Repeatability and Diurnal Variation of Tear Ferning Test

Ali M. Masmali,*, Jarallah M. Al-Bahlal, Gamal A. El-Hiti, Saeed Akhtar, Christine Purslow, Paul J. Murphy, Turki Almubrad

*a Cornea Research Chair (CRC), Department of Optometry, College of Applied Medical Sciences, King Saud University, P.O. Box 10219, Riyadh 11433, Saudi Arabia

b School of Optometry and Vision Sciences, Cardiff University, Cardiff, UK
c School of Optometry and Vision Science, University of Waterloo, Waterloo, Canada

The authors have no conflicts of interest to disclose.

* Corresponding author at: Cornea Research Chair (CRC), Department of Optometry, College of Applied Medical Sciences, King Saud University, P.O. Box 10219, Riyadh 11433, Saudi Arabia.
Tel.: +966 11 4693547; fax: +966 11 4693536
E-mail address: amasmali@ksu.edu.sa (A. Masmali).

Number of Tables: 2
Number of Figures: 4
Fund has been received from the College of Applied Medical Sciences Research Canter and the Deanship of Scientific Research at King Saud University.
Objectives: To investigate tear ferning test repeatability between sessions by observing changes in the tear fern pattern during the day.

Methods: Twenty-three healthy young adults (15 male and 8 female), ranging in age from 20 to 32 years (mean±SD: 22.9±3.3 years) without signs or symptoms of dry eye disease, ocular disease or contact lens wear, were enrolled in the study. Schirmer I, tear break up time (TBUT) test and McMonnies questionnaire were used to screen volunteers. Schirmer I and TBUT tests were applied to both eyes in each subject. Four samples of tear fluid were collected from the right eye of each subject using glass capillaries, at set intervals during a single day (9am, 11am, 2pm and 4pm). The tear ferning (TF) patterns obtained from samples were classified according to the Masmali TF grading scale, to increments of 0.1.

Results: The median values obtained from the McMonnies, Schirmer and TBUT tests were 4.0±2.0, 30.0±7.0mm (OD), and 16.0±10.0s (OD), respectively. There were no statistically significant differences between the TF grades for tear samples collected at different times of the day (Wilks’ Lambda, $p = 0.351$). The majority (84.8%) of TF grades were between 0.0 and 1.5; the remaining 15.2% of subjects had TF between grades 1.6 and 1.9. The overall mean grade for the tear ferning was 1.1±0.3. There were small, insignificant correlations between TF grades and the McMonnies questionnaire ($r = 0.130$) and TBUT ($r = 0.248$), and a negligible correlation with Schirmer test ($r = -0.046$).

Conclusions: The results found no significant differences within the tear ferning for tear samples collected at different times of the day, suggesting that there is little diurnal variation evident.
Keywords: Tear ferning; non-dry eye subjects; Masmali grading scale; Schirmer test; Tear break up time; McMonnies questionnaire

INTRODUCTION

Tear production is very important for clear vision and eye health. Dry eye patients suffer from discomfort, such as sensitivity to light, stinging, burning, blurriness and grittiness, or complain of scratchy and itchy eyes.\textsuperscript{1−3} The multiple causes of dry eye make its diagnoses and treatment challenging.\textsuperscript{4} Moreover, the current available methods for the diagnosis of dry eye are far from perfect, with poor correlations between signs and symptoms, and between diagnostic tests.\textsuperscript{5}

The ideal test should be simple to use, repeatable, sensitive and specific to dry eye disease, and should ideally correlate with symptoms. Several clinical tests focus on examination of tear film quantity (volume), stability, or quality (composition). Tear volume assessment can be carried out using the Schirmer’s test\textsuperscript{6} or by tear meniscus measurement.\textsuperscript{7} The Schirmer’s test is the most common method for the evaluation of tear production,\textsuperscript{8−10} but its invasive approach makes it liable to reflex tearing.\textsuperscript{11} The phenol red thread test (PRT) can also be used, and has advantages over Schirmer’s test in being more comfortable for the patient, requiring less time and there is no need for anaesthesia\textsuperscript{8}, but there is still a question on what exactly the thread is measuring – whether it is the basal secretion rate\textsuperscript{12} or perhaps related to wetting characteristics of the thread.\textsuperscript{13} Tear meniscus measurement has the advantage of being non-invasive, depending on technique, but the test lacks universal cut-off values for normative data.\textsuperscript{2}
Tear film stability can be assessed by measuring tear break-up time (TBUT).\textsuperscript{14} However, further studies are needed to refine the sensitivity, specificity and reproducibility of the test.\textsuperscript{2} Tear clearance assessment can be evaluated by the fluorescein clearance test.\textsuperscript{15,16} The test evaluates reflex tears, basal tears and tear clearance simultaneously with the advantage of being relatively easy to perform and inexpensive.\textsuperscript{17} However, low specificity and sensitivity for tear evaluation and reflex tears production are disadvantages.\textsuperscript{17,18} Non-invasive tear break-up time (NITBUT) can assess tear stability, but it has not been confirmed whether this test is evaluating changes in tear stability from changes to the lipid layer or to the overall tear film.\textsuperscript{19}

Some aspects of the tear film chemical properties can be assessed using tear osmolarity.\textsuperscript{20–22} Osmolarity is a measure of the solute concentration, particularly of ions such as sodium and potassium, in the tear film, and is expressed by the unit mOsm/L. A reduction in tear volume by increased evaporation of decreased production may result in hyper-osmolarity. The TearLab\textsuperscript{TM} osmolarity system (TearLab\textsuperscript{TM} Corp., San Diego, California) can measure the osmolarity of tears efficiently, but the cost associated with the running of this test is high, and repeatability requires multiple testing.\textsuperscript{23}

An alternative for assessing tear film composition is to use tear ferning (TF), which has showed good specificity and sensitivity.\textsuperscript{24,25} Bodily fluids, when allowed to dry on a glass slide at room temperature and low humidity, produce ferns of specific patterns.\textsuperscript{26} The process of the TF test involves the use of a glass capillary tube to collect a sample of tears from the inferior tear meniscus.\textsuperscript{25,27} The sample is expelled from the capillary tube and the tears are allowed to dry in air at room temperature.\textsuperscript{26,28} The ferning patterns produced are then observed under light microscopy\textsuperscript{29} at magnification levels ranging from 10–100X.\textsuperscript{30,31}
In 1984, Rolando suggested a tear ferning (TF) grading scale consisting of four types (I–IV), in which Types I and II were more commonly observed in normal eye subjects, while, Types III and IV were typically observed in dry eye patients. Recently, the Masmali 5-point TF grading scale has been developed which overcomes some of the limitations associated with the Rolando scale. The Masmali TF grading scale was found to have good validity in describing TF patterns, with Grades ≥2 classified as abnormal. With using this new grading scale, the TF test has the potential to be practiced in the clinic and can be used as a support for other dry eye tests.

This paper reports on a study that investigates one aspect of the validity of the TF test: testing the repeatability of tear ferning pattern during different times of the day, using the Masmali grading scale.

METHODS

Subjects

Twenty-three healthy young adults (15 male and 8 female) who ranged in age from 20 to 32 years (mean±SD: 22.9±3.3 years) were recruited from King Saud University students and staff in Riyadh, Saudi Arabia. Ethical approval was obtained from the College of Applied Medical Science Research Centre, King Saud University. This study followed the tenets of the Declaration of Helsinki, in which informed consent was obtained from the subjects after an explanation of the nature and possible consequences of the study. Subjects were then examined with routine slit lamp biomicroscopy examination to assess the anterior part of the eye and to confirm the absence of ocular diseases. At this point volunteers also completed the McMonnies’ questionnaire to exclude dry eye patients. Dry eye was diagnosed for a score >14.5. In
addition, Schirmer I and tear break-up time (TBUT) tests were applied for both eyes of each subject to assess exclusion criteria.

A single tear sample (first sample: 9am) was collected prior to the Schirmer test screening to avoid bias, and after applying Schirmer’s test, ten minutes was allowed to expire prior to TBUT assessment. All subjects were examined in the same laboratory, where room temperature remained stable at 23°C and 40% humidity (one room was selected for this study and temperature and humidity were checked every day during the study). Subjects spent the day in the building at room temperature, and were examined indoors between 9am and 4pm. All tear samples were collected from the subjects by the same investigator using the same method and under the same condition.

The TearFlo™ Schirmer filter paper strips were purchased from Contacare Ophthalmics and Diagnostics (Gujarat, India) and were applied to both eyes at the same time; a value above 10 mm was considered as normal. The tear break-up time (TBUT) was performed three times in each eye and the average time was recorded. The cut-off value for dry eye was <10 seconds.

The study design was masked to avoid any bias. The McMonnies’ questionnaire, slit-lamp examination, Schirmer’s test and tear collections were completed by one investigator, and the imaging of the tear ferning patterns slides and the grading of the ferning patterns was completed by another investigator, who was blind to the subject’s other test results.

Tear collection

The tear samples were collected at four different times during the day (9am, 11am, 2pm and 4pm). Each sample (1µl) was collected from the lower meniscus of the right eye only using a glass capillary tube (10µl, Drummond Scientific Company, USA) and allowed to dry on a clean,
unused glass slide for 10 minutes under normal room temperature (23°C) and humidity (40%).

Samples were immediately observed under digital microscope (Olympus DP72) with 10X magnification. Each ferning pattern observed was graded using the Masmali TF grading scale in 0.1 increments to improve grade refinement.

Statistical Analysis

Data were collated using Excel (Microsoft Office 2010) and analysed using SPSS software (IBM Software, version 20). Data were examined for normality using Kolmogorov-Smirnov tests and were found to be normally distributed (Kolmogorov-Smirnov, p >0.05) for TF grades and not normally distributed (Kolmogorov-Smirnov, p <0.05) for McMonnies, Schirmer and TBUT tests. The mean±standard deviation (SD) was used to describe the results from TF grades, while the median±inter-quartile range (IQR) was used to describe the results for McMonnies, Schirmer and TBUT tests. The parametric test (one-way repeated measures ANOVA) was used to compare TF grade at different time points. Since the data collected from both eyes for Schirmer and TBUT were correlated (Schirmer's test: Spearman's rho= 0.52; TBUT: Spearman's rho= 0.74), the measurements for the right eye only were used. In normal eye studies, it has been recommended that when the data from both eyes is highly correlated only one eye per participant can be used. Spearman’s correlation was used to investigate the relationship between all data obtained (McMonnies, Schirmer, TBUT and TF grades). Correlation test was used to study the relationship between TF grade, McMonnies, Schirmer and TBUT results. Correlation coefficients were graded as: small (0.10 to 0.29), medium (0.30 to 0.49) and large (0.50 to 1.00). The Coefficient of variation between the four sessions was calculated using the formula (100 X SD)/overall mean).
RESULTS

The median (±IQR) score for the McMonnies questionnaire was 4.0±2.0. The median (±IQR) values obtained from the Schirmer and TBUT tests were 30.0±7.0 mm (OD) and 16.0±10.0 s (OD), respectively.

Tear Ferning

There were no significant differences between the TF grades for the four samples, collected at different sessions and different times during the day, within each subject (Wilks’ Lambda, $p = 0.351$), and there were no statistically significant differences between the pair-wise comparisons of any two samples (Table 1).

Table 1 here

The mean±SD TF grading pattern for the four samples collected from each subject at different times during the day is shown in Figure 1. The average coefficient of variation was 0.30% and the cohort range was 0.05% to 1.6%.

Figure 1 here

As an example, the tear ferning patterns for the four samples collected from one subject at 9am (A), 11am (B), 2pm (C) and 4pm (D), illustrated in Figure 2, showed no significant differences.

Figure 2 here
The Bland–Altman plot showing the mean differences between the four sessions and the ±2SD limits of agreement for all subjects is presented in Figure 3.

The tear fern grading scale results for the right eye only showed that the majority (84.8%) of TF grades were between 0.0 and 1.5, with the remaining 15.2% of subjects having TF grades between 1.6 and 1.9. The mean tear ferning grade for all samples collected during the day was in the range of 1.0–1.1 (mean±SD: 1.1±0.3), based on the Masmali TF grading scale.\textsuperscript{33} It was found that the most observed tear ferning patterns (76.1%) corresponding to grades between 0.6 and 1.0. The TF grading scale range percentages are shown in Figure 4.

There were small, but not significant, correlations between the TF grades and the McMonnies questionnaire (Spearman; $r = 0.130$) and TBUT (Spearman; $r = 0.248$), and a negligible negative correlation with Schirmer test ($r = -0.046$). A medium (and significant) correlation was found between McMonnies questionnaire and Schirmer’s test, with a Spearman’s correlation ($r$) of 0.461 (Table 2).
Tear ferning has been reported to have potential to become a simple clinical test that can evaluate the quality of tear compositions. By drying a small tear sample on a clean glass slide to produce a tear ferning pattern, aspects of tear composition, especially of electrolyte and macromolecule concentration, can be observed. Tear ferning has its origins in examining the quality of mucins from mucous secreting tissues, but work by Rolando showed its potential for assessing tear film quality. A significant development was the availability of the Rolando tear fern scale to grade the ferning pattern produced. More recently, in response to weaknesses in the design of the Rolando scale, the Masmali scale was developed. With this new scale, there is potential for tear ferning to become a more regularly included test for the tear film clinician.

However, to make a clinical test useful, its repeatability must be known, and should be within acceptable limits. Indeed, the validity of any measurement is absent when it is totally unrepeatable. The results from this study show good repeatability, with no significant differences in the TF patterns between the four tear samples collected from one eye at different times in the day (9am, 11am, 2pm and 4pm), using the Masmali scale. This matches the results of a previous study investigating repeatability with the Rolando scale, which found no significant difference between tear samples collected at only two times of the day (once in the morning and once in the afternoon). However, this study has improved over the previous study, by having four samples for comparison (two samples at different times in the morning and two samples at different times in the afternoon) rather than only two samples during the day, as well as using the Masmali TF grading scale to classify the ferning patterns.

A previous study found similar levels of good repeatability, where no significant difference in tear fern pattern was found between five tear samples collected from one eye over
the same session, and where no significant difference was found between five drops dried from a single tear sample.\textsuperscript{35} The average grade observed also matches previous results for a normal cohort using the Masmali grading scale.\textsuperscript{36} The most observed grade was Grade 1 and the mean was Grade 1.1.

Repeatability of the ferning pattern produced from a tear sample can be potentially influenced by the collection method, and also by the grading scheme.\textsuperscript{48} Norn \textsuperscript{48} studied the repeatability of two tear sample collection methods - the use of glass rods sampling produced high variability (a coefficient of variation of 99–128\%), and while lower variability results were obtained by using capillary tubes (coefficient of variation: 35\%) for sampling a random volume, and (coefficient of variation: 6.4\%) for collecting a standardized tear volume, these coefficients are still high. In contrast, the use of the Masmali grading scale in this study showed excellent repeatability for the tear ferning test with a 0.30\% average coefficient of variation.

This study has a limitation that it has been done only on healthy subjects, and dry eye subjects may show different result. A significant diurnal variation of visual function and ocular surface physiology,\textsuperscript{49} and of tear osmolarity\textsuperscript{50} have been found in dry eye subjects. So it could be assumed that variation in a dry eye cohort may produce some variability and so the next study that needs doing is to repeat this one using a cohort of dry eye subjects. This study also used fluorescein BUT, and non-invasive TBUT would reveal different characteristics of the tear film, which might be helpful in assessing correlation of tear ferning with other clinical tests for dry eye.

The results from this study show that tear ferning has good repeatability, and that the use of the Masmali grading scale, in a healthy subject cohort, will produce consistent grading results. It has also shown that a tear sample collected a different time points will produce a similar
ferning pattern. These results support the tear ferning test and suggest that it has potential for clinical and research use, as part of a routine tear film examination.

ACKNOWLEDGEMENTS

The authors extend their appreciation to the College of Applied Medical Sciences Research Center and the Deanship of Scientific Research at King Saud University for its funding of this research.

REFERENCES


Figures Legend

FIG. 1. The mean±SD TF grade for the four samples collected from each subject at different time during the day.

FIG. 2. Tear ferning patterns of the four samples collected from one subject at 9am (A), 11am (B), 2pm (C) and 4pm (D), showing no significant differences (Grade 0).

FIG. 3. Bland–Altman plot showing the mean differences between the four sessions and the ±2SD limits of agreement for all subjects.

FIG. 4. Percentages of the TF grades range during the day.
### TABLE 1. Mean Differences and Confidence Interval for Repeatability of TF Grades

<table>
<thead>
<tr>
<th>Tear Samples</th>
<th>Mean Differences</th>
<th>Sig.</th>
<th>95% Confidence Interval of the Differences</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Lower</td>
</tr>
<tr>
<td>1</td>
<td>−0.004</td>
<td>1</td>
<td>−0.266</td>
</tr>
<tr>
<td>2</td>
<td>0.035</td>
<td>1</td>
<td>−0.184</td>
</tr>
<tr>
<td>3</td>
<td>0.135</td>
<td>0.797</td>
<td>−0.115</td>
</tr>
<tr>
<td>4</td>
<td>0.004</td>
<td>1</td>
<td>−0.258</td>
</tr>
<tr>
<td>2</td>
<td>0.039</td>
<td>1</td>
<td>−0.234</td>
</tr>
<tr>
<td>3</td>
<td>0.139</td>
<td>0.598</td>
<td>−0.095</td>
</tr>
<tr>
<td>4</td>
<td>−0.035</td>
<td>1</td>
<td>−0.253</td>
</tr>
<tr>
<td>3</td>
<td>−0.039</td>
<td>1</td>
<td>−0.312</td>
</tr>
<tr>
<td>4</td>
<td>0.100</td>
<td>1</td>
<td>−0.160</td>
</tr>
<tr>
<td>4</td>
<td>−0.135</td>
<td>0.797</td>
<td>−0.385</td>
</tr>
<tr>
<td>3</td>
<td>−0.139</td>
<td>0.598</td>
<td>−0.374</td>
</tr>
<tr>
<td>2</td>
<td>−0.100</td>
<td>1</td>
<td>−0.360</td>
</tr>
</tbody>
</table>

### TABLE 2. Correlation Between TF Grade, McMonnies Score, Schirmer and TBUT Tests

<table>
<thead>
<tr>
<th>Test/Correlation</th>
<th>TF</th>
<th>McMonnies</th>
<th>Schirmer</th>
<th>TBUT</th>
</tr>
</thead>
<tbody>
<tr>
<td>TF</td>
<td>Spearman's Correlation</td>
<td>1</td>
<td>0.130</td>
<td>−0.046</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td></td>
<td>0.553</td>
<td>0.834</td>
<td>0.254</td>
</tr>
<tr>
<td>N</td>
<td>23</td>
<td>23</td>
<td>23</td>
<td>23</td>
</tr>
<tr>
<td>McMonnies</td>
<td>Spearman's Correlation</td>
<td>0.130</td>
<td>1</td>
<td>0.461&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>0.553</td>
<td></td>
<td>0.027</td>
<td>0.403</td>
</tr>
<tr>
<td>N</td>
<td>23</td>
<td>23</td>
<td>23</td>
<td>23</td>
</tr>
<tr>
<td>Schirmer (OD)</td>
<td>Spearman's Correlation</td>
<td>−0.046</td>
<td>0.461&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>0.834</td>
<td></td>
<td>0.027</td>
<td>0.389</td>
</tr>
<tr>
<td>N</td>
<td>23</td>
<td>23</td>
<td>23</td>
<td>23</td>
</tr>
<tr>
<td>TBUT (OD)</td>
<td>Spearman's Correlation</td>
<td>0.248</td>
<td>−0.183</td>
<td>−0.189</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>0.254</td>
<td></td>
<td>0.403</td>
<td>0.389</td>
</tr>
<tr>
<td>N</td>
<td>23</td>
<td>23</td>
<td>23</td>
<td>23</td>
</tr>
</tbody>
</table>

<sup>a</sup> Correlation is significant at the 0.05 level.
Figure 4
Figure 3
Figure 1

Figure 2