



Published in final edited form as:

Cornea. 2020 January ; 39(1): 92–98. doi:10.1097/ICO.0000000000002148.

## Efficacy of the Fluorescein Tear Breakup Time (TBUT) Test in Dry Eye

Jerry R. Paugh, OD, PhD<sup>1</sup>, June Tse, OD, MS<sup>2</sup>, Tiffany Nguyen, OD, MS<sup>3</sup>, Alan Sasai, OD<sup>1</sup>, Elaine Chen, OD<sup>1</sup>, Melinda Thomas De Jesus, OD, MS<sup>4</sup>, Justin Kwan, OD<sup>5</sup>, Andrew Loc Nguyen, PhD<sup>6</sup>, Marjan Farid, MD<sup>7</sup>, Sumit Garg, MD<sup>7</sup>, James V. Jester, PhD<sup>7,8</sup>

<sup>1</sup>Southern California College of Optometry at Marshall B. Ketchum University, Fullerton, California, USA

<sup>2</sup>Private Practice, Hillsboro, Oregon, USA

<sup>3</sup>Private Practice, Dover, Delaware, USA

<sup>4</sup>Private Practice, West Covina, California USA

<sup>5</sup>Private Practice, Niles, Illinois, USA

<sup>6</sup>Department of Mathematics, California State University Fullerton, California, USA

<sup>7</sup>Gavin Herbert Eye Institute, University of California, Irvine, California, USA

<sup>8</sup>Biomedical Engineering, University of California, Irvine, California, USA

### Abstract

**Purpose**—To examine the effect of volume and method on fluorescein tear breakup time (TBUT) values, and to evaluate test efficacy in an independent sample free of selection bias.

**Methods**—Subjects were assessed using a battery of dry eye tests. *Efficacy study*: subjects were randomized to the Dry Eye Test (DET), Standard Strip (SS) and liquid NaFl, on separate days. A masked examiner measured TBUTs from video recordings. *Verification study*: Subjects were investigated for efficacy using volumes of 5.0 and 2.0 µl microliters of NaFl for TBUT.

**Results**—*Efficacy study*: 46 subjects completed the study. Log-transformed TBUTs were significantly different, normals vs. dries, for all 3 methods (all p values < 0.001). AUCs, cut-points, sensitivity and specificity were: 1) DET: 0.873, 4.4 secs, 0.97 and 0.67, respectively, 2) 2.0 microliters: 0.901, 3.22 secs, 0.90 and 0.87, respectively, and 3) SS: 0.912, 3.42 secs, 0.97 and 0.80, respectively. *Verification study*: Data splitting analysis for the 2.0 µl data (n = 174 dries and 97 normals) generated an AUC of 0.917, a cut-point of 6.05 seconds for sensitivity of 0.87 and specificity of 0.81. The 5.0 µl sample yielded an AUC of 0.940, with sensitivity and specificity of 0.92 and 0.83, respectively at a cut-point of 5.5 seconds.

**Corresponding Author:** Jerry R. Paugh, O.D., Ph.D., Southern California College of Optometry at Marshall B. Ketchum University, 2575 Yorba Linda Blvd., Fullerton, CA 92831, Telephone: 714-449-7487, jpaugh@ketchum.edu, Facsimile: 714-992-7809.

The authors have no conflicts of interest for the subject matter of this paper.

The authors have no proprietary or commercial interest in any materials discussed in this article.

**Conclusions**—Little difference in TBUT was found using the three clinical methods with video recording. Analysis using liquid NaFl suggests that the TBUT test has excellent diagnostic accuracy and that a cut-point of 5.3 to 6.0 seconds is the optimum to differentiate normals from dry eye persons.

### Keywords

tear breakup time; tear stability; breakup time; fluorescein; dry eye; test efficacy

---

### Introduction

The measurement of tear film stability is an important assessment for both diagnosis and treatment monitoring in dry eye. Loss of normal tear stability may be a downstream consequence of alterations in the ocular surface and/or the composition of the tear layers,<sup>1</sup> but is a hallmark of the disease state.

Wolffsohn et al. in the DEWS II Diagnostic report<sup>2</sup> recommend non-invasive breakup time assessment of tear stability. However, non-invasive measurements of tear stability are still not widely used, perhaps due to the cost of the instruments, as evidenced by several surveys of optometrists and ophthalmologists.<sup>3-5</sup> In contradistinction, fluorescein breakup time is still widely used by dry eye practitioners,<sup>3-5</sup> including the researchers of the Sjogren's International Collaborative Clinical Alliance (SICCA)<sup>6</sup> and is recommended as a dry eye syndrome diagnostic test by the Academy of Ophthalmology Preferred Practice Pattern.<sup>7</sup> Taken together, this evidence demonstrates that tear film breakup time (TBUT) using sodium fluorescein remains a key evaluation in dry eye diagnosis.

Despite the commonality of its use, there have been relatively few formal investigations of the test efficacy and associated cut-points for tear stability to differentiate dry from normal individuals. Moreover, there are several important issues that relate to test efficacy determination. These include the method of classification for normals and dry eye groups, inclusion of both major dry eye sub-types (i.e., aqueous tear deficient (ATD) and meibomian gland dysfunction, (MGD), selection and spectrum bias,<sup>8</sup> and the method of fluorescein dye instillation.

Lemp et al.,<sup>9</sup> using 5.0 µl of 2.0% liquid dye suggested a cut-point of < 10 seconds to differentiate normals from drys. The reported sensitivity and specificity was 84.4% and 45.3%, respectively.<sup>9</sup> Downie,<sup>10</sup> using the Dry Eye Test (DET) of Korb<sup>11</sup> found sensitivity and specificity (using mean fluorescein breakup time values) of 63% and 72%, respectively, with a cut-point of 7.5 seconds. Contrasted against these cut-points are numerous mean values in normals less than 10 seconds.<sup>11-16</sup>

The concentration and volume of fluorescein dye instilled for the breakup time test has been a subject of discussion, whether by wetted strip or micropipette. Finnemore et al.<sup>17</sup> pointed out that significant variability can occur in tear film fluorescence with clinical methods and demonstrated hypofluorescence (i.e., delayed maximum fluorescent intensity due to dye self-absorption) with varying concentrations of instilled dye (1.0 µl of 2.0 and 5.0% sodium fluorescein<sup>17</sup>).

Korb et al.<sup>11</sup> compared TBUTs in presumably normal subjects using a standard fluorescein strip to a modified (narrower) dye impregnated strip, the Dry Eye Test (DET) both wetted with one drop of non-preserved saline. Despite the variable dye volume instilled, the mean values of TBUT for the standard and narrow strip were approximately 7.4 to 7.5 seconds, and not significantly different. However, Korb and co-workers<sup>11</sup> found less variability of TBUT and less ocular sensation for the DET compared to a standard fluorescein strip. Johnson and Murphy,<sup>18</sup> compared a standard strip to varying small volumes of fluorescein dye (1.0, 2.7 and 7.4  $\mu\text{l}$ ) in young, presumably normal subjects. They found no difference between the standard strip and 1.0  $\mu\text{l}$  volumes, but some lengthening of TBUT when the volume increased to 2.7  $\mu\text{l}$  from 1.0  $\mu\text{l}$ .

Taken together, these data suggest that large differences in the mean values of TBUT are not found when using a strip vs. controlled volumes of liquid dye. However, no single study has evaluated the several methods in the same group of subjects, that includes the major dry eye sub-types.

It was the purpose of this investigation to first determine test efficacy of the TBUT test using sodium fluorescein without selection bias using a randomized and masked study design (efficacy study). These preliminary cut-points were then verified using a second, independent clinical sample as recommended by the initial Dry Eye Workshop Diagnostic Subcommittee report<sup>8</sup> (verification study).

## Materials and Methods

### Subjects

Enrolled subjects included normals (i.e., non-dry eye subjects), those with aqueous tear deficiency (ATD) and meibomian gland dysfunction (MGD). This was primarily a clinic-based sample. The study adhered to the tenets of the Declaration of Helsinki and prior to participation subjects provided written, informed consent. Human subjects approval was granted by the Institutional Review Boards of the Southern California College of Optometry at Marshall B. Ketchum University and the University of California at Irvine.

Major inclusion criteria were: adult subjects, > age 18 years, with either no dry eye (normals) or mild-to-severe dry eye. Criteria for the dry eye diagnosis in both studies included ocular surface staining, and either evidence of aqueous tear deficiency or meibomian gland dysfunction. For the efficacy study, two of three dry eye questionnaires had to demonstrate symptoms of ocular dryness in addition to the objective dry eye tests (see below). Tear stability was not used as a qualifying criterion in either study, avoiding selection bias.<sup>8</sup>

Major exclusion criteria were: 1. Extraocular or intraocular surgery within the past twelve months. LASIK and other kerato-refractive procedure patients could qualify if the most recent surgery or enhancement was 12 or more months prior. 2. Evidence of active ocular allergy or infection. 4. Greater than mild eyelid deformity (e.g., ectropion, ptosis). 5. Use of concomitant ocular medications except dry eye treatments, as long as there was no instillation on the examination day. 6. Use of systemic medications unless they were on a

stable dosing regimen for more than 30 days prior to examination. 7. Uncontrolled diabetes. 8. Individuals who were pregnant or lactating. 9. Any type of contact lens wear (soft, hybrid or rigid gas permeable contact lenses).

### Classification of Dry Eye

Classification of the subjects as normal or dry was comparable for both the efficacy and verification studies. TBUT data were masked from the individual who assigned subjects to normal, aqueous tear deficient or meibomian gland dysfunction sub-types. Symptoms were used to categorize the subjects in addition to the objective clinical tests, detailed below. Classification was based on the observed values from the worst eye of the two, which was the eye with the greatest total staining. We did not have tear osmolarity available to us to globally classify the subjects, nor meibography for a majority of the study subjects. Non-invasive breakup time could not be used as a measure of tear stability as any measure of tear stability measurement would have caused selection bias as tear stability was the parameter under examination.<sup>8</sup>

**Questionnaires**—For the efficacy study, symptom surveys included the Women’s Health Questionnaire (WHQ), modified Schein, and the Ocular Surface Disease Index (OSDI); two of three surveys had to be dry for subjects to be considered dry. For the verification study, either the Schein or the OSDI survey was used.

By the WHQ, subjects were considered dry if both dryness and irritation either often or constantly were reported, or there was a report of dry eye diagnosis by a clinician.<sup>19</sup> For the modified Schein instrument, subjects were considered dry if their score was greater than 7.0 (integer scale).<sup>20</sup> The cut-point for OSDI score was > 12 for a subject to be considered dry.<sup>21-23</sup>

**Corneal and conjunctival Staining**—Sullivan et al.<sup>24</sup> in a multi-center study present extensive data in normals and dry subjects for staining using the NEI system.<sup>25</sup> Using independent component analysis, they reported small (0.4 and 1.1 average corneal and conjunctival staining values, respectively) for normals (n = 75 subjects) and slightly greater (1.7 and 2.6 mean values for cornea and conjunctiva, respectively) for mild/moderate subjects (n = 175). We used a combined corneal and conjunctival score limit of 6 (0 – 33 total scale) to define normals, and > 6.0 to define definite dry eye.

**Meibomian Gland Function and Structure**—We used gentle meibomian gland expression (vs. digital pressure) for the lower eyelids using a cotton-tipped applicator. Secretions were scored according to the Bron 0 – 3 scale.<sup>26, 27</sup> Subjects were considered to have MGD if their secretion score for the worst eye was 1.0.<sup>27</sup> Meibomian gland atrophy or dropout was determined by meiboscopy along the entire lower eyelids of both eyes.<sup>28</sup> We counted missing whole or half glands and estimated gland loss on a 0 – 4 scale were 0 = no loss, grade 1 = 25% loss, grade 2 = 50% loss, grade 3 = 75% loss, and grade 4 = 100%. Subjects were considered to have atrophy consistent with MGD if their gland loss was 1.0.<sup>27</sup>

**Overall Classification**—Subjects were considered to have MGD if they had positive symptoms, NEI staining > 6.0, secretion grade or dropout 1.0, and Schirmer without anesthetic > 5 mm of wetting in 5 minutes. Subjects were classified as having aqueous tear deficiency (ATD) if they had positive symptoms, NEI staining > 6.0, and Schirmer without anesthetic 5 mm of wetting in 5 minutes. These subjects could have signs of MGD, as has been reported,<sup>28</sup> but the critical criterion beyond symptoms and staining was the Schirmer I result.

### Study Procedures

Subjects for the efficacy and verification studies were examined similarly. Tests included: tear meniscus height (reticule eyepiece), white light interferometry (Yokoi et al. scale<sup>29</sup>), TBUT (verification study only, but not used for assignment as normal or dry to prevent selection bias), corneal staining (sodium fluorescein; of either 5.0 µl of 2.0% or 2.0 µl of 1.0% wt./vol., micropipette) and conjunctival staining (lissamine green; strip, wetted, soak for 5 seconds,<sup>2</sup> excess shaken off), meibomian gland expression, lower eyelid meiboscopy for gland dropout,<sup>28</sup> (both as above) and the Schirmer I test without anesthesia (strip moved at 2 minutes if no wetting was observed<sup>30</sup>). A yellow Tiffen filter was used for both staining assessment and the TBUT studies to enhance observation.

### Fluorescein Breakup Time Determination, Efficacy Study

Subjects were qualified for the efficacy study as above, without the TBUT data. The subjects were then scheduled for three follow-up visits to determine TBUT, within the hours of 12:00 and 6:00 PM to obviate changes due to diurnal variation.<sup>31</sup>

Only one eye of each subject was used for the TBUT determinations. Subjects were not allowed to use artificial tears on the day of the TBUT determination.

Two examiners assessed the TBUT using the three dye instillation methods (see below) on separate days. To enforce masking of the measurements, one examiner instilled each portion of dye without the second examiner observing the instillation. TBUTs were immediately recorded using a slit lamp biomicroscope and video recording system. The three dyes were instilled on the superior bulbar conjunctiva. The subject was asked to blink naturally<sup>2</sup> three times following instillation to distribute the dye evenly. A yellow filter was used in all cases to enhance observation of the first change (occurrence of a dark spot or definite break) in the dye-impregnated tear film. Following each recording the subject closed their eyes for a 30 second “rest” period to re-establish tear stability homeostasis since preliminary data from our clinic (unpublished) suggested that TBUT becomes more rapid if breakup times are collected too close together.

A Nikon D5100 SLR, 16.2 megapixel camera, attached to a Nikon FS-2 photo biomicroscope, with 32 MB flash memory was used to record TBUTs at each visit for later TBUT determination. The biomicroscope lamp intensity was set at maximum, with a broad beam, and “scanning” was prohibited.<sup>12</sup> The video recordings were measured for TBUT using a stopwatch using 3-6 separate recordings. The measuring examiner was masked to the

method of dye instillation and to the status of the subject (i.e., whether normal or dry). The order of dye instillation method was randomized. The three methods were as follows.

- Dry Eye Test (DET) strips (0.12 mg NaFl per strip; one drop of non-preserved saline to wet; excess fluid NOT shaken off; AMCON laboratories, St. Louis, MO)
- Standard strips (1.0 mg NaFl per strip; one drop of non-preserved saline to wet; strip shaken to remove excess dye; BioGlo™, HUB Pharmaceuticals, Rancho Cucamonga, CA)
- 2.0µl of 1.0% small volume sodium fluorescein with pipette (Greenpark Compounding Pharmacy, Houston, TX)

### **Fluorescein Breakup Time Determination, Verification Study**

Retrospectively, clinical data were combined from several studies wherein consistent methodology (i.e., eligibility criteria and testing, as above) and sub-type assignment criteria were used. This allowed examination of sodium fluorescein volume effects on TBUT and increased sample size for comprehensive test efficacy evaluation. Most data were collected by a single investigator (JRP) and the remainder from similarly trained investigators.

### **Data Analysis**

For both studies, the principal outcome parameter was the receiver operating characteristic curve and associated area under the curve (AUC). Normality was checked and either non-parametric tests or log transformations undertaken (e.g., for TBUTs) prior to statistical testing. The general approach for specific tear film parameters was to undertake a one-way ANOVA to determine overall differences, followed by Tukey pair-wise comparisons for sub-group analysis. Tests were considered significant if the p value was less than 0.05.

For the verification study, subject data from an independent sample from the same clinic were evaluated. The validation approach for test efficacy was that of data splitting<sup>32</sup> since the data set was large and could be randomly split into two samples. Approximately 60% of the initial sample was used to establish the ROC curve and cut-points (model building). The remaining data were used for model checking (also called cross-validation). The predictive accuracy of the model-building portions of the samples were checked with the remaining portions of each larger data set.

## **Results**

### **Efficacy Study**

Forty-six subjects completed the efficacy study; 15 normals, 11 ATDs and 20 MGDs. The median age for the normal group was 54.0 years and 61.5 years for the combined dry eye group ( $p = 0.119$ , Mann-Whitney U test). There were more females in the dry eye groups combined (22 of 31) compared to the normal group (6 of 9;  $p = 0.02$ , Fisher exact test).

The descriptive and inferential statistics for the three TBUT methods are summarized in Table 1. In all cases the normals were significantly different (i.e., longer TBUTs) compared

to the two dry eye sub-type groups with all three TBUT methods (One-way ANOVA,  $p < 0.001$ ). By sub-type, ATD vs. normals and MGDs vs. normals were all significantly different (all  $p$  values  $< 0.001$ ) for all three methods (Tukey pairwise comparisons).

### Receiver Operating Characteristic (ROC) Analysis and Cut-Points

The combined ATD and MGD dry eye sub-type values for TBUT ( $n = 31$  subjects) were compared against the normals ( $n = 15$  subjects) for test efficacy. ROC curve analysis was used to determine the test sensitivity and test specificity of each TBUT method according to the table of results for various values. In general, the cut-point with the greatest sensitivity and specificity was considered first. If the sensitivity was compromised by attempting to maximize specificity, the cut-point with the greater sensitivity was selected. For some patients there is poor correlation between dry eye signs and symptoms, and maximizing detection of dry eye is important clinically.

The data in Table 1 demonstrate efficacy (sensitivity and specificity) of near or greater than 70% for all three TBUT methods, a level suggested for an effective diagnostic test.<sup>33</sup> Moreover, the AUCs for all three methods are near or greater than 0.90, which suggests outstanding discrimination.<sup>34</sup>

### TBUT and the Effect of Method

The efficacy study sample, which used the video observation method, allowed comparison of TBUT data among three methods in normals ( $n = 15$ ) and dry eye subjects ( $n = 31$  total). For normals, there was no method effect ( $p = 0.728$ ) and statistical power was 0.90 or greater for log levels of difference of 0.05, 0.08 or 0.10), indicating adequate sample size to observe a difference should one exist. For the combined dry eye subjects there was a method effect ( $p = 0.004$ ), with the DET different compared to the standard strip ( $p = 0.010$ ; Tukey comparison) and liquid dye (2.0  $\mu$ l of 1.0%;  $p = 0.012$ ; Tukey comparison). Additional analysis by dry eye sub-type (ATDs ( $n = 11$ ) and MGDs ( $n = 20$ )) suggested that the dry eye differences for DET vs. the standard strip and liquid dye were found only for the MGD group; general linear model  $p = 0.002$ ; Tukey significant for liquid dye vs. DET ( $p = 0.002$ ) and standard strip vs. DET  $p = 0.037$ . In the MGD group, no difference was found for standard strip vs. liquid dye ( $p = 0.479$ ). We conclude that there are method differences in this sample, but only for the DET in the MGD dry eye sub-type and not in normals nor the ATD dry eye sub-type.

### Verification Study

For the independent sample analysis, dry and normal subjects were analyzed for the three methods. These data are summarized in Table 2.

In the normal group, no significant differences were found in log-transformed TBUT values by instillation method (general linear model; normals,  $p = 0.157$ ). However, there were differences among methods in the dry group ( $p < 0.001$ ) and Tukey pair-wise comparisons demonstrated differences for each method compared to the others ( $p = 0.002$ , 5.0 to 2.0  $\mu$ l;  $p < 0.001$ , video vs. 2.0  $\mu$ l, and  $p < 0.001$ , video vs. 5.0  $\mu$ l). Due to these differences, and the



markedly shorter TBUTs found using the video method, we further analyzed only the 2.0 and 5.0  $\mu\text{l}$  TBUT data for test efficacy.

### ROC Curves and Test Efficacy

ROC curves were developed for the combined 2.0 and 5.0  $\mu\text{l}$  sample since the TBUTs were clinically similar (Table 2) and for the 2.0  $\mu\text{l}$  and 5.0  $\mu\text{l}$  data separately. For the combined 2.0 and 5.0  $\mu\text{l}$  sample, a total of 182 dry eye and 91 normal subjects were available for model building; the ROC curve is shown in Figure 1. The cut-points, sensitivity, specificity and associated predictive accuracy (normals and drys) are shown in Table 3. The AUC was 0.92 (95% CI, 0.74 – 0.87; excellent discrimination<sup>34</sup>) for a cut-point of 5.3 seconds, Figure 1, which is at near-equal sensitivity and specificity (Table 3). At this cut-point, 83% of normals were correctly predicted ( $n = 30$  subjects) and 78% of dry eye subjects correctly predicted ( $n = 54$  subjects). Additional cut-points in Table 3 demonstrate the trade-offs for sensitivity and specificity. Test sensitivity and specificity values greater than 0.70 are considered efficacious for dry eye tests.<sup>33</sup>

An analysis similar to that for the combined 2.0 and 5.0  $\mu\text{l}$  sample was conducted solely on the 2.0  $\mu\text{l}$  sample. A total of 114 dry subjects and 62 normals subjects were used for model-building. Conversely, 60 dry subjects and 35 normals subjects were used for model-checking. The ROC curve for model-building is shown in Figure 1 for a cut-point of 6.1 seconds (AUC = 0.92). At this cut-point, 68% of normals were correctly predicted ( $n = 24$  subjects) and 83% of dry eye subjects correctly predicted ( $n = 50$  subjects).

The 5.0  $\mu\text{l}$  sample, 30 normals and 77 dry eye subjects, was not large enough to allow data splitting, so the entire sample was considered for test efficacy. The optimum cut-point was 5.5 seconds, with sensitivity of 0.92 and specificity of 0.83 and AUC of 0.94; Figure 1.

### Correlations to clinical tests

Using the combined 2.0 and 5.0  $\mu\text{l}$  data, statistically significant correlations ( $p < 0.001$ ), albeit of modest Pearson values were found for several parameters and TBUT, including total eye NEI staining score (Pearson's  $r = -0.36$ ), Schirmer score ( $r = 0.285$ ), meibomian gland secretion score ( $r = -0.209$ ) and gland dropout via meiboscopy ( $r = -0.137$ ). The TBUT and NEI staining data are shown in Figure 2.

## Discussion

### The Effect of instillation method and dye volume

To date there have been very few investigations of TBUT using fluorescein without selection bias, with well-defined subject classification and using ROC curve analysis. We have attempted to establish test efficacy and cut-points for the breakup time test without bias when using routine clinical methods. One of the most significant questions in measurement of tear stability is whether the amount of sodium fluorescein matters and whether methods with reasonably small volumes of dye affect the measurement.

The reports of Korb et al.,<sup>11</sup> using a standard strip and a reduced dye concentration strip, and Johnson and Murphy<sup>18</sup> using a standard strip and several small liquid volumes found few



differences in fluorescein breakup time in apparently normal subjects. In contradistinction, in the efficacy study we have examined the effect of dye instillation method in normals and the two major sub-types of dry eye using rigorous study design (i.e., lack of selection bias, investigator masking, randomization, and evidence based sub-type classification).

We found no difference among the methods of Standard Strip, DET and 2.0  $\mu$ l of 1.0% liquid dye in normals. In the dry eye group, shorter TBUT in the DET test results, compared to liquid dye and standard strip, were found only for the MGD subjects, but standard strip and liquid dye results were not different. Abdul-Fattah et al.,<sup>35</sup> using an *in vitro* model eye, found differing amounts of fluorescein dye instilled for the standard strip and DET-size strip (0.000073 and 0.000022 g, respectively). Finnemore and co-workers<sup>17</sup> found dye self-absorption clinically when 1.0  $\mu$ l of 10.0 % fluorescein (i.e., 0.001 gram fluorescein) was instilled into apparently normal eyes, but not when 1.0  $\mu$ l of 2.0 % dye was instilled (i.e., 0.00002 gram). In the present study the standard strip vs. DET did give statistically, although not clinically meaningful, differences in TBUT only for the MGD subjects (Table 1). It is possible that despite the likely self-absorption effect of a greater dye delivery, the normal procedural delay in TBUT determination of 1 – 2 minutes prior to measurement may have allowed adequate dilution time for a reliable TBUT determination. We note that the 2.0  $\mu$ l of 1.0 % sodium fluorescein concentration effectively delivers 0.00002 grams to the eye, similar to the DET, which dye amount was shown not to suffer self-absorption.<sup>17</sup>

Taken together, the reports of Korb et al.,<sup>11</sup> Johnson and Murphy,<sup>18</sup> and the data of our study suggest little difference in TBUTs among a standard strip and small volumes of dye instilled using a micropipette. Moreover, our data for the verification study suggest no statistical difference in TBUT values for the 5.0  $\mu$ l (2% concentration) compared to the 2.0  $\mu$ l volume (1.0% concentration) for the normal subjects, although statistical differences in the dry group were found. Examining the data of Table 2 for the non-video values suggests that meaningful clinical differences in TBUT for normals and dry eye subjects do not exist.

### **Diagnostic cut-point and recommend procedure for fluorescein breakup time determination**

ROC curve analysis in both studies demonstrates outstanding discrimination<sup>34</sup> of the fluorescein breakup time test to differentiate normal from dry individuals. All AUC values were 0.87 or greater (Table 1, Figure 1). Moreover, the sensitivities and specificities were all greater than 0.70, the level recommended for an effective dry eye test,<sup>33</sup> and the predicted accuracy from the model-building data was quite high (Table 3 and Results). Taken together, these analyses confirm the utility of using fluorescein breakup time as an important test in the diagnosis of dry eye, both MGD and aqueous tear deficiency.

The cut-points for differentiation of the two methods were slightly variable, but consistent over a narrow range of cut-points from 5.3 to 6.1 seconds (Table 3, Figure 1). We recommend the greater cut-point of 6.0 seconds to adequately classify the definite dry eye individuals, which is in agreement with Lemp et al.<sup>9</sup> who used osmolarity with a battery of tests and severity classification to characterize their sample.

The DEWS II report<sup>2</sup> recommended three natural blinks for the fluorescein breakup time test, with which we agree. Three natural blinks (vs. rapid or blepharospasm hard blinks) distribute the fluorescein through the tear film and allow normal tear thickness to be attained. Use of a yellow filter enhances the observation of the first change (i.e., a dark spot) in the film. Although we use micropipettes in our studies for consistency, standard strips provide TBUT values clinically equivalent to liquid and DET strip results (efficacy study). We found the DET difficult to impart consistent amounts of dye to the eye; often we had to re-instill dye using a fresh strip. The standard strips and the micropipette volumes were much easier for instillation and were very comfortable for the subjects.

In summary, using a video recording method for fluorescein breakup time does not demonstrate major differences in TBUT in dry eye or normal individuals for the three clinical methods examined. In a larger sample of dry and normal subjects, liquid volumes of either 2.0 or 5.0  $\mu$ l do not demonstrate clinically meaningful differences (Table 2) although statistical differences were observed in the dry eye sample only. We conclude that within these ranges of concentration and volume that dye concentration makes little difference in the TBUT values obtained using a consistent technique. Moreover, we believe the diagnostic cut-point to differentiate drys from normals should be in the range of 5.3 to 6.0 seconds for optimum test sensitivity and specificity. The fluorescein tear breakup time test is a valid and useful screening test for dry eye. When used in conjunction with additional diagnostic tests to determine etiology and severity appropriate treatment for dry eye disease can be effected.

## Acknowledgments

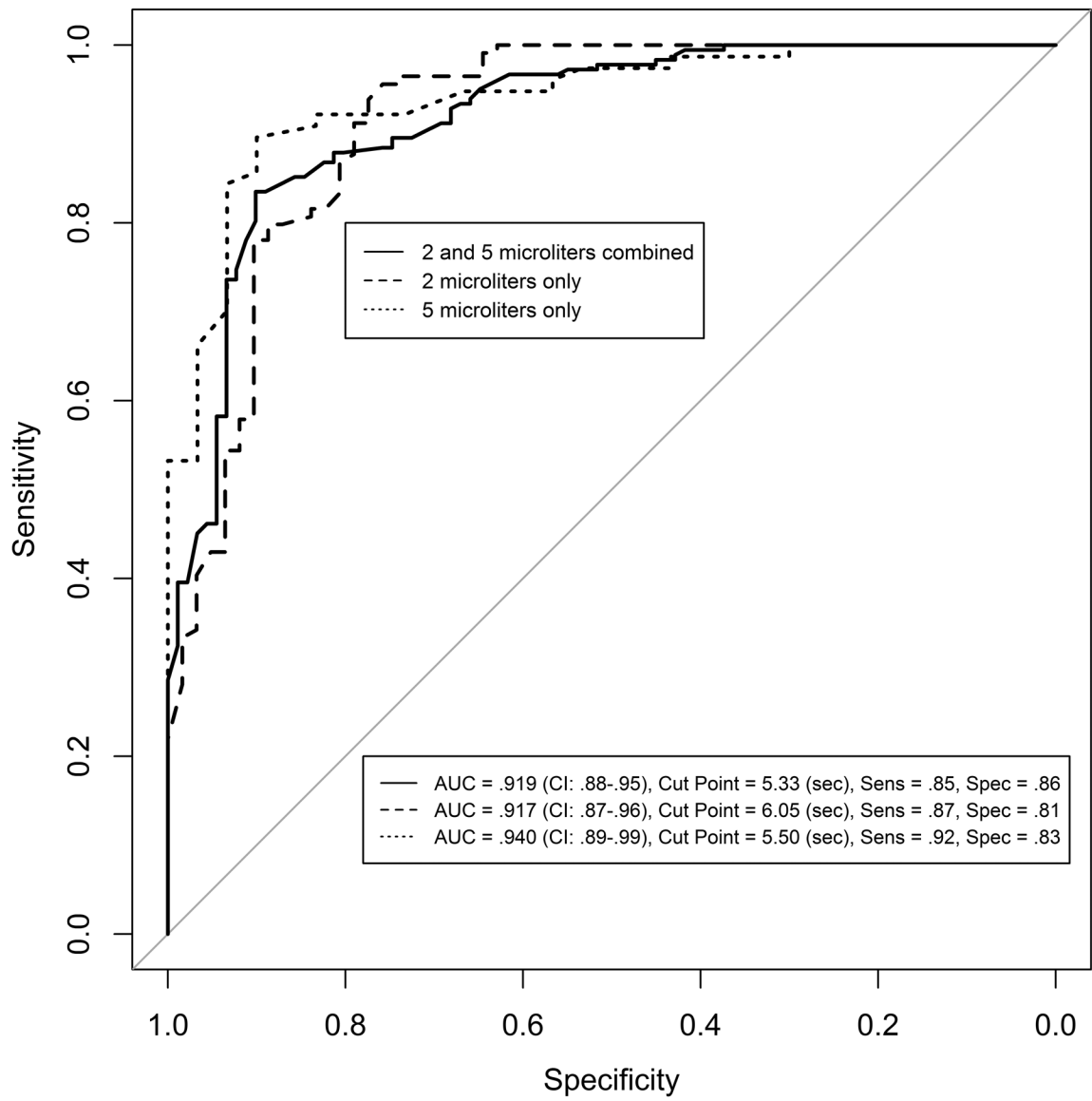
Supported in part by an Investigator Initiated Trial grant from Alcon Laboratories (JP), NEI EY021510 (JJ), supported by an unrestricted grant from Research to Prevent Blindness, Inc., (RPB-203478), and the Skirball Program in Molecular Ophthalmology.

## References

1. Bron AJ, de Paiva CS, Chauhan SK , et al. TFOS DEWS II pathophysiology report. *Ocul Surf* 2017; 15:438–510. [PubMed: 28736340]
2. Wolffsohn JS, Arita R, Chalmers R , et al. TFOS DEWS II Diagnostic Methodology report. *Ocul Surf* 2017; 15:539–574. [PubMed: 28736342]
3. Cardona G, Seres C, Quevedo L , et al. Knowledge and use of tear film evaluation tests by spanish practitioners. *Optom Vis Sci* 2011; 88:1106–1111. [PubMed: 21642888]
4. Xue AL, Downie LE, Ormonde SE , et al. A comparison of the self-reported dry eye practices of New Zealand optometrists and ophthalmologists. *Ophthalmic Physiol Opt* 2017; 37:191–201. [PubMed: 28211183]
5. Downie LE, Rumney N, Gad A , et al. Comparing self-reported optometric dry eye clinical practices in Australia and the United Kingdom: is there scope for practice improvement? *Ophthalmic Physiol Opt* 2016; 36:140–151. [PubMed: 26890703]
6. Whitcher JP, Shiboski CH, Shiboski SC , et al. A simplified quantitative method for assessing keratoconjunctivitis sicca from the Sjogren's Syndrome International Registry. *Am J Ophthalmol* 2010; 149:405–415. [PubMed: 20035924]
7. Akpek EK, Amescua G, Farid M , et al. Dry Eye Syndrome Preferred Practice Pattern(R). *Ophthalmology* 2018; 126:pp. 286–334.
8. Bron AJ, Abelson MB, Ousler G , et al. Methodologies to diagnose and monitor dry eye disease: report of the diagnostic methodology subcommittee of the International Dry Eye Workshop (2007). . *The Ocular Surface* 2007; 5:108–152. [PubMed: 17508118]

9. 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 [PubMed: 21310379]
10. Downie LE. Automated Tear Film Surface Quality Breakup Time as a Novel Clinical Marker for Tear Hyperosmolarity in Dry Eye Disease. *Invest Ophthalmol Vis Sci* 2015; 56:7260–7268. [PubMed: 26544794]
11. Korb D, 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. [PubMed: 11685057]
12. Cho P B B. Review of the tear break-up time and a closer look at the tear break-up time of Hong Kong Chinese. *Optom Vis Sci* 1993; 70:30–38. [PubMed: 8430006]
13. Shimazaki J, Sakata M, Tsubota K. Ocular surface changes and discomfort in patients with Meibomian gland dysfunction. *Arch Ophthalmol* 1995; 113:1266–1270. [PubMed: 7575257]
14. Abelson MB, Ousler GW, Nally LA , et al. Alternative reference values for tear film break up time in normal and dry eye populations. *Lacrimal Gland, Tear Film, and Dry Eye Syndromes 3: Basic Science and Clinical Relevance* 2002; 506, Part B:1121–1125.
15. Lam H, Bleiden L, De Paiva CS , et al. Tear cytokine profiles in dysfunctional tear syndrome. *Am J Ophthalmol* 2009; 147:198–205. [PubMed: 18992869]
16. Arita R, Itoh K, Maeda S , et al. Proposed diagnostic criteria for obstructive meibomian gland dysfunction. *Ophthalmol* 2009; 116:2058–2063.
17. Finnemore VM, Korb DR, Greiner JV , et al. Fluorescein dye concentration as a factor in tear film fluorescence in Sullivan D, Dartt DA, Meneray MA ed. *Lacrimal gland, tear film, and dry eye syndromes 2: basic science and clinical relevance*. Bermuda: Plenum Press, 1998; 875–878.
18. Johnson M, Murphy PJ. The effect of instilled fluorescein solution volume on the values and repeatability of TBUT measurements. *Cornea* 2005; 24:811–817. [PubMed: 16160497]
19. Gulati A, Sullivan R, Buring JE , et al. Validation and repeatability of a short questionnaire for dry eye syndrome. *Am J Ophthalmol* 2006; 142:125–131. [PubMed: 16815260]
20. Paugh JR, Nguyen AL. Validation of the modified Schein dry eye questionnaire. *Invest Ophthalmol Vis Sci* 2006; 47:ARVO Abstract, Program #276.
21. Schiffman RM, Christianson MD, Jacobsen G , et al. Reliability and validity of the Ocular Surface Disease Index. *Arch Ophthalmol* 2000; 118:615–621. [PubMed: 10815152]
22. Miller KL, Walt JG, Mink DR , et al. Minimal clinically important difference for the ocular surface disease index. *Arch Ophthalmol* 2010; 128:94–101. [PubMed: 20065224]
23. Tian L, Qu JH, Zhang XY , et al. Repeatability and Reproducibility of Noninvasive Keratograph 5M Measurements in Patients with Dry Eye Disease. *J Ophthalmol* 2016; 2016:8013621. [PubMed: 27190639]
24. Sullivan BD, Whitmer D, Nichols KK , et al. An objective approach to dry eye disease severity. *Invest Ophthalmol Vis Sci* 2010; 51:6125–6130. [PubMed: 20631232]
25. Lemp M Report of the National Eye Institute/Industry workshop on clinical trials in dry eyes. *CLAO J* 1995; 21:221–232. [PubMed: 8565190]
26. Bron AJ, Benjamin L and Snibson GR. Meibomian gland disease. Classification and grading of lid changes. *Eye* 1991; 5:395–411. [PubMed: 1743355]
27. Foulks GN, Bron AJ. Meibomian gland dysfunction: a clinical scheme for description, diagnosis, classification and grading. *The Ocular Surface* 2003; 1:107–126. [PubMed: 17075643]
28. Pflugfelder SC, Tseng SC, Sanabria O, et al. Evaluation of subjective assessments and objective diagnostic tests for diagnosing tear-film disorders known to cause ocular irritation. *Cornea* 1998; 17:38–56. [PubMed: 9436879]
29. Yokoi N, Takehisa Y, Kinoshita S. Correlation of tear lipid layer interference patterns with the diagnosis and severity of dry eye. *Am J Ophthalmol* 1996; 122:818–824. [PubMed: 8956636]
30. Mackie IA, Seal DV. The questionably dry eye. *Br J Ophthalmol* 1981; 65:2–9. [PubMed: 7448154]
31. Patel S, Bevan R and Farrell JC. Diurnal variation in tear film stability. *Am J Optom Physiol Opt* 1988; 65:151–154. [PubMed: 3364523]

32. Kutner MH, Nachtsheim CJ, Neter J , et al. Applied Linear Statistical Models. New York: McGraw-Hill Companies, 2005; p. 372.
33. Tomlinson A, Bron AJ, Korb DR , et al. The international workshop on meibomian gland dysfunction: report of the diagnosis subcommittee. Invest Ophthalmol Vis Sci 2011; 52:2006–2049. [PubMed: 21450918]
34. Hosmer DW, Lemeshow S Chapter 5: Assessing the Fit of the Model. Applied Logistic Regression. New York: John Wiley & Sons, Inc, 2000; p. 162.
35. Abdul-Fattah AM, Bhargava HN, Korb DR , et al. Quantitative in vitro comparison of fluorescein delivery to the eye via impregnated paper strip and volumetric techniques. Optom Vis Sci 2002; 79:435–438. [PubMed: 12137398]



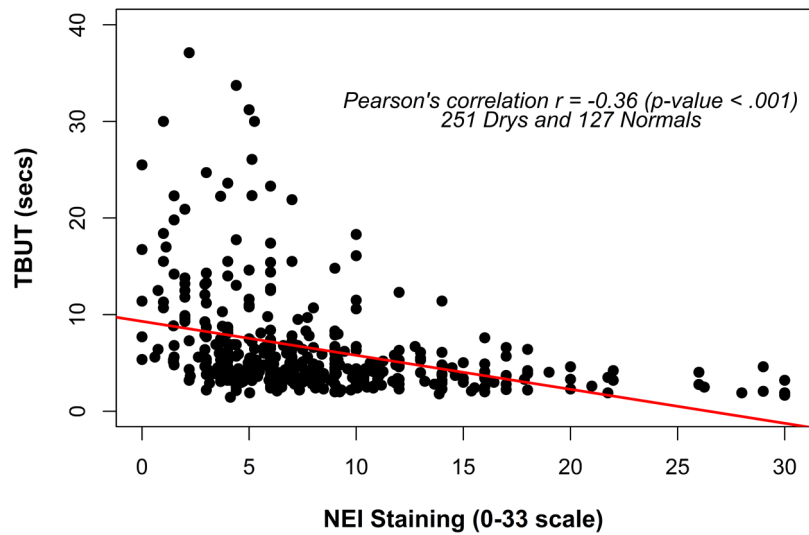
**Figure 1.**

Receiver Operating Characteristic (ROC) curves for Tear Breakup Time.

**Solid line:** model - building ROC curve for combined 2.0 and 5.0  $\mu\text{l}$  dye volumes (1.0% and 2.0 % wt/vol sodium fluorescein, respectively); 69 normal and 99 dry eye subjects.

**Large dashed line:** model – building ROC curve for 2.0  $\mu\text{l}$ , 1.0% dye; 62 normal and 114 dry eye subjects.

**Dotted line:** ROC curve for 5.0  $\mu\text{l}$ , 2.0% dye; all data (30 normals and 77 dry eye subjects).



**Figure 2.** Correlation of TBUT and NEI staining (corneal and conjunctival scores combined, 0 – 33 scale), in the subjects' worst eye (i.e., the most heavily stained of the right or left eyes). N = 232 total normal and dry subjects.

**Table 1:**

Video Tear Break-Up Time: Descriptive Statistics\* and ROC Analysis

| Test Method                                | Subject Category        | Mean (median) <sup>‡</sup> | Optimum Cut-Point <sup>‡</sup> | Test Sensitivity | Test Specificity | AUC (95% CI)               |
|--|-------------------------|----------------------------|--------------------------------|------------------|------------------|----------------------------|
| <b>Dry Eye Test</b> (one drop, not shaken) | <b>Normals</b> (n = 15) | <b>5.8 secs</b> (4.8)      | <b>4.4</b>                     | <b>0.97</b>      | <b>0.67</b>      | <b>0.873</b> (0.76 – 0.98) |
|  | <b>ATDs</b> (n = 11)    | <b>2.0</b> (1.8)           |                                |                  |                  |                            |
|  | <b>MGDs</b> (n = 20)    | <b>2.9</b> (2.7)           |                                |                  |                  |                            |
| <b>Liquid Dye</b> (2.0 µl of 1.0% NaFl)    | <b>Normals</b> (n = 15) | <b>5.7</b> (5.2)           | <b>3.2</b>                     | <b>0.90</b>      | <b>0.87</b>      | <b>0.90</b> (0.80 – 1.0)   |
|  | <b>ATDs</b> (n = 11)    | <b>2.2</b> (1.9)           |                                |                  |                  |                            |
|  | <b>MGDs</b> (n = 20)    | <b>2.1</b> (1.9)           |                                |                  |                  |                            |
| <b>Standard Strip</b> (one drop, shaken)   | <b>Normals</b> (n = 15) | <b>5.3</b> (5.4)           | <b>3.4</b>                     | <b>0.97</b>      | <b>0.80</b>      | <b>0.91</b> (0.81 – 1.0)   |
|  | <b>ATDs</b> (n = 11)    | <b>1.6</b> (1.6)           |                                |                  |                  |                            |
|  | <b>MGDs</b> (n = 20)    | <b>2.3</b> (2.0)           |                                |                  |                  |                            |

\* Raw values for TBUT were log-transformed prior to statistical analysis

<sup>‡</sup> in seconds<sup>‡</sup> Cut-point is the best balance of sensitivity and specificity



**Table 2:**

Tear Break-Up Time by Method: Descriptive Statistics

| Test Method                            | Subject Category   | Mean<br>(SD, sec) | Median<br>(sec) | Log<br>Transformed<br>Mean<br>(secs) |
|--|--------------------|-------------------|-----------------|--------------------------------------|
| 2.0 $\mu$ l of 1.0% NaFl               | Normals (n = 97)   | 11.3 (7.1)        | 9.27            | 2.26                                 |
|  | All Drys (n = 174) | 4.39 (1.8)        | 4.00            | 1.41                                 |
| 5.0 $\mu$ l of 2.0% NaFl               | Normals (n = 30)   | 10.04 (6.7)       | 7.30            | 2.14                                 |
|  | All Drys (n = 77)  | 3.68 (1.44)       | 3.50            | 1.24                                 |
| Video Method: 2.0 $\mu$ l of 1.0% NaFl | Normals (n = 10)   | 7.14 (2.4)        | 6.22            | 1.92                                 |
|  | All Drys (n = 32)  | 2.20 (1.3)        | 1.9             | 0.65                                 |

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 3:**Cut-points for varying sensitivity and specificity, combined 2.0 and 5.0  $\mu$ l data

| Cut-Point (seconds) | Sensitivity | Specificity | Normals Correctly Predicted*<br>Number (%) | Dry Subjects Correctly Predicted†<br>Number (%) |
|---------------------|-------------|-------------|--|---|
| 6.05                | 0.91        | 0.70        | 30 (83.3)                                  | 57 (82.6)                                       |
| 5.85                | 0.90        | 0.75        |  |   |
| 5.5                 | 0.88        | 0.81        | 30 (83.3)                                  | 54 (78.3)                                       |
| 5.41                | 0.87        | 0.82        |  |   |
| <b>5.33</b>         | <b>0.85</b> | <b>0.86</b> | 30 (83.3)                                  | 54 (78.3)                                       |
| 5.25                | 0.84        | 0.88        |  |   |
| 5.10                | 0.84        | 0.90        | 32 (88.9)                                  | 52 (75.4)                                       |
| 4.85                | 0.78        | 0.91        |  |   |

\* 36 total normal subjects in the model-checking data set

† 69 total dry eye subjects in the model-checking data set

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript