Optimal methodology for lid wiper epitheliopathy identification

Chris W. Lievens, OD, MS[a]; Yvonne Norgett, PhD[b]; Nancy Briggs, PhD[c]; Peter M. Allen, PhD[d]; Marta Vianya-Estopa, PhD[b]

[a]The Eye Center, Southern College of Optometry Memphis, TN, USA

[b]Department of Vision & Hearing Sciences, Anglia Ruskin University, Cambridge, UK

[c]Mark Wainwright Analytical Centre, University New South Wales, Sydney, Australia

[d]Vision and Eye Research Unit, Anglia Ruskin University, Cambridge, UK

Corresponding author:

Chris W. Lievens

1245 Madison Avenue

Memphis, TN 38104

clievens@sco.edu

Commercial Relationship Disclosure: Financial research support from Alcon (CL), Allergan (CL) and Transitions (CL) in the past three years.

Funding disclosure: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

1. Abstract

Purpose: Lid wiper epitheliopathy (LWE) is a clinical sign that has been associated with dry eye disease. This study used a semi-automated method to identify the effect of drop instillation and post-dye viewing time on the absorption of lissamine green (LG) and sodium fluorescein (NaFl) on the upper eyelid in order to ascertain the optimal identification for LWE assessment.

Methods: In 37 participants with LWE, 1-drop of 1% LG (10μL) was applied to the superior bulbar conjunctiva in the right eye, and photographs of the lid margin were taken 1, 3, and 5 minutes after instillation. Measurements were repeated in the same eye following instillations of 2-drops of 1% LG. The same procedures were followed for application of 2% NaFl (2μL) to the left eye. Staining area was determined using software to detect and measure dye-stained images. Analysis used a linear mixed model with fixed effects of time, number of drops and their interaction.

Results: For LG, multivariate analysis showed that time of drop instillation was significant ($p=0.0091$) as was the area of staining in the 2-drop versus 1-drop condition ($p<0.0001$). For NaFl, there was a significant effect of time ($p<0.0001$), drops ($p<0.0001$), and a time/drops interaction ($p<0.0134$), suggesting that both time and number of drops are important.

Conclusion: A single drop of dye is insufficient to reveal the full extent of LWE staining. A 2-drop instillation is recommended and observation is recommended between 1-5 minutes (LG) and between 3-5 minutes (NaFl).

Key words: Lid Wiper; Epitheliopathy; Dry Eye; Lissamine Green; Sodium Fluorescein

1. Introduction

The prevalence of dry eye disease (DED) has been reported to be as low as 5% and as high as 50% [1–3]. Given the challenges associated with DED that result in ocular discomfort, reduced quality of life, lost productivity and rising healthcare costs, advances in the understanding and corresponding management are important [3]. It has been theorized that compromise to the epithelium of the wiping system results in symptomatology much like that in DED [1,4]. Additional research and investigation into lid wiper epitheliopathy (LWE) is needed.

The lid wiper is in constant contact with the ocular surface and travels across the surface with thousands of blinks per day. It is susceptible to mechanical trauma when there is inadequate lubrication as in the case of DED [4]. Despite this anatomical phenomenon, LWE is a relatively new finding in the clinical investigation of the anterior eye and was first reported as a clinical feature and associated with DED in 2002 [5]. It is surprising that aberrant changes to the eyelid anatomy have only been recently described since it is critical in maintaining ocular surface integrity [2]. Anatomically, the marginal conjunctiva begins at the inner-lid border and presents as a thickened and cushioned epithelium. This zone is in apposition to the globe and it is the structure that physically wipes the bulbar surface and distributes the thin tear film over the surface of the eye [2]. LWE is generally accepted to be the result of inadequate ocular lubrication and excessive friction [6]. When the tear film is disturbed, ocular drying can ensue in the interblink interval, and the frictional forces can then become aberrant and result in surface insult [7].

LWE is visualized by everting the eyelid after a dye has been instilled and observing the area proximal to the eyelashes [6]. An observable line at the mucocutaneous junction, called the line of Marx, is present in all eyelids and any further staining (beyond the line of Marx) of the tissue in the palpebral marginal conjunctiva can be evidence of LWE. Visual distinction of the line of Marx in its anatomical position versus the adjacent and immediately proximal lid wiper region is important to note when assessing this area [6]. Recent investigations of LWE using semi-automated methods have included the line of Marx when measuring the lid wiper area [8–10]. Since 2012 there have been several publications around LWE reporting differing methods of dye utilization [6]. Studies vary

in their use of either lissamine green (LG) or sodium fluorescein (NaFl), and in the time point after dye instillation for assessment. Rose Bengal (RB) tends to sting upon instillation and negatively affects cell vitality [11] and has become less widely used in recent years [6] even though substantial studies assessing patient tolerance of RB are lacking [12]. The current published evidence provides a large degree of inconsistency with respect to LWE identification, assessment, symptom linkage, and grading of severity. When LWE was described in 2002 by Korb et al. [5], staining of the lid wiper had not previously been reported, so it is not surprising that dye usage and severity grading varied in the subsequent literature. There have been several unique approaches to staining the lid wiper anatomy. They are summarized, in comparison to the Korb 2002 protocols [5] and presented in Table 1. Recently, a review by Begley et al. [13] highlights that variations in the use of ophthalmic dyes in terms of concentration, timing of observation and the methodology of instillation has negative ramifications for reliable and consistent ocular surface interpretation and evaluation.

Table 1: Literature review regarding staining techniques

Study	Dye Instillation Protocol	Timing of Assessment
Korb Protocol A: Korb, et al.[5] Korb, et al.[1]	40 µL of 2% NaFl solution. 2 instillations 5 minutes apart. RB paper strip wetted with 2 50 µL drops of sterile saline	View lid wiper 1 minute after drop instillations. Not specifically mentioned, but implies that a 1 minute wait time was used prior to lid wiper evaluation.
Korb Protocol B: Multiple reports since 2010[4,14,14–20]	40 µL of mixed solution with equal volumes of 2% NaFl and 1% LG. 2 instillations 5 minutes apart	View lid wiper 1 minute after drop instillations.
Kunnen et al.[10]	2-drops of 20 µL 2% NaFl and 1% LG instilled 5 minutes apart	View lid wiper 1 minute after drop instillations.

Jalbert et al.[21]	LG paper strip submerged in 200 µL of saline. 20 µL of resulting dye instilled (2-drops 5 minutes apart)	View lid wiper 30 seconds after drop instillations.
Navascues-Cornago et al.[8]	LG paper strip submerged in an unspecified amount of saline. Unspecified amount applied to eye.	Unspecified time of visual inspection.
Navascues-Cornago et al.[22]	LG paper strip submerged in an unspecified amount of saline for 1 minute. Unspecified amount applied to eye. 2 instillations 4 minutes apart.	View lid wiper 1 minute after drop instillations.
Satjawatcharaphong et al.[23]	LG paper strip wetted with unspecified amount saline. Unspecified amount applied to eye. 2 instillations 5 minutes apart.	View lid wiper 1 minute after drop instillations.
Varikooty et al.[24]	NaFl 2% 10 µL, 2-drops 1 minute apart. LG 1% 15 µL 2-drops 1 minute apart.	View lid wiper 3 minutes after drop instillations.
Varikooty et al.[25]	NaFl paper strip wetted with unspecified amount of saline. 2-drops 3 minutes apart.	View lid wiper 3 minutes after drop instillations.

	LG paper strip wetted with unspecified amount of saline. 2-drops 3 minutes apart.	

In addition to the variability in dye utilization for optimal observation, a recent study found that observers inaccurately estimate LWE staining using Korb's grading as compared to a semi-objective technique [9]. Two abstracts have described an automated process to detect and grade LWE using digital images [9,17]. A third process has been described that uses an open-access computer program but it is not automated and relies completely on the user for consistency [8]. The lid wiper region itself is poorly defined and so it is logical that an observer would have difficulty making a visual estimation of the percentage of the region affected, and the corresponding grading of LWE. Given that there exists widespread variability in all aspects of LWE clinical evaluation and research methodology, the aim of this study was to optimize the clinical identification of LWE in a consistent and semi-automated manner. The present study observed the absorption of LG and NaFl on the upper eyelid as a function of time and dosage using a semi-automated approach, with the goal of establishing: (1) the best time point, post-dye instillation, to view LWE; and (2) the effect of drop instillation (1 vs 2) to visualize LWE. Additionally, inter- and intra-examiner repeatability of measurements were investigated to check the validity of the semi-automated system used.

2. Material and Methods

Participants and experimental protocol

Participants were recruited from the Southern College Optometry (SCO; Memphis, TN, USA) patient base. Participants were financially compensated for their time and travel expenses. The study was approved by the Institutional Review Board of SCO and conformed to the tenets of the Declaration of Helsinki. Ethical approval was additionally obtained from Anglia Ruskin University (Cambridge, United Kingdom). Written informed

consent was obtained after explanation of the study and possible consequences of taking part.

The inclusion criteria included age ≥ 18 years and presence of LWE in both eyes. LWE determination was made by visual inspection of the lid wiper region 1 minute after a single drop of LG was instilled [4]. Exclusion criteria included habitual contact lens wearers in an extended wear modality (routinely sleeping in lenses overnight). Candidates with any anterior segment infection, inflammation, disease, or abnormality (within the previous 7 days) and/or those currently using systemic or ocular medications that would typically contraindicate contact lens wear were also excluded. Finally, candidates who were monocular or had known allergies to the ophthalmic dyes used in this study were not enrolled. 57 participants were screened for enrollment and 20 were excluded who did not have LWE.

The experimental protocol is outlined in Table 2. For each participant, all data were collected in a single visit. Baseline slit lamp biomicroscopy and digital photography were performed on each eye using the same unit, BI900 LED Slit Lamp, with EyeSuite Imaging (Haag-Streit, Bern, SUI). Baseline assessments of the cornea, bulbar conjunctiva, palpebral conjunctiva, and upper eyelid margin were made for each eye. The Brien Holden Vision Institute Grading Scale (formerly referred to as the Cornea and Contact Lens Research Unit Grading scale) was used to assess clinical findings for the anterior eye segment.

Table 2. Summary of experimental protocol; RE, right eye. LE, left eye; LG, lissamine green; NaFl, sodium fluorescein

Step	Description
1	Participant demographic data recorded
2	Medical history and ocular history recorded
3	Medication use recorded
4	LogMAR (RE/LE)

5	Biomicroscopy (slit lamp) anterior segment findings (RE/LE) • Examination of cornea, bulbar conjunctiva, palpebral conjunctiva, upper eyelid margin at (1) baseline and (2) conclusion of visit
6	Dosing of single drop 1% LG (10 μL) in **RE** superior bulbar conjunctiva and photography at 1, 3, and 5 minutes post-LG instillation
7	Dosing of single drop 2% NaFl (2 μL) in **LE** superior bulbar conjunctiva and photography at 1, 3, and 5 minutes post-NaFl instillation
8	Dosing of 2-drops 1% LG (10 μL each), one minute apart, in **RE** superior bulbar conjunctiva and photography at 1, 3, and 5 minutes post-LG instillation
9	Dosing of 2-drops 2% NaFl (2 μL), one minute apart, in **LE** superior bulbar conjunctiva and photography at 1, 3, and 5 minutes post-NaFl instillation

The right eye (RE) was used for all comparisons for LG usage while the left eye (LE) was used for all comparisons regarding NaFl usage. The effect of drop instillation and post-dye absorption for each dye was measured, rather than differences between eyes and so data from both eyes were therefore used. All dyes were instilled via a MicroPette Plus Single-Channel Variable Volume Pipettor, 2-20 μL volume, (Scilogex, Rocky Hill, CT, USA) to assure exact dosages. Separate pipettors were used for LG and NaFl instillation. For single drop instillation, 1% LG (10 μL applied RE) or 2% NaFl (2 μL applied LE) was applied to the superior bulbar conjunctiva [26,27]. The eyelid was carefully everted using a cotton-tipped applicator before each photograph (attention was made to not applanate the lid margin, causing iatrogenic staining). Photographs of the lid margin were taken at 1, 3, and 5 minutes post-dosing. Participants were instructed to blink normally after dye instillations (during the wait time prior to photography).

In order to examine staining after instillation of two drops in each dye, 2-drops of 1% LG (10 μL each) were applied, one minute apart, to the superior bulbar conjunctiva in the RE, and photographs of the lid margin were taken at 1, 3, and 5 minutes following instillation of the second drop. Similarly, 2-drops of 2% NaFl (2 μL each) were applied to the superior bulbar conjunctiva in the LE, and photographs were taken at the same time points described above. Care was taken in the photographs to capture the entire upper lid

margin, canthus to canthus. Because the eyes are tested in succession, and given that LG and NaFl staining has been reported to fade after several minutes, a washout period of 20-25 minutes was allotted between individual eye assessments to allow for dye clearance [28].

Image Analysis

This study uses ADCIS (Advanced Concepts in Imaging Software, Saint Contest, FR) described by Varikooty et al. [17] Images of the everted lid (resolution of 2000*1000 digitized on 8 bits, 12x magnification, Haag-Streit BI900 LED Slit Lamp system and Canon EOS 60D digital camera) were captured in raw mode, and then converted into tiff-format images. The software is designed to automatically detect LWE when using LG dye (see examples in Figure 1). Once this dyed area is detected, the software automatically segments the area and processes a series of computed measures (shape and intensity of the automatically detected regions). A manual option to annotate captured images prior to analysis was used to detect NaFl staining (Figure 1), as the system was not designed to automatically detect NaFl. In cases in which the software 'misses' areas of LG, the user had the ability to manually edit segmentation results to add missing areas, delete artifacts, and amend area edges. The ADCIS image processing algorithm carries multiple steps to deliver image optimization and analysis including image transformation, top-hat transformation and Otsu thresholding [17]. As LWE may have different presentations (continuous and non-continuous staining), the calculated area of lid wiper staining (mm^2) used for analysis includes all stained regions as well as the Line of Marx. This approach is consistent with two previous studies using alternative semi-automated methodologies [8,9].

To assess repeatability of the ADCIS semi-automated measurement technique, intra- and inter-observer repeatability measurements were reanalysed on a separate occasion. 37 images of LG-stained and NaFl-stained LWE (2-drops, 3-minute observation time point) were re-analysed through the ADCIS system by the same investigator 6 months after the initial processing. Data from the initial analysis were withheld and the images were evaluated in a different order. In this comparison it is possible to assess the similarity of the duplicate measures obtained through the ADCIS procedure. Similarly, inter-observer

repeatability was assessed with assistance from a separate masked grader with the same 37 images analysed with the ADCIS system.

Figure 1. Area of lid wiper epitheliopathy (LWE) staining

Data Analysis

Statistical analyses were performed using SAS software v9.4 (Cary, NC, USA). Examination of the dataset revealed normality in the raw variable (area of staining). The data were analyzed using a linear mixed model with fixed effects of time, number of drops and their interaction. Assumptions of normality and homogeneity of variance for the residuals were examined. Changes in the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) were examined to select the best-fitting covariance structure [29,30]. In addition, separate variance-covariance matrices for each drop condition were examined. This process resulted in choosing a structure (compound symmetric with heterogeneous variance) estimated for each drop condition separately. The repeatability of the ADCIS method was assessed through plotting Bland-Altman schematics.

3. Results

Thirty-seven (37) participants completed the present study. Mean age was 25 years (range 23-30 years), sex split was 73% female and 27% male, participant ethnicity was 78% Caucasian, 11% Asian, 8% African American and 3% Hispanic.

Area of LWE staining with LG

For LG, the multivariate covariance analysis showed that the time of drop instillation was significant ($p=0.0091$) (see Table 3 and Figure 2). The LG drops condition analysis showed a significantly larger area of staining with 2 drops than 1 drop ($p<0.0001$). The group*time interaction was not significant ($p=0.92$), indicating that the change in area

of LG staining over time is not different for the 1- and 2-drops conditions (see the parallel curves in Figure 2). Table 4 demonstrates all of the point to point time/drop comparisons. The 2-drop condition was significantly different from the 1-drop condition, with 2-drops of LG revealing a greater area of staining than 1-drop at all of the time points. When 2 drops are used, the time point of clinical observation can occur at 1, 3 or 5 minutes post-drops to reveal maximum uptake of dye.

Table 3. Main effects of drop dosing versus time. LG, lissamine green; NaFl, sodium fluorescein

Effect	LG		NaFL	
	F	P	F	p
Time	5.02	**0.0091**	22.58	**<0.0001**
Drops	28.85	**<0.0001**	74.56	**<0.0001**
Time*Drops	0.08	0.9192	4.58	**0.0134**

Figure 2. Staining area (mm^2) 1, 3, and 5 minutes following instillation of lissamine green (LG).

Area of LWE staining with NaFl

For NaFl, the multivariate covariance analysis (Table 3) revealed a significant effect of time ($p<0.0001$), drops ($p<0.0001$) and time*drops ($p=0.0134$), suggesting that both conditions are important for area of NaFl staining. As shown in Figure 3, when 1-drop of NaFl was used, there was a steep increase in LWE area of staining over time and the greatest area of staining was observed 5 minutes post-instillation. Similarly, when 2-drops of NaFl were used, there was an increase in staining area particularly when comparing 1 and 5 minutes. Table 4 shows statistically significant differences for all time point paired comparisons, with the exception of the 2-drop condition (1 minute versus 3 minutes $p=0.055$, and 3 minutes versus 5 minutes $p=0.525$). For NaFl, two drops should be used to reveal the maximal dye staining area. Additionally, a 3 minute to 5 minute wait time is necessary to reveal the maximum uptake of dye.

252

253 **Figure 3. Staining area (mm^2) 1, 3, and 5 minutes following instillation of sodium**
254 **fluorescein (NaFl).**

255

256 **Table 4. Pairwise comparisons of staining area, 1- and 2- drop conditions.** LG,
257 lissamine green; NaFl, sodium fluorescein

Drop comparison	Drop comparison		Mean Difference	P
LG 1-drop	LG 1-drop	1 minute vs 3 minutes	0.325	0.104
		1 minute vs 5 minutes	-0.043	0.871
		3 minutes vs 5 minutes	0.369	0.091
LG 2-drops	LG 2-drops	1 minute vs 3 minutes	0.438	0.095
		1 minute vs 5 minutes	0.124	0.699
		3 minutes vs 5 minutes	0.313	0.116
LG 1-drop	LG 2-drops	1 minute vs 3 minutes	2.465	**<0.0001**
		1 minute vs 5 minutes	2.151	**<0.0001**
		3 minutes vs 5 minutes	-1.826	**0.0003**
		3 minutes vs 1 minute	-1.702	**<0.0001**
		5 minutes vs 1 minute	-2.07	**<0.0001**
		5 minutes vs 3 minutes	2.508	**<0.0001**
		1 minute vs 1 minute	-2.027	**<0.0001**
		3 minutes vs 3 minutes	-2.140	**<0.0001**
		5 minutes vs 5 minutes	-2.195	**<0.0001**
NaFl 1-drop	NaFl 1-drop	1 minute vs 3 minutes	1.060	**0.0004**
		1 minute vs 5 minutes	1.939	**<0.0001**
		3 minutes vs 5 minutes	-0.879	**0.0027**
NaFl 2-drops	NaFl 2-drops	1 minute vs 3 minutes	0.551	0.055
		1 minute vs 5 minutes	0.732	**0.0117**
		3 minutes vs 5 minutes	-0.181	0.525

NaFl 1-drop	NaFl 2-drops	1 minute vs 3 minutes	2.534	**<0.0001**
		1 minute vs 5 minutes	9.59	**<0.0001**
		3 minutes vs 5 minutes	-1.654	**<0.0001**
		3 minutes vs 1 minute	-0.923	**0.0017**
		5 minutes vs 1 minute	-0.044	0.8767
		5 minutes vs 3 minutes	0.595	**0.0389**
		1 minute vs 1 minute	-1.983	**<0.0001**
		3 minutes vs 3 minutes	-1.474	**<0.0001**
		5 minutes vs 5 minutes	-0.776	**0.0077**

Image analysis repeatability

Intra-observer agreement was examined for LG and NaFl at the 2-drop 3-minute time mark. The 2-drop condition was chosen for this analysis as it revealed a greater area of staining for both LG and NaFl and the 3-minute mark was chosen since it had the greatest area of staining for LG. Analysis of mean staining (2-drops at the 3-minute mark) six months after initial analysis (same observer) was performed. A Bland-Altman comparison was performed for repeatability and agreement amongst the measures at the two time points (see Figures 4 and 5). The mean difference between time analysis repeats for the area of staining for LG was -0.10mm^2 and 95% limits of agreement between repeats were between -1.10mm^2 and 0.89mm^2, whereas the mean difference for NaFl was -0.20mm^2 and 95% limits of agreement between repeats were between -1.67mm^2 and 1.28mm^2. The data were slightly skewed to the top of the Bland-Altman plot, indicating that there was a tendency for a higher degree of area staining to be measured at time point 2 (the latter time point) versus time point 1 (the earlier time point) for both LG and NaFl. Next, inter-observer agreement was examined for LG and NaFl (see Figures 6 and 7). Analysis of mean staining was examined six months after initial analysis, but with a different observer. Bland-Altman comparisons were again performed. The plots show the repeatability for the area of staining using 2-drops of LG and 2-drops of NaFl at the 3-minute observation time. The mean difference between time analysis repeats for the area of

staining for LG was -0.15mm^2 and 95% limits of agreement between repeats were between -1.53mm^2 and 1.23mm^2, whereas the mean difference for NaFl was -0.30mm^2 and 95% limits of agreement between repeats were between -3.03mm^2 and 2.43mm^2.

Figure 4. Intra-observer Analysis LG. Bland-Altman plots for lid wiper epitheliopathy area of staining using 2-drops of LG at the 3 minute observation time point

Figure 5. Intra-observer Analysis NaFl. Bland-Altman plots for lid wiper epitheliopathy area of staining using 2-drops of NaFl at the 3 minute observation time point

Figure 6. Inter-observer Analysis LG. Bland-Altman plots for lid wiper epitheliopathy area of staining using 2-drops of LG at the 3 minute observation time point

Figure 7. Inter-observer Analysis NaFl. Bland-Altman plots for lid wiper epitheliopathy area of staining using 2-drops of NaFl at the 3 minute observation time point

4. Discussion

This study used a semi-automated method to compare the area of LWE tissue staining on the upper eyelid with a controlled volume of LG and NaFl to identify the optimal drop instillation and post-dye instillation viewing time. Values of area of lid wiper stained with lissamine green are similar to other published values, albeit somewhat higher, reflecting the fact that participants recruited for this study already showed LWE [8,9]. The findings of this study suggest that for both vital dyes (LG and NaFl) repeat instillation (2-drop) is necessary to reveal the full extent of LWE staining. Optimal viewing time is critical for NaFl as clinicians need to wait 5 minutes if using 1-drop whereas a faster observation time of 3-5 minutes can be achieved when using 2-drops. This, clearly illustrates that assessment time is critical when making such clinical observations. In agreement with these results, the latest TFOS DEWS II report recommends waiting 3-6 min after repeat instillation using separate fluorescein strips when observing lid wiper

epitheliopathy [27]. Similarly, this research showed that when using 2-drops of LG, optimal observation can occur between 1 and 5 minutes. This is also in agreement with the recommendation included in the latest TFOS DEWS II report that suggest that when using a strip wetted with 1 drop of saline observation should occur after between 1 and 4 minutes [27]. Based on these findings, to make efficient use of clinical time 1-minute observation time is recommended for LG and 2-drop instillation when viewing LWE.

The use of a semi-automated software to detect and measure dye stained-eye images offers superior guidance when compared to subjective evaluations of LWE. Without attention to a specific protocol for staining of the lid wiper, LWE can be easily underestimated and overlooked and with respect to research, difficult to compare results across different studies. Prior to the present investigation there were currently no clear guidelines to identify LWE (see Table 1). Wolffsohn et al. [31] found that eyecare practitioners (ECPs) took 7 minutes, on average, to assess anterior eye health including LWE evaluation suggesting that new evidence-based protocols are needed to ensure clinicians do not overlook signs of ocular discomfort within the time-constraints of clinical practice [31]. There have been conflicting reports to discern whether LWE is a true pathology or an acceptable physiological variation of the lid wiper, [6,32] possibly due to variations in assessment of LWE. Given the recommendation by TFOS DEWS II to evaluate the lid wiper area for insult and the possible linkage to DED, a consistent and reliable approach is indicated [27,32].

In 2005, Korb et al. [1] gave possible reasons why LWE had gone unnoticed for such a long time in research and clinical care. First, eyelids are unlikely to be everted during routine anterior segment evaluations. Second, when eyelids are everted, the lid margins are generally not inspected (instead, visual attention is typically directed to the tarsus). In 2015 Wolffsohn et al. [31], surveyed worldwide ECPs and found that only 26% of respondents evaluated LWE. Third, LWE requires vital dye staining to show compromise of the tissue, but vital dye staining prior to lid eversion is uncommon in practice, and when it is used, the most common approach is to use paper dye impregnated strips that deliver an inconsistent amount of dye to the eye, which may or may not be adequate. Fourth, the time of the vital dye to adsorb affected tissue may be a critical

component [1]. The current study was focused on the third and fourth factors. A similar study looked at how the timing of clinical inspection can influence findings when using NaFl in observations of the tear meniscus and found that too much delay in making observations with NaFl can yield misleading results due to a reduction in the clinical presentation [33]. The data presented here suggest that making observations without sufficient dye instillations may similarly not yield the full extent of LWE.

Of interest is the usage of automated or semi-automated methods to quantify staining. Previous staining technique protocols have been developed based on subjective grading by human observers. Recently, Kunnen et al. [9] used a semi-automated method and noticed that human observers underestimate the width and overestimate the height of LWE[10]. Although computer image analysis may not be feasible in clinical practice presently, it should be considered where LWE staining is a variable in a research study. Semi-automated methods have the ability to derive results which can inform clinical practice. The present study evaluated intra-observer and inter-observer repeatability using a semi-automated method. Interestingly, the greatest variability, shown by the spread of the 95% limits of agreement, was observed between observers (i.e. inter-observer) when evaluating LWE using NaFl. Clinicians need to be aware that NaFl stains much of the upper lid tarsus and can be visually distracting when clinically evaluating LWE. Thus, it was not surprising to find less agreement during the NaFL Bland-Altman analysis. As noted by Efron [6], recent investigations have increasingly used LG only over NaFl as the preferred dye. Yet, this is the first study to justify the superior capabilities of LG in terms of agreement between observers. Evidence gained from semi-automated methods has helped in highlighting critical clinical aspects to support clinicians and researchers grading and assessing the lid wiper region. It has been noted that LWE grading in its present form may not be able to discriminate minute changes at the lid margin which could partly explain the variability of previous work [34]. This would be an impetus for the use of computer assisted assessment tools as was done in this study. Future work should additionally focus on the development of grading scales to assess LWE optimally.

It is worth noting that the present study used a pipettor for the application of vital dye in comparison to the more widely used paper-impregnated strips in clinical practice.

The pipettor was used to ensure repeatability of dosage. 10µL of LG was previously shown to have good inter-observer reliability when compared to doses of 5 and 20µL [28]. A concern when using paper strips is potential variability when using various product brands. A 2018 study found significant differences in comparing four different brands of LG-strips [35]. It has been suggested that in clinical trials, dye concentration, volume and duration of contact to the ocular surface is critical in determining damage [36]. Inadequate volume and concentration can result in a weak staining pattern that can be underestimated [36]. Additionally, in research pipetted volumes (even when paper-impregnated strips are used) can lead to more reproducible findings [36].

Careful considerations were made regarding optimal experimental design to investigate all factors. Currently, no longitudinal data exist for appearance of LWE. Thus, comparing dosages at different time points might have introduced bias to the results. As it was critical to ensure baseline LWE remained stable, testing on separate days was considered inappropriate as the potential variability of LWE measurements could introduce a confounding factor in the study. As a result, all measurements were taken in a single visit with sequential staining in each eye, allowing for a 20-25 minute interval between the 1-drop and 2-drop conditions. It has previously been reported that LG staining fades rapidly after 4 minutes [28] and the interval used in this study largely exceeded the optimal viewing times recommended in the TFOS DEWS II report [27]. This caveat needs to be taken into consideration in future investigations. The order of dye/dosage instillation was not randomized to allow for maximum clearance of dye. In addition, the same order of drops was used in every subject and the RE (LG staining) was always tested before the LE (NaFl staining). The investigator was aware of which eye was being graded but the semi-automated assessment would have minimized any bias in the results. The upper eyelid was chosen for the present work as it may be more related to symptomatology in clinical care and the cause of upper- and lower-lid LWE may be from different etiologies [7,32,37,38]. Participants were instructed to blink normally between dye instillations in an attempt to allow for spontaneous blinking and prevent voluntary blinking. This process afforded dye uptake and not excessive dye clearance.

Shaw et al. recently reported that repeated eversion of the upper eyelid increases LWE when using LG [39], thus, suggesting a potential link between eyelid manipulation and staining in LWE. However, it is worth noting significant differences between the present study and this work. First, the fact that Shaw et al. did not control the volume of LG instilled and secondly the fact that nine eversions (9 x 15 seconds) were carried out with 3 minute breaks [39]. In contrast, the present study controlled the delivery of LG, included a total of 6 eversions per eye and allowed a longer interval between eversions. Shaw et al. reported a continuous increase in LG area of staining which could have been impacted by repeated instillations of dye with every eversion (i.e. LG was instilled 9 times and could have contributed to an increased measurement of LWE) [39]. As shown in Figure 3, the present data show a decline in the area of LG staining for both 1-drop and 2-drop instillations at the 5 minute time point. Finally, it is also worth noting that the participants included in Shaw et al.'s work had no initial LWE and were predominantly of Asian background. In contrast, the present study enrolled participants with known LWE and were predominantly Caucasian. As Shaw et al. pointed out, Asian eyelids have different geometries and have been reported to have increased LWE [39]. Future investigations should further evaluate the role of ethnicity on LWE.

The grading scales used by Korb et al. require clinicians to mentally measure lid wiper staining in two dimensions while removing the line of Marx [1,4,5]. Korb's process was initially based on an estimated linear distance of LWE (measured in mm), scaled to a 0-3 score and then averaged for two dyes [5]. Years later, the process evolved to use dyes in combination with a grading approach that averaged horizontal length of LWE with sagittal width of LWE [4]. A mental process, like this, is challenging and may contribute to difficulty and inconsistency when making clinical observations and assessments of severity.

A consistent and more precise methodology to examine LWE can facilitate meta-analyses and help identify the role (if any) that LWE plays in DED. For the use of LG and NaFl, two drops are recommended in order to fully and efficiently reveal LWE. When 2-drops of the respective dyes are instilled, LG-stained eyes can be examined after a 1-minute wait, whereas 3 to 5 minutes are needed for NaFl.

5. Acknowledgements

The authors acknowledge Scott M. Paluszkiewicz, PhD, for his editorial support and Alison Moore for her work as an independent grader in the ADCIS system.

6. References

[1] Korb DR, Herman JP, Greiner J V, Scaffidi RC, Finnemore VM, Exford JM, et al. Lid wiper epitheliopathy and dry eye symptoms. Eye Contact Lens 2005;31:2–8. https://doi.org/10.1097/01.ICL.0000140910.03095.FA.

[2] Knop E, Korb DR, Blackie CA, Knop N. The lid margin is an underestimated structure for preservation of ocular surface health and development of dry eye disease. Dev Ophthalmol 2010;45:108–22. https://doi.org/10.1159/000315024.

[3] Nelson JD, Craig JP, Akpek EK, Azar DT, Belmonte C, Bron AJ, et al. TFOS DEWS II Introduction. Ocul Surf 2017;15:269–75. https://doi.org/10.1016/J.JTOS.2017.05.005.

[4] Korb DR, Herman JP, Blackie CA, Scaffidi RC, Greiner J V, Exford JM, et al. Prevalence of lid wiper epitheliopathy in subjects with dry eye signs and symptoms. Cornea 2010;29:377–83. https://doi.org/10.1097/ICO.0b013e3181ba0cb2.

[5] Korb DR, Greiner J V, Herman JP, Hebert E, Finnemore VM, Exford JM, et al. Lid-wiper epitheliopathy and dry-eye symptoms in contact lens wearers. CLAO J 2002;28:211–6. https://doi.org/10.1097/01.ICL.0000029344.37847.5A.

[6] Efron N, Brennan NA, Morgan PB, Wilson T. Lid wiper epitheliopathy. Prog Retin Eye Res 2016;53:140–74. https://doi.org/10.1016/j.preteyeres.2016.04.004.

[7] McMonnies CW. An examination of the relationship between ocular surface tear osmolarity compartments and epitheliopathy. Ocul Surf 2015;13:110–7. https://doi.org/10.1016/j.jtos.2014.07.002.

[8] Navascues-Cornago M, Maldonado-Codina C, Gupta R, Morgan PB. Characterization of Upper Eyelid Tarsus and Lid Wiper Dimensions. Eye Contact Lens 2016;42:289–94. https://doi.org/10.1097/ICL.0000000000000230.

458 [9] Kunnen CME, Wolffsohn JS, Ritchey ER. Comparison of subjective grading of lid
459 wiper epitheliopathy with a semi-objective method. Cont Lens Anterior Eye
460 2018;41:28–33. https://doi.org/10.1016/j.clae.2017.09.008.
461 [10] Kunnen, Carolina; Percy, Lazon De La Lara; Holden, Brien A; Papas EB.
462 Automated assessment of lid margin lissamine green staining. vol. 35. C.V. Mosby
463 Co; 2014.
464 [11] Tseng SC. Evaluation of the ocular surface in dry-eye conditions. Int Ophthalmol
465 Clin 1994;34:57–69.
466 [12] Doughty MJ. Rose bengal staining as an assessment of ocular surface damage and
467 recovery in dry eye disease—A review. Contact Lens Anterior Eye 2013;36:272–80.
468 https://doi.org/10.1016/j.clae.2013.07.008.
469 [13] Begley C, Caffery B, Chalmers R, Situ P, Simpson T, Nelson JD. Review and
470 analysis of grading scales for ocular surface staining. Ocul Surf 2019;17:208–20.
471 https://doi.org/10.1016/j.jtos.2019.01.004.
472 [14] Korb DR, Herman JP, Finnemore VM, Exford JM, Blackie CA. An evaluation of the
473 efficacy of fluorescein, rose bengal, lissamine green, and a new dye mixture for
474 ocular surface staining. Eye Contact Lens 2008;34:61–4.
475 https://doi.org/10.1097/ICL.0b013e31811ead93.
476 [15] Pult H, Purslow C, Berry M, Murphy PJ. Clinical tests for successful contact lens
477 wear: relationship and predictive potential. Optom Vis Sci 2008;85:E924-9.
478 https://doi.org/10.1097/OPX.0b013e3181888909.
479 [16] Pult H, Purslow C, Murphy PJ. The relationship between clinical signs and dry eye
480 symptoms. Eye 2011;25:502–10. https://doi.org/10.1038/eye.2010.228.
481 [17] Varikooty, J., Lay, B., Jones L. Optimization of assessment and grading for lid wiper
482 epitheliopathy. Optom Vis Sci 2012:88.
483 [18] Best N, Drury L, Wolffsohn JS. Predicting success with silicone-hydrogel contact
484 lenses in new wearers. Contact Lens Anterior Eye 2013;36:232–7.
485 https://doi.org/10.1016/j.clae.2013.02.013.
486 [19] Shiraishi A, Yamaguchi M, Ohashi Y. Prevalence of upper- and lower-lid-wiper
487 epitheliopathy in contact lens wearers and non-wearers. Eye Contact Lens

2014;40:220–4. https://doi.org/10.1097/ICL.0000000000000040.

[20] Guthrie SE, Jones L, Blackie CA, Korb DR. A comparative study between an oil-in-water emulsion and nonlipid eye drops used for rewetting contact lenses. Eye Contact Lens 2015;41:373–7. https://doi.org/10.1097/ICL.0000000000000138.

[21] Jalbert I, Madigan MC, Shao M, Ng J, Cheng J, Wong D, et al. Assessing the human lid margin epithelium using impression cytology. Acta Ophthalmol 2012;90:e547–52. https://doi.org/10.1111/j.1755-3768.2012.02482.x.

[22] Navascues-Cornago M, Morgan PB, Maldonado-Codina C. Lid Margin Sensitivity and Staining in Contact Lens Wear Versus No Lens Wear. Cornea 2015;34:808–16. https://doi.org/10.1097/ICO.0000000000000448.

[23] Satjawatcharaphong P, Ge S, Lin MC. Clinical outcomes associated with thermal pulsation system treatment. Optom Vis Sci 2015;92:e334–41. https://doi.org/10.1097/OPX.0000000000000670.

[24] Varikooty J, Srinivasan S, Subbaraman L, Woods CA, Fonn D, Simpson TL JL. Variations in observable lid wiper epitheliopathy (LWE) staining patterns in wearers of silicone hydrogel lenses. Contact Lens Anterior Eye 2015;38:471–6.

[25] Varikooty J, Srinivasan S, Jones L. Atypical manifestation of upper lid margin staining in silicone hydrogel lens wearers with symptoms of dry eye. Contact Lens Anterior Eye 2008;31:44–6. https://doi.org/10.1016/j.clae.2007.07.001.

[26] Bron AJ, Argüeso P, Irkec M, Bright F V. Clinical staining of the ocular surface: mechanisms and interpretations. Prog Retin Eye Res 2015;44:36–61. https://doi.org/10.1016/j.preteyeres.2014.10.001.

[27] Wolffsohn JS, Arita R, Chalmers R, Djalilian A, Dogru M, Dumbleton K, et al. TFOS DEWS II Diagnostic Methodology report. Ocul Surf 2017;15:539–74. https://doi.org/10.1016/j.jtos.2017.05.001.

[28] Hamrah P, Alipour F, Jiang S, Sohn J-H, Foulks GN. Optimizing evaluation of Lissamine Green parameters for ocular surface staining. Eye (Lond) 2011;25:1429–34. https://doi.org/10.1038/eye.2011.184.

[29] Akaike H. Information theory and an extension of the maximum likelihood principle. In: Petrov BN FCF, editor. Proc. 2nd Int. Symp. Inf. Theory, Budapest: 1998, p.

199–213. https://doi.org/10.1007/978-1-4612-1694-0_15.

[30] Schwarz G. Estimating the Dimension of a Model. Ann Stat 1978;6:461–4. https://doi.org/10.1214/aos/1176344136.

[31] Wolffsohn JS, Naroo SA, Christie C, Morris J, Conway R, Maldonado-Codina C, et al. Anterior eye health recording. Contact Lens Anterior Eye 2015;38:266–71. https://doi.org/10.1016/j.clae.2015.03.001.

[32] Li W, Yeh TN, Leung T, Yuen T, Lerma M, Lin MC. The Relationship of Lid Wiper Epitheliopathy to Ocular Surface Signs and Symptoms. Investig Opthalmology Vis Sci 2018;59:1878. https://doi.org/10.1167/iovs.17-23639.

[33] Markoulli M, Isa NA PE. Temporal Characteristics of Sodium Fluorescein in the Tear Meniscus. Optom Vis Sci 2017;94:166–73.

[34] Muntz A, Subbaraman LN, Craig JP, Jones L. Cytomorphological assessment of the lid margin in relation to symptoms, contact lens wear and lid wiper epitheliopathy. Ocul Surf 2020;18:214–20. https://doi.org/10.1016/j.jtos.2019.12.001.

[35] Delaveris A, Stahl U, Madigan M, Jalbert I. Comparative performance of lissamine green stains. Contact Lens Anterior Eye 2018;41:23–7. https://doi.org/10.1016/j.clae.2017.11.002.

[36] Foulks GN. Challenges and pitfalls in clinical trials of treatments for dry eye. Ocul Surf 2003;1:20–30.

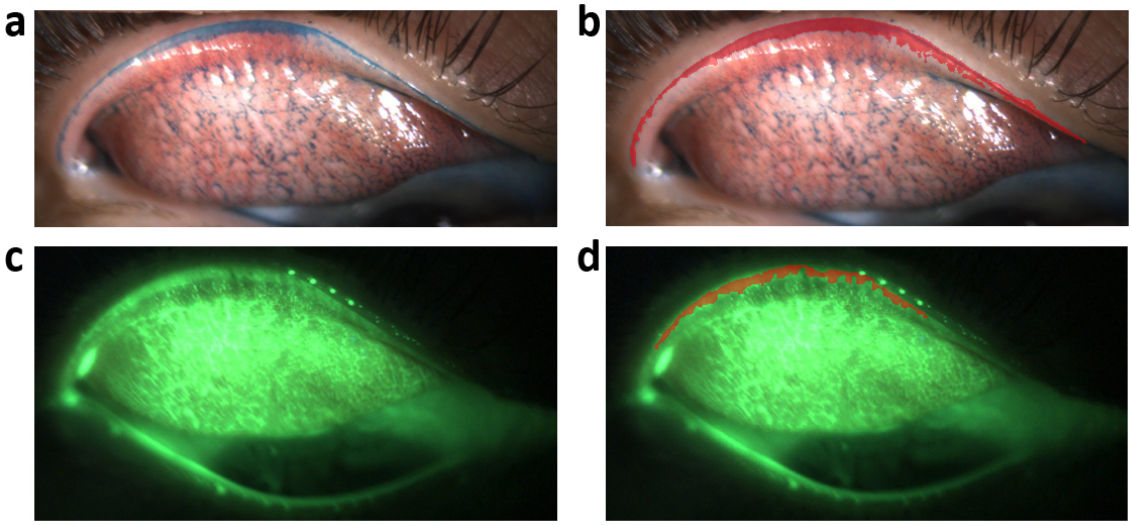
[37] Siddireddy JS, Vijay AK, Tan J, Willcox M. The eyelids and tear film in contact lens discomfort. Contact Lens Anterior Eye 2018;41:144–53. https://doi.org/10.1016/j.clae.2017.10.004.

[38] Shiraishi A, Yamaguchi M, Ohashi Y. Prevalence of upper- and lower-lid-wiper epitheliopathy in contact lens wearers and non-wearers. Eye Contact Lens 2014;40:220–4. https://doi.org/10.1097/ICL.0000000000000040.

[39] Shaw A, Collins M, Huang J, Nguyen HMP, Kim Z, Lee G, et al. Lid wiper epitheliopathy: The influence of multiple lid eversions and exposure time. Cont Lens Anterior Eye 2019;42:304–10. https://doi.org/10.1016/j.clae.2018.09.003.

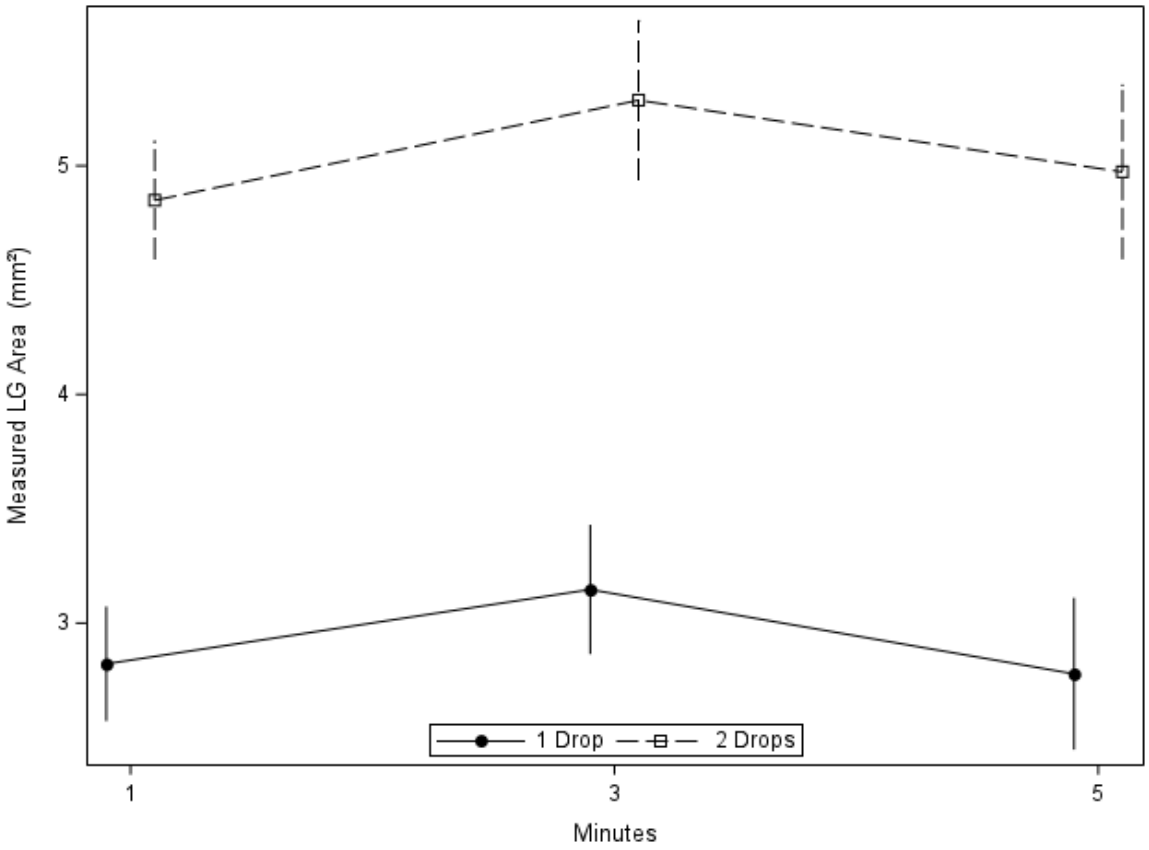
548 **7. Figure Legend**

549 Figure 1. Area of lid wiper epitheliopathy (LWE) staining showing: a) Area of LWE
550 staining using LG before semi-objective processing, b) Area of LWE staining using LG
551 staining after semi-objective processing, c) Area of LWE staining using NaFl before semi-
552 objective processing and d) Area of LWE staining using NaFl after semi-objective
553 processing.



554

555 Figure 2. Staining area (mm^2) 1, 3, and 5 minutes following instillation of lissamine green
556 (LG). Values presented are mean ± standard error.



557

Figure 3. Staining area (mm^2) 1, 3, and 5 minutes following instillation of sodium fluorescein (NaFl). Values presented are mean ± standard error.

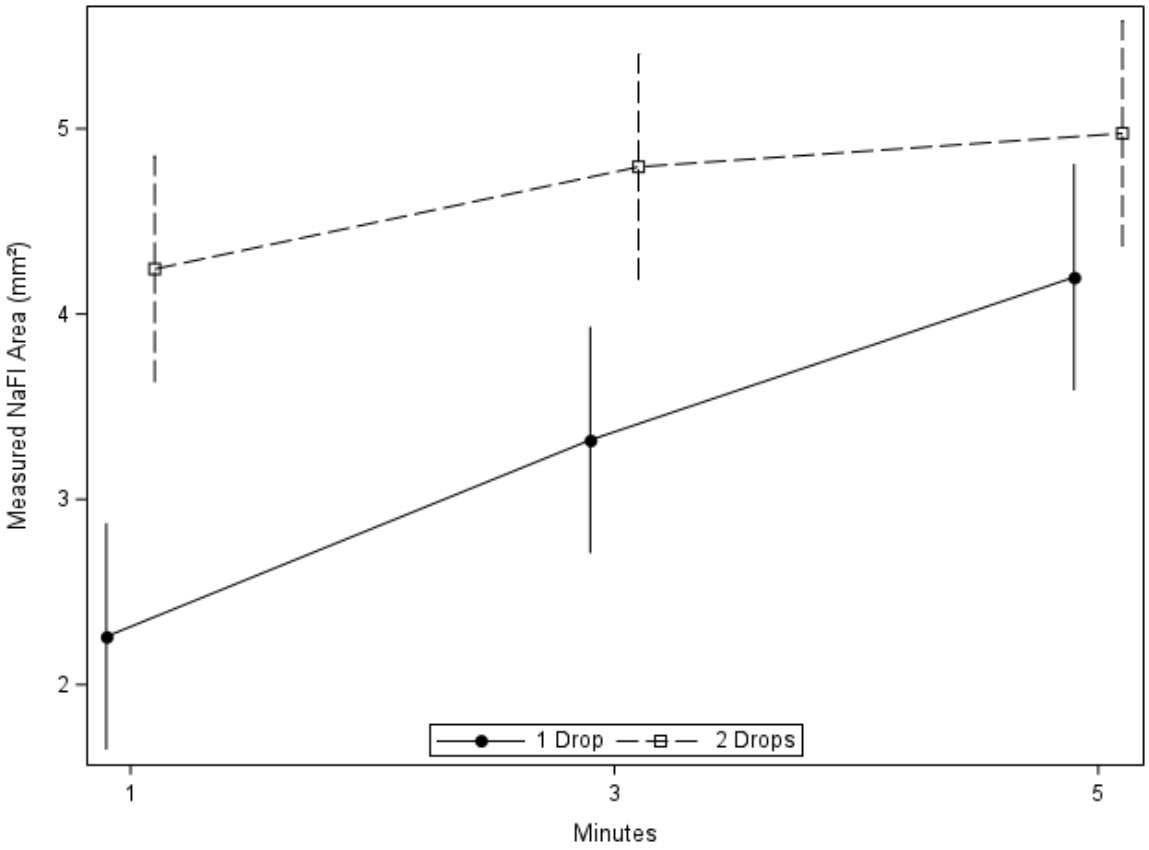


Figure 4. Intra-observer Analysis LG. Bland-Altman plots for lid wiper epitheliopathy area of staining using 2-drops of LG at the 3 minute observation time point (ADCIS system by same investigator with the same images at initial processing and 6 months later). The solid line shows the mean difference between repeats (-0.10 mm2) and the dashed lines show the 95% limits of agreement (-1.10 to 0.89 mm2). LG, lissamine green.

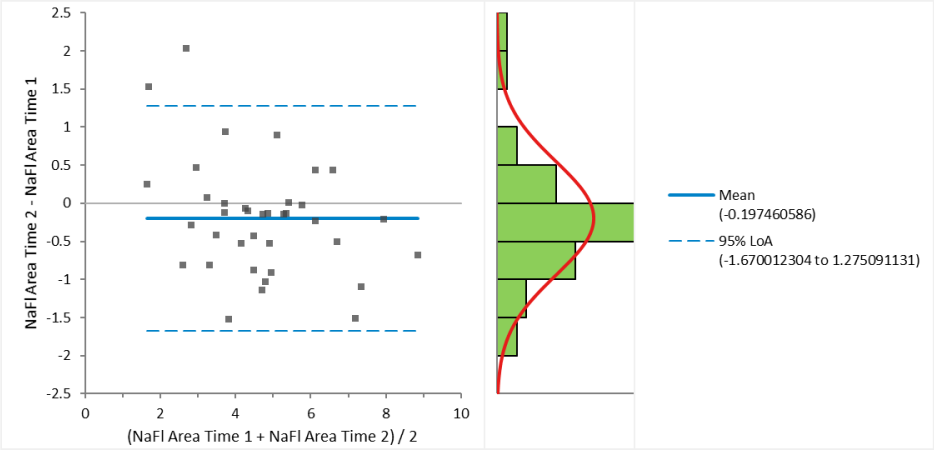


Figure 5. Intra-observer Analysis NaFl. Bland-Altman plots for lid wiper epitheliopathy area of staining using 2-drops of NaFl at the 3 minute observation time point (ADCIS system by same investigator with the same images at initial processing and 6 months later). The solid line shows the mean difference between repeats (-0.20 mm2) and the dashed lines show the 95% limits of agreement (-1.67 to 1.28 mm2). NaFl, sodium fluorescein.

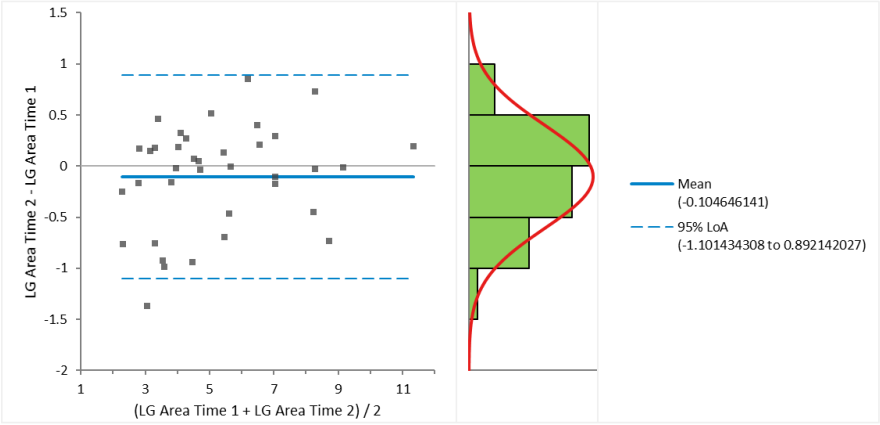


Figure 6. Inter-observer Analysis LG. Bland-Altman plots for lid wiper epitheliopathy area of staining using 2-drops of LG at the 3 minute observation time point (ADCIS system with a new investigator with the same images at initial processing and 6 months later). The solid line shows the mean difference between repeats (-0.15 mm2) and the dashed lines show the 95% limits of agreement (-1.53 to 1.23 mm2). LG, lissamine green.

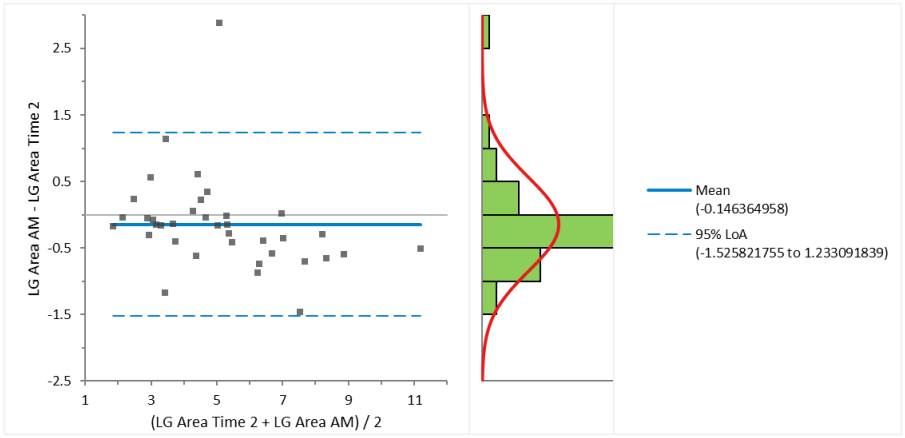


Figure 7. Inter-observer Analysis NaFl. Bland-Altman plots for lid wiper epitheliopathy area of staining using 2-drops of NaFl at the 3 minute observation time point (ADCIS system with a new investigator with the same images at initial processing and 6 months later). The solid line shows the mean difference between repeats (-0.30 mm2) and the dashed lines show the 95% limits of agreement (-3.03 to 2.43 mm2). NaFl, sodium fluorescein.