Clinically Relevant Immune-Cellular Metrics of Inflammation in Meibomian Gland Dysfunction

Yureeda Qazi,[1] Ahmad Kheirkhah,[1] Caroline Blackie,[2] Monique Trinidad,[1] Candice Williams,[1] Andrea Cruzat,[1] Donald R. Korb,[2] and Pedram Hamrah[1,3]

[1]Ocular Surface Imaging Center, Cornea Service, Massachusetts Eye and Ear Infirmary, Department of Ophthalmology, Harvard Medical School, Boston, Massachusetts, United States
[2]Korb and Associates, Boston, Massachusetts, United States
[3]Center for Translational Ocular Immunology and Cornea Service, New England Eye Center/Tufts Medical Center, Department of Ophthalmology, Tufts University School of Medicine, Boston, Massachusetts, United States

Correspondence: Pedram Hamrah, Cornea Service, New England Eye Center/Tufts Medical Center, Department of Ophthalmology, Tufts University School of Medicine, 800 Washington Street, Boston, MA 02111, USA; phamrah@tuftsmedicalcenter.org, p_hamrah@yahoo.com.

Submitted: August 23, 2018
Accepted: October 18, 2018

Citation: Qazi Y, Kheirkhah A, Blackie C, et al. Clinically relevant immune-cellular metrics of inflammation in meibomian gland dysfunction. *Invest Ophthalmol Vis Sci*. 2018;59:6111–6123. https://doi.org/10.1167/iovs.18-25571

Purpose. To determine the reliability and clinical relevance of in vivo confocal microscopy (IVCM)-based immune-cellular metrics of palpebral conjunctival inflammation in meibomian gland dysfunction (MGD).

Methods. Sixteen MGD patients and 13 reference controls included in this cross-sectional, retrospective study, had an ocular surface exam, symptom assessment (Ocular Surface Disease Index questionnaire [OSDI]), and palpebral conjunctival IVCM imaging. Bland-Altman analyses, intraclass correlation coefficient (ICC_a), Lin's concordance correlation coefficient (ρ_c), receiver operating characteristic (ROC) analyses, and correlations were performed. Clinical outcome measures were symptom severity (OSDI scores), tear break-up time (TBUT), and corneal fluorescein staining (CFS grade).

Results. Compared to controls, patients with MGD had variable symptom severity (average OSDI score: 48.3 ± 7.6, $P = 0.0008$, range: 8.3–85.42), shorter TBUT (6.8 ± 0.9 seconds, $P = 0.002$), comparable corneal staining (0.31 ± 0.19, $P = 0.20$), and greater conjunctival inflammation (epithelial immune cells [EIC]: 477.8 ± 54.2 vs. 123.3 ± 17.2 cells/mm^2, $P < 0.0001$; intraglandular immune cells [IGIC]: 41.9 ± 3.3% vs. 20.33 ± 7.3%, $P < 0.01$). Immune-cellular metrics had high inter- and intraobserver agreement (ρ_c: 0.86–0.94; ICC_a and Cronbach's α: 0.85–0.97, $P < 0.0001$). EIC correlated positively with OSDI (r_s: 0.49, $P = 0.03$), while both EIC and IGIC correlated inversely with TBUT (r_s: −0.47, −0.45, $P < 0.05$), and had high accuracy in detecting inflammation (ROC area under the curve [AUC]: 0.97 and 0.89, $P \leq 0.001$).

Conclusions. EIC and IGIC are increased in highly symptomatic patients with MGD that have minimal corneal staining, and correlate with symptoms and clinical signs. EIC and IGIC may provide reliable and clinically relevant metrics of inflammation.

Keywords: MGD, inflammation, confocal microscopy, dry eye

The recently highlighted role of inflammation in the pathogenesis of meibomian gland dysfunction (MGD)[1] makes the detection and assessment of inflammation vital for the evaluation of patients with dry eye disease (DED) and assessment of therapeutic efficacy of anti-inflammatory therapies in these patients. The development of objective cellular metrics becomes especially necessary since an estimated 40% of patients with MGD either are asymptomatic,[1] or have symptoms but an unremarkable clinical examination, forming the complex and challenging group of dry eye patients with symptom–sign disparity.[2–5] Our group recently studied such MGD patients that presented with persistent symptoms of discomfort in the presence of a normal slit-lamp examination after treatment with gland expression procedures for MGD. We demonstrated that these MGD patients with refractory symptoms despite an improved examination had clinically non-apparent inflammation of the palpebral conjunctiva, evident on laser scanning in vivo confocal microscopy (IVCM).[5] Compared to MGD patients that had improved symptoms following anti-inflammatory treatment, patients with refractory symptoms had a 3-fold increase in palpebral conjunctival epithelial immune cell density (EIC),[5] despite the absence of apparent clinical signs on slit-lamp examination, furthering our understanding of the pathogenesis underlying symptom–sign disparity in MGD patients.

Among the quantitative methods currently available for evaluating the palpebral conjunctiva and meibomian glands, only IVCM allows direct visualization[6–8] and quantitation of immune cells (ICs). Recently, Zhou and Robertson[9] used immunohistochemistry and IVCM of human eyelid sections to assess the validity of previously published papers to date. They demonstrated that previous quantitative and qualitative IVCM studies that presumably analyzed meibomian glands to establish diagnostic utility, classifications and grading of MGD, correlations to signs and symptoms, and meibomian gland changes in systemic and other ocular diseases,[7,10–24] were actually studying rete ridges at the dermal–epidermal junction of the free lid margin, which also have an acinar appearance on

IVCM.[9] Further, Knop and colleagues[25] had shown similar findings through hematoxylin and eosin sections and IVCM when describing the lid wiper and mucocutaneous junction. Thus, to date, IVCM has been applied to quantitatively assess the lid margin, focusing on dendritiform cells in superficial epithelia and around the acinar-like cross sections of the rete ridges, not the meibomian glands,[6,10–14,16–18,26] in addition to one qualitative report of meibomian glands within the substantia propria.[8] Therefore, in alignment with the evidence, we believe that this study now demonstrates quantitative, layer-by-layer assessment of inflammation in both the palpebral conjunctiva and meibomian glands. Consequently, this is, to our knowledge, the first detailed report on the diagnostic utility, reliability, and clinical relevance of IVCM-based palpebral conjunctival epithelial, stromal (substantia propria), and meibomian glandular metrics of inflammation.

In the current study, we developed, validated, and correlated quantitative novel immune-cellular metrics for the assessment of inflammation, which is often nonobvious in MGD. We developed and tested a 10-metric panel focused on layer-by-layer quantitation of palpebral conjunctival inflammation in MGD by IVCM. We established inter- and intraobserver agreement, consistency, reproducibility, and concordance for each of the 10 in vivo metrics. Clinical relevance of these parameters was assessed by correlation with clinical signs (tear break-up time [TBUT], corneal fluorescein staining grade [CFS]), and symptom severity (Ocular Surface Disease Index score [OSDI]). Palpebral conjunctival IVCM-based immune-cellular metrics of inflammation proved to be consistent, reproducible, accurate, and clinically relevant in MGD. Palpebral conjunctival EIC and intraglandular immune cells (IGIC) correlated directly with symptom severity (OSDI) and inversely with tear film stability (TBUT). Receiver operating characteristic (ROC) analyses illustrated very good to excellent accuracy of EIC and IGIC in detecting eyelid inflammation. EIC and IGIC may therefore serve as clinically relevant endpoints of inflammation in evaluating MGD patients for palpebral conjunctival inflammation and monitoring their response to treatment in clinical trials and clinical practice.

Materials and Methods

Study Design and Patient Population

We conducted a retrospective, cross-sectional, observational, controlled study, investigating both eyes of 16 MGD patients seen at the Cornea Service of the Massachusetts Eye and Ear Infirmary (MEEI), Boston, Massachusetts, between 2011 and 2013 as part of routine care. This study was approved by the MEEI institutional review board (IRB), was Health Insurance Portability and Accountability Act (HIPAA) compliant, and adhered to the tenets of the Declaration of Helsinki. One eye of 13 healthy, asymptomatic, age- and sex-matched individuals from our normative research database were used as reference controls.

With respect to inclusion criteria, patients included in this study had a clinical diagnosis of MGD based on clinical symptoms of ocular discomfort and either lid findings on clinical slit-lamp examination (irregular lid margin, telangiectasia, Meibomian Glands Yielding Liquid Secretion score) or a TBUT of less than 10 seconds (s). Refractory MGD patients with persistent symptoms despite clinical improvement in TBUT after prior treatment (TBUT $>$ 10 s) who were referred to us were also included.

As to exclusion criteria, neuropathic corneal pain was ruled out in all MGD patients, both clinically (positive response to topical anesthetic in any patient with complaints of corneal pain), and through imaging of corneal subbasal nerves using IVCM. Hence, our study population did not have neuropathic corneal pain, which can be a confounding factor in the assessment of patients with symptom–sign disparity.

All reference controls in this study were healthy, asymptomatic individuals who had no CFS and a TBUT $>$ 10 s. All controls were drawn from an IRB-approved prospective normative study database that enrolled healthy subjects after having a complete history and ocular examination comprising CFS, TBUT, anterior segment examination, applanation tonometry, and corneal sensation. Only subjects who had an unremarkable examination were recruited to the database.

All associated imaging was performed at the MEEI Ocular Surface Imaging Center (OSIC). Medical charts and study forms were reviewed for medical history and details of anterior segment slit-lamp examination findings including TBUT and CFS. Patient symptom severity scores from the OSDI questionnaire were also reviewed.[27–29] Two masked observers analyzed IVCM images of the palpebral conjunctiva acquired as part of routine patient care.

IVCM

Laser scanning IVCM (Heidelberg Retinal Tomograph 3 with the Rostock Cornea Module [HRT3/RCM]; Heidelberg Engineering GmbH, Heidelberg, Germany) had been used to capture layer-by-layer en face images of the palpebral conjunctival epithelium, using the method described below. The HRT3/RCM IVCM provided a field of view of 400 $\times$ 400 μm, with a lateral resolution of 1 to 2 μm and an axial resolution of 4 μm.

Both eyes of all MGD patients and one eye of healthy asymptomatic controls were topically anesthetized using one drop of 0.5% proparacaine hydrochloride (Alcaine; Alcon, Fort Worth, TX, USA) per eye followed by application of hydroxypropyl methylcellulose 0.3% gel (GenTeal Gel; Alcon Laboratories, a Novartis Company, Fort Worth, TX, USA). Prior to imaging each patient, the base of a sterile, disposable polymethylmethacrylate cap (Tomo-Cap; Heidelberg Engineering) was filled halfway with hydroxypropyl methylcellulose 0.3% (GenTeal Gel) and mounted on the exterior of the HRT3/RCM optical lens. To enhance optical coupling, a drop of hydroxypropyl methylcellulose 0.3% was also placed on the external surface of the Tomo-Cap at its tip. The seated patient's chin and forehead were carefully and comfortably placed firmly against the chin and forehead rests. The tip of a cotton swab was placed on the skin of the lower eyelid, parallel to and below the lid margin. With gentle downward pressure on the cotton swab, the lower eyelid was everted and the camera was manually advanced until the gel on the tip of the Tomo-Cap made contact with the palpebral conjunctival mucosa. The camera was moved as far laterally as possible while maintaining contact with the palpebral conjunctiva. Once epithelial and goblet cells of the palpebral conjunctiva were visualized, the depth of the scan was set to 0 μm (reference depth) and scanning was started in sequence mode, which provided a dynamic movie mode of up to 100 frames at variable depths and an acquisition speed of 30 frames/second.

At each point (lateral, central, medial) along the palpebral conjunctiva images were acquired, starting at a depth of 5 μm (epithelium) and advancing deeper into the tissue up to a depth of 200 μm into the substantia propria. Transition from the epithelium into the substantia propria was identified by the presence of a fibrous, amorphous background punctuated with clusters of circular acinar-like structures and vessels. With further advancement into the substantia propria,

Investigative Ophthalmology & Visual Science

Table 1. IVCM-Based Metrics for Quantitative Evaluation of Palpebral Conjunctival Epithelium, Substantia Propria, and Glands in MGD-Associated Ocular Inflammation

Parameter	Unit of Measurement	Significance
Epithelial immune cell density, EIC	Cells/mm^2	Location and extent of inflammation; status of immune activation on the surface of palpebral conjunctiva
Stromal immune cell density, SIC	Cells/mm^2	Location and extent of inflammation; status of immune activation in the substantia propria of palpebral conjunctiva
Periglandular immune cells, PGIC	No. of cells within 20-μm radius around the gland	Location and extent of inflammation; status of immune activation in the substantia propria of palpebral conjunctiva
Intraglandular immune, IGIC	% of glandular ductule luminal area occupied by immune cells	Extent and infiltration of inflammation; degree of plugging and occlusion of glandular ductules with visible immune cellular content
Luminal width of ductules with visible intraglandular content, IGIC$^+$W	μm	Morphologic alterations in glandular architecture
Luminal length of ductules with visible intraglandular content, IGIC$^+$L	μm	Morphologic alterations in glandular architecture
Thickness of hyperreflective luminal ring in ductules with visible intraglandular content, IGIC$^+$r	μm	Reactive proliferative or atrophic activity within glandular ductule
Luminal width of ductules without visible intraglandular content, IGIC$^-$W	μm	Morphologic alterations in glandular architecture
Luminal length of ductules without visible intraglandular content, IGIC$^-$L	μm	Morphologic alterations in glandular architecture
Thickness of hyperreflective luminal ring in ductules without visible intraglandular content, IGIC$^-$r	μm	Reactive proliferative or atrophic activity within glandular ductule

glandular ductules were visualized. Image of the palpebral conjunctiva at depths greater than 200 μm could not be visualized. Once the inferolateral region of the everted lower eyelid had been imaged, the depth of focus was receded back to the epithelium, and the lens with the Tomo-Cap was glided medially to the center of the everted lower eyelid. Before beginning image acquisition at the new position, reference depth was reset to 0 μm and the region was imaged again. Upon completing imaging of the central palpebral conjunctiva, the lens was advanced to the most medial (nasal) position possible without losing contact with the palpebral conjunctival epithelium. At the new position, the reference depth was reset to 0 μm upon visualization of the epithelial cells, before proceeding to image the deeper layers. Once the three regions (lateral, central, medial) at the inferior edge of the everted eyelid in the horizontal plane had been surveyed, the lens was advanced superiorly along the conjunctiva, moving toward the fornix, and repositioned at the lateral margin in the new horizontal plane. The steps described above were repeated until the majority of the everted lower eyelid had been imaged. A total of 5 to 12 sequence scans of 100 images each were acquired per eyelid and stored on a network computer.

Patients who could tolerate eversion of the upper eyelid with a cotton swab received an additional drop of topical anesthetic and hydroxypropyl methylcellulose 0.3% gel. The same procedure described for lower eyelid imaging was applied to the everted upper eyelid, with the exception that after completing image acquisition at each of the three points (lateral, central, medial) along the initial horizontal plane, the camera was now advanced inferiorly along the eyelid toward the fornix. For upper eyelid imaging, an assistant helped keep the eyelid everted with the tip of a cotton swab, securing the lashes against the supraorbital arch.

Image Analysis

Two masked observers made all the measurements using ImageJ software (https://imagej.nih.gov/ij/; provided in the public domain by National Institutes of Health, Bethesda, MD, USA). A single observer selected representative images for analysis. The criteria for selection of images included good focus for structures of interest, visualization of the structure(s) of interest, absence of motion artifacts, and without regions of overexposure or hyperreflectivity that may make differentiating structures difficult. Three images per parameter per eye were analyzed for each subject totaling quantitative analysis of 1320 images. For patients, measurements from both eyes of each patient were averaged to represent a single sample. All results were reported as a mean of the measurements made by both the observers and expressed as mean ± standard error of mean (SEM) unless noted otherwise.

Palpebral conjunctival images were analyzed for EIC, stromal IC (SIC), IGIC, and gland ductular luminal dimensions (Table 1; Fig. 1). For en face images, the whole frame (400 × 400 μm) was analyzed, counting all cells within the frame and expressing density as the number of cells/mm^2. For oblique images, immune cells were counted only within the region of interest (ROI). The ROI was selected and measured using the polygon tool in ImageJ. Immune cells were identified as hyperreflective, polymorphous (dendritiform, nondendritiform, spindle-shaped) cellular structures within the epithelium and substantia propria, ranging in size from 5 to 20 μm.[30] Intraductular immune cell analysis was performed on all glandular ductules within each frame that had visible, hyperreflective, intraluminal content. Luminal occlusion by IGIC within ductules was reported as mean percentage occlusion. Ductular dimensions of internal luminal width (W) and length (L) were measured for both glandular ductules with (IGIC$^+$) and without (IGIC$^-$) visible, hyperreflective, intraluminal

Investigative Ophthalmology & Visual Science

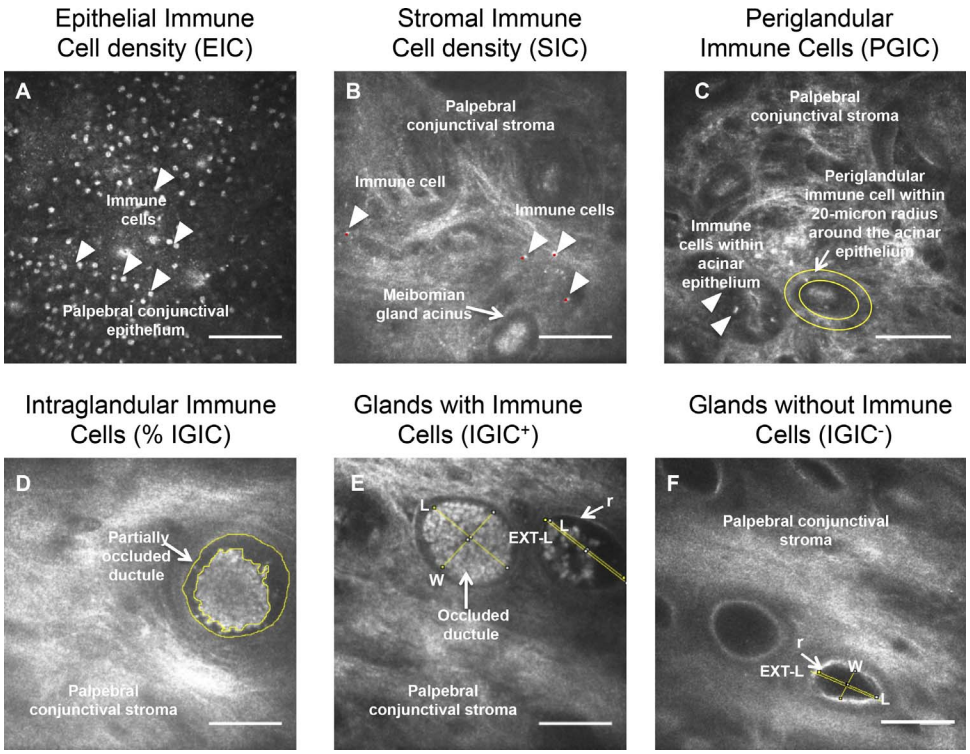


Figure 1. Image analysis of palpebral conjunctival in vivo confocal micrographs using ImageJ. En face in vivo confocal micrographs were analyzed using ImageJ. Immune cells were identified as hyperreflective, polymorphous structures ranging in size from 5 to 20 μm (*red dots*, *arrowheads*; **A**, **B**). EIC (**A**) and SIC (**B**) were expressed as number of cells per mm^2; PGIC (**C**) were identified as the number of cells within a 20-μm radius around the glands; IGIC (**D**) were quantified as the percentage intraluminal area occupied by immune cells; $IGIC^+$ (**E**) and $IGIC^-$ (**F**) were assessed for lumen width (*W*), lumen internal length (*L*), glandular external length (*EXT-L*), and thickness of luminal hyperreflective ring presumed to be ductular epithelium (*r*) (*arrow*; $r = [EXT\text{-}L - L]/2$). *Scale bars*: 100 μm.

content. All ductules within each frame were analyzed. Thickness of the internal luminal ring (r) lining the ductule was calculated for each of these ductules per frame as described below.

Epithelial Immune Cell (EIC) Density. In the analysis of EIC (Fig. 1A), particular attention was paid to avoid counting goblet cells and epithelial cells that typically appear larger and less hyperreflective than immune cells.[31] All cells were counted within an en face frame, or within the ROI in oblique frames, and reported as cells/mm^2.

Stromal Immune Cell (SIC) Density. In the analysis of SIC (Fig. 1B), particular attention was paid not to count immune cells within acinar epithelium and lumens of ductules, blood and lymphatic vessels, as they were not the interstitial matrix of the palpebral conjunctiva. All cells were counted within an en face frame, or within the ROI in oblique frames, and reported as cells/mm^2.

Periglandular Immune Cells (PGIC). All cells within a 20-μm radius around the acinar epithelium for each acinus were counted and averaged per frame (Fig. 1C). Careful attention was paid not to include intraepithelial immune cells located within the acinar epithelium (Fig. 1C). Results were reported as the mean number of cells per 20-μm radius of the external acinar epithelium.

Intraglandular Immunce Cells (IGIC). The intraluminal area occupied by cells within each glandular ductule and the total luminal area of each glandular ductule (Fig. 1D) were measured for all ductules within every frame and expressed as mean percentage occlusion.

Luminal Width of Ductules With ($IGIC^+$W) and Without ($IGIC^-$W) Visible Intraglandular Content. The internal luminal width (W) of every glandular ductule with and without visible, hyperreflective, intraglandular content within each frame was measured (Figs. 1E, 1F), averaged, and reported in microns. The caliper was placed perpendicularly to that used for measuring internal length (L) of the ductules.

Luminal Length of Ductules With ($IGIC^+$L) and Without ($IGIC^-$L) Visible Intraglandular Content. The internal luminal length (L) of every glandular ductule with and without visible, hyperreflective, intraglandular content within each frame was measured (Figs. 1E, 1F) averaged, and reported in microns. The caliper was placed perpendicularly to that used for measuring internal width (W) of the ductules.

Thickness of Internal Luminal Hyperreflective Ring (r) in Ductules With ($IGIC^+$r) and Without ($IGIC^-$r) Visible Intraglandular Content. After placing calipers measuring the internal luminal length of the glandular ductule (L), the same caliper was extended at each end to encompass the external length (EXT-L) of the ductule to ensure that the plane of measurement was constant (Figs. 1E, 1F). To compute thickness of the internal luminal hyperreflective ring (r), presumed to be ductular epithelium, this equation was applied: (external length − luminal length)/2. These measurements were performed for all glandular ductules within each frame, with and without visible intraglandular content. The results were averaged per frame and reported in microns.

Statistical Analysis

Normality of data was determined using the Shapiro-Wilk normality test based on which either parametric (Student's independent samples 2-tailed *t*-test) or nonparametric tests (Mann-Whitney *U* test) were applied for intergroup comparisons. Effect size was computed for all parameters that had a

Investigative Ophthalmology & Visual Science

Table 2. Demographics and Clinical Characteristics of Controls and Patients With MGD

	Age, y, Mean ± SEM	Sex, M:F	OSDI Score Mean ± SEM (Range)	Tear Film Osmolarity, mOsm/L, Mean ± SEM	TBUT, s, Mean ± SEM	CFS Grade Mean ± SEM	*n*, Patients
Controls	43.23 ± 3.99	6:7	12.00	303.8 ± 4.2	10.97 ± 0.69	0	13
MGD	54.63 ± 3.83	2:14	49.48 ± 6.97 (8.30–85.42)	303.9 ± 6.0	6.84 ± 0.93	0.31 ± 0.19	16
P value	0.051	0.09	0.0008	0.99	0.002	0.20	
U value	NA	NA	13		33.5	39	

Both controls and patients with MGD were age- and sex-matched. *P* value for the difference in sex ratio between the groups represents the *F* statistic for Fisher's exact test. MGD patients had a shorter TBUT (seconds, $P = 0.002$), and a wide range of symptom severity the mean of which was significantly higher than that for asymptomatic, healthy controls (OSDI score, $P = 0.0008$). There were no differences in the tear film osmolarity ($P = 0.20$) and CFS grade (NEI grading system, $P = 0.99$) between the two groups. Data are reported as mean ± SEM. A *P* value less than 0.05 was considered statistically significant.

normal distribution (parametric) using Cohen's d, Glass's delta, and Hedge's g statistics. Post hoc power was determined for all the parameters. Sex-matching was evaluated by Fisher's exact test. A *P* value less than 0.05 was considered statistically significant.

Interobserver reliability was determined for each of the 10 imaging parameters by the Bland-Altman analysis and plots (agreement),[32] intraclass correlation coefficient of agreement (ICC_a, reproducibility), Cronbach's alpha (α, internal consistency), and Lin's concordance correlation coefficient (ρ_c, reproducibility with bias correction). ICC_a was computed using a 2-way mixed model and stepped for agreement. For LCCC, a prewritten syntax (httP =//gjyp.nl/marta/Lin.sps; Marta Garcia-Granero; in the public domain) was modified for our datasets and run in the Statistical Program for Social Sciences (IBM Corp. Released 2013; IBM SPSS Statistics for Windows, Version 22.0; Armonk, NY, USA). As part of Bland-Altman analysis, in addition to the 95% limits of agreement (LoA), statistical significance and 95% confidence interval (CI) of the mean interobserver difference (bias) were also computed using the one-sample *t*-test (2-tailed). To detect proportional bias in the differences in measurements between observers, linear regression analysis was applied. If a proportional bias was detected, log transformation was performed, after which the results were back-transformed.[33] Correlation between imaging-based immune-cellular metrics (EIC, IGIC, PGIC) and clinical parameters (OSDI, TBUT, CFS) was determined using Spearman's rank-order correlation coefficient (r_s) and coefficient of determination (r^2). Observed power analyses were performed for all the correlation coefficients.

The diagnostic utility of EIC, and IGIC was determined by performing ROC curve analyses including area under the curve (AUC), its 95% confidence limits, and accuracy ratio (AR). As part of ROC curve analysis, sensitivity, specificity, positive and negative predictive values (PPV, NPV), positive and negative likelihood ratios (LR+, LR−), accuracy, diagnostic odds ratio (DOR), and Youden's index (J) were calculated to determine potential cutoff values. Since SPSS cannot calculate confidence intervals for proportions, we used an online calculator (httP =//vassarstats.net/clin1.html; in the public domain) to calculate the 95% confidence limits for the measures of diagnostic accuracy.

Results

Demographics and Clinical Profile

Both patients with MGD and normal controls were age- (controls: 43.2 ± 3.9 years, MGD: 54.6 ± 3.8 years, $P = 0.051$) and sex-matched (*F* test statistic: 0.092). Our sample of patients with MGD had wide variation in symptom severity ranging from being asymptomatic to severely symptomatic (OSDI score range: 8.3–85.42, mean OSDI score: 49.48 ± 6.97, $P = 0.0008$), a significant reduction in TBUT (controls: 10.9 ± 0.7 seconds, MGD: 6.8 ± 0.9 seconds, $P = 0.002$), negative Schirmer test (10.8 ± 2.3 mm), insignificant CFS (controls: 0, MGD: 0.31 ± 0.19, $P = 0.20$), and normal values for tear film osmolarity (controls: 303.8 ± 4.2 mOsm/L, MGD: 303.9 ± 6.0 mOsm/L, $P = 0.99$). Details of the demographic profile and clinical characteristics of the study groups are listed in Table 2.

Qualitative Assessment of MGD-Associated Palpebral Conjunctival Inflammation by IVCM

Qualitatively, there were several notable differences between the palpebral conjunctival tissue and glands of healthy controls, and patients with MGD (Fig. 2). As compared to controls (Fig. 2A), the palpebral conjunctiva in patients with MGD had greater infiltration of EIC (Fig. 2B), seen as an increased number of hyperreflective polymorphous structures. Similarly, progressing to the deeper layers of the palpebral conjunctiva, in comparison to controls (Fig. 2C), patients with MGD (Fig. 2D) had increased speckling of the substantia propria, caused by an increase in SIC. Fibrosis was also seen around acini. Some ICs were found immediately adjacent to the external acinar epithelia (Figs. 2E, 2F). These PGICs were distinct from intraepithelial immune cells previously described in the literature,[34] which lie specifically within the epithelium of the acini; such intraepithelial immune cells were also observed in our sample of MGD patients (Fig. 2F). Of particular interest was the presence of hyperreflective, nondendritic, circular structures within the lumens of conjunctival glandular ductules (presumed meibomian glands) both in controls, and patients with MGD (Figs. 2G, 2H). This intraglandular content was increased in MGD patients as compared to controls, with greater occlusion and distension of the ductular lumen in MGD (Fig. 2H).

In Vivo Metrics of Palpebral Conjunctival and Glandular Inflammation in MGD Using IVCM

Immune-Cellular Metrics. Of the four metrics (EIC, SIC, PGIC, IGIC) devised to quantify immune cells in the palpebral conjunctival epithelium, substantia propria, and glands, EIC and IGIC demonstrated the most significant differences between healthy controls and MGD (Table 3); EIC and IGIC had very large effect sizes (EIC: 2.0–5.3, IGIC: 0.9–1.5), and an observed power of 100%. As compared to controls (EIC: 123.3 ± 17.2 cells/mm^2, IGIC: 20.33 ± 7.3%), MGD patients had a near 4-fold increase in epithelial immune cells (EIC: 477.8 ± 54.2 cells/mm^2, $P < 0.0001$), and a 2-fold increase in gland intraductular occlusion (IGIC: 41.9 ± 3.3%, $P < 0.01$). In comparison to controls (PGIC: 0.9 ± 0.4 cells/20-μm radius), PGIC nearly

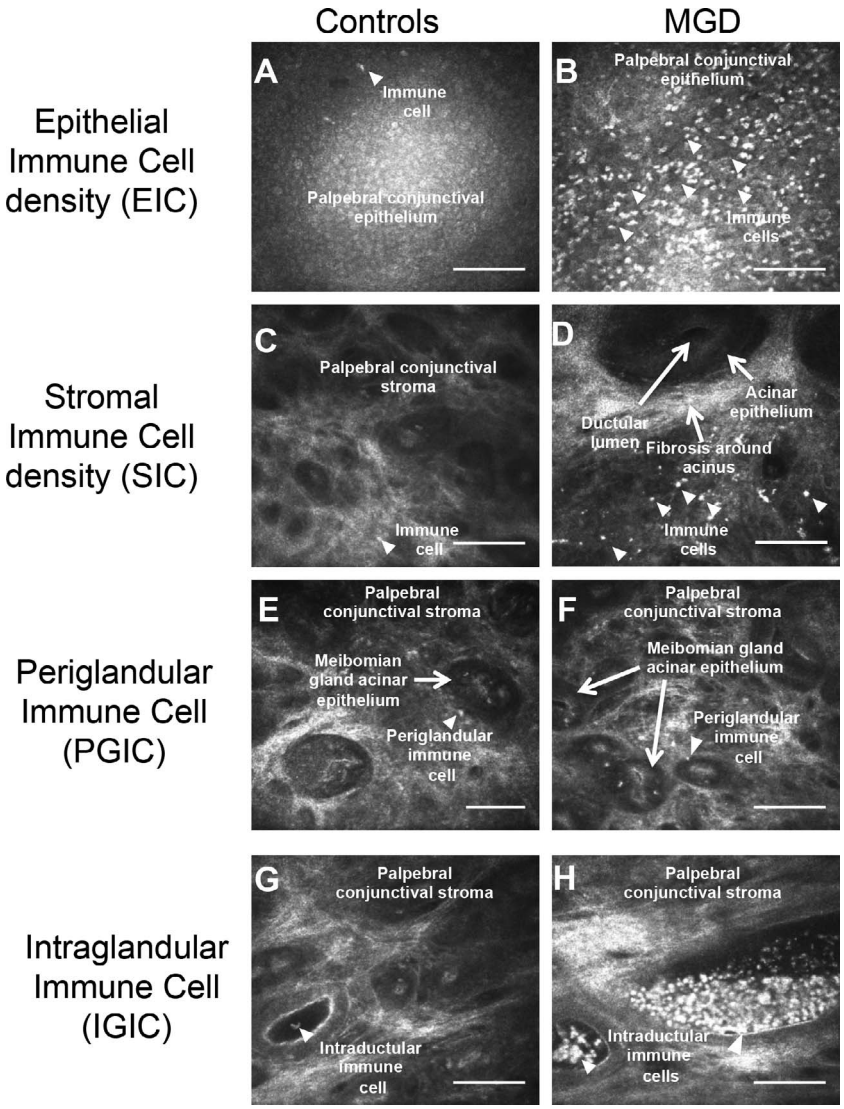


Figure 2. IVCM of the palpebral conjunctiva in controls and patients with MGD-associated palpebral conjunctival inflammation. Representative en face in vivo confocal micrographs of healthy controls (**A**, **C**, **E**, **G**) and patients with MGD (**B**, **D**, **F**, **H**), illustrating increased immune cells in the palpebral conjunctival epithelium (**B**) and stroma with periacinar fibrosis (**D**), periglandular cells around the acinar epithelium (**F**), and increased occlusion of presumed meibomian gland ductules leading to altered ductular dimensions in MGD (**H**). *Scale bars*:100 µm.

doubled in MGD (PGIC: 2.6 ± 0.5 cells/20-µm radius, $P = 0.07$, power: 67%), and this increase approached but did not achieve statistical significance. Among the glandular ductules without visible cellular content, those in MGD had larger luminal dimensions as compared to healthy controls (IGIC¯W: 12.3 ± 0.9 vs. 6.4 ± 3.1 µm, $P = 0.03$; IGIC¯L: 24.1 ± 1.4 vs. 14.7 ± 5.9 µm, $P = 0.03$). These dimensions also had large effect sizes (IGIC¯W: 0.9–1.2, IGIC¯L: 0.7–1.2; Table 3) and an observed power of 99% to 100%. Furthermore, when all glands with visible IGIC (controls and MGD patients combined) were compared to all glands without visible IGIC (controls and MGD patients combined), it was clear that glands with occluded ductules not only had significantly larger luminal dimensions (width: 38.7 ± 5.5 vs. 10.8 ± 1.2 µm, $P < 0.0001$; length: 53.5 ± 6.9 vs. 21.7 ± 1.9 µm, $P < 0.0001$), but also had thinning of the ductular luminal hyperreflective ring (r), presumed to be ductular epithelium (r: 4.3 ± 0.3 vs. 5.4 ± 0.4 µm, $P = 0.04$).

Inter- and Intraobserver Agreement and Reliability of IVCM-Based Metrics of Inflammation. Bland-Altman analysis (Table 4) showed that for all immune-cellular metrics (EIC, SIC, PGIC, IGIC), the mean difference (bias) between interobserver measurements was small (bias: 6.9%–14.9%, EIC < IGIC < SIC < PGIC) and not statistically significant (P = 0.13–0.43). Bland-Altman plots (Fig. 3) illustrated good interobserver agreement for each of these parameters except PGIC, where the 95% LoA were too large to be acceptable for this parameter. All immune cellular measurements were without proportional bias. EIC, SIC, and IGIC had excellent internal consistency (Cronbach's α: 0.93–0.97; Table 4), and high combined inter- and intraobserver agreement (ICC_a: 0.93–0.97; Table 4). ρ_c, which corrects for chance among interobserver measurements, confirmed that among the immune-cellular metrics, EIC, SIC, and IGIC had the strongest reproducibility free of chance (EIC > SIC > IGIC, ρ_c: 0.86–0.94; Table 4), whereas interobserver reproducibility for PGIC was considered substantial but not perfect (ρ_c: 0.72; Table 4).

Glandular Metrics. Bland-Altman analyses (Table 4) showed that the mean difference (bias) between interobserver measurements for glandular ductule luminal width and length was significant for the smaller ductules without IGIC (IGIC¯W,

Table 3. Differences in Immune-Cellular Metrics of Palpebral Conjunctival and Glandular Inflammation in Healthy Controls and MGD Using IVCM

Parameters	Controls, Mean ± SEM	MGD, Mean ± SEM	*P* Value*	Effect Size of Parametric Data: Cohen's d Based on μ	Glass's Delta Based on σ	Hedge's g Based on *n*	Post Hoc Power, %
EIC, cells/mm^2	123.3 ± 17.2	477.8 ± 54.2	***<0.0001***	2.2	5.3	2.0	100
SIC, cells/mm^2	36.7 ± 10.2	66.9 ± 15.9	0.19	–	–	–	27
PGIC, cells/20-μm radius	0.97 ± 0.4	2.6 ± 0.5	0.07	–	–	–	67
IGIC, % occlusion	20.33 ± 7.3	41.9 ± 3.3	**<0.01**	0.9	1.5	0.9	100
IGIC$^+$W, μm	46.7 ± 17.5	35.8 ± 4.2	0.86	–	–	–	45
IGIC$^+$L, μm	62.1 ± 23.3	50.3 ± 4.6	0.68	–	–	–	43
IGIC$^+$r, μm	3.7 ± 0.4	4.5 ± 0.3	0.18	0.04	0.03	0.06	45
IGIC$^-$W, μm	6.4 ± 3.1	12.3 ± 0.9	0.03	1.1	0.9	1.2	99
IGIC$^-$L, μm	14.7 ± 5.9	24.1 ± 1.4	0.03	0.9	0.7	1.2	100
IGIC$^-$r, μm	4.3 ± 0.01	5.6 ± 0.5	0.33	1.0	131	0.8	45

Parameters compared between healthy controls and MGD-associated palpebral conjunctival inflammation were EIC, SIC, PGIC, IGIC, IGIC$^+$W, IGIC$^+$L, IGIC$^+$r, IGIC$^-$W, IGIC$^-$L, and IGIC$^-$r.

* Student's independent samples *t*-test (2-tailed) and Mann-Whitney *U* test (2-tailed) were applied as appropriate to determine the significance of differences in means between controls and patients with MGD. Means are reported as mean ± SEM. A *P* value less than 0.05 was considered statistically significant. For immune-cellular metrics, statistically significant *P* values are presented in bold, and a very highly statistically significant *P* value is italicized. Effect sizes were determined for all parametric data based on specific differences between the groups, comparing means (μ; Cohen's d), standard deviations (σ; Glass's d), and sample sizes (*n*; Hedge's g). Hedge's g is superior to Cohen's d for sample sizes < 20. Effect size: 0.2 (small effect), 0.5 (medium effect), 0.8 (large effect). Effect sizes cannot be determined for nonparametric datasets. Post hoc power analysis was consistent with observed *P* values for each of the variables.

IGIC$^-$L, $P < 0.05$), but not the larger ones with IGIC (IGIC$^+$W, IGIC$^+$L, $P > 0.1$), even though all these parameters had high internal consistency (Cronbach's α: 0.83–0.96; Table 4). However, both the combined inter- and intraobserver agreement (ICC_a: 0.69, 0.78; Table 4) and Lin's concordance correlation coefficient (ρ_c: 0.52, 0.63) were low, confirming that IGIC$^-$W and IGIC$^-$L have low interobserver agreement and low reproducibility, most likely due to the small size of these ductular measurements, hence low tolerance for deviation between observers. The Bland-Altman plot (Supplementary Figs. S1, S2) also illustrated large LoA for IGIC$^-$W and IGIC$^-$L, which would be unacceptable for glandular ductules with small luminal dimensions. r had proportional bias and displayed a significant interobserver mean difference in measurements ($P < 0.01$) for both IGIC$^+$r and IGIC$^-$r even after back-transformation. This underscores the importance of appropriate tests of reliability as correlations and concordance coefficients alone can be misleading.

Correlation of IVCM-Based Metrics of Inflammation With Symptom Severity, TBUT, and CFS

Of the four immune cellular parameters, EIC and IGIC both showed moderate strength of direct correlation with OSDI (EIC-OSDI r_s: 0.49, $P = 0.03$, IGIC-OSDI r_s: 0.48, $P = 0.05$; Figs. 4A, 4B; Table 5), and inverse correlation with TBUT (EIC-TBUT r_s: –0.47, $P = 0.02$, IGIC-TBUT r_s: –0.45, $P = 0.04$; Figs. 4C, 4D; Table 5), making them clinically relevant endpoints of palpebral conjunctival and glandular inflammation. While IGIC could not attain statistical significance in its correlation with OSDI, we do believe that it displays an important trend that may have been compromised by a lower power (observed power 49.1%; Table 5). EIC and IGIC accounted for up to 24% and 22% of variability in OSDI scores and TBUT, respectively. CFS explained another 27% of the variation in OSDI. None of the epithelial and glandular immune cellular parameters correlated with CFS (r_s: 0.01–0.29, $P = 0.15$–0.96; Table 5). Neither SIC nor PGIC demonstrated any relationships with OSDI, TBUT, and CFS.

Diagnostic Utility of EIC and IGIC in Detecting Inflammation

ROC plots and analyses (Fig. 5) suggested that EIC and IGIC were excellent to very good parameters to detect inflammation, respectively. The ROC plots had large AUCs for both EIC (AUC: 0.97, $P < 0.001$, 95% CI, 0.91–1.00) and IGIC (AUC: 0.89, $P = 0.001$, 95% CI, 0.75–1.00), indicating very good to excellent diagnostic accuracy for palpebral conjunctival and glandular inflammation. Additionally, ARs were also high for both EIC (AR: 0.94) and IGIC (AR: 0.79). ROC plot analysis with estimations for sensitivity, specificity, PPV, NPV, LR+, LR– and accuracy, provided potential cutoffs for EIC and IGIC to determine the presence of palpebral conjunctival tissue inflammation. EIC ≥ 195.8 cells/mm^2 provided high sensitivity (0.94, 95% CI 0.68–0.99) and specificity (0.92, 95% CI 0.59–0.99), high PPV (0.94, 95% CI 0.68–0.99) and NPV (0.92, 95% CI 0.59–0.99), and a strong ability to rule in (LR+: 11.25, 95% CI 1.72–73.78) and rule out (LR–: 0.07, 95% CI 0.01–0.4616) the presence of palpebral conjunctival epithelial inflammation. IGIC ≥ 21.2% had excellent sensitivity (100%, 95% CI 0.75–1.00) with good specificity (78%, 95% CI 0.40–0.96), good PPV (0.88, 95% CI 0.62–0.98) and excellent NPV (1.00, 95% CI 0.56–1.00), and a strong ability to rule out (LR–: 0.00) but not rule in (LR+: 4.50, 95% CI 1.33–15.28) the presence of intraglandular inflammation. DOR and J, both of which are independent of disease prevalence and measure the discriminative power of a test, were high for both EIC (DOR: 1.3, J: 0.85) and IGIC (DOR: 1.7, J: 0.78), suggesting their potential use as discriminative tests for ruling in or ruling out inflammation.

Discussion

In an era of decelerated dry eye drug development, inconsistent clinical trial endpoints, and frustration among ophthalmologists and optometrists caring for patients with symptom–sign disparity, there is a growing need for objective, reliable, and clinically relevant measures of detecting, assessing, and monitoring disease severity in patients with ocular surface disease, such as DED or MGD.[35] Since inflammation contrib-

Investigative Ophthalmology & Visual Science

TABLE 4. Inter- and Intraobserver Agreement and Reliability of Immune-Cellular Metrics Measuring Inflammation of the Palpebral Conjunctiva and Glands Using IVCM

In Vivo Metric	Bland-Altman Analysis*					Intraclass Correlation				Interclass Correlation		
	Bias, Mean ± SD	95% CI	*P* Value	% Bias	LoA	Cronbach's α†	ICC_a	95% CI	*P* Value	ρ_c	95% CI	*n*
EIC, cells/mm^2	23.76 ± 84.45	−8.98, 56.51	**0.13**	**7.21**	−141.76, 189.29	**0.97**	**0.97**	0.93, 0.99	<0.001	**0.94**	0.87, 0.97	28
SIC, cells/mm^2	−4.41 ± 21.83	−12.88, 4.05	**0.29**	**8.39**	−47.19, 38.37	**0.96**	**0.96**	0.91, 0.98	<0.001	**0.92**	0.83, 0.96	28
PGIC, cells/ 20-μm radius	−0.26 ± 1.41	0.92, 0.4	**0.43**	**14.93**	−3.02, 2.51	**0.85**	**0.85**	0.61, 0.94	<0.001	**0.72**	0.44, 0.88	20
IGIC, % occlusion	2.12 ± 11.10	−2.6, 6.8	**0.36**	**6.91**	−19.62, 23.87	**0.93**	**0.93**	0.84, 0.97	<0.001	**0.86**	0.71, 0.94	24
IGIC$^+$W, μm	−1.48 ± 11.21	−6.3, 3.4	**0.53**	**3.83**	−23.46, 20.49	**0.95**	**0.95**	0.89, 0.98	<0.001	**0.90**	0.79, 0.96	23
IGIC$^+$L, μm	−4.38 ± 13.13	−10.1, 1.3	**0.12**	**7.94**	−30.11, 21.36	**0.96**	**0.96**	0.89, 0.98	<0.001	**0.91**	0.82, 0.96	23
IGIC$^+$r, μm	1.44 ± 2.10	0.43, 2.45	0.008	33.32	−2.67, 5.56	0.43	0.31	−0.37, 0.69	0.16	0.17	−0.13, 0.45	19
IGIC$^-$W, μm	−5.38 ± 4.94	−7.84, −2.92	<0.001	49.63	−15.07, 4.31	0.85	0.69	−0.13, 0.90	<0.001	0.52	0.23, 0.72	18
IGIC$^-$L, μm	−4.79 ± 8.92	−9.2, −0.36	0.04	20.95	−22.26, 12.68	0.83	0.78	0.41, 0.92	0.001	0.63	0.30, 0.82	18
IGIC$^-$r, μm	0.97 ± 1.21	0.27, 1.67	0.01	16.00	−1.41, 3.34	**0.94**	**0.87**	0.41, 0.96	<0.001	0.95	0.88, 0.98	14

EIC, SIC, PGIC, IGIC, IGIC$^+$W, IGIC$^+$L, IGIC$^+$r, IGIC$^-$W, IGIC$^-$L, and IGIC$^-$r were assessed for inter- and intraobserver agreement. The hyperreflective ring of the acinar ductules is presumed to be ductular epithelium.

* Bland-Altman analysis determined the mean interobserver difference in measurements (bias), 95% CI, and LoA. Bias was calculated as the mean of the absolute differences in measurements between the two observers and reported as mean ± standard deviation (SD). Percent interobserver bias (range: 3.83%–14.93%) was negligible for all immune-cellular metrics, EIC, SIC, PGIC, and IGIC ($P = 0.12$–0.43). Values in bold represent the metrics with negligible interobserver bias. LoA were calculated as bias ± 1.96 SD.

† Cronbach's α is a measure of internal consistency (excellent: ≥ 0.9, good: 0.7–0.89, acceptable: 0.6–0.69, poor: 0.5–0.59, unacceptable: <0.5), whereas ICC_a measures reproducibility. ρ_c is a robust measure of interobserver reproducibility and agreement incorporating correction for bias (almost perfect: 0.81–1.0, substantial: 0.61–0.80, moderate: 0.41–0.60, fair: 0.21–0.40). A *P* value less than 0.05 was considered statistically significant. Values in bold represent metrics with good to excellent internal consistency, and substantial to almost perfect interobserver reproducibility.

utes to the pathogenesis of MGD,[36,37] and given the paucity of quantitative metrics that reproducibly correlate with both symptoms and clinical signs in DED-associated inflammation, we conducted a rigorous analysis of 10 exploratory IVCM-based immune-cellular metrics for the evaluation of inflammation in MGD. Of these 10 metrics, 4 were focused specifically on quantifying inflammation in each layer of the palpebral conjunctiva (EIC, SIC, PGIC) as well as within meibomian glands (IGIC). It is noteworthy to highlight that LR+ > 10 and LR− < 0.1 are considered features of diagnostic tests that have an impact on clinical decision making. Likelihood ratios are essential to clinically meaningful diagnostic tests because they are an indicator of the number of times (more or less) a patient with the disease, in this case, inflammation, is likely to have the disease in comparison to individuals free of the disease.[38–42] Both EIC and IGIC exhibited such strengths in the detection of palpebral conjunctival and glandular inflammation. Despite strong ROC-based accuracy values for EIC (0.93) and IGIC (0.92), it is important to be cognizant of the fact that the range specified for EIC and IGIC in our study may be prone to some unreliability given the small sample size of our study. Therefore, the ranges provided to diagnose palpebral conjunctival epithelial and intraglandular inflammation in patients with MGD should be used for guiding clinicians and investigators for now rather than being considered definitive. Larger studies are warranted to confirm our findings and to provide more absolute values.

Many of the available tests for the assessment of MGD or DED severity struggle to demonstrate correlation with both symptoms and clinical signs,[1] making the interpretation of their clinical usefulness difficult both in clinical practice and as endpoints in clinical trials. With regard to evaporative DED (EDE), there are certain tests like tear lipid layer interferometry that have shown correlation with TBUT[43]; however, the issues of symptom–sign discordance and low correlation between clinical signs and tests in MGD persist.[3,4,44] In contrast, EIC and IGIC both exhibited modest but significant inverse correlation with TBUT, and direct correlation with OSDI, accounting for up to 24% and 22% of variability in OSDI and TBUT, respectively. It is therefore not surprising that intraductal probing of meibomian glands both provides both symptomatic relief to patients and also considerably improves TBUT.[45–47] However, it is unlikely that any disease-related distension of glands and ducts will revert after probing. We observed that among patients with MGD, even ductules that seemed apparently void of visible immune cellular content had significantly larger luminal dimensions than their counterparts in asymptomatic, healthy individuals. This may indicate that any pressure-induced distension of the meibomian gland tract remains irreversible even when the occlusion is relieved.

The MGD patients in our study were interesting in that they had symptoms of dry eye with minimal to no corneal staining despite shortened TBUT. The lack of correlation of any palpebral conjunctival metrics of inflammation with CFS, yet moderately strong and significant correlations with OSDI and TBUT, suggests that corneal epithelial integrity is not necessarily affected by palpebral conjunctival inflammation. Therefore, in a symptomatic patient with MGD, an absence of corneal staining does not automatically rule out underlying clinically nonapparent palpebral conjunctival inflammation. Thus, evaluation by IVCM may aid in the assessment of symptomatic MGD patients with an incongruent corneal slit-lamp examination, which can then be differentiated from dry eye patients with symptom–sign discordance due to a central neurologic process.[44] These findings emphasize that outcome measure selection needs to be based not only on the type of DED but on imaging-based metrics as well.

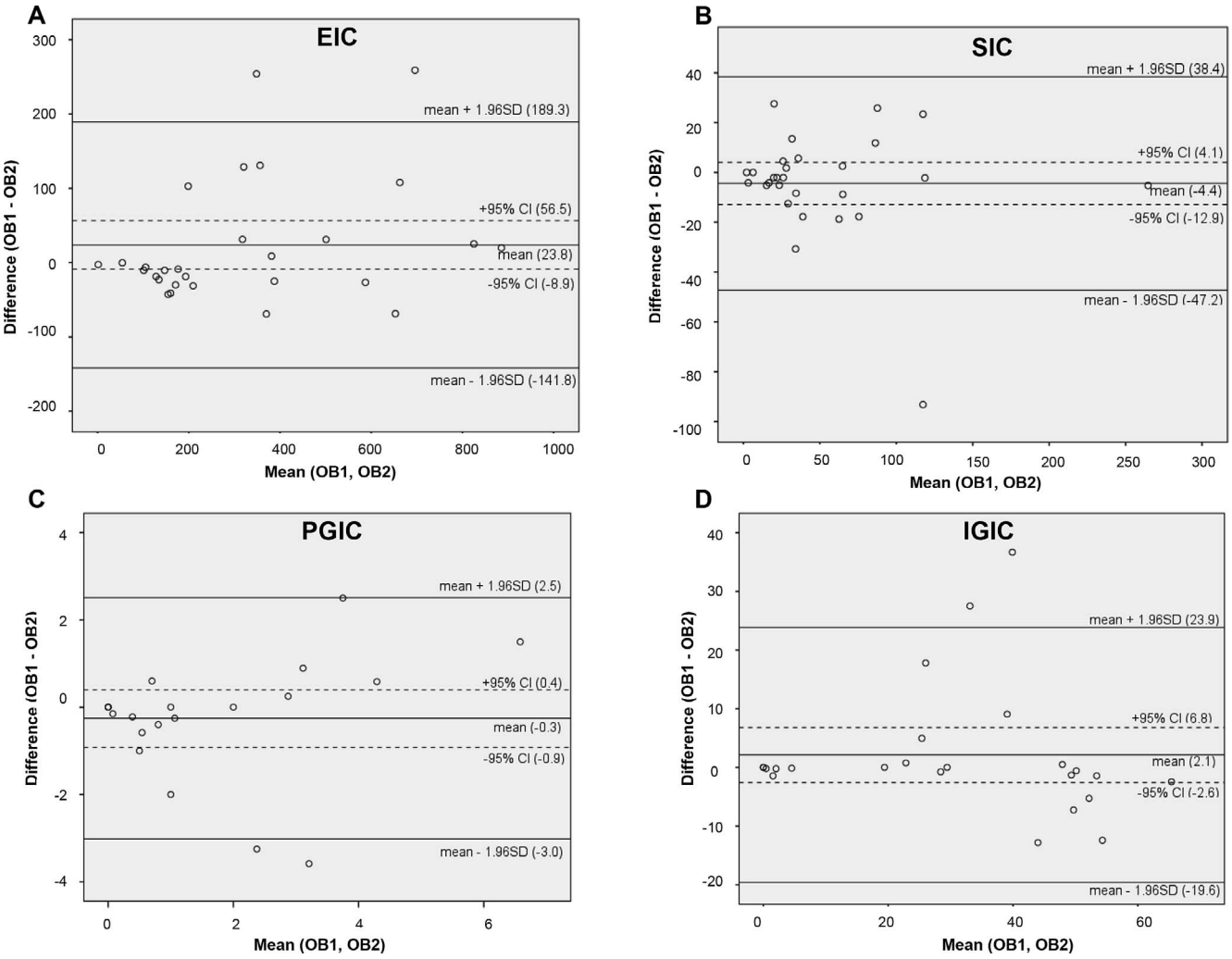


Figure 3. Bland-Altman plots assessing interobserver agreement for in vivo immune-cellular metrics of palpebral conjunctival inflammation. Each plot was generated using the interobserver difference in measurements (*y*-axis) against the mean of measurements for both observers (*x*-axis). The mean difference between interobserver measurements (bias) is represented by the *middle solid line*, accompanied by its 95% CI (*broken lines*) and 95% LoA (mean ± 1.96 SD; *top* and *bottom solid lines*). All metrics assessing immune cells, both in the palpebral conjunctival tissue (EIC, SIC, PGIC; **A**–**C**) and meibomian glands (IGIC; **D**), demonstrated very good interobserver agreement without proportional bias.

Imaging-based biomarkers of inflammation could provide additional information facilitating characterization of the cellular milieu in the tissue. This insight into the inflammatory state of the tissue may potentially enhance patient care by way of appropriate patient stratification for anti-inflammatory treatment regimens and measurement of therapeutic efficacy of drugs as part of both routine care and clinical trials in ocular surface disease. Among the many issues in selection of clinical endpoints and biomarkers for clinical trials have been the overbearing presence of inconsistent reproducibility, and lack of standardization and validation. Looking toward imaging outcome measures as surrogates of clinical disease is a promising approach. The past decade has seen a steady rise in the application of imaging biomarkers across specialties within ophthalmic clinical trials.[48] While this trend is encouraging, a systematic, nationwide strategy focusing on establishing centralized image reading centers or the use of artificial intelligence needs to be institutionalized. Such establishments will help to ensure consistent methodology for standardization of metrics and subsequent validation of potential imaging biomarkers that can be stratified for types of ocular surface disease.

Furthermore, we believe that imaging biomarkers EIC and IGIC may be promising surrogate biomarkers of inflammation in MGD since they fulfil the criteria for surrogate biomarkers.[49–53] Not only do EIC and IGIC correlate with the clinical endpoints of symptom severity (OSDI) and clinical sign (TBUT), which is necessary but not sufficient,[50] but they are also directly involved in the pathophysiological cascade of inflammation that is involved in the pathogenesis in MGD.[54,55]

An interesting observation made in the course of this study was the morphology of cellular content within the glands; the cells appeared round, and consistent with both size and appearance of lymphocytes,[30] highly suggestive of T cells. The role of T cell–mediated immunity in the pathogenesis of a subset of dry eye disease, aqueous-deficient dry eye disease (ADDE), is strongly established in the literature,[56–58] providing a cue for the role of adaptive immunity in possibly other subsets of DED as well. Nien and colleagues[34] reported the presence of $CD45^+$ inflammatory cells in the substantia propria of the palpebral conjunctiva with infiltration of the acinar epithelium in patients with MGD. Similarly, we observed frequent speckling of acinar epithelium, presumably immune cells. More recently, Reyes et al.[59] identified neutrophils

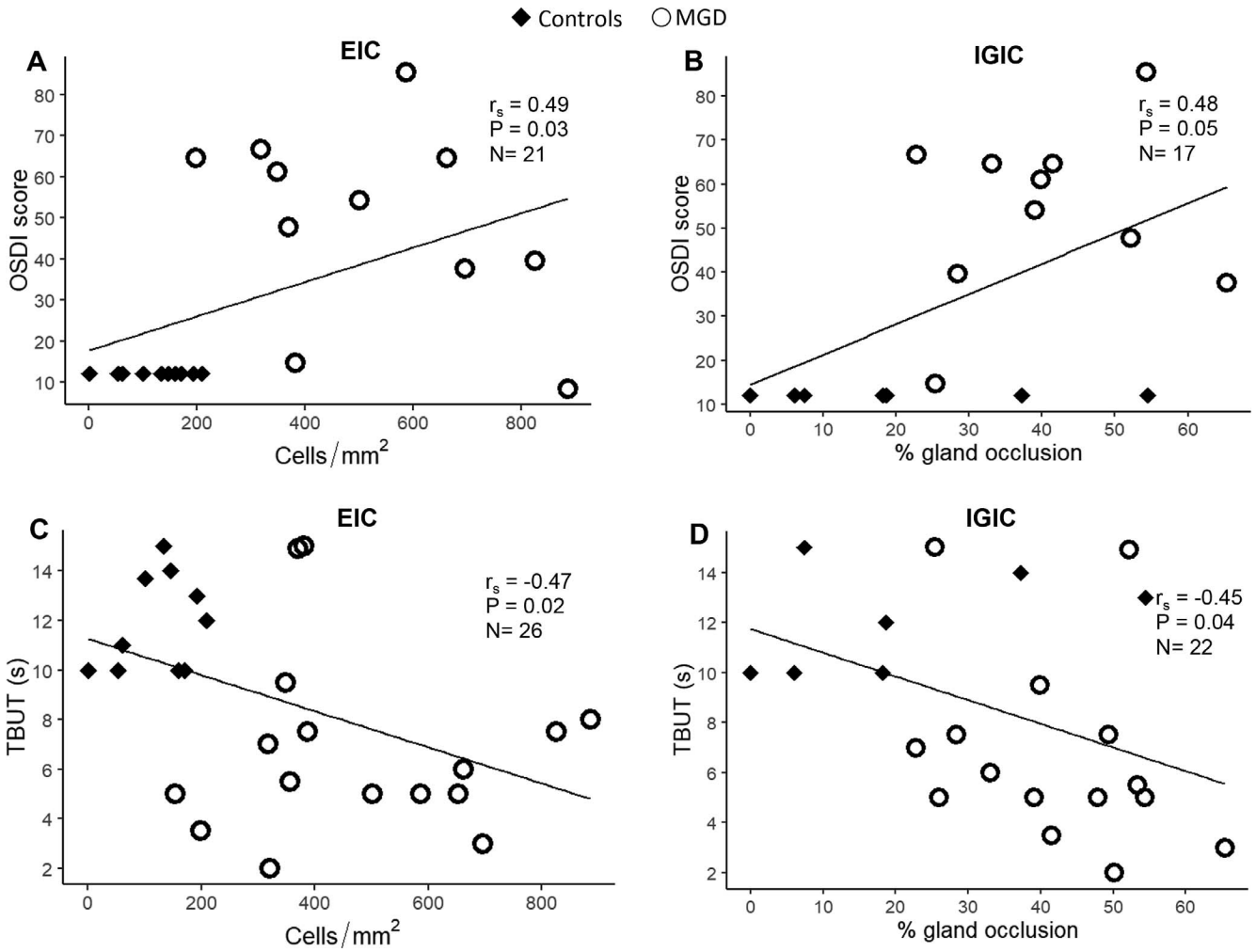


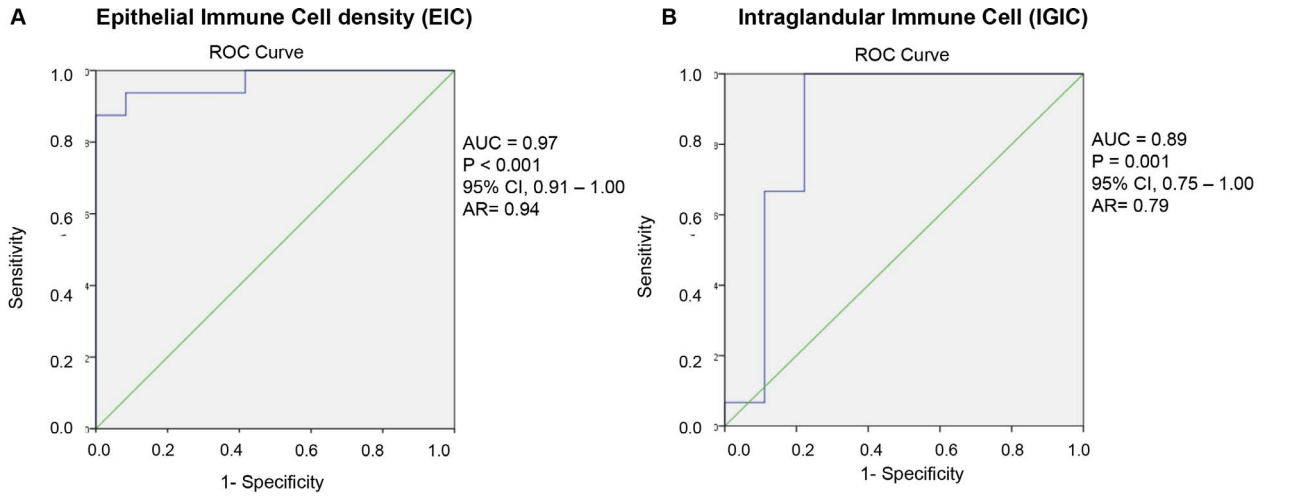
Figure 4. Correlation of immune-cellular metrics with symptom severity and clinical signs in MGD-associated palpebral conjunctival inflammation. r_s was used to determine the strength of correlation between in vivo confocal immune cell parameters that were significantly altered in MGD, and both clinical symptom severity (OSDI scores) and TBUT. Both palpebral conjunctival EIC (**A**, **C**) and percentage occlusion of presumed meibomian IGIC (Figs. 3B, 3D) showed direct correlation with OSDI (**A**, **B**), and inverse correlation with TBUT (**C**, **D**). $P < 0.05$ was considered statistically significant. *N* represents the number of subjects available with both measurements.

Table 5. Clinical Relevance of IVCM-Based Metrics of Palpebral Conjunctival Epithelial and Glandular Inflammation

Correlated Parameters	Correlation Coefficient, r_s	Coefficient of Determination, r^2	*P* Value	Observed Power, %	*n*
EIC, cells/mm^2–OSDI score	**0.49**	**0.24**	**0.03**	61.3	21
IGIC, % occlusion–OSDI score*	**0.48**	**0.23**	**0.05**	49.1	17
PGIC, cells/20-µm radius–OSDI score	0.43	0.18	0.09	37.0	16
EIC, cells/mm^2–TBUT	**−0.47**	**0.22**	**0.02**	68.3	26
IGIC, % occlusion–TBUT	**−0.45**	**0.20**	**0.04**	55.2	22
PGIC, cells/20-µm radius)–TBUT	−0.21	0.04	0.36	13.7	21
EIC, cells/mm^2–CFS grade	0.29	0.08	0.21	22.3	19
IGIC, % occlusion–CFS grade	0.39	0.15	0.15	28.6	15
PGIC, cells/20-µm radius–CFS grade	0.01	0.00	0.96	4.2	14
TBUT–OSDI score	**−0.56**	**0.31**	**0.007**	77.9	22
CFS grade–OSDI score	**0.52**	**0.27**	**0.02**	63.9	19
TBUT–CFS grade	**−0.55**	**0.30**	**0.01**	71.4	20

EIC, PGIC, and IGIC were correlated with symptom severity (OSDI score), tear film stability (TBUT, seconds), and corneal epithelial integrity (NEI corneal fluorescein staining grade, CFS) using r_s and r^2. EIC and IGIC showed significant correlation with OSDI and TBUT, but not CFS. EIC and IGIC predicted up to 24% variability in OSDI scores, and 22% variability in TBUT. There was symptom–sign correlation in our study sample. Of the 13 controls and 16 MGD patients, images with glands were identified in 7 and 15 subjects, respectively.

* Correlation of IGIC with OSDI approached but did not attain statistical significance, likely due to lower observed power of 49%. *n* represents the total number of subjects. A *P* value less than 0.05 was considered statistically significant, which are shown in bold along with their corresponding r_s and r^2 values.



EIC (cells/mm^2)	Sens	Spec	PPV	NPV	LR+	LR-	Accuracy
≥184.7	1.00	0.83	0.88	1.00	6.00	0.00	0.93
≥ 195.8	0.94	0.92	0.94	0.92	11.25	0.07	0.93
≥ 203.9	0.88	0.92	0.93	0.85	10.50	0.14	0.89

IGIC (%)	Sens	Spec	PPV	NPV	LR+	LR-	Accuracy
≥21.1	1.00	0.78	0.88	1.00	4.50	0.00	0.92
≥ 24.1	0.93	0.78	0.88	0.88	4.20	0.09	0.88
≥ 25.7	0.87	0.78	0.87	0.78	3.90	0.17	0.83

Figure 5. ROC curves and analyses illustrating the diagnostic utility of EIC and IGIC in the assessment of palpebral conjunctival inflammation. AUC and AR for both EIC and IGIC indicate that these immune-cellular metrics (EIC, IGIC) are very good to excellent diagnostic aids for palpebral conjunctival inflammation. ROC analysis: sensitivity (sens), specificity (spec), PPV, NPV, LR+, and LR–. EIC $\geq$ 195.8 cells/mm^2 and IGIC $\geq$ 21.1% may prove to be precise and accurate thresholds in screening for and confirming eyelid inflammation. Diagnostic accuracy grades based on AUC: excellent (AUC 0.9–1.0), very good (AUC 0.8–0.9), good (0.7–0.8), sufficient (AUC 0.6–0.7), bad (AUC 0.5–0.6), test not useful (AUC $<$ 0.5).[66]

causing obstruction of meibomian glands in a mouse model of chronic inflammation associated with ocular allergy. This finding suggests that perhaps the innate immune system plays an early role in meibomian gland inflammation, inciting the adaptive arm of the immune response through myeloid and plasmacytoid dendritic cells.[60]

In patients with MGD, we also observed subepithelial fibrosis of the substantia propria around acini and ductules. Based on this finding and our earlier work,[61] we hypothesized that periductular fibrosis may be a possible mechanism behind obstructive forms of MGD. This band of constriction around acinar ductules is likely broken, yielding a "pop" sound during meibomian gland probing,[45] relieving the stricture. Our hypothesis has been supported by colleagues with similar findings,[62] which has led to a new classification system for obstructive MGD.[63] MGD is frequently seen in individuals older than 60 years,[34,64] in whom it is postulated that decreased cell cycling and meibocyte differentiation with increasing age leads to glandular dysfunction.[34] However, it appears that the process of palpebral conjunctival goblet cell loss and reduction in meibomian acinar units begins as early as the early 40s, or perhaps even earlier.[11] This contrasts with earlier studies in which even individuals older than 70 years did not show changes in meibomian glands, ducts, and the quality of meibum.[65]

IVCM of the palpebral conjunctiva necessitates technical training of imaging personnel, which can be time-consuming to begin with, but with experience over time it may become practical to do in an outpatient setting. In addition to technical training, careful image selection and quantitation of EIC and IGIC also require training, close monitoring, and stringent quality control, which may not be readily available at many centers with laser IVCM. Hence, there is benefit in establishing centralized national or international reading centers that can validate and standardize imaging surrogate biomarkers of ocular surface disease in addition to providing quick, accurate, and reliable imaging reports. EIC and IGIC have already been shown to be robust metrics in the detection of inflammation in patients with MGD. Our group initially reported findings of a series of five such MGD patients who were symptomatic despite an unremarkable slit-lamp examination.[5] In a research setting, however, we believe that these in vivo metrics quantifying palpebral conjunctival epithelial and intraglandular inflammation may supersede other clinical endpoints either in the detection and grading of inflammation toward patient inclusion criteria, or, detecting response to therapy even in short-duration clinical trials. With the very large effect sizes observed for EIC and IGIC, smaller sample sizes will be required for research studies, making clinical trials much more affordable. Moreover, EIC and IGIC may also be useful in quantifying and grading severity of the inflammatory component of disease.

Our study provides validated real-world evidence for the use of EIC and IGIC in detecting palpebral conjunctival epithelial and intraglandular inflammation in evaporative DED. The next steps forward would be to test these metrics of inflammation as clinical endpoints in a randomized, controlled clinical trial targeting inflammation in MGD. It would be interesting to see if EIC, SIC, and/or IGIC may be more responsive to anti-inflammatory therapy than symptom severity and clinical signs such as CFS, TBUT, tear film osmolarity, and meibum scores. Since these are immune cellular parameters, we hypothesize that the metrics defined in our paper will respond earlier and to a larger degree than clinical tests, which are macroscopic, leading to shorter and consequently less expensive clinical trials.

Investigative Ophthalmology & Visual Science

Acknowledgments

Supported by National Institutes of Health K08-EY020575 (PH), L30-EY021919 (PH), Falk Medical Research Foundation (PH), Research to Prevent Blindness Career Development Award (PH). The authors alone are responsible for the content and writing of the paper.

Disclosure: **Y. Qazi**, P; **A. Kheirkhah**, None; **C. Blackie**, None; **M. Trinidad**, None; **C. Williams**, None; **A. Cruzat**, None; **D.R. Korb**, None; **P. Hamrah**, P

References

1. Bron AJ, Tomlinson A, Foulks GN, et al. Rethinking dry eye disease: a perspective on clinical implications. *Ocul Surf*. 2014;12:S1–S31.
2. Nichols KK, Nichols JJ, Mitchell GL. The lack of association between signs and symptoms in patients with dry eye disease. *Cornea*. 2004;23:762–770.
3. Sullivan BD, Crews LA, Messmer EM, et al. Correlations between commonly used objective signs and symptoms for the diagnosis of dry eye disease: clinical implications. *Acta Ophthalmol*. 2014;92:161–166.
4. Cuevas M, Gonzalez-Garcia MJ, Castellanos E, et al. Correlations among symptoms, signs, and clinical tests in evaporative-type dry eye disease caused by meibomian gland dysfunction (MGD). *Curr Eye Res*. 2012;37:855–863.
5. Qazi Y, Kheirkhah A, Blackie C, et al. In vivo detection of clinically non-apparent ocular surface inflammation in patients with meibomian gland dysfunction-associated refractory dry eye symptoms: a pilot study. *Eye (Lond)*. 2015;29:1099–1110.
6. Kobayashi A, Yoshita T, Sugiyama K. In vivo findings of the bulbar/palpebral conjunctiva and presumed meibomian glands by laser scanning confocal microscopy. *Cornea*. 2005;24:985–988.
7. Matsumoto Y, Sato EA, Ibrahim OM, Dogru M, Tsubota K. The application of in vivo laser confocal microscopy to the diagnosis and evaluation of meibomian gland dysfunction. *Mol Vis*. 2008;14:1263–1271.
8. Efron N, Al-Dossari M, Pritchard N. In vivo confocal microscopy of the palpebral conjunctiva and tarsal plate. *Optom Vis Sci*. 2009;86:E1303–E1308.
9. Zhou S, Robertson DM. Wide-field in vivo confocal microscopy of meibomian gland acini and rete ridges in the eyelid margin. *Invest Ophthalmol Vis Sci*. 2018;59:4249–4257.
10. Ban Y, Ogawa Y, Ibrahim OM, et al. Morphologic evaluation of meibomian glands in chronic graft-versus-host disease using in vivo laser confocal microscopy. *Mol Vis*. 2011;17:2533–2543.
11. Wei A, Hong J, Sun X, Xu J. Evaluation of age-related changes in human palpebral conjunctiva and meibomian glands by in vivo confocal microscopy. *Cornea*. 2011;30:1007–1012.
12. Ibrahim OM, Matsumoto Y, Dogru M, et al. In vivo confocal microscopy evaluation of meibomian gland dysfunction in atopic-keratoconjunctivitis patients. *Ophthalmology*. 2012;119:1961–1968.
13. Villani E, Mantelli F, Nucci P. In-vivo confocal microscopy of the ocular surface: ocular allergy and dry eye. *Curr Opin Allergy Clin Immunol*. 2013;13:569–576.
14. Matsumoto Y, Shigeno Y, Sato EA, et al. The evaluation of the treatment response in obstructive meibomian gland disease by in vivo laser confocal microscopy. *Graefes Arch Clin Exp Ophthalmol*. 2009;247:821–829.
15. Villani E, Canton V, Magnani F, Viola F, Nucci P, Ratiglia R. The aging meibomian gland: an in vivo confocal study. *Invest Ophthalmol Vis Sci*. 2013;54:4735–4740.
16. Villani E, Ceresara G, Beretta S, Magnani F, Viola F, Ratiglia R. In vivo confocal microscopy of meibomian glands in contact lens wearers. *Invest Ophthalmol Vis Sci*. 2011;52:5215–5219.
17. Villani E, Beretta S, De Capitani M, Galimberti D, Viola F, Ratiglia R. In vivo confocal microscopy of meibomian glands in Sjogren's syndrome. *Invest Ophthalmol Vis Sci*. 2011;52:933–939.
18. Ibrahim OM, Matsumoto Y, Dogru M, et al. The efficacy, sensitivity, and specificity of in vivo laser confocal microscopy in the diagnosis of meibomian gland dysfunction. *Ophthalmology*. 2010;117:665–672.
19. Di Staso S, Agnifili L, Cecannecchia S, Di Gregorio A, Ciancaglini M. In vivo analysis of prostaglandins-induced ocular surface and periocular adnexa modifications in patients with glaucoma. *In Vivo*. 2018;32:211–220.
20. Agnifili L, Mastropasqua R, Fasanella V, et al. Meibomian gland features and conjunctival goblet cell density in glaucomatous patients controlled with prostaglandin/timolol fixed combinations: a case control, cross-sectional study. *J Glaucoma*. 2018;27:364–370.
21. Lin T, Gong L. In vivo confocal microscopy of meibomian glands in primary blepharospasm: a prospective case-control study in a Chinese population. *Medicine (Baltimore)*. 2016;95:e3833.
22. Liang H, Randon M, Michee S, Tahiri R, Labbe A, Baudouin C. In vivo confocal microscopy evaluation of ocular and cutaneous alterations in patients with rosacea. *Br J Ophthalmol*. 2017;101:268–274.
23. Fasanella V, Agnifili L, Mastropasqua R, et al. In vivo laser scanning confocal microscopy of human meibomian glands in aging and ocular surface diseases. *Biomed Res Int*. 2016;2016:7432131.
24. Hong J, Yu Z, Cui X, et al. Meibomian gland alteration in patients with primary chronic dacryocystitis: an in vivo confocal microscopy study. *Curr Eye Res*. 2015;40:772–779.
25. Knop E, Knop N, Zhivov A, et al. The lid wiper and mucocutaneous junction anatomy of the human eyelid margins: an in vivo confocal and histological study. *J Anat*. 2011;218:449–461.
26. Villani E, Canton V, Magnani F, Viola F, Nucci P, Ratiglia R. The aging meibomian gland: an in vivo confocal study. *Invest Ophthalmol Vis Sci*. 2013;54:4735–4740.
27. Dougherty BE, Nichols JJ, Nichols KK. Rasch analysis of the Ocular Surface Disease Index (OSDI). *Invest Ophthalmol Vis Sci*. 2011;52:8630–8635.
28. Ozcura F, Aydin S, Helvaci MR. Ocular surface disease index for the diagnosis of dry eye syndrome. *Ocul Immunol Inflamm*. 2007;15:389–393.
29. Schiffman RM, Christianson MD, Jacobsen G, Hirsch JD, Reis BL. Reliability and validity of the Ocular Surface Disease Index. *Arch Ophthalmol*. 2000;118:615–621.
30. Guthoff RF, Zhivov A, Stachs O. In vivo confocal microscopy, an inner vision of the cornea – a major review. *Clin Exp Ophthalmol*. 2009;37:100–117.
31. Villani E, Beretta S, Galimberti D, Viola F, Ratiglia R. In vivo confocal microscopy of conjunctival roundish bright objects: young, older, and Sjogren subjects. *Invest Ophthalmol Vis Sci*. 2011;52:4829–4832.
32. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;1:307–310.
33. Bland JM, Altman DG. Measuring agreement in method comparison studies. *Stat Methods Med Res*. 1999;8:135–160.
34. Nien CJ, Massei S, Lin G, et al. Effects of age and dysfunction on human meibomian glands. *Arch Ophthalmol*. 2011;129:462–469.
35. Roy NS, Wei Y, Kuklinski E, Asbell PA. The growing need for validated biomarkers and endpoints for dry eye clinical research. *Invest Ophthalmol Vis Sci*. 2017;58:BIO1–BIO19.
36. Suzuki T, Teramukai S, Kinoshita S. Meibomian glands and ocular surface inflammation. *Ocul Surf*. 2015;13:133–149.

37. Mizoguchi S, Iwanishi H, Arita R, et al. Ocular surface inflammation impairs structure and function of meibomian gland. *Exp Eye Res*. 2017;163:78–84.
38. Akobeng AK. Understanding diagnostic tests 2: likelihood ratios, pre- and post-test probabilities and their use in clinical practice. *Acta Paediatr*. 2007;96:487–491.
39. Carneiro AV. Diagnostic characteristics of tests: sensitivity, specificity, predictive values and likelihood ratios. *Rev Port Cardiol*. 2011;30:551–558.
40. Perera R, Heneghan C. Making sense of diagnostic tests likelihood ratios. *Evid Based Med*. 2006;11:130–131.
41. Deeks JJ, Altman DG. Diagnostic tests 4: likelihood ratios. *BMJ*. 2004;329:168–169.
42. McClure P. Likelihood ratios: determining the usefulness of diagnostic tests. *J Hand Ther*. 2001;14:304–305.
43. Isreb MA, Greiner JV, Korb DR, et al. Correlation of lipid layer thickness measurements with fluorescein tear film break-up time and Schirmer's test. *Eye (Lond)*. 2003;17:79–83.
44. Shtein RM, Harper DE, Pallazola V, et al. Discordant dry eye disease (an American Ophthalmological Society Thesis). *Trans Am Ophthalmol Soc*. 2016;114:T4.
45. Maskin SL. Intraductal meibomian gland probing relieves symptoms of obstructive meibomian gland dysfunction. *Cornea*. 2010;29:1145–1152.
46. Sik Sarman Z, Cucen B, Yuksel N, Cengiz A, Caglar Y. Effectiveness of intraductal meibomian gland probing for obstructive meibomian gland dysfunction. *Cornea*. 2016;35:721–724.
47. Ma X, Lu Y. Efficacy of intraductal meibomian gland probing on tear function in patients with obstructive meibomian gland dysfunction. *Cornea*. 2016;35:725–730.
48. Villani E, Massaro D, Scaramuzzi M, Hamrah P, Medeiros FA, Nucci P. Decade-long profile of imaging biomarker use in ophthalmic clinical trials. *Invest Ophthalmol Vis Sci*. 2017;58:BIO76–BIO81.
49. FDA-NIH Biomarker Working Group. *BEST (Biomarkers, EndpointS, and other Tools) Resource*. Silver Spring, MD: Food and Drug Administration; Bethesda, MD: National Institutes of Health; 2016–2018.
50. Villani E, Vujosevic S. Foreword: biomarkers and surrogate endpoints in ophthalmic clinical research. *Invest Ophthalmol Vis Sci*. 2017;58:BIOi–BIOii.
51. Man in 't Veld AJ. Surrogate end points in clinical trials. *Blood Press Suppl*. 1997;2:120–123.
52. Fleming TR, DeMets DL. Surrogate end points in clinical trials: are we being misled? *Ann Intern Med*. 1996;125:605–613.
53. Ellenberg SS. Surrogate end points in clinical trials. *BMJ*. 1991;302:63–64.
54. Baudouin C, Messmer EM, Aragona P, et al. Revisiting the vicious circle of dry eye disease: a focus on the pathophysiology of meibomian gland dysfunction. *Br J Ophthalmol*. 2016;100:300–306.
55. Knop E, Knop N, Millar T, Obata H, Sullivan DA. The international workshop on meibomian gland dysfunction: report of the subcommittee on anatomy, physiology, and pathophysiology of the meibomian gland. *Invest Ophthalmol Vis Sci*. 2011;52:1938–1978.
56. De Paiva CS, Villarreal AL, Corrales RM, et al. Dry eye-induced conjunctival epithelial squamous metaplasia is modulated by interferon-gamma. *Invest Ophthalmol Vis Sci*. 2007;48:2553–2560.
57. Stern ME, Gao J, Schwalb TA, et al. Conjunctival T-cell subpopulations in Sjogren's and non-Sjogren's patients with dry eye. *Invest Ophthalmol Vis Sci*. 2002;43:2609–2614.
58. El Annan J, Chauhan SK, Ecoiffier T, Zhang Q, Saban DR, Dana R. Characterization of effector T cells in dry eye disease. *Invest Ophthalmol Vis Sci*. 2009;50:3802–3807.
59. Reyes NJ, Yu C, Mathew R, et al. Neutrophils cause obstruction of eyelid sebaceous glands in inflammatory eye disease in mice. *Sci Transl Med*. 2018;10 :eaas9164.
60. Ohbayashi M, Manzouri B, Flynn T, et al. Dynamic changes in conjunctival dendritic cell numbers, anatomical position and phenotype during experimental allergic conjunctivitis. *Exp Mol Pathol*. 2007;83:216–223.
61. Qazi Y, Korb D, Blackie C, Hamrah P. Identification of novel subclinical parameters for the management of MGD. *27th Biennial Cornea Conference*. Boston, Massachusetts; 2011.
62. Maskin SL. Intraductal meibomian gland probing relieves symptoms of obstructive meibomian gland dysfunction. *Cornea*. 2010;29:1145–1152.
63. Maskin SL, Testa WR. Growth of meibomian gland tissue after intraductal meibomian gland probing in patients with obstructive meibomian gland dysfunction. *Br J Ophthalmol*. 2018;102:59–68.
64. Pult H. Relationships between meibomian gland loss and age, sex, and dry eye [published online ahead of print February 12, 2018]. *Eye Contact Lens*. doi:10.1097/ICL.0000000000000467.
65. Hykin PG, Bron AJ. Age-related morphological changes in lid margin and meibomian gland anatomy. *Cornea*. 1992;11:334–342.
66. Simundic AM. Measures of diagnostic accuracy: basic definitions. *eJIFCC*. 2009;19:203–211.

Investigative Ophthalmology & Visual Science