. Complications, history of contact lens wear or eye surgery, the presence of ocular allergy, and concomitant medications were noted as background information. The subjects were assessed for lid margin and meibomian gland findings, as well as for their experience of subjective symptoms.

Lid margin findings were evaluated for the upper and lower eyelids with the use of a slit-lamp microscope. Telangiectasia was assessed on a scale from 0 to 3: 0 = no findings; 1 = mild telangiectasia; 2 = moderate telangiectasia or redness; 3 = severe telangiectasia or redness. Mucocutaneous junction was assessed on a scale from 0 to 3: 0 = Marx line (ML) courses on the skin side of the meibomian orifice (MO) line and does not touch MOs at all; 1 = parts of ML touch MOs; 2 = ML courses through MOs; 3 = ML courses along the eyelid margin side of MOs.[21] Irregularity, plugging, foaming, and thickness were assessed on a scale from 0 to 2: 0 = no findings; 1 = mild findings; 2 = severe findings.

Corneal and conjunctival staining were scored from 0 to 9.[22] The tear film breakup time (BUT) was measured 3 times consecutively after the instillation of fluorescein, and the mean value was adopted. Tear fluid production was evaluated with the Schirmer test without anesthesia.

Meibomian glands were evaluated for the upper and lower lids with the use of a noncontact meibography system attached to a slit-lamp microscope. Partial or complete loss of meibomian glands was scored on a scale from 0 to 3 (meiboscore), as described previously.[23] The extent of meibomian gland dropout was determined on a scale from 0 to 2 based on the number of affected glands: 0 = none; 1 = small number; 2 = large number.

The degree of ease with which meibum could be expressed at the central area of both upper and lower eyelid was evaluated semiquantitatively on a scale from 0 to 3: 0 = clear meibum readily expressed; 1 = cloudy meibum expressed with mild pressure; 2 = cloudy meibum expressed with more than moderate pressure; 3 = meibum could not be expressed even with strong pressure.[16]

Images of lid margins and orifices of meibomian glands at the upper and lower eyelids of 1 eye were obtained with a digital camera and meibography system attached to a slit-lamp microscope (SL-D701 DC4 BG-5; Topcon Japan, Tokyo, Japan).[23] The images of the lid margins and those of meibomian glands were both acquired at 10× magnification for the full length of each eyelid within a single photograph. Four images (2 of each eyelid) were collected for each subject.

• **NEW GRADING SYSTEM:** We developed grading scales for MGD with the use of printed images. All images were acquired by an ophthalmologist (R.A. or R.S.) with a specialty in MGD. Four signs for lid margin findings detected with a slit-lamp microscope and 2 signs of meibomian glands detected by meibography were selected for development of the grading scales. Three MGD experts (R.A., R.S., and S.F.), each participating at a different institution, developed draft grading scales for

TABLE 2. Distribution of Severity for Lid Margin and Meibomian Gland Findings in Subjects Evaluated for Establishment of Grading Scales for Meibomian Gland Dysfunction

	MGD Patients (N = 38)		Controls (N = 20)	
Parameter	Upper	Lower	Upper	Lower
Telangiectasia				
0	3 (7.9)	4 (10.8)	17 (89.5)	18 (94.7)
1	14 (36.8)	21 (56.8)	2 (10.5)	1 (5.3)
2	14 (36.8)	8 (21.6)	0 (0)	0 (0)
3	7 (18.4)	4 (10.8)	0 (0)	0 (0)
Missing data	0	1	1	1
Mucocutaneous junction				
0	3 (8.3)	4 (11.8)	18 (94.7)	18 (94.7)
1	12 (33.3)	12 (35.3)	1 (5.3)	1 (5.3)
2	11 (30.6)	12 (35.3)	0 (0)	0 (0)
3	10 (27.8)	6 (17.6)	0 (0)	0 (0)
Missing data	2	4	1	1
Irregularity				
0	22 (57.9)	19 (50.0)	18 (94.7)	18 (94.7)
1	14 (36.8)	14 (36.8)	1 (5.3)	1 (5.3)
2	2 (5.3)	5 (13.2)	0 (0)	0 (0)
Missing data	0	0	1	1
Plugging				
0	5 (13.2)	9 (23.7)	18 (94.7)	19 (100)
1	18 (47.4)	21 (55.3)	1 (5.3)	0 (0)
2	15 (39.5)	8 (21.1)	0 (0)	0 (0)
Missing data	0	0	1	1
Foaming				
0	37 (97.4)	34 (91.9)	19 (100)	19 (100)
1	1 (2.6)	2 (5.4)	0 (0)	0 (0)
2	0 (0)	1 (2.7)	0 (0)	0 (0)
Missing data	0	1	1	1
Thickness				
0	29 (76.3)	30 (78.9)	19 (100)	19 (100)
1	7 (18.4)	7 (18.4)	0 (0)	0 (0)
2	2 (5.3)	1 (2.6)	0 (0)	0 (0)
Missing data	0	0	1	1
Meiboscore				
0	1 (2.7)	0 (0)	12 (60.0)	16 (80.0)
1	11 (29.7)	15 (40.5)	8 (40.0)	4 (20.0)
2	17 (45.9)	10 (27.0)	0 (0)	0 (0)
3	8 (21.6)	12 (32.4)	0 (0)	0 (0)
Missing data	1	1	0	0

MGD = meibomian gland dysfunction.
Data are n (%).

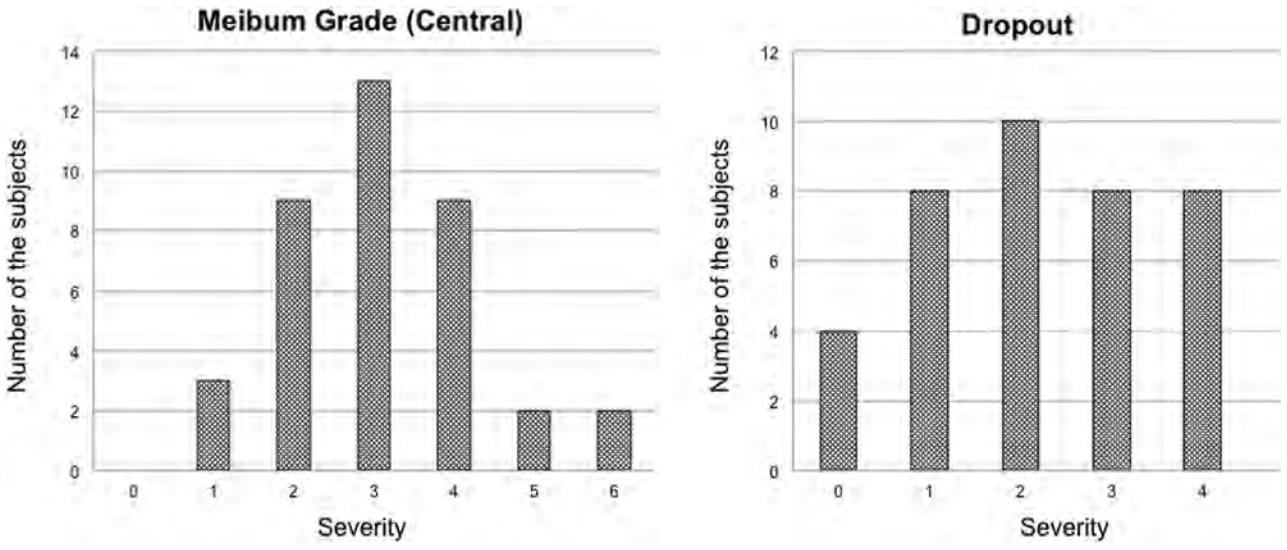

FIGURE 1. Distribution of the severity of meibum grade (Left) and meibomian gland dropout (Right) for the enrolled meibomian gland dysfunction patients (n = 38). Meibum grade was scored on a scale of 0-3 for each of the upper and lower eyelids and then summed. Meibomian gland dropout was scored on a scale of 0-2 for each of the upper and lower eyelids and then summed.

TABLE 3. Proposed Grading Scales for Meibomian Gland Dysfunction

Abnormal Lid Margin Findings of Vascularity
0 = No or slight redness in lid margin conjunctiva and no telangiectasia crossing meibomian gland orifices
1 = Redness in lid margin conjunctiva and no telangiectasia crossing meibomian gland orifices
2 = Redness in lid margin conjunctiva and telangiectasia crossing meibomian gland orifices with a distribution of less than half of the full length of the lid
3 = Redness in lid margin conjunctiva and telangiectasia crossing meibomian gland orifices with a distribution of half or more of the full length of the lid
Plugging of Gland Orifices
0 = No plugging of gland orifices
1 = Fewer than 3 pluggings of gland orifices
2 = Three or more pluggings of gland orifices with a distribution of less than half of the full length of the lid
3 = Three or more pluggings of gland orifices with a distribution of half or more of the full length of the lid
Lid Margin Irregularity
0 = No lid margin irregularity
1 = Fewer than 3 lid margin irregularities with shallow notching
2 = Three or more lid margin irregularities or deep notching
Lid Margin Thickening
0 = No lid margin thickening
1 = Lid margin thickening with or without localized rounding
2 = Lid margin thickening with diffuse rounding
Partial Glands
0 = No partial glands
1 = Fewer than 3 partial glands
2 = Three or more partial glands and fewer than 3 partial glands with loss of half or more of the full length
3 = Three or more partial glands with loss of half or more of the full length
Gland Dropout
0 = No gland dropout
1 = Fewer than 3 gland dropouts
2 = Three or more gland dropouts

MGD with key conceptual components based on morphologic and anatomic criteria described in previous reviews.[13,14] These draft grading scales were then evaluated with a preliminary validation test in which each of the 3 MGD experts classified the printed images independently according to the draft scales, and the

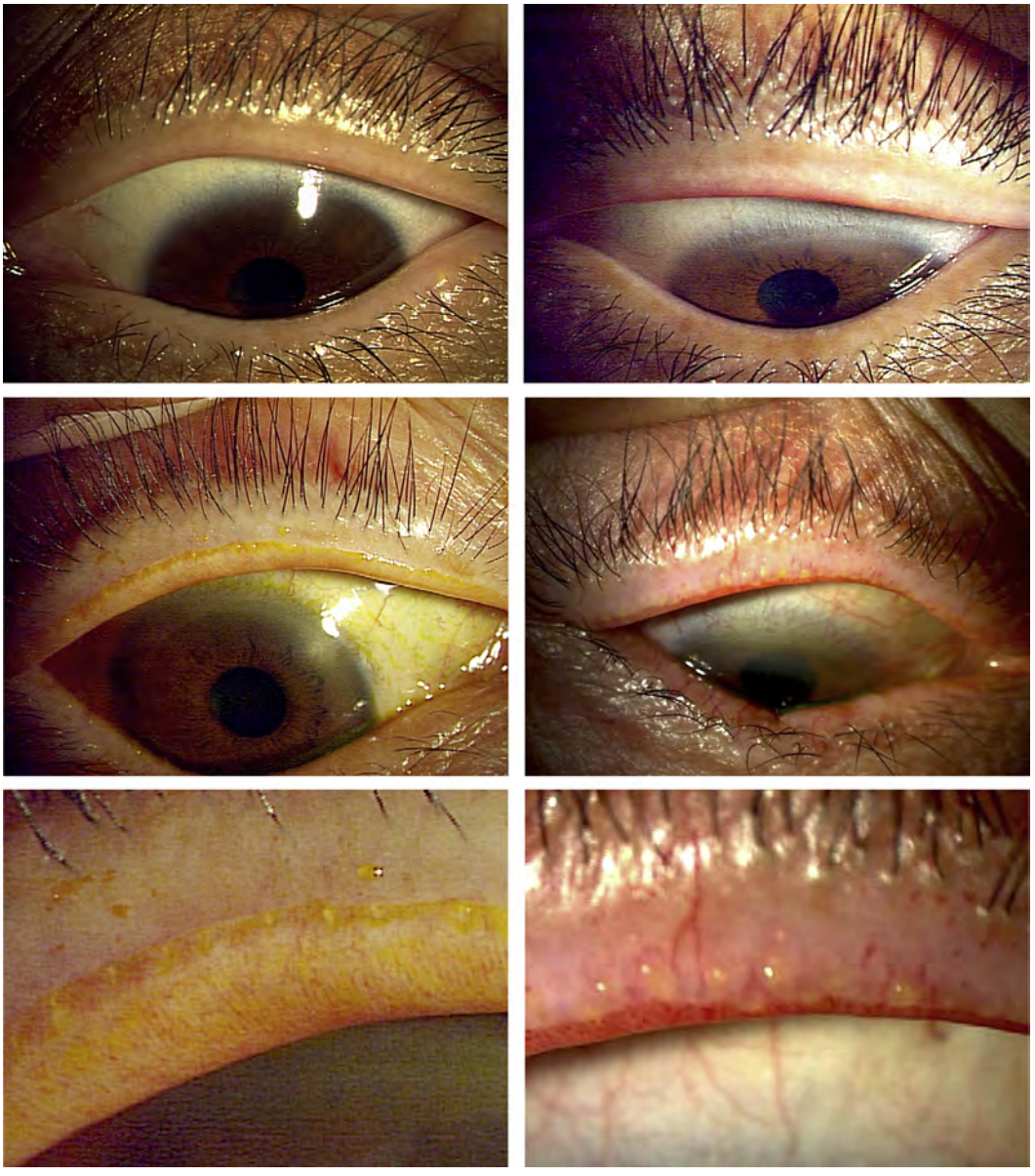

FIGURE 2. Representative images of lid margin findings of vascularity. Grade 0 (Upper left), grade 1 (Upper right), grade 2 (Middle left), and grade 3 (Middle right) are based on the degree of redness in the lid margin and the distribution of telangiectasia crossing meibomian gland orifices. Note that telangiectasia is observed but does not cross the gland orifices in half of the eyelid in grade 2 (Bottom left), whereas it crosses the gland orifices in half or more of the eyelid in grade 3 (Bottom right).

consistency for each scale was determined. The 3 experts then reviewed the results and adjusted the grading scales accordingly. The 3 experts then reviewed the results and adjusted the grading scales accordingly with the advice of another expert. The final versions proceeded to the validation test as the proposed grading scales.

• **VALIDATION TESTING:** Validation testing was performed to evaluate the robustness of the proposed grading scales after an interval of >2 weeks since their development. The raters received printed images of the upper and lower eyelids of the original 58 subjects together with printed grading scales and representative images. Each printed image had a randomly assigned number for analysis. Raters individually classified each image according to the grading scales. The test was performed by each rater at a separate site. After the test, the assigned numbers of the classified images for each grading scale were recorded for statistical analysis.

Interrater and intrarater reliability. Three MGD experts (R.A., R.S., and S.F.) performed the validation test, and

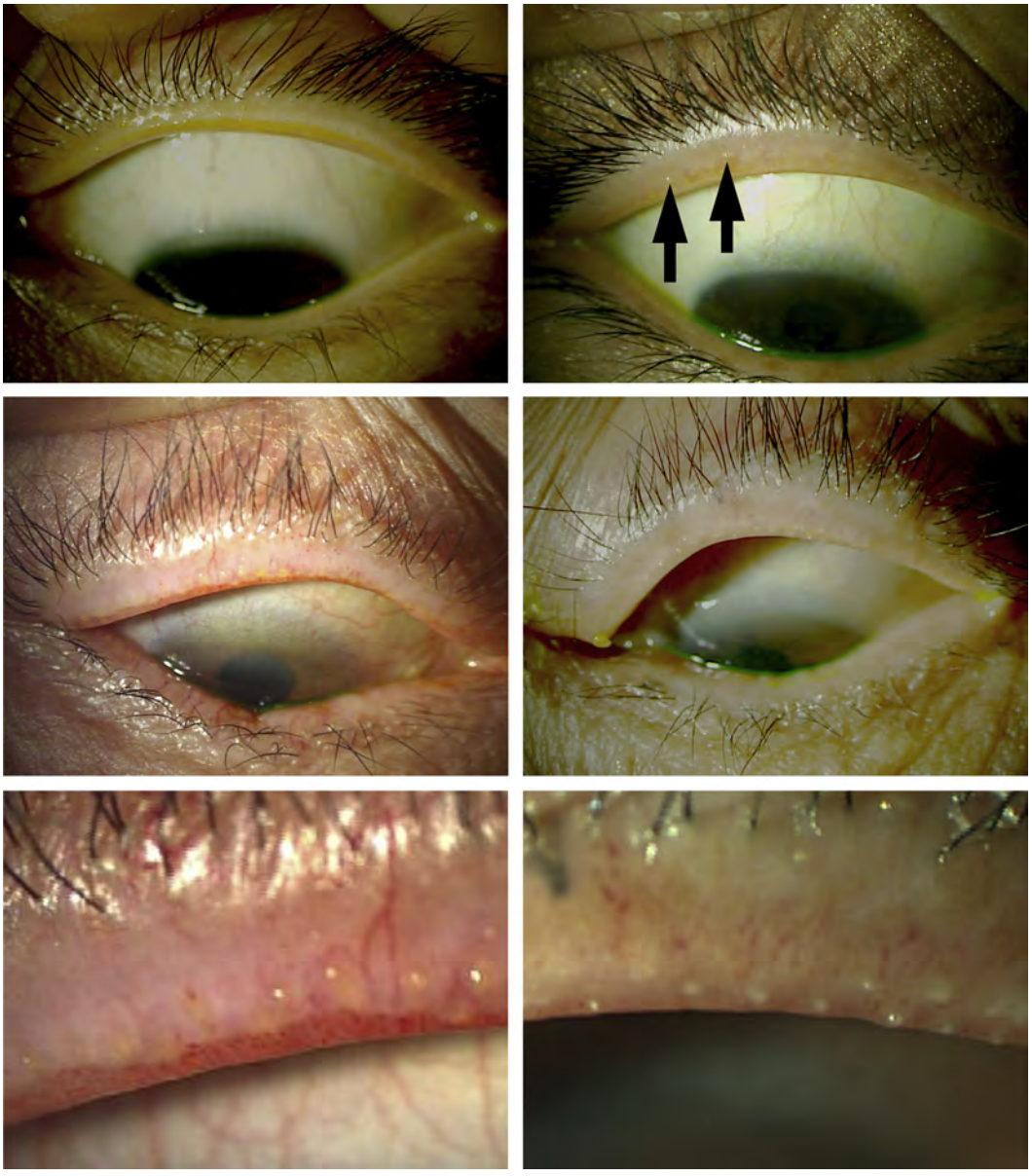

FIGURE 3. Representative images of plugging of gland orifices. Grade 0 (Upper left), grade 1 (Upper right), grade 2 (Middle left), and grade 3 (Middle right) are based on the number and distribution of abnormal findings for meibomian gland orifices such as capping, pouting, and ridge. Arrows indicate plugging (Upper right). Plugging is apparent but shows a distribution of less than half of the lid length in grade 2 (Bottom left), whereas it is distributed along half or more of the lid length in grade 3 (Bottom right).

consistency among their results for each grading scale was evaluated.

Two of the MGD experts who participated in the determination of interrater reliability (R.A. and S.F.) performed the test a second time after an interval of >2 weeks. Consistency between the results of the first and second tests for each grading scale was evaluated.

Effect of clinical experience. Three general ophthalmologists (non–MGD experts) who had been certified for 3, 5, or 8 years, as well as 3 non-physicians, also performed the validation test. An introduction and explanation of the grading scales were given before the test by an MGD expert (R.A.). Consistency between the results for each group of raters and those for an MGD expert (first test performance by R.A.) was evaluated as interrater reliability.

Additional validation. For confirmation of the intrarater and interrater reliability of the grading scales, 2 MGD experts (N.M. and Tohru Sakimoto [Nihon University]) who did not contribute to their development performed

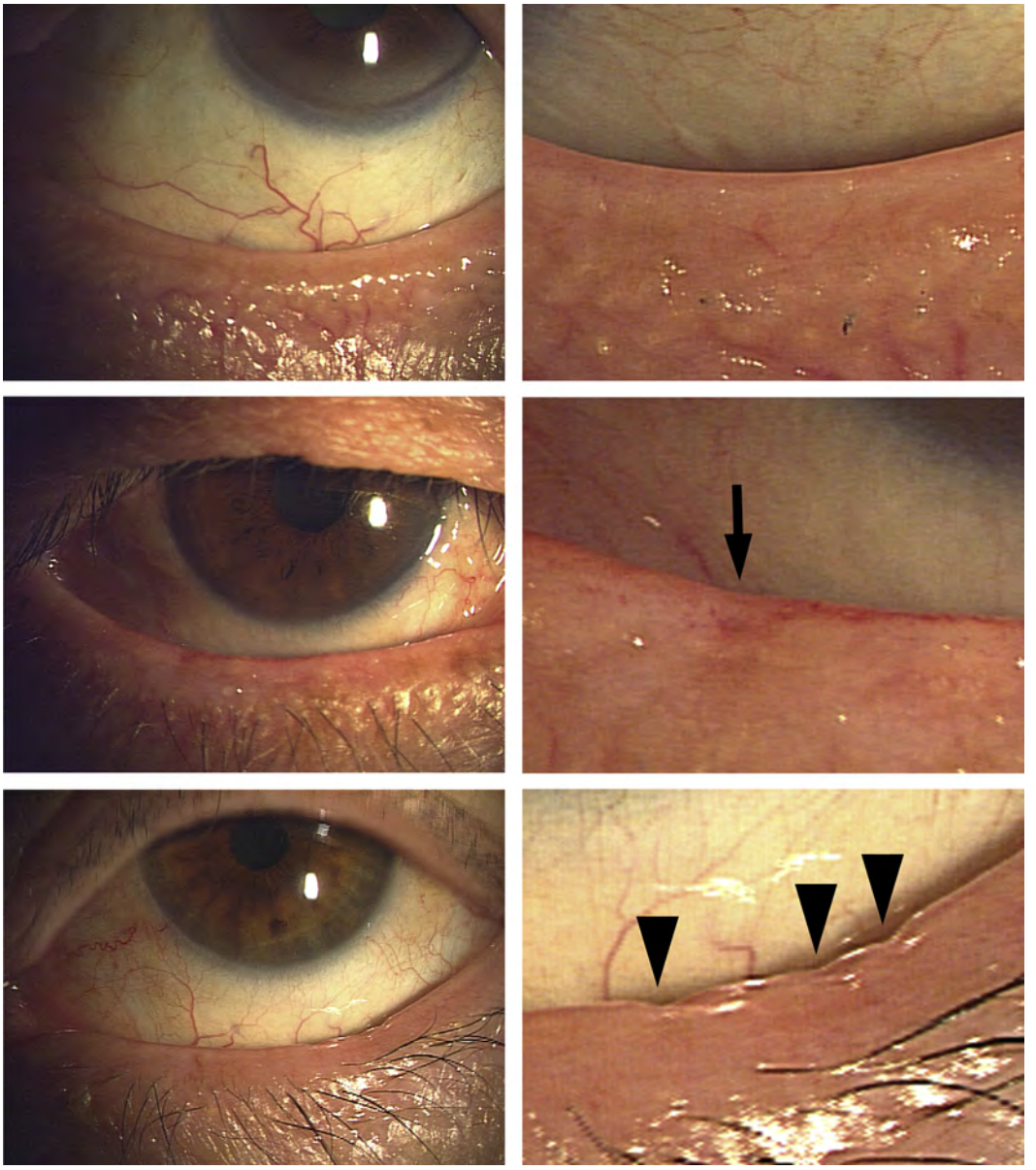

FIGURE 4. Representative images of lid margin irregularity. Grade 0 (Top left), grade 1 (Middle left), and grade 2 (Bottom left) are based on the number of irregularities and the form of notching. Notching is not observed in grade 0 (Top right), is shallow in grade 1 (Middle right, arrow), and is deep in grade 2 (Bottom right, arrowheads).

the validation test as described above but with images of the upper and lower eyelids of 18 additional MGD patients that were not used for scale development. Those images of additional MGD patients were obtained at Itoh Clinic.

• **DIAGNOSTIC ABILITY OF THE PROPOSED GRADING SCALES:** For evaluation of the diagnostic ability of the proposed grading scales, the MGD patients and control subjects were classified on the basis of the validation test results for an MGD expert (first test performance by R.A.). The total score of the upper and lower eyelids was used for the calculation. Receiver operating characteristic (ROC) curves and the area under the curve (AUC) were calculated for each grading scale.

• **STATISTICAL ANALYSIS:** Data were collected into an Excel 2010 worksheet and were analyzed with the use of SAS version 9.2 software (SAS Institute, Cary, North Carolina, USA). Weighted kappa values and 95% confidence intervals (CIs) for each grading scale were calculated to evaluate consistency in test performance. Median values were used to summarize the results for weighted kappa values for each test.

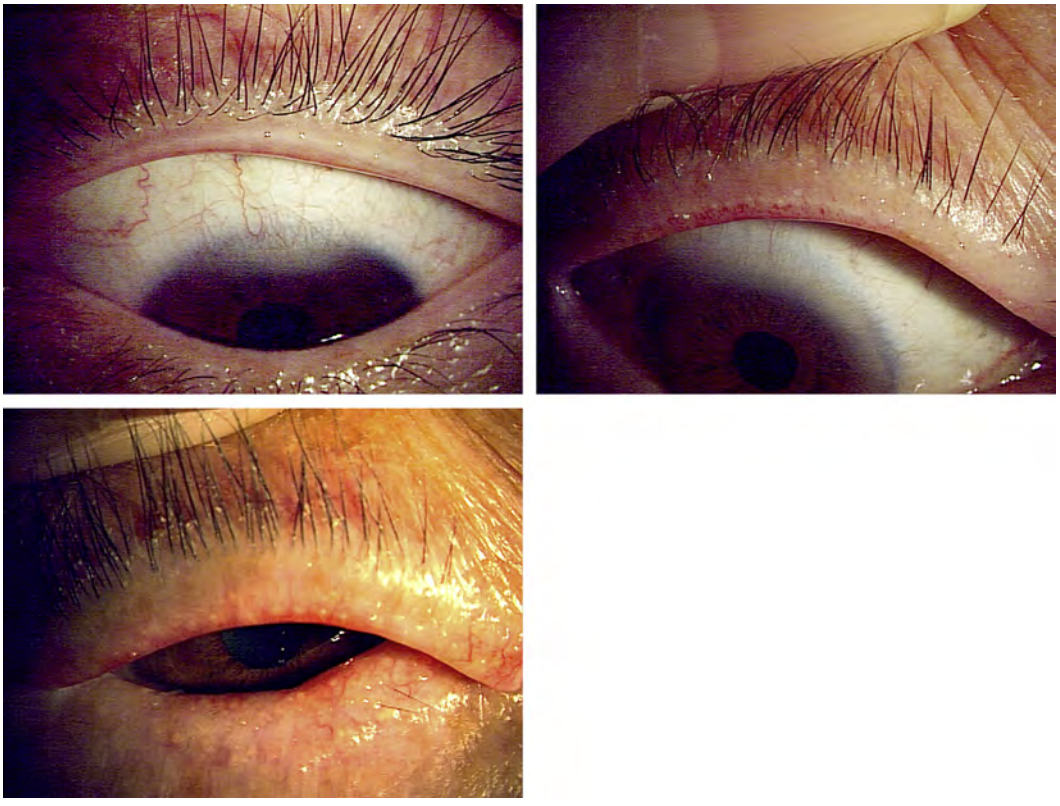

FIGURE 5. Representative images of lid margin thickening. Grade 0 (Upper left), grade 1 (Upper right), and grade 2 (Bottom left) are based on the presence of thickening and rounding.

RESULTS

THIRTY-EIGHT MGD PATIENTS (13 MEN AND 25 WOMEN; mean age ± SD, 66.9 ± 15.0 years) and 20 control subjects (8 men and 12 women; 64.5 ± 6.7 years) were eventually enrolled. All subjects underwent clinical assessment and image collection, and their clinical characteristics are shown in Table 1.

For the MGD patients, lid margin and meibomian gland findings encompassed the range of severity levels and appeared to reflect the distribution of findings encountered in the clinic (Table 2, Figure 1). For the control subjects, almost all findings were normal (Table 2). With the exception of 6 missing images, 4 images acquired from each subject (total of 226 images) were used to develop the new grading system.

• **NEW GRADING SYSTEM:** We developed 6 grading scales for MGD based on images of the upper and lower eyelids of 38 MGD patients and 20 control subjects (Table 3, Supplemental Figure; Supplemental Material available at AJO.com). Abnormal lid margin findings of vascularity (Figure 2), plugging of gland orifices (Figure 3), lid margin irregularity (Figure 4), and lid margin thickening (Figure 5) were evaluated in the full-length images of each eyelid obtained with a slit-lamp microscope. Partial glands (Figure 6) and gland dropout (Figure 7) were evaluated by noncontact meibography for meibomian glands in the middle two-thirds of each eyelid, given that it is difficult to capture and examine glands at the ends of each eyelid in a single photograph (Figure 6).

Each grading scale is based on specific features. Abnormal lid margin findings of vascularity are based on 2 key components: the degree of redness at the lid margin and the distribution of telangiectasia crossing meibomian gland orifices. Plugging of gland orifices includes abnormal findings for meibomian gland orifices such as capping, pouting, and ridge and is evaluated on the basis of the number and distribution of these abnormal findings. Lid margin irregularity is evaluated on the basis of the number of lid margin irregularities and form of notching. Lid margin thickening is evaluated based on the presence of thickening and rounding. Partial glands are defined as meibomian glands showing partial loss from the orifice or fornix and are evaluated on the basis of their number and length (Figure 6). Gland dropout is evaluated on the basis of the number of meibomian glands with complete loss from orifice to fornix (Figure 7).

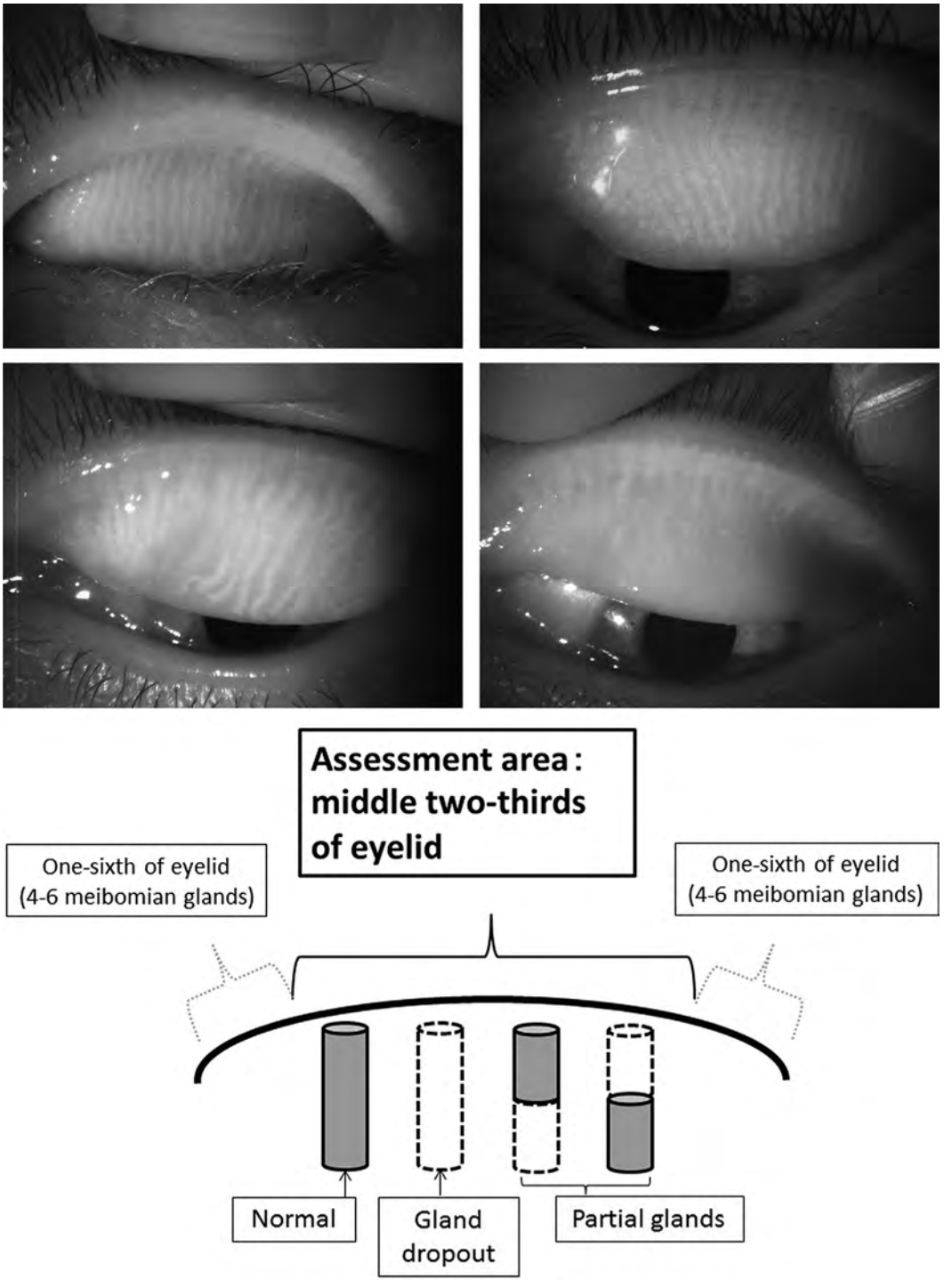

FIGURE 6. Representative images of partial glands. Grade 0 (Upper left), grade 1 (Upper right), grade 2 (Middle left), and grade 3 (Middle right) are based on the number and extent of partial meibomian glands, which are defined as glands with partial loss between the orifice and fornix. The scheme illustrates the definition of partial glands and gland dropout (Bottom), both of which are assessed over the middle two-thirds of each eyelid.

• **VALIDATION TESTING:** Evaluation of interrater reliability for performance of a validation test for the 6 grading scales by 3 MGD experts yielded a range (median) of weighted kappa values of 0.36-0.81 (0.68) for the upper eyelid and 0.57-0.87 (0.65) for the lower eyelid, as shown in Table 4, referred with agreement.[24]

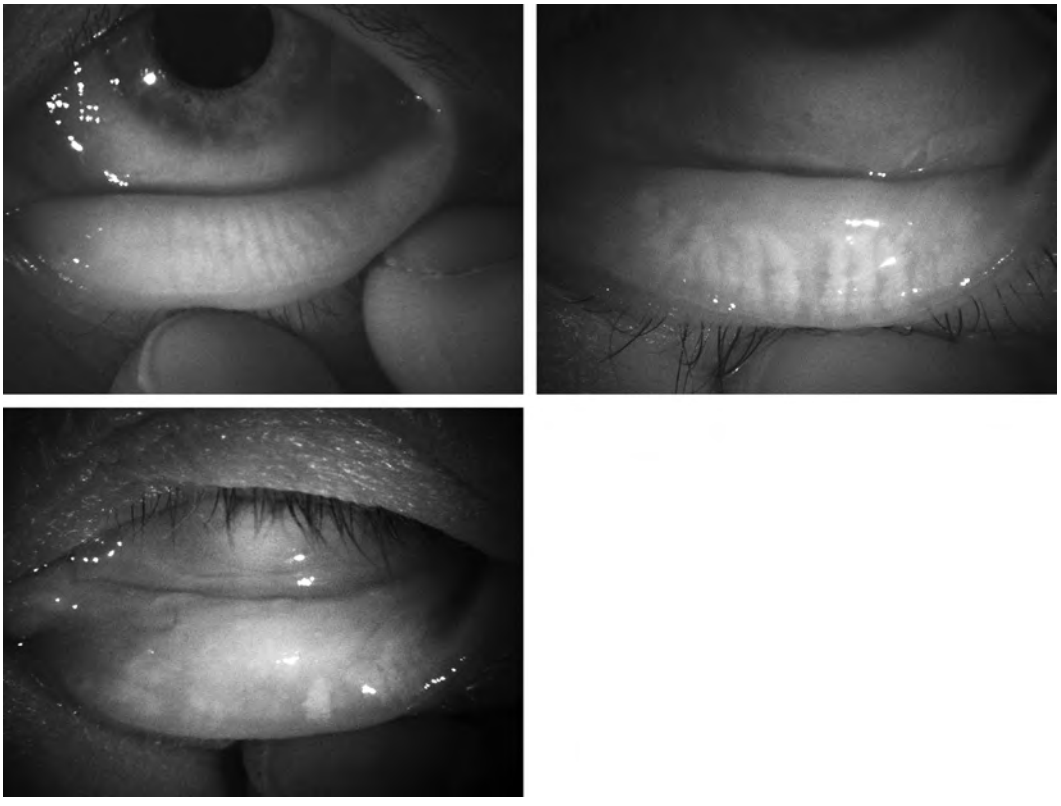

FIGURE 7. Representative images of gland dropout. Grade 0 (Upper left), grade 1 (Upper right), and grade 2 (Bottom left) are based on the number of meibomian glands showing complete loss from orifice to fornix.

The range (median) of weighted kappa values for evaluation of intrarater reliability for 2 MGD experts who performed the test twice was 0.49-0.92 (0.79) for the upper eyelid and 0.67-0.93 (0.85) for the lower eyelid (Table 4). With regard to the effect of clinical experience, evaluation of interrater reliability yielded a range (median) of weighted kappa values of 0.42-0.66 (0.59) for the upper eyelid and 0.41-0.73 (0.65) for the lower eyelid for 3 general ophthalmologists and of 0.35-0.68 (0.53) for the upper eyelid and 0.30-0.77 (0.63) for the lower eyelid for 3 non-physicians (Table 4). For confirmation of the reliability of the grading scales, the validation test was performed by 2 additional MGD experts who did not contribute to their development and with images obtained from a different group of 18 MGD patients that were also not used for scale development. The range (median) of weighted kappa values for evaluation of intrarater reliability for the 2 additional MGD experts who performed the test twice was 0.47-0.92 (0.65) for the upper eyelid and 0.49-0.91 (0.80) for the lower eyelid (Table 5). The evaluation of interrater reliability yielded a range (median) of weighted kappa values of 0.49-0.92 (0.53) for the upper eyelid and 0.30-0.74 (0.51) for the lower eyelid (Table 5).

• **DIAGNOSTIC ABILITY OF THE PROPOSED GRADING SCALES:** Finally, with regard to the diagnostic ability of the proposed grading scales, generation of ROC curves revealed that gland dropout showed the greatest AUC with a value of 0.78 (95% CI, 0.66-0.90), followed by partial glands (AUC = 0.72; 95% CI, 0.58-0.86), plugging of gland orifices (AUC = 0.70; 95% CI, 0.55-0.85), lid margin thickening (AUC = 0.69; 95% CI, 0.54-0.84), lid margin irregularity (AUC = 0.59; 95% CI, 0.46-0.73), and abnormal lid margin findings of vascularity (AUC = 0.54; 95% CI, 0.37-0.71) (Figure 8).

DISCUSSION

WE HAVE HEREIN DEVELOPED 6 NEW GRADING SCALES FOR typical clinical findings in MGD and have evaluated the reliability of these scales with the use of 226 images obtained from 58 subjects. The new grading scales were based on real clinical findings and were found to be robust and to show better than moderate agreement in most instances on validation testing. Whereas evaluation of dynamic features of meibomian gland function such as meibum expressibility

TABLE 4. Weighted Kappa Values and 95% Confidence Intervals for Intrarater and Interrater Variability in Initial Evaluation of Grading Scales for Meibomian Gland Dysfunction

Grading Scale	Intrarater				Interrater							
	MGD Expert 1 (R.A.)		MGD Expert 2 (S.F.)		MGD Experts		General Ophthalmologists		Non-physicians	
	Upper	Lower	Upper	Lower	Upper	Lower	Upper	Lower	Upper	Lower
Abnormal lid margin findings of vascularity	0.74 (0.60-0.88)	0.86 (0.77-0.95)	0.80 (0.67-0.93)	0.90 (0.80-1.00)	0.81 (0.71-0.91)	0.75 (0.65-0.86)	0.59 (0.43-0.74)	0.69 (0.59-0.80)	0.68 (0.54-0.82)	0.68 (0.55-0.81)
Plugging of gland orifices	0.49 (0.27-0.71)	0.77 (0.58-0.95)	0.72 (0.56-0.88)	0.67 (0.48-0.85)	0.58 (0.42-0.74)	0.63 (0.41-0.86)	0.59 (0.44-0.74)	0.41 (0.17-0.65)	0.35 (0.18-0.52)	0.30 (0.14-0.45)
Lid margin irregularity	0.67 (0.40-0.94)	0.87 (0.70-1.00)	0.87 (0.77-0.97)	0.76 (0.58-0.95)	0.36 (0.15-0.57)	0.57 (0.29-0.86)	0.42 (0.19-0.65)	0.62 (0.31-0.92)	0.46 (0.25-0.67)	0.58 (0.25-0.92)
Lid margin thickening	0.75 (0.62-0.89)	0.82 (0.70-0.94)	0.82 (0.70-0.95)	0.75 (0.62-0.87)	0.66 (0.51-0.81)	0.66 (0.50-0.81)	0.60 (0.46-0.74)	0.46 (0.32-0.61)	0.60 (0.46-0.74)	0.44 (0.32-0.57)
Partial glands	0.77 (0.58-0.95)	0.86 (0.75-0.97)	0.86 (0.74-0.97)	0.87 (0.77-0.98)	0.71 (0.50-0.91)	0.64 (0.46-0.82)	0.47 (0.35-0.60)	0.72 (0.60-0.83)	0.44 (0.28-0.59)	0.68 (0.56-0.81)
Gland dropout	0.92 (0.86-0.99)	0.93 (0.87-1.00)	0.90 (0.82-0.98)	0.83 (0.74-0.93)	0.73 (0.58-0.88)	0.87 (0.80-0.94)	0.66 (0.51-0.81)	0.73 (0.59-0.88)	0.60 (0.45-0.75)	0.77 (0.65-0.88)

MGD = meibomian gland dysfunction.

Kappa values of <0.01, 0.01-0.20, 0.21-0.40, 0.41-0.60, 0.61-0.80, and 0.81-1.00 correspond to poor, slight, fair, moderate, substantial, and almost-perfect agreement, respectively.[24]

and color is informative with regard to assessment of MGD, such evaluation is difficult to record as dynamic results. The new grading scales with representative images should thus prove helpful for the diagnosis of MGD in a consistent manner and be particularly useful for the evaluation of MGD in multicenter clinical studies.

The reliability of the new grading scales is likely attributable at least in part to the clear definitions and descriptions on which they are based, with avoidance of imprecise terms such as mild, moderate, or severe. The scales thus allow an objective classification of morphologic and anatomic findings. Moreover, the grading scales are applicable to both upper and lower eyelids. Given that the degree of morphologic change of meibomian glands has been found to differ between the upper and lower eyelids in MGD patient and dry eye patient populations, examination of both eyelids is essential for proper evaluation of the pathologic condition of such individuals.[25,26]

The new grading scales showed a high level of consistency among MGD experts, general ophthalmologists, and non-physicians. These findings indicate that any ophthalmologist should be able to diagnose MGD and classify its severity in a consistent manner with the grading scales. This ability is likely to prove especially useful in clinical studies or trials, with the grading scales allowing investigators at different sites to enroll appropriate subjects and to evaluate the efficacy of medication in a consistent manner and thus allowing the results obtained at the different sites to be compared directly.

Our evaluation of the new grading scales revealed that consistency among raters was lowest for plugging of gland orifices and lid margin irregularity. Plugging of gland orifices is usually assessed by determining the expressibility of meibum, and lid margin irregularity is usually assessed with a slit-lamp microscope by changing the angle of the light relative to the patient's eyelids. The consistent grading of these clinical findings based only on images might thus be expected to be more difficult than that for the other assessed signs.

We also investigated the diagnostic ability of the new grading scales for the enrolled MGD patients and control subjects. Construction of ROC curves and calculation of AUC values revealed that the scales for both partial glands and gland dropout showed sufficient diagnostic ability, suggesting that both signs are MGD specific and should be evaluated separately. These findings are consistent with the results of a previous study.[20] On the other hand, the AUC for abnormal lid margin findings of vascularity was lowest among the grading scales. The grading scale for this sign appears to be affected not only by MGD but also by aging or other factors, such as workplace exposure to dust particles, urban living, and cosmetics.[15] As far as we are aware, this is the first study to evaluate diagnostic ability of partial glands and gland dropout separately on the

TABLE 5. Weighted Kappa Values and 95% Confidence Intervals for Intrarater and Interrater Variability in Additional Evaluation of Grading Scales for Meibomian Gland Dysfunction

	Intrarater				Interrater	
	Additional MGD Expert 1 (N.M.)		Additional MGD Expert 2 (T.S.)		Additional MGD Experts	
Grading Scale	Upper	Lower	Upper	Lower	Upper	Lower
Abnormal lid margin findings of vascularity	0.92 (0.83-0.99)	0.91 (0.83-0.99)	0.66 (0.52-0.80)	0.84 (0.73-0.95)	0.92 (0.83-0.99)	0.58 (0.43-0.73)
Plugging of gland orifices	0.74 (0.61-0.88)	0.84 (0.73-0.95)	0.47 (0.31-0.66)	0.70 (0.57-0.83)	0.50 (0.33-0.66)	0.30 (0.15-0.45)
Lid margin irregularity	0.82 (0.65-0.99)	0.80 (0.67-0.93)	0.64 (0.31-0.97)	0.49 (0.26-0.72)	0.49 (0.30-0.68)	0.72 (0.54-0.91)
Lid margin thickening	0.52 (0.34-0.70)	0.49 (0.32-0.66)	0.50 (0.29-0.71)	0.80 (0.67-0.94)	0.55 (0.39-0.72)	0.41 (0.25-0.56)
Partial glands	0.72 (0.60-0.85)	0.80 (0.66-0.93)	0.50 (0.34-0.66)	0.74 (0.61-0.87)	0.52 (0.37-0.67)	0.43 (0.28-0.57)
Gland dropout	0.82 (0.71-0.93)	0.81 (0.68-0.94)	0.49 (0.33-0.65)	0.82 (0.71-0.94)	0.54 (0.40-0.68)	0.74 (0.61-0.87)

MGD = meibomian gland dysfunction.

Kappa values of <0.01, 0.01-0.20, 0.21-0.40, 0.41-0.60, 0.61-0.80, and 0.81-1.00 correspond to poor, slight, fair, moderate, substantial, and almost-perfect agreement, respectively.[24]

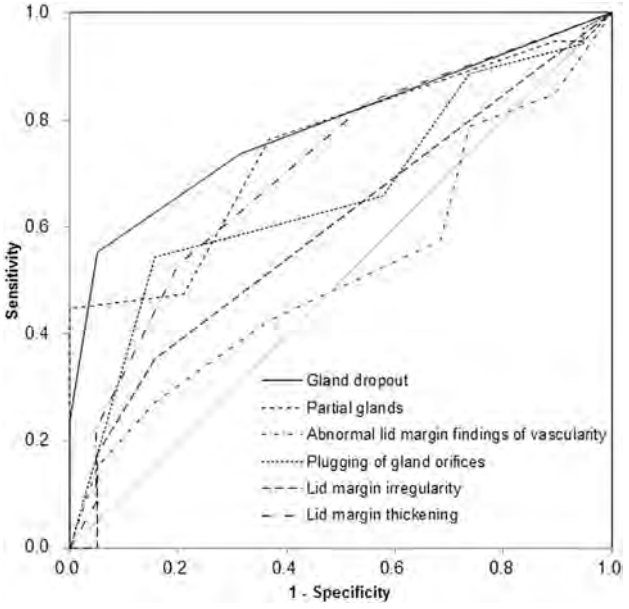

FIGURE 8. Receiver operating characteristic curves for the 6 new grading scales.

basis of AUC values. A previous study examined the diagnostic ability of the total score for 4 eyelid signs (vascular engorgement, plugged meibomian gland orifices, irregular lid margin, and anterior or posterior replacement of the mucocutaneous junction) by calculation of AUC values.[21] We evaluated the AUC for each of these signs (with the exception of mucocutaneous junction) separately. These results can provide useful information for interpretation of scores obtained with the grading scales in clinical practice.

In the present study, we evaluated only 1 eye of each subject; therefore we are unable to address the bilaterality of MGD characteristics. MGD is one of the most common causative conditions of dry eye.[27] Application of the grading scales to both eyes of an individual would provide information on the bilaterality of MGD or dry eye.

There is a general consensus regarding the definition of MGD severity based on meibum score, symptoms, and lid margin findings.[11,14] However, the ability to evaluate MGD severity more accurately with reliable objective scales based on standardized criteria is needed. Our proposed new grading system can be used to define MGD severity. It should be possible to classify MGD patients according to disease severity with the use of such a grading system, including multiple static and dynamic evaluation procedures.

There are some limitations to the present study. First, the study was based on static evaluation of photographs and thus did not take into account dynamic observations such as meibum expressibility, including meibum quality and quantity. Given that evaluation of meibum expressibility is one of the most informative procedures regarding the pathogenesis of MGD, movie-based assessment of meibum expressibility and other dynamic features should be performed in the future. Second, the new grading scales were developed with the use of images obtained in a cross-sectional analysis. We were therefore not able to evaluate whether they are suitable for detection of changes in clinical findings. It will be necessary to confirm that the grading scales are appropriate for evaluation of medical treatment in longitudinal studies. Third, we also evaluated the reliability of the grading scales with the same set of images. Given that plugging of gland orifices and lid margin irregularity can be assessed accurately only by physical examination with observation in 3 dimensions, the assessment of these signs from 2-dimensional images may be problematic. There is thus still

room for improvement of the proposed clinical scoring for MGD, but, despite its basis in static evaluation, it is potentially sufficient for clinical application. Its further improvement will be required for it to become a standard procedure for MGD diagnosis.

In conclusion, we have developed new grading scales for MGD that are sufficiently reliable for use by any ophthalmologist. Further studies are warranted to determine whether the grading scales are suitable for evaluation of the efficacy of MGD management.

FUNDING/SUPPORT: THIS STUDY WAS SUPPORTED BY SANTEN PHARMACEUTICAL CO LTD (OSAKA, JAPAN). THE SPONSOR participated in study design; in the collection, analysis, and interpretation of data; in the writing of the report; and in the decision to submit the article for publication. Financial disclosures: Reiko Arita has received personal fees from Santen Pharmaceutical Co Ltd (Osaka, Japan), Otsuka Pharmaceutical Co Ltd (Tokushima, Japan), Kiribai Chemical Co Ltd (Osaka, Japan), Kao Corporation (Tokyo, Japan), Japan Focus Company Ltd (Tokyo, Japan), TOPCON Japan (Tokyo, Japan), and TearScience (Morrisville, NC). Naoyuki Morishige has received personal fees from Santen Pharmaceutical Co Ltd (Osaka, Japan), Otsuka Pharmaceutical Co Ltd (Tokushima, Japan), Kao Corporation (Tokyo, Japan), Japan Focus Company Ltd (Tokyo, Japan), Pfizer Japan Inc (Tokyo, Japan), Wakamoto Co Ltd (Tokyo, Japan), and Astellas Pharma Inc (Tokyo, Japan). Shima Fukuoka has received personal fees from Santen Pharmaceutical Co Ltd (Osaka, Japan). Izumi Minoura, Kei Asai, Tateki Goto, Takahiro Imanaka, and Masatsugu Nakamura are employees of Santen Pharmaceutical Co Ltd (Osaka, Japan). The following author has no financial disclosures: Rika Shirakawa. All authors attest that they meet the current ICMJE criteria for authorship.

REFERENCES

1. Mishima S, Maurice DM. The oily layer of the tear film and evaporation from the corneal surface. *Exp Eye Res* 1961;1:39–45.
2. Lemp MA. Report of the National Eye Institute/Industry Workshop on Clinical Trials in Dry Eyes. *CLAO J* 1995; 21(4):221–232.
3. Mathers WD. Ocular evaporation in meibomian gland dysfunction and dry eye. *Ophthalmology* 1993;100(3):347–351.
4. Shimazaki J, Sakata M, Tsubota K. Ocular surface changes and discomfort in patients with meibomian gland dysfunction. *Arch Ophthalmol* 1995;113(10):1266–1270.
5. Lee SH, Tseng SC. Rose bengal staining and cytologic characteristics associated with lipid tear deficiency. *Am J Ophthalmol* 1997;124(6):736–750.
6. Mathers WD, Shields WJ, Sachdev MS, et al. Meibomian gland dysfunction in chronic blepharitis. *Cornea* 1991; 10(4):277–285.
7. Korb DR, Henriquez AS. Meibomian gland dysfunction and contact lens intolerance. *J Am Optom Assoc* 1980;51(3): 243–251.
8. Henriquez AS, Korb DR. Meibomian glands and contact lens wear. *Br J Ophthalmol* 1981;65(2):108–111.
9. Zengin N, Tol H, Gunduz K, et al. Meibomian gland dysfunction and tear film abnormalities in rosacea. *Cornea* 1995; 14(2):144–146.
10. Ong BL. Relation between contact lens wear and meibomian gland dysfunction. *Optom Vis Sci* 1996;73(3):208–210.
11. Nelson JD, Shimazaki J, Benitez-del-Castillo JM, et al. The International Workshop on Meibomian Gland Dysfunction: Report of the Definition and Classification Subcommittee. *Invest Ophthalmol Vis Sci* 2011;52(4):1930–1937.
12. Tomlinson A, Bron AJ, Korb DR, et al. The International Workshop on Meibomian Gland Dysfunction: Report of the Diagnosis Subcommittee. *Invest Ophthalmol Vis Sci* 2011;52(4):2006–2049.
13. Bron AJ, Benjamin L, Snibson GR. Meibomian gland disease. Classification and grading of lid changes. *Eye* 1991;5(4): 395–411.
14. Foulks GN, Bron AJ. Meibomian gland dysfunction: a clinical scheme for description, diagnosis, classification and grading. *Ocul Surf* 2003;1(3):107–126.
15. Hykin PG, Bron AJ. Age-related morphological changes in lid margin and meibomian gland anatomy. *Cornea* 1992; 11(4):334–342.
16. Shimazaki J, Goto E, Ono M, et al. Meibomian gland dysfunction in patients with Sjogren syndrome. *Ophthalmology* 1998; 105(8):1485–1488.
17. Pflugfelder SC, Tseng SC, Sanabria O, et al. Evaluation of subjective assessments and objective diagnostic tests for diagnosing tear-film disorders known to cause ocular irritation. *Cornea* 1998;17(1):38–56.
18. Nichols JJ, Berntsen DA, Mitchell GL, Nichols KK. An assessment of grading scales for meibography images. *Cornea* 2005;24(4):382–388.
19. Amano S, Arita R, Kinoshita S, et al. The Japan Dry Eye Society Meibomian Gland Dysfunction Working Group. Definition and diagnostic criteria for meibomian gland dysfunction. *Atarashii Ganka (J Eye)* 2010;27(5): 627–631.
20. Arita R, Itoh K, Maeda S, et al. Proposed diagnostic criteria for obstructive meibomian gland dysfunction. *Ophthalmology* 2009;116(11):2058–2063.
21. Yamaguchi M, Kutsuna M, Uno T, et al. Marx line: fluorescein staining line on the inner lid as indicator of meibomian gland function. *Am J Ophthalmol* 2006;141(4):669–675.
22. Van Bijsterveld OP. Diagnostic tests in the sicca syndrome. *Arch Ophthalmol* 1969;82(1):10–14.
23. Arita R, Itoh K, Inoue K, Amano S. Noncontact infrared meibography to document age-related changes of the meibomian glands in a normal population. *Ophthalmology* 2008; 115(5):911–915.
24. Landis JR, Koch GC. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159–174.
25. Pult H, Riede-Pult BH, Nichols JJ. Relation between upper and lower lids' meibomian gland morphology, tear film, and dry eye. *Optom Vis Sci* 2012;89(3):E310–E315.
26. Koh S, Ikeda C, Fujimoto H, et al. Regional differences in tear film stability and meibomian glands in patients with aqueous-deficient dry eye. *Eye Contact Lens* 2016;42(4):250–255.
27. Lemp MA, Crews LA, Bron AJ, Foulks GN, Sullivan BD. Distribution of aqueous-deficient and evaporative dry eye in a clinic-based patient cohort: a retrospective study. *Cornea* 2012;31(5):472–478.