Cornea

Automated Tear Film Surface Quality Breakup Time as a Novel Clinical Marker for Tear Hyperosmolarity in Dry Eye Disease

Laura Elizabeth Downie

Department of Optometry and Vision Sciences, Faculty of Medicine, Dentistry & Health Sciences, University of Melbourne, Parkville, Victoria, Australia

PURPOSE. To assess the diagnostic performance of a novel, automated, noninvasive measure of tear film stability derived from Placido disc videokeratography, the tear film surface quality breakup time (TFSQ-BUT), as a clinical marker for diagnosing dry eye disease (DED) relative to a standard of tear hyperosmolarity.

METHODS. This prospective, cross-sectional study involved 45 participants (28 DED, 17 controls). Symptoms (Ocular Surface Disease Index) and signs (tear osmolarity, TFSQ-BUT, tear breakup time measured with sodium fluorescein [NaFl-BUT], ocular surface staining and Schirmer test with topical anesthesia) of DED were assessed. Three measures of TFSQ-BUT and NaFl-BUT were taken per eye; "first," "average," and "shortest" BUT were analyzed separately. Optimal diagnostic cutoff values were determined using the Youden Index. The repeatability and agreement of the TFSQ-BUT was compared with two clinicians who manually assessed noninvasive BUT (CNI-BUT). Repeatability of methods was assessed using the geometric coefficient of variation (gCoV, %). Agreement between methods was considered with Bland-Altman analysis.

RESULTS. Eyes with DED had significantly shorter TFSQ-BUTs than controls ($P < 0.05$). There was a significant, moderate correlation between both shortest and average TFSQ-BUT and NaFl-BUT ($r = 0.35$, $P = 0.02$ and $r = 0.38$, $P = 0.01$, respectively). The receiver-operator characteristic (ROC) curve for shortest TFSQ-BUT showed an area under the curve of 0.92 ($P < 0.0001$). Shortest TFSQ-BUT with a criterion of 12.1 seconds had a sensitivity of 82% and specificity of 94% for diagnosing DED against tear hyperosmolarity. Automated TFSQ-BUT showed less variability (gCoV = 9.4%, 95% confidence interval [CI]: 7.1%–14.0%) than CNI-BUT (gCoV = 27.0%, 95% CI: 19.62%–41.06%, $P < 0.05$).

CONCLUSIONS. Automated TFSQ-BUT is a repeatable, noninvasive clinical marker with both high sensitivity and specificity for tear hyperosmolarity.

Keywords: dry eye, diagnosis, tear film, tear breakup time, TBUT, tear osmolarity, hyperosmolarity, noninvasive, NI-TBUT, ROC curve

Dry eye disease (DED) is a "multifactorial disease of the tears and ocular surface that results in symptoms of discomfort, visual disturbance, and tear film instability. It is accompanied by increased osmolarity of the tear film and inflammation of the ocular surface."[1] Dry eye disease is underwritten by perturbations to the lacrimal functional unit (LFU), consisting of the lacrimal gland and its accessory glands, cornea, conjunctiva, meibomian glands, eyelids, and their associated sensory and motor nerves.[2] Under physiologic conditions, the integrated LFU regulates the secretion, distribution, and clearance of tears, in response to endogenous and exogenous factors, to preserve ocular surface integrity.[2] Disruption to one or more components of the LFU can result in a loss of tear homeostasis, thereby leading to tear film dysfunction.

Tear hyperosmolarity is a key feature of DED.[1,3,4] Reduced aqueous production and/or excessive tear evaporation decrease(s) tear film volume and increase(s) tear protein and electrolyte concentration.[5,6] A meta-analysis of 16 studies identified a reference value of 316 mOsmol/L to be specific

for diagnosing clinically significant DED, providing a sensitivity (true positive rate) of 73% and specificity (true negative rate) of 90%.[7] Most studies included in this analysis used laboratory-based, freezing-point depression techniques to measure tear osmolarity,[7] which is regarded as the "gold standard" method of assessment. This technique differs from electrical impedance osmometry, used by the osmolarity system (TearLab; TearLab Corp., San Diego, CA, USA). While measurements using both methods to quantify tear osmolarity from human tear samples have been reported to correlate,[8] differences in measures have also been described for assaying standard solutions of known osmolarity.[9,10] Measurement reliability with the TearLab system can be influenced by a range of factors, such as ambient temperature,[11] which may account for differences in findings between studies.

A number of studies have reported a lack of correlation between tear osmolarity measures and traditional dry eye diagnostic tests, including the Schirmer test, sodium fluorescein tear breakup time (NaFl-BUT), and corneal staining with vital

dyes.[12-16] However, such findings, which speak to the complexity surrounding the clinical diagnosis of DED,[17] are not unanticipated. In contrast to tear osmolarity, which is objective, quantitative, and can be measured through acquiring a minute (50 nL) tear volume, most traditional diagnostic tests for DED are invasive (artificially disrupting tear film status) and rely on subjective assessment. A quantitative, automated, clinical parameter that noninvasively assesses tear film integrity might therefore be expected to more closely correlate with tear film osmolarity measurements and serve as a useful surrogate diagnostic marker for DED.

In this regard, a relevant potential candidate is noninvasive tear breakup time (NI-TBUT), which is intended to quantify natural tear film stability.[18] Although tear hyperosmolarity and tear instability are both key components in the definition of DED,[1] to the author's knowledge, this is the first study to have directly considered their potential relationship. This study tests a hypothesis that a novel, automated noninvasive measure of tear stability, herein termed the tear film surface quality breakup time (TFSQ-BUT) derived from dynamic-area high-speed Placido disc videokeratography, is a useful surrogate marker for diagnosing DED, relative to a current standard of tear hyperosmolarity.

METHODS

This research project was conducted in accordance with the tenets of Declaration of Helsinki and was approved by the University of Melbourne Human Research Ethics Committee (Health Sciences subcommittee).

Study Participants

This cross-sectional study was undertaken at the University of Melbourne EyeCare Clinic and involved 45 adults (29 female, 16 male) who provided written informed consent to participate. Exclusion criteria included any of the following: contact lens wear, anterior segment surgery, keratorefractive procedures, history of ocular disease (other than DED), pregnancy, topical medications (other than ocular lubricants), punctal plug placement or ocular trauma, and within 6 months of enrollment. Participants underwent a comprehensive dry eye clinical examination including (listed in order of testing): symptom assessment (Ocular Surface Disease Index, OSDI), tear osmolarity measurement (TearLab, TearLab Corp.), assessment of non-invasive tear break-up time (TFSQ-BUT), full slit lamp examination (including sodium fluorescein tear break up time (NaFl-BUT) corneal NaFl staining and conjunctival lissamine green (LG) staining, both measured using the Oxford scale[19]) and the Schirmer test with topical anesthesia. All assessments were performed in the same clinical environment.

Potential participants ($n = 60$) were screened to determine their eligibility, which involved criteria relating to both dry eye clinical symptoms and signs. A clinical definition of DED was considered as an OSDI score of ≥ 18,[20] tear osmolarity ≥ 316 mOsmol/L and at least one (of three) consecutive NaFl-BUTs less than 10 seconds, in at least one eye. This tear osmolarity criterion has a reported predictive diagnostic accuracy of 89% for clinically significant DED.[7] Control participants had an OSDI score of ≤ 12 and bilateral tear osmolarity < 308 mOsmol/L. Potential participants who both satisfied these inclusion criteria and were not deemed ineligible to participate from the specified exclusion criteria ($n = 28$ DED, $n = 17$ controls) were included in this study. DED participants were also being screened for possible participation in a dry eye intervention study. For all participants, the eye with higher tear

osmolarity was defined as the 'study eye,' being a measure of DED severity,[21] and used for all subsequent analyses.

Tear Osmolarity

Tear osmolarity was measured bilaterally from the inferior lateral tear meniscus using the TearLab (TearLab Corp.) system. The instrument was calibrated daily, according to the manufacturer's instructions. Room temperature was maintained between 20°C and 24°C. Participants using lubricant drops were instructed not to instill these for at least 2 hours prior to the study visit; this was confirmed prior to taking measurements. The same diagnostic pen was used for all assessments. Right eyes were measured first.

Automated TFSQ-BUT

Dynamic-area, high-speed Placido disc videokeratography was performed using the E300 corneal topographer (Medmont International Pty Ltd., Victoria, Australia). Participants were instructed to focus on the central fixation target, gently blink twice and then to suppress blinking during the capture period. Using a frame rate of four photokeratoscopic images per second (4 Hz), a video was captured of the reflected Placido disc mires for up to 23 seconds post-blink. Three measures were taken on each eye, alternating between eyes, with right eyes measured first. Participants were advised to blink freely between each measurement.

The E300 corneal topography system noninvasively analyzes changes in tear film stability by analyzing the structure of the reflected Placido disc image. The software calculates TFSQ values at 300 radial analysis points along each of the 32 rings, using an approach similar to the block-feature TFSQ indicator described by Alonso-Caneiro and colleagues.[22] However, whereas this paper calculates the block-feature TFSQ from raw image data, the Medmont software leverages the existing E300 topographical analysis algorithm to extract the ring location data from the image. Significantly, the topographical analysis algorithm is able to identify and eliminate images with excessive movement and to correctly identify ring reflections in areas that contain shadows from eyelashes. The local TFSQ value at a given analysis point is calculated by finding the SD of the radial distances to the next innermost ring for $n = 8$ points (9.6°) either side of the analysis point.

First, the average ring width at the point i is calculated as:

$$w_{av_i} = \frac{\sum_{j=i-n}^{i+n} w_j}{n}$$

The standard deviation of the widths of the surrounding $n = 8$ points is then calculated by:

$$SD_i = \sqrt{\frac{\sum_{j=i-n}^{i+n} (w_j - w_{av_i})^2}{n}}$$

Finally, the local TFSQ value is calculated as a dimensionless value:

$$TFSQ_i = 100 \left(\frac{SD_i}{w_{av_i}} \right)^2$$

A local TFSQ value of 0.30 or greater corresponds to visible distortion in the ring pattern. The following novel parameters, based upon the TFSQ index,[22] are then defined: (1) TFSQ-Area (%): the percentage of the tear film examination area (defined
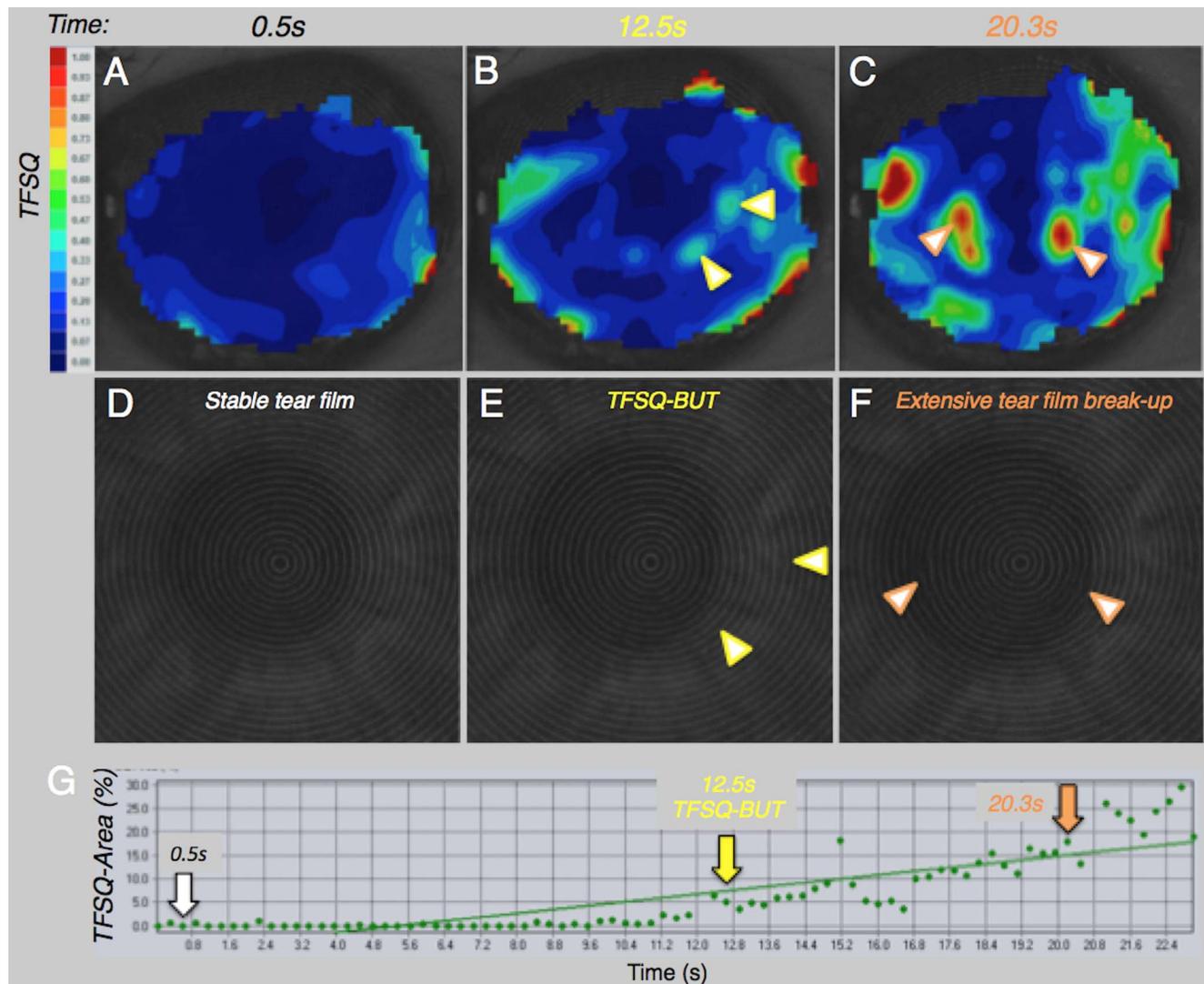
FIGURE 1. Representative TFSQ index color maps, taken at (**A**) 0.5, (**B**) 12.5, and (**C**) 20.3 seconds post-blink. Increasing tear instability is appreciable by a progressive increase in the proportion of warmer colors (*yellow* and *red*) in the TFSQ map. (**D**) Enlargements of the Placido disc rings at each time-point highlight the absence of distortion at 0.5 seconds post-blink. (**E**) Early tear film breakup at 12.5 seconds. (**F**) Extensive distortion of the mires by 20.3 seconds. (**G**) A regression plot shows the change in TFSQ-Area (%) over time post-blink (seconds). In this capture, the TFSQ-BUT is 12.5 seconds.

by a 7 mm diameter) with a TFSQ index greater than 0.30; and (2) TFSQ-BUT: time (in seconds) at which the TFSQ-Area (%) is calculated to be at least 5.0% in two consecutive photo-keratoscopic images.

Figure 1 shows representative TFSQ index color maps, taken at 0.5, 12.5, and 20.3 (Figs. 1A–C) seconds post-blink. Increasing tear instability is visually appreciable by a progressive increase in the proportion of warmer colors (yellow and red) in the TFSQ map. Enlargements of the Placido disc rings show an absence of distortion at 0.5 seconds post-blink (Fig. 1D), early tear film breakup at 12.5 seconds (Fig. 1E), and extensive distortion of the mires by 20.3 seconds (Fig. 1F).

Sodium Fluorescein Tear Breakup Time (NaFl-BUT)

NaFl-BUT was measured using Dry Eye Test (DET) NaFl strips (Amcon Laboratories, Inc., St. Louis, MO, USA), consisting of a slim applicator tip impregnated with 0.12 mg of NaFl to instill approximately 1 µL of fluid into the eye. Strips of DET NaFl

have shown improved measurement reliability and enhanced precision compared with conventional strips.[23] A single DET was moistened with non-preserved saline and applied to the superior bulbar conjunctiva. One minute after NaFl instillation, NaFl-BUT was measured at a slit lamp biomicroscope using ×10 magnification with cobalt blue illumination and a Wratten 12 yellow-barrier filter. Participants were asked to gently blink twice and then to hold their eyes open for as long as possible; at the end of the second blink, a stopwatch was started and then stopped once the first tear break-up was noted. Right eyes were measured first; three NaFl-BUTs were recorded for each eye.

Total Ocular Surface Staining

Corneal NaFl staining was examined using a slit lamp biomicroscope (×16 magnification with cobalt blue illumination and a Wratten 12 yellow-barrier filter). Nasal and temporal conjunctival LG staining was assessed, 1 minute after

TABLE 1. Study Participant Characteristics

Characteristic	Group	
	Control, $n = 17$	DED, $n = 28$
Age, y	45.5 ± 14.6 (21–74)	41.2 ± 14.4 (18–67)
Sex, male:female	7:10	9:19
Tear osmolarity, mOsmol/L	296.2 ± 9.9 (276–307)	325.0 ± 17.8 (317–400)***
Total ocular surface staining score	0.6 ± 0.2 (0–2.5)	1.2 ± 0.2 (0–4.2)*
Schirmer with topical anesthesia, mm/5 min	15.8 ± 1.8 (6–28)	9.0 ± 1.5 (0–24)***

Findings relating to clinical parameters (i.e., tear osmolarity, "total ocular surface staining score" and Schirmer test) are from the eye with highest tear osmolarity. The "total ocular surface staining" score (ranging from 0.0–15.0) is defined as the sum of the NaFl corneal, LG temporal conjunctival and LG nasal conjunctival scores, each graded using the five-step Oxford scale.[19] Asterisks show statistically significant differences between treatment groups. Numbers in square brackets show range of values.

* $P < 0.05$.
*** $P < 0.001$.

instillation of LG (GreenGlo, Sigma Pharmaceuticals, North Liberty, IA, USA). Staining was assessed using the five-point Oxford scale[19] with 0.1 grading increments. A "total ocular surface staining" score (ranging from 0.0–5.0), defined as the sum of the NaFl corneal, LG temporal conjunctival and LG nasal conjunctival scores, was calculated.

Schirmer Test

One drop of 0.5% proxymetacaine hydrochloride (Alcon Laboratories, New South Wales, Australia) was instilled into each inferior conjunctival sac. After four minutes, the folded edge of a sterile Schirmer strip (EagleVision, Memphis, TN, USA) was gently inserted between the middle and lateral third of each lower lid margin. Participants were instructed to close their eyes in the dimly lit room. After 5 minutes, the length of strip wetting was recorded in millimeters.

Repeatability and Agreement of the Automated TFSQ-BUT

A subset of Placido-disc videos ($n = 20$) was used for repeatability and agreement tests on the automated TFSQ-TBUT. Each 4-Hz video was split into two 2-Hz videos, using a custom function (supplied by Medmont Pty Ltd., Victoria, Australia), that allocated alternate even and odd numbered images within the video to separate exams. Two research clinicians viewed the videos of the Placido ring reflections (in the absence of the TFSQ color map overlay) and manually assessed the noninvasive tear breakup time, in seconds, (termed the clinician-derived noninvasive break-up time, CNI-BUT), as indicated by the first observation of distortion of the Placido rings. Following an initial presentation of videos ($n = 20$) derived from even-numbered images, videos consisting of odd-numbered images were randomly presented for a repeat assessment. All assessments were undertaken in one sitting, under normal room illumination. Automated TFSQ-BUTs were also calculated for each video.

Intramethod repeatability, being the variability in automated TFSQ-BUT and CNI-BUT between repeated trials, was examined using a geometric coefficient of variation (gCoV, %) as described by Hopkins[24] and Vaz.[25] Intermethod (automated versus clinician-derived noninvasive BUT) variability was examined using Bland-Altman analysis.[26] The mean difference (bias) and limits of agreement (LoA, defined as bias ± two SDs of the mean difference) were calculated. The method described by Carkeet,[27] which takes into account sample size,

was used to calculate exact 95% confidence limits (CLs) for the LoAs.

Statistical Analysis

Data were analyzed using spreadsheet software (Microsoft Excel; Microsoft Office for Mac 2011, version 14.4.1, Microsoft Corp., Redmond, WA, USA) and graphing software (GraphPad Prism 5; GraphPad Software, San Diego, CA, USA). Descriptive statistics are summarized as mean ± SD. A Kolmogorov-Smirnov test was used to assess for normality of continuous variables. Comparisons between groups were undertaken using either a *t*-test or Mann-Whitney *U* test, as appropriate. A χ^2 test was used to compare discrete variables. Pairwise correlations between TFSQ-BUT and NaFl-BUT parameters were explored using Spearman's correlation coefficient (*r*). An alpha value of 0.05 was adopted for statistical significance.

To investigate the diagnostic capacity of a number of TFSQ-BUT and NaFl-BUT parameters (i.e., shortest, average and first of the three measured BUTs using each method), the sensitivity and specificity for each parameter was determined against a standard diagnostic criterion for DED (tear hyperosmolarity of ≥316 mOsmol/L).[7] Receiver-operator characteristic (ROC) curves, showing sensitivity versus false positive rate (1 – specificity), were plotted and used to evaluate the discriminative capacity of the BUT parameters. The area under the ROC curve (AUC) was calculated to provide a measure of the overall performance of each BUT parameter (i.e., a measure of diagnostic accuracy). Optimal diagnostic cutoff values for BUT parameters were determined using the Youden index, which measures the distance of each point on the ROC curve from the identity (diagonal) line, for each criterion; the maximum value is considered the criterion of interest.[28] Likelihood ratios (LRs), being the gradient of the ROC curve at the cutoff criterion, were calculated. Likelihood ratios represent the ratio between the probability of a positive test result given the presence of disease and the probability of a positive test result given the absence of disease.

RESULTS

Participant Characteristics

Demographic and dry eye clinical characteristics of participants are summarized in Table 1. Control ($n = 17$) and DED ($n = 28$) participants had similar age and sex distributions. Consistent with the enrollment criteria applied to each study population, participants with DED had significantly higher tear

TABLE 2. Summary of TFSQ-BUT and NaFl-BUT Parameters in Each Group

BUT Parameter, s	Control, $n = 17$	DED, $n = 28$
Shortest TFSQ-BUT	19.4 ± 5.3	7.9 ± 4.9***
Average TFSQ-BUT	21.3 ± 2.7	13.4 ± 5.2***
First TFSQ-BUT	21.2 ± 3.5	13.9 ± 7.4***
Shortest NaFl-BUT	8.5 ± 5.3	5.9 ± 3.0
Average NaFl-BUT	12.7 ± 7.9	7.3 ± 4.2*
First NaFl-BUT	11.8 ± 8.5	6.7 ± 3.7*

Data are shown for the eye with highest tear osmolarity. Asterisks show statistically significant differences between groups.
* $P < 0.05$.
** $P < 0.01$.
*** $P < 0.001$.

osmolarity (325 ± 17.8 vs. 296 ± 9.9 mOsmol/L; $P < 0.001$), greater ocular surface staining (total ocular surface staining scores: 1.2 ± 0.2 vs. 0.6 ± 0.2, $P < 0.05$) and lower Schirmer test scores (9.0 ± 1.5 vs. 15.8 ± 1.8, $P < 0.001$) than control participants.

Comparison of TFSQ-BUT and NaFl-BUT Parameters

Table 2 summarizes group findings for TFSQ-BUT and NaFl-BUT parameters. There was no significant order effect for BUT measures in either study group (data not shown). For both methods, parameters for the shortest, average, and first of three repeated BUT measures are shown. Overall, noninvasive TFSQ-BUTs were longer than NaFl-BUTs in both groups. For each TFSQ-BUT parameter, DED eyes showed significantly shorter BUTs than controls. For the NaFl-based tests, significant differences between groups were found for the average (of three) consecutive NaFl-BUT measures and the first NaFl-BUT, with both parameters being shorter in DED eyes.

Pairwise correlations between TFSQ-BUT and NaFl-BUT parameters were explored over the full dataset (Table 3). Moderately positive correlations were observed between shortest TFSQ-BUT and NaFl-BUT ($r = 0.35$, $P < 0.05$) and average TFSQ-BUT and NaFl-BUT ($r = 0.38$, $P = 0.05$).

Diagnostic Performance of TFSQ-BUT and NaFl-BUT Parameters for DED

ROC curves showing the discriminative capacity of the TFSQ-BUT and NaFl-BUT parameters for tear hyperosmolarity in DED are shown in Figures 2A and 2B, respectively. Overall, TFSQ-BUT parameters showed significantly better diagnostic performance than NaFl-BUT parameters. The ROC curves for both shortest and average TFSQ-BUT had an AUC of 0.92 ($P < 0.001$; Table 4). Of the NaFl-BUT parameters, first NaFl-BUT trended

TABLE 3. Summary of Correlation Between TFSQ-BUT and NaFl-BUT Parameters

BUT Parameter (TFSQ-BUT vs. NaFl-BUT)	Correlation Coefficient, r (95% CI)
Shortest	0.35 (0.06, 0.59)*
Average	0.38 (0.09, 0.61)*
First	0.22 (−0.09, 0.49)

Data are shown for the eye with highest tear osmolarity. Asterisks show statistically significant correlations; * $P < 0.05$.

toward the best diagnostic performance, with an AUC of 0.72 ($P = 0.013$; Table 4), however, this difference was overall not significant compared with shortest and average NaFl-BUTs ($P > 0.05$).

Optimal cutoff values for diagnosing tear hyperosmolarity in DED were determined (Table 4). Shortest TFSQ-BUT showed the best discriminative capacity with a sensitivity of 82% and specificity of 94%, using a criterion of 12.1 seconds. Average TFSQ-BUT had a sensitivity of 77% and specificity of 94% with a cutoff of 16.1 seconds. Of the fluorescein-based tests, first NaFl-BUT showed the best diagnostic power (sensitivity: 78%, specificity: 72% using a criterion of 8.0 seconds).

Plots showing the relationship between tear osmolarity and noninvasive (shortest TFSQ-BUT) and fluorescein-derived (first NaFl-BUT) parameters with best diagnostic capacity for DED are shown in Figures 3A and 3B, respectively. Consistent with shortest TFSQ-BUT having both higher sensitivity and specificity than first NaFl-BUT, there were overall fewer false positives and false negative classifications using this parameter, with a criterion of 12.1 seconds.

Repeatability and Agreement of the Automated TFSQ-BUT

Intramethod repeatability was determined for both the automated TFSQ-BUT and CNI-BUT. Automated TFSQ-BUT had significantly ($P < 0.05$) less intramethod variability (gCoV = 9.4%, 95% CI: 7.1%–14.0%) than CNI-BUT (gCoV = 27.0%, 95% CI: 19.62%–41.06%).

Intermethod agreement was assessed using a Bland-Altman analysis for the automated TFSQ-BUT and CNI-BUT. CNI-BUT showed a bias toward shorter breakup times by 2.0 seconds; however, this was not statistically significant. The 95% LoA was −3.7 to 7.8 seconds (Fig. 4); 95% CLs for the LoA, factoring in sample size, were: upper bound (6.39–10.48 seconds), lower bound (−6.40 to −2.31 seconds).

Discussion

This paper reports that a novel, automated noninvasive measure of tear film stability derived from Placido disc videokeratography (TFSQ-BUT), has sufficient discriminative capacity to be a valuable marker of tear hyperosmolarity in DED. Using a cutoff value of 12.1 seconds, the shortest TFSQ-BUT (of three repeat measures) showed high sensitivity (82%) and specificity (94%) for diagnosing moderate to severe DED against tear hyperosmolarity; this diagnostic performance was superior to traditional NaFl-BUT measures. A modest correlation ($r \approx 0.4$) was evident between TFSQ-BUT and NaFl-BUT parameters. Automated measures of TFSQ-BUT were significantly more repeatable than clinician-derived estimates of noninvasive tear breakup time, supporting the utility of this examiner independent measure in clinical settings.

The diagnosis of DED is recognized to be challenging, necessitating the undertaking of multiple clinical techniques in order to try and obtain a complete impression of a patient's ocular surface health and tear film integrity.[17] Tear hyperosmolarity, particularly with the application of freezing-point depression techniques, has been reported to have superior overall diagnostic capacity for DED compared with a range of standard diagnostic tests.[1] Despite this, a number of barriers to its widespread clinical adoption, even with the clinically-practicable system (TearLab Corp.) that utilizes electrical impedance, have been described.[29] Tear osmolarity assessment requires specialized instrumentation, daily calibration of this equipment, and test card consumables. Obtaining reliable tear osmolarity measurements necessitates the temperature of the
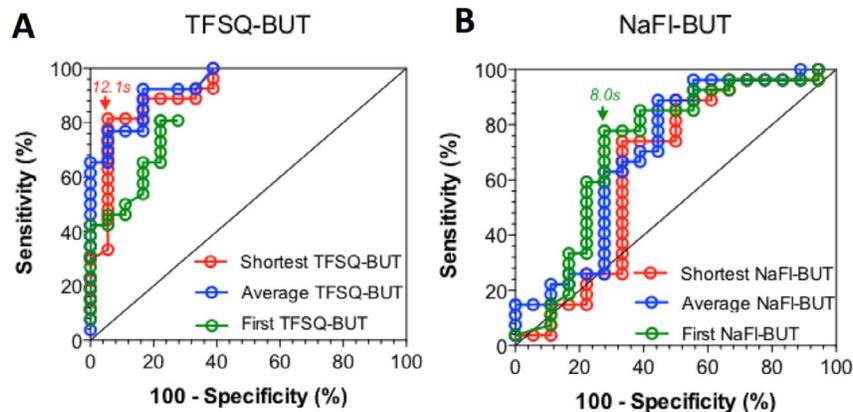
FIGURE 2. ROC curves showing the performance of (**A**) TFSQ-BUT and (**B**) NaFl-BUT parameters for diagnosing dry eye disease, measured against a standard of tear hyperosmolarity. Optimal cutoff values are shown. (**A**) Shortest TFSQ-BUT with a cutoff of 12.1 seconds. (**B**) First NaFl-BUT with a cutoff of 8.0 seconds.

clinical environment to be appropriately regulated.[11] Adequate control of such variables is critical for the reliability of the assay and was carefully maintained in this study. Therefore, clinicians utilizing tear osmolarity measures for diagnosing DED need to be aware of the importance of controlling for such factors.

Recent studies of the self-reported clinical practice behaviors of eye care clinicians in multiple demographics show that NaFl-BUT remains as one of the most common diagnostic tests for DED.[30-34] Clinical preferences for NaFl-BUT measurements are not surprising; the test can be rapidly undertaken and is inexpensive to perform. However, a major disadvantage of this test is that the instillation of fluid into the eye destabilizes the tear film.[35] Measurement of NaFl-BUT can also be influenced by factors such as pH, the volume of fluorescein instilled, slit lamp illumination technique and the clinician's expertise.[17,18] Using microquantities of NaFl, as was undertaken for this study, improves the reproducibility of NaFl-BUT measurements.[36] With such volumes, a NaFl-BUT of 5 seconds or less has been reported to differentiate between dry eye and control populations; however, sensitivity and specificity calculations for this cutoff value were not provided.[37]

This study reports that when measuring three consecutive NaFl-BUTs (with a 1 μL NaFl volume), using the first measure with a cutoff of 8.0 seconds has 78% sensitivity and 72% specificity for diagnosing DED against tear hyperosmolarity. These diagnostic performance values are similar to those

previously reported for NaFl-BUT parameters against different diagnostic parameters for DED.[3,38] Although the ROC characteristics showed a trend toward first NaFl-BUT having superior discriminative capacity relative to average and shortest NaFl-BUT, this was not significant ($P > 0.05$). These findings support the conclusions of Papas,[39] who noted the first determination of NaFl-BUT to be as good as any measure in a sequence of consecutive readings.

In recent years, there has been a research-driven trend to develop relatively less invasive measures of tear stability.[18] Clinical devices that currently exist for this purpose include the Tearscope (Keeler Ophthalmic Instruments, USA) and OCULUS Keratograph (Oculus, Inc., Wetzlar, Germany). While the Tearscope relies upon the subjective determination of BUT, the OCULUS Keratograph provides an automated quantification. Evaluation of the OCULUS Keratograph has identified the need for a calibration offset for its noninvasive BUT measurements to be comparable with other devices[40]; whether a newer release of this instrumentation (OCULUS Keratograph 5M) has led to such improvements remains to be determined. Although the Tearscope has been reported to be more reliable than other techniques, such as slit lamp observation or viewing videokeratoscope mires, measurement variability is significant[41] and this effect is exaggerated with multiple examiners.[42] This study supports these findings, reporting significantly poorer repeatability for CNI-BUT measures than automated

TABLE 4. Summary Table of ROC Curve Parameters

Parameter	AUC (95% CI)	Criterion, s	Sensitivity, % (95% CI)	Specificity, % (95% CI)	Youden Index	LR
Shortest TFSQ-BUT	**0.92*** (0.83–1.00)**	**12.1**	**81.5 (61.9, 93.7)**	**94.4 (72.7, 99.9)**	**0.76**	**14.7**
Average TFSQ-BUT	0.92*** (0.87–1.00)	16.1	76.9 (56.4, 91.0)	94.4 (72.7, 99.9)	0.71	13.9
First TFSQ-BUT	0.81*** (0.69–0.94)	18.8	69.2 (48.2, 85.7)	77.8 (52.4, 93.6)	0.47	3.1
Shortest NaFl-BUT	0.65 (0.46–0.83)	6.3	74.1 (53.7, 88.9)	66.7 (41.0, 86.7)	0.41	2.2
Average NaFl-BUT	0.70* (0.53–0.87)	7.5	63.0 (42.4, 80.6)	72.2 (46.5, 90.3)	0.35	2.3
First NaFl-BUT	**0.72* (0.55–0.89)**	**8.0**	**77.8 (57.7, 91.4)**	**72.2 (46.5, 90.3)**	**0.55**	**2.8**

Data are shown for the eye with highest tear osmolarity. Asterisks show statistically significant differences between groups. Items in bold highlights the two best diagnostic criteria (based upon Youden index) for the noninvasive (TFSQ) and sodium fluorescein-based (NaFl) tear breakup methods. AUC, area under the curve; CI, confidence interval; LR, likelihood ratio; ROC, receiver operator characteristic; s, seconds.

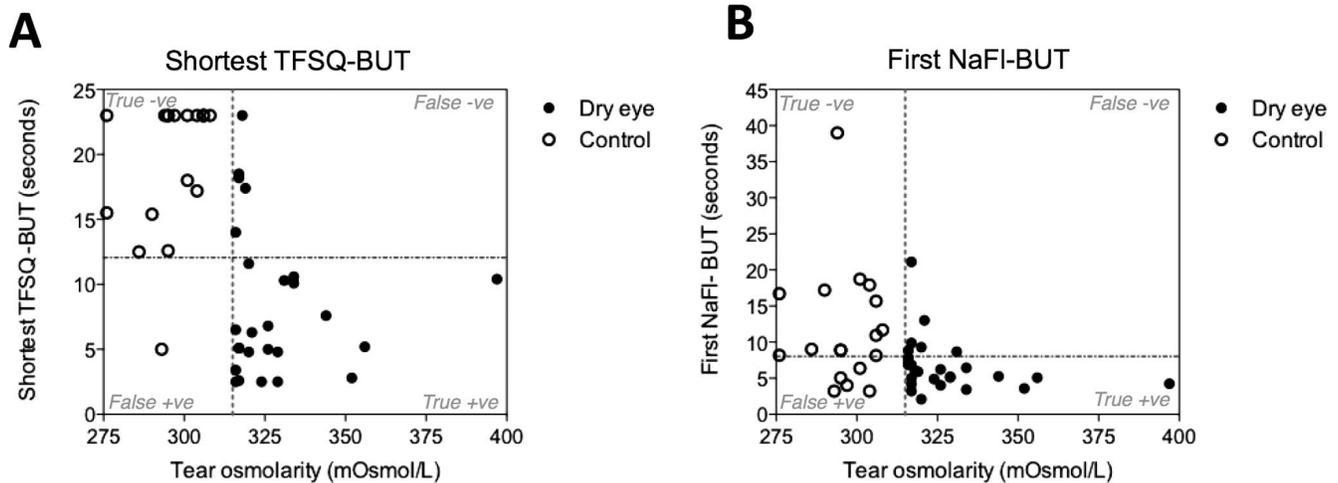* $P < 0.05$.
** $P < 0.01$.
*** $P < 0.001$.

FIGURE 3. Plots showing the relationship between tear osmolarity and (**A**) shortest TFSQ-BUT and (**B**) NaFl-BUT, respectively, for all participants ($n = 45$). The *vertical dotted line* represents the tear osmolarity criterion for diagnosis of dry eye disease (316 mOsmol/L). The *horizontal dotted lines* indicate the optimal diagnostic criterion for each parameter: (**A**) 12.1 seconds and (**B**) 8.0 seconds.

TFSQ-BUTs. The degree of repeatability of the automated TFSQ-BUT (gCoV = 9.4%) was found to be similar to other contemporary ophthalmologic devices, including those used to quantify anterior segment biometry,[43] ocular bulbar redness,[44] and iridocorneal angle.[45]

Given that tear instability and hyperosmolarity coexist in DED,[1] there is strong scientific rationale for an interrelationship between their clinical measures. The present study appears to be the first to have directly considered this potential association. Shortest TFSQ-BUT was found to have the best overall discriminative capacity for tear hyperosmolarity. Interestingly, this parameter had superior diagnostic power to the first TFSQ-BUT measure. This finding may reflect the known variability in clinical expression of DED, being a hallmark of the condition that complicates diagnosis.[46] Sampling three noninvasive measures of tear stability, rather than a single measure, may increase the likelihood of capturing at least one "abnormal" result in an eye with DED. A healthy tear film would be predicted to demonstrate a consistently stable tear profile, and therefore less variation in consecutive TFSQ-BUT measures.

A consideration when interpreting the findings from this study is that the control and DED populations were intentionally defined to achieve an unambiguous classification of tear film quality, as either "normal" (control) or "abnormal" (DED). The average tear osmolarity value of the DED population (325 mOsmol/L) is consistent with moderate to severe DED, rather than earlier stages of the disease.[47] In this respect, the findings can be considered to reflect the diagnostic capacity of the TFSQ-BUT parameter to distinguish between asymptomatic healthy controls from those with potentially more severe expressions of DED, as defined by tear osmolarity. Interestingly, despite the reported level of tear hyperosmolarity in the DED group, the overall extent of ocular surface staining in this group was low (mean total ocular surface staining score of 1.2 out of 15.0) and Schirmer test scores were consistent with borderline, rather than severe, aqueous deficiency (mean of 9.0 mm in 5 minutes). Indeed, these findings are in agreement with previous work that has demonstrated a lack of consistent relationships between common signs and symptoms of DED.[21,48] Further investigation, in particular a longitudinal clinical study, would be of value to assess the diagnostic value of the TFSQ-BUT in marginal cases of DED.

In conclusion, this study demonstrates the clinical utility of a novel noninvasive tear breakup time measure, the shortest TFSQ-BUT, as a surrogate marker for tear hyperosmolarity in people with moderate to severe DED.
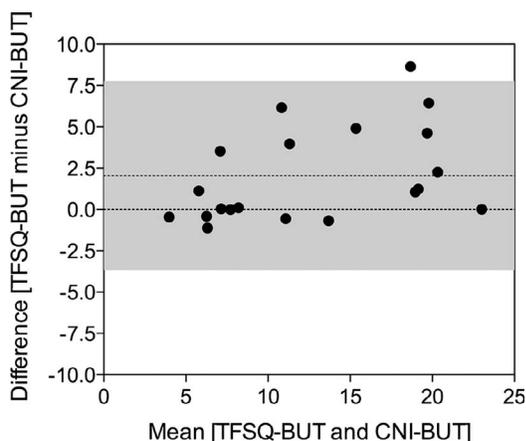

FIGURE 4. Bland-Altman plot comparing noninvasive tear break up time (seconds) performed by the automated TFSQ-BUT and subjectively assessed by clinicians (CNI-BUT). The *dotted line* shows the bias (2.0 seconds) for the comparison between the two methods. The *grey shaded area* highlights the LoA; 95% CLs for the LoA, factoring in sample size, were: upper bound (6.4 to 10.5 seconds), lower bound (−6.4 to −2.3 seconds).

Investigative Ophthalmology & Visual Science

References

1. The definition and classification of dry eye disease. Report of the definition and classification subcommittee of the International Dry Eye WorkShop. *Ocul Surf.* 2007;5:75–92.

2. Stern ME, Beuerman RW, Fox RI, Gao J, Mircheff AK, Pflugfelder SC. The pathology of dry eye: the interaction between the ocular surface and lacrimal glands. *Cornea.* 1998;17:584–589.

3. Methodologies to diagnose and monitor dry eye disease. Report of the diagnostic methodology subcommittee of the International Dry Eye WorkShop. *Ocul Surf.* 2007;5: 108–152.

4. Bron AJ, Tomlinson A, Foulks GN, et al. Rethinking dry eye disease: a perspective on clinical implications. *Ocul Surf.* 2014;12:S1–S31.

5. Bron AJ, Tiffany JM, Yokoi N, Gouveia SM. Using osmolarity to diagnose dry eye: a compartmental hypothesis and review of our assumptions. *Adv Exp Med Biol.* 2002;506:1087–1095.

6. Foulks GN. The correlation between the tear film lipid layer and dry eye disease. *Surv Ophthalmol.* 2007;52:369–374.

7. Tomlinson A, Khanal S, Ramaesh K, Diaper C, McFadyen A. Tear film osmolarity: determination of a referent for dry eye diagnosis. *Invest Ophthalmol Vis Sci.* 2006;47:4309–4315.

8. Tomlinson A, McCann LC, Pearce EI. Comparison of human tear film osmolarity measured by electrical impedance and freezing point depression techniques. *Cornea.* 2010;29: 1036–1041.

9. García-Resúa C, Pena-Verdeal H, Miñones M, Gilino J, Giraldez MJ, Yebra-Pimentel E. Accuracy of two osmometers on standard samples: electrical impedance technique and freezing point depression technique. *Proc SPIE Int Soc Opt Eng.* 2013;8785ET.

10. Pena-Verdeal H, García-Resúa C, Miñones M, Giraldez MJ, Yebra-Pimentel E. Accuracy of a freezing point depression technique osmometer. *Optom Vis Sci.* 2015;92:e273–e283.

11. Downie LE, Vingrys AJ. Accuracy of laboratory assays in ophthalmic practice [published online ahead of print October 15, 2015[. *JAMA Ophthalmol.* doi:10.1001/jamaophthalmol.2015.3856.

12. Najafi L, Malek M, Valojerdi AE, Khamseh ME, Aghaei H. Dry eye disease in type 2 diabetes mellitus; comparison of the tear osmolarity test with other common diagnostic tests: a diagnostic accuracy study using STARD standard. *J Diabetes Metab Disord.* 2015;14:39.

13. Schargus M, Ivanova S, Kakkassery V, Dick HB, Joachim S. Correlation of tear film osmolarity and 2 different MMP-9 tests with common dry eye tests in a cohort of non-dry eye patients. *Cornea.* 2015;34:739–744.

14. Fuerst N, Langelier N, Massaro-Giordano M, et al. Tear osmolarity and dry eye symptoms in diabetics. *Clin Ophthalmol.* 2014;8:507–515.

15. Alves M, Reinach PS, Paula JS, et al. Comparison of diagnostic tests in distinct well-defined conditions related to dry eye disease. *PLoS One.* 2014;9:e97921.

16. Bunya VY, Fuerst NM, Pistilli M, et al. Variability of tear osmolarity in patients with dry eye. *JAMA Ophthalmol.* 2015; 133:662–667.

17. Downie LE, Keller PR. A pragmatic approach to dry eye diagnosis: evidence into practice [published online ahead of print September 29, 2015]. *Optom Vis Sci.* doi:10.1097/OPX.0000000000000653.

18. Sweeney DF, Millar TJ, Raju SR. Tear film stability: a review. *Exp Eye Res.* 2013;117:28–38.

19. Bron AJ, Evans VE, Smith JA. Grading of corneal and conjunctival staining in the context of other dry eye tests. *Cornea.* 2003;22:640–650.

20. 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.

21. Lemp MA, Bron AJ, Baudouin C, et al. Tear osmolarity in the diagnosis and management of dry eye disease. *Am J Ophthalmol.* 2011;151:792–798.e791.

22. Alonso-Caneiro D, Turuwhenua J, Iskander DR, Collins MJ. Diagnosing dry eye with dynamic-area high-speed video-keratoscopy. *J Biomed Opt.* 2011;16:076012.

23. Korb DR, Greiner JV, Herman J. Comparison of fluorescein break-up time measurement reproducibility using standard fluorescein strips versus the Dry Eye Test (DET) method. *Cornea.* 2001;20:811–815.

24. Hopkins WG. Measures of reliability in sports medicine and science. *Sports Med.* 2000;30:1–15.

25. Vaz S, Falkmer T, Passmore AE, Parsons R, Andreou P. The case for using the repeatability coefficient when calculating test-retest reliability. *PLoS One.* 2013;8:e73990.

26. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet.* 1986;1:307–310.

27. Carkeet A. Exact parametric confidence intervals for bland-altman limits of agreement. *Optom Vis Sci.* 2015;92:e71–e80.

28. Youden WJ. Index for rating diagnostic tests. *Cancer.* 1950;3: 32–35.

29. Khanal S, Millar TJ. Barriers to clinical uptake of tear osmolarity measurements. *Br J Ophthalmol.* 2012;96:341–344.

30. Cardona G, Seres C, Quevedo L, Auge M. Knowledge and use of tear film evaluation tests by spanish practitioners. *Optom Vis Sci.* 2011;88:1106–1111.

31. Downie LE, Keller PR, Vingrys AJ. An evidence-based analysis of Australian optometrists' dry eye practices. *Optom Vis Sci.* 2013;90:1385–1395.

32. Turner AW, Layton CJ, Bron AJ. Survey of eye practitioners' attitudes towards diagnostic tests and therapies for dry eye disease. *Clin Exp Ophthalmol.* 2005;33:351–355.

33. Graham JE, McGilligan VE, Berrar D, et al. Attitudes towards diagnostic tests and therapies for dry eye disease. *Ophthalmic Res.* 2010;43:11–17.

34. Smith J, Nichols KK, Baldwin EK. Current patterns in the use of diagnostic tests in dry eye evaluation. *Cornea.* 2008;27: 656–662.

35. Mengher LS, Bron AJ, Tonge SR, Gilbert DJ. Effect of fluorescein instillation on the pre-corneal tear film stability. *Curr Eye Res.* 1985;4:9–12.

36. Pult H, Riede-Pult BH. A new modified fluorescein strip: Its repeatability and usefulness in tear film break-up time analysis. *Contact Lens Anterior Eye.* 2012;35:35–38.

37. Abelson MB, Ousler GW III, Nally LA, Welch D, Krenzer K. Alternative reference values for tear film break up time in normal and dry eye populations. *Adv Exp Med Biol.* 2002; 506:1121–1125.

38. Vitali C, Moutsopoulos HM, Bombardieri S. The European Community Study Group on diagnostic criteria for Sjogren's syndrome. Sensitivity and specificity of tests for ocular and oral involvement in Sjogren's syndrome. *Ann Rheum Dis.* 1994;53:637–647.

39. Papas E. Tear break-up time: clinical procedures and their effects. *Ophthalmic Physiol Opt.* 1999;19:274–275.

40. Best N, Drury L, Wolffsohn JS. Clinical evaluation of the OCULUS Keratograph. *Contact Lens Anterior Eye*. 2012;35: 171–174.

41. Elliott M, Fandrich H, Simpson T, Fonn D. Analysis of the repeatability of tear break-up time measurement techniques on asymptomatic subjects before, during and after contact lens wear. *Contact Lens Anterior Eye*. 1998;21:98–103.

42. Nichols JJ, Nichols KK, Puent B, Saracino M, Mitchell GL. Evaluation of tear film interference patterns and measures of tear break-up time. *Optom Vis Sci*. 2002;79:363–369.

43. Hernandez-Camarena JC, Chirinos-Saldana P, Navas A, et al. Repeatability, reproducibility, and agreement between three different Scheimpflug systems in measuring corneal and anterior segment biometry. *J Refract Surg*. 2014;30:616–621.

44. Downie LE, Keller PR, Vingrys AJ. Assessing ocular bulbar redness: a comparison of methods. *Ophthalmic Physiol Opt*. In press. doi:10.1111/opo.12245.

45. Romkens HC, Beckers HJ, Frusch M, Berendschot TT, de Brabander J, Webers CA. Reproducibility of anterior chamber angle analyses with the swept-source optical coherence tomography in young, healthy Caucasians. *Invest Ophthalmol Vis Sci*. 2014;55:3999–4004.

46. Foulks GN, Pflugfelder SC. New testing options for diagnosing and grading dry eye disease. *Am J Ophthalmol*. 2014;157:1122–1129.

47. Sullivan BD, Whitmer D, Nichols KK, et al. An objective approach to dry eye disease severity. *Invest Ophthalmol Vis Sci*. 2010;51:6125–6130.

48. 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.