

## **Tear production levels and dry eye disease severity in a large Norwegian cohort**

Mazyar Yazdani<sup>1,2</sup>, Xiangjun Chen<sup>2,3,4,5</sup>, Behzod Tashbayev<sup>2,4</sup>, Øygunn A. Utheim<sup>1,2</sup>, Sten Ræder<sup>2</sup>,  
Neil Lagli<sup>6</sup>, Aleksandar Stojanovic<sup>7</sup>, Darlene A. Dartt<sup>8</sup>, Tor P. Utheim<sup>\*1,2,3,4,5,9,10,11,12</sup>

<sup>1</sup>Department of Medical Biochemistry, Oslo University Hospital, Oslo, Norway; <sup>2</sup>The Norwegian Dry Eye Clinic, Oslo, Norway; <sup>3</sup>Department of Ophthalmology, Arendal Hospital, Norway; <sup>4</sup>Department of Oral Surgery and Oral Medicine, Faculty of Dentistry, University of Oslo, Norway; <sup>5</sup>Faculty of Health Sciences, National Centre for Optics, Vision and Eye Care, University College of Southeast Norway, Norway; <sup>6</sup>Department of Clinical and Experimental Medicine; Ophthalmology, Linköping University, Linköping, Sweden; <sup>7</sup>Department of ophthalmology, University Hospital North Norway, Tromsø, Norway; <sup>8</sup>Schepens Eye Research Institute/Massachusetts Eye and Ear, Harvard Medical School, Boston, MA, United States of America; <sup>9</sup>Department of Plastic and Reconstructive Surgery, Oslo University Hospital, Oslo, Norway; <sup>10</sup>Department of maxillofacial surgery, Oslo University Hospital, Oslo, Norway; <sup>11</sup>Department of Ophthalmology, Stavanger University Hospital, Norway; <sup>12</sup>Department of Clinical Medicine, Faculty of Medicine, University of Bergen, Bergen, Norway.

\* Corresponding author: Tor P. Utheim

Email: utheim2@gmail.com

Telephone: +47 93085870

Corresponding address: Department of Medical Biochemistry, Oslo University Hospital, Ullevål, Kirkeveien 166, 0450 Oslo, Norway.

## **Abstract**

**Purpose:** To determine if the Schirmer I test (without anesthesia) cut-off value is a predictor of dry eye severity in a large Norwegian cohort of dry eye disease (DED) patients, which are grouped into six levels of tear production.

**Methods:** Patients (n=1090) with DED of different etiologies received an extensive dry eye work-up: osmolarity (Osm), tear meniscus height (TMH), tear film break-up time (TFBUT), ocular protection index (OPI), ocular surface staining (OSS), Schirmer I test (ST), meibum expressibility (ME) and meibum quality (MQ). Classification of dry eye severity level (DESL) and diagnosis of meibomian gland dysfunction (MGD) were also included. The cohort was divided into six groups: below and above cut-off values of 5 (groups 1 and 2), 10 (groups 3 and 4) and 15 mm (groups 5 and 6) of ST. Mann-Whitney test and Chi-Square test were used for group comparison of parameters ( $P \leq 0.05$ ).

**Results:** The groups 1, 3, and 5 had values indicating more severe DED than the groups 2, 4, 6 with significant difference in DESL, Osm, TFBUT, OPI, OSS and TMH. Regardless of the choice of cut-off values, there was no statistically significant difference in ME, MQ and MGD between groups below and above selected cut-off value. When gender difference was considered in each group, significant difference was only observed for DESL (groups 2, 4 and 5), TFBUT (groups 2, 4 and 5), OPI (groups 2 and 6) and ME (group1).

**Conclusions:** Schirmer I is a robust discriminator for DESL, Osm, TFBUT, OPI, OSS and TMH, but not for ME, MQ and MGD. Patients with lower tear production levels presented with more severe DED at all three defined cut-off values. Interestingly, the differences in the mean values of DESL were minimal although statistically significant. Thus, the clinical value of different Schirmer levels appears to be limited.

**Keywords:** Dry eye disease, Schirmer I test, tear production levels, clinical tests, large population of patients.

## **Introduction**

Dry eye disease (DED) is a multifactorial disorder, which can be caused by an alteration in the quality or quantity of tear film's three layers<sup>1</sup>. DED can result from primary factors such as lacrimal gland atrophy and secondary factors such as pathological changes in the eyelids, conjunctiva, or cornea along with other influencing factors, including immunological processes, neurotransmitters, hormones, pharmaceuticals, contact lenses and environmental pollution<sup>2-4</sup>. In addition to these causes, several risk factors have been proposed as contributing to the development of the disease: gender (i.e. female), increasing age, therapies (e.g. postmenopausal estrogen and radiation), dietary deficiencies (e.g. vitamin A and omega-3 fatty acids), and systemic diseases (e.g. hepatitis C)<sup>5-8</sup>.

Several clinical tests are available for diagnostic evaluations, clinical trials and follow-up examinations, including Schirmer I test (ST), measurement of tear osmolarity (Osm), tear film break-up time (TFBUT), ocular surface staining (OSS), tear meniscus height (TMH), meibum expressibility (ME) and meibum quality (MQ). Among them, ST is the most widely used method for assessment of aqueous tear production. It was first introduced by Schirmer in 1903 as installed piece of striped and marked (35 by 5 mm) blotting paper on lower eyelid margin with or without anesthesia to collect tears. ST suffers from several drawbacks such as poor reproducibility, sensitivity and specificity, long testing time (i.e. 5 min), potential for evaporative loss, high variability, discomfort, possibility of uneven absorption of tear by paper strip, and no well-defined cut-off value<sup>9-13</sup>. Nevertheless, ST without topical anaesthesia is considered a valid ophthalmological test for patients with severe DED<sup>14</sup>. To assess whether the ST can be utilized as a discriminator for other objective tests, we investigated clinical parameters in a large Norwegian cohort of 1090 dry eye disease (DED) patients grouped into six levels of tear production. The six groups were below and above cut-off values of 5 mm (groups 1 and 2), 10 mm (groups 3 and 4) and 15 mm (groups 5 and 6) of the ST.

## **Materials and Methods**

### **Patients**

One thousand and ninety subjects (average age:  $52.86 \pm 16.03$  years, range: 8-95, 796 females and 294 males) diagnosed with DED of different etiologies were consecutively included in the study at the Norwegian Dry Eye Clinic between 2012 and 2016. All subjects received an extensive dry eye work-up.

All examinations were carried out at the same clinic by the same ophthalmologist during normal working hours between 9 AM and 3 PM. The use of the data for the study has been reviewed by The Regional Committee for Medical & Health Research Ethics, Section C, South East Norway (REC). REC found the research project "Evaluation of data from the Norwegian Dry Eye Clinic" to be outside the remit of the Act on Medical and Health Research (2008) and therefore could be implemented without its approval. Prior to data collection, informed consent was obtained from all participants. All procedures performed in this study were in compliance with the Declaration of Helsinki.

### **Clinical Evaluation**

All subjects underwent a comprehensive ophthalmic examination by the following stepwise tests: osmolarity (Osm), tear meniscus height (TMH), tear film break-up time (TFBUT), ocular protection index (OPI), ocular surface staining (OSS), Schirmer I test (ST), meibum expressibility (ME), and meibum quality (MQ), classification of dry eye severity level (DESL), and diagnosis of meibomian gland dysfunction (MGD).

The tear quality evaluation was performed using TFBUT after instillation of 5  $\mu$ l 2% fluorescein sodium. Ocular surface staining with fluorescein was recorded after Oxford Grading Scheme: 0-15; range of corneal staining: 0-5)<sup>15</sup>. Schirmer test was performed without topical anaesthesia in 5 minutes. The calculation of OPI was based on the ratio of the TFBUT divided by the blink interval<sup>16</sup>. TMH was measured using a slit lamp. ME was recorded based on the number of secreted glands of the lower eyelid viewed at the slit lamp when light pressure applied by cotton tips on central five MGs (0=all five glands expressible; 1=three to four glands expressible; 2=one to two glands expressible; and 3=no gland expressible). For MQ, the central eight glands of the lower eyelid were scored on a scale of 0 to 3 for each gland (0=clear; 1=cloudy; 2=cloudy with debris (granular); and 3=thick, like toothpaste). The sum of the central eight glands was then calculated (total score range, 0-24). DESL was assessed according to the Behrens, et al.<sup>17</sup>, modified by 2007 International Dry Eye Workshop criteria<sup>18</sup> (Table 1) and diagnosis of meibomian gland dysfunction (MGD) was based on the suggestions by the international workshop on MGD<sup>19</sup>. Only 520 subjects were

examined for tear film osmolarity measurement using TearLab Osmolarity System (TearLab Corp, San Diego, CA).

### **Statistical Analysis**

Data from right eye were used for statistical analysis. The cohort was divided into six groups: below and above cut-off values (5, 10 and 15 mm) of ST for comparisons. Mann-Whitney test and Chi-Square test were used for group comparison of parameters using SPSS version 21 (SPSS Inc., Chicago, IL, USA). Chi-Square test was used to determine if females were different from males in each group. Data are presented as mean and standard error. *P* values less than or equal to 0.05 were considered significant.

## Results

Table 2 presents the subject demographics (average age:  $52.86 \pm 16.03$  yrs, range: 8-95). The number of females (n=796) was higher than males (n=294). In all groups, except age group 0-19, females outnumbered males. The highest number of patients belonged to age group 40-59 (n=430). In contrast, age groups 80-99 (n=34) and 0-19 (n=9) presented the lowest number of patients, respectively.

Table 3 shows the distribution of Schirmer test wetting length. The overall average was  $14.67 \pm 0.308$  mm. Among participants, 20%, 45% and 62% had Schirmer test of  $<5$  mm,  $\leq 10$  mm and  $\leq 15$  mm, respectively. The percentage of fully soaked strip ( $\geq 30$  mm) was 12.7%.

The analysis revealed that groups 1, 3, and 5 (i.e., individuals with values below cut-off levels of  $\leq 5$  mm,  $\leq 10$  mm, and  $\leq 15$  mm, respectively) had values indicating more severe DED with significant difference of  $p \leq 0.05$  for DESL, Osm, TFBUT, OPI, OSS and TMH compared to the corresponding groups (2, 4, and 6) (Table 4). Regardless of the choice of cut-off values, there was no statistically significant difference in ME, MQ and MGD between the groups. The Chi-Square test indicated significant difference between six tear production level groups for TMH ( $p < 0.001$ ) in all three defined cut-off values but there was no significant difference for MGD (Table 4).

When gender difference was considered in each group (Table 5), Chi-Square test indicated significant difference only for DESL (groups 2, 4 and 5), TFBUT (groups 2, 4 and 5), OPI (groups 2 and 6) and ME (group1).

## Discussion

This study revealed that patients with lower tear production levels measured by ST presented with more severe DED than patients with higher tear production. All the three defined cut-off values ( $\leq 5$ ,  $\leq 10$  and  $\leq 15$  mm) gave rise to significant difference for DESL, Osm, TFBUT, OPI, OSS and TMH, but not for ME, MQ and MGD, in the six tear production level groups. There was no clear pattern in the gender difference of each group and significant difference was only observed for DESL (groups 2, 4 and 5), TFBUT (groups 2, 4 and 5), OPI (groups 2 and 6) and ME (group 1).

An association between ST and other commonly used clinical DED tests has previously been published. For example, three out of four statistical analyses performed by Nichols, et al.<sup>20</sup> revealed a significant relationship between the ST (cut-off  $\leq 5$  mm) and TFBUT (cut-off  $\leq 10$  s) in DED patients (n=75) compared to other objective tests. These two parameters were significantly correlated ( $r=0.40$ ) and lower ST results were found to predict lower TFBUT in the stepwise, multivariate logistic regression models. Also, Wilcoxon rank sum analysis showed that subjects with an abnormal TFBUT had lower mean Schirmer scores.

Sullivan, et al.<sup>21</sup> determined correlation coefficients between ST (cut-off  $< 7$  mm) and other tests: Osm (0.05), TFBUT (0.08), fluorescein corneal staining (0.14), lissamine green conjunctival staining (0.13) and Bron/Foulks meibomian gland grading (0.05) in 344 subjects with (262) and without (82) DED. In order to compare their results with an independent data set, the authors used the data collected from a previously published study in Germany with 200 subjects (184 DED and 16 health controls)<sup>22</sup>. Accordingly, correlation coefficients were reported between ST and other clinical parameters, such as Osm (0.00), TFBUT (0.06), fluorescein corneal staining (0.03), lissamine green conjunctival staining (0.03) and Bron/Foulks meibomian gland grading (0.03). These results indicate that this test of aqueous deficiency can be utilized as a discriminator for some other objective tests.

Among all the parameters investigated in the current study, ST was unable to significantly discriminate only those related to meibomian gland function (ME, MQ and MGD). Patients with MGD have an unstable tear lipid layer but are generally considered to have normal tear production. This instability in the lipid layer may cause a decrease in tear volume, which in turn results in higher rates of evaporations and eventually damage to the ocular surface<sup>23</sup>. In contrast, Tung, et al.<sup>24</sup> observed a different pattern in MGD wherein higher tear volume correlates with worse ocular surface disease. Thus, the authors hypothesized that changes in tear composition such as hyperosmolarity, lactoferrin concentration and inflammatory mediators may play a more important role in causing corneal epithelial damage in MGD than increased evaporation. This may also be the case in our study, as ST could not significantly determine meibomian gland parameters in DED

patients. Therefore, measuring tear production may not be an efficient discriminator for ME, MQ, and MGD tests.

Historically, a gradual decrease in cut-off value for ST has been observed. For example, Schirmer<sup>25</sup>, De Roeth<sup>26</sup> and Van Bijsterveld<sup>27</sup> used 15, 10 and 5.5 mm as normal cut-off values in their studies, respectively. The lack of well-defined cut-off values for the available DED tests such as ST complicates research making it difficult to compare results from different studies. The choice of cut-off value by researchers has been based on the purpose of their studies, e.g. increased sensitivity for screening and specificity for therapeutic purposes. Irrespective of whether the cut-off value is defined as 5, 10 or 15 mm, the present study showed that ST is a robust discriminator for some clinical DED tests in a cohort of 1090 DED individuals.

For ST without anesthesia, the application of 10 mm cut-off value was considered sufficiently high not to overlook DED for screening purposes, but a proper value should be used to increase specificity in comparing inter-institutional data and evaluating efficacy of drugs even though sensitivity decreases to some extent<sup>28</sup>. However, Lee and Hyun<sup>29</sup> suggested that there is no available normal cut-off value for ST due to the high number of false positive or false negative values, leading to the ratio of 48.4% misdiagnosis in their study. They concluded that ST was not a reproducible test to differentiate DED patients (n=15) from normal individuals (n=110) and therefore analogized ST to the toss of a coin in the diagnosis of DED. In contrast, Van Bijsterveld's (1969) cut-off value of 5.5 mm was suggested to be acceptable because of impossibility in ruling out both false positivity and false negativity in ST test due to its inaccuracy<sup>28,30,31</sup>. In the current study, a possible explanation for obtaining significant differences using all three defined cut-off values might be related to the sample size. We recruited 1090 subjects in contrast to other studies with a significantly lower number of individuals. Clearly, larger sample size increases statistical robustness<sup>32</sup>. Thus, by increasing the number of subjects, the conclusions may change and researchers should be mindful of this when designing their studies.

In response to the limitations of ST, such as invasive nature, unavoidable variability, poor reproducibility, long testing time (i.e. 5 min), low sensitivity and specificity, alternative strategies have been proposed by researchers. A study evaluating the diagnostic usefulness of ST by using the Japanese diagnostic criteria on DED concluded that the combination of ST and TFBUT increased the diagnostic predictability<sup>28</sup>. Other authors prefer to use less invasive tests, e.g. phenol red thread test<sup>33</sup> or non-invasive methods, such as thermography, keratography and Hartmann-Shack wavefront sensor<sup>34-36</sup>. Nevertheless, based on our results, ST should be chosen as a clinical standard for measuring tear production until a more accurate test appears.

**Conclusion**

Schirmer test is a robust discriminator for DESL, Osm, TFBUT, OPI, OSS and TMH, but not for ME, MQ and MGD. Our results demonstrate that patients with lower tear production levels presented with more severe DED, irrespective of different cut-off levels (5, 10 or 15 mm). Interestingly, the differences in the mean values of DESL were minimal although statistically significant. Thus, the clinical value of different Schirmer levels appears to be limited. No clear pattern of sex differences of results were observed in the groups.

**Declaration of Interests**

We declare the authors have no competing interests as defined by Current Eye Research, or other interests that might be perceived to influence the interpretation of the article. The authors alone are responsible for the content and writing of the paper.

**Acknowledgement**

This work was supported by Oslo University Hospital, Norway, and the Norwegian Dry Eye Clinic, Oslo, Norway.

## References

1. Rahman MQ, Chuah KS, Macdonald ECA, Trusler JPM, Ramaesh K. The effect of pH, dilution, and temperature on the viscosity of ocular lubricants[mdash]shift in rheological parameters and potential clinical significance. *Eye* 2012;26(12):1579-1584.
2. Bron A, Benjamin L, Snibson G. Meibomian gland disease. Classification and grading of lid changes. *Eye* 1991;5(Pt 4):395-411.
3. 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(6):584-589.
4. Brewitt H, Sistani F. Dry eye disease: The scale of the problem. *Surv Ophthalmol.* 2001;45, Supplement 2:S199-S202.
5. Schaumberg DA, Buring JE, Sullivan DA, Dana MR. Hormone replacement therapy and dry eye syndrome. *JAMA* 2001;286(17):2114-2119.
6. Zegans ME, Anninger W, Chapman C, Gordon SR. Ocular manifestations of hepatitis C virus infection. *Curr Opin Ophthalmol.* 2002;13(6):423-427.
7. Colev M, Engel H, Mayers M, Markowitz M, Cahill L. Vegan diet and vitamin A deficiency. *Clin Pediatr.* 2004;43(1):107-109.
8. Miljanović B, Trivedi KA, Dana MR, Gilbard JP, Buring JE, Schaumberg DA. Relation between dietary n-3 and n-6 fatty acids and clinically diagnosed dry eye syndrome in women. *Am J Clin Nutr.* 2005;82(4):887-893.
9. Bron AJ, Abelson MB, Ousler G, Pearce E, Tomlinson A, Yokoi N, Smith JA, Begley C, Caffery B, Nichols K. Methodologies to diagnose and monitor dry eye disease. *Ocul Surf.* 2007;5(2):108-152.
10. Nelson P. A shorter Schirmer tear test. *Optom Mon.* 1982;73:568-9.
11. Cho P, Yap M. Schirmer test. II. A clinical study of its repeatability. *Optom Vis Sci.* 1993;70(2):157-159.
12. Saleh T, McDermott B, Bates A, Ewings P. Phenol red thread test vs Schirmer's test: A comparative study. *Eye* 2006;20(8):913-915.
13. Senchyna M, Wax MB. Quantitative assessment of tear production: A review of methods and utility in dry eye drug discovery. *J Ocul Biol Dis Infor.* 2008;1(1):1-6.
14. Tsubota K, Xu KP, Fujihara T, Katagiri S, Takeuchi T. Decreased reflex tearing is associated with lymphocytic infiltration in lacrimal glands. *J Rheumatol.* 1996;23(2):313-320.
15. Bron AJ, Evans VE, Smith JA. Grading of corneal and conjunctival staining in the context of other dry eye tests. *Cornea* 2003;22(7):640-50.
16. Ousler III GW, Hagberg KW, Schindelar M, Welch D, Abelson MB. The ocular protection index. *Cornea* 2008;27(5):509-513.

17. Behrens A, Doyle JJ, Stern L, Chuck RS, McDonnell PJ, Azar DT, Dua HS, Hom M, Karpecki PM, Laibson PR. Dysfunctional tear syndrome: a Delphi approach to treatment recommendations. *Cornea* 2006;25(8):900-907.
18. Lemp MA, Foulks GN. The definition and classification of dry eye disease: Report of the definition and classification subcommittee of the international dry eye workshop (2007). *Ocul Surf.* 2007;5(2):75-92.
19. Geerling G, Tauber J, Baudouin C, Goto E, Matsumoto Y, O'Brien T, Rolando M, Tsubota K, Nichols KK. The international workshop on meibomian gland dysfunction: Report of the subcommittee on management and treatment of meibomian gland dysfunction. *Invest Ophthalmol Vis Sci.* 2011;52(4):2050-2064.
20. Nichols KK, Nichols JJ, Lynn Mitchell G. The relation between tear film tests in patients with dry eye disease. *Ophthalmic Physiol Opt.* 2003;23(6):553-560.
21. Sullivan BD, Crews LA, Messmer EM, Foulks GN, Nichols KK, Baenninger P, Geerling G, Figueiredo F, Lemp MA. Correlations between commonly used objective signs and symptoms for the diagnosis of dry eye disease: Clinical implications. *Acta Ophthalmol.* 2014;92(2):161-166.
22. Messmer E, Bulgen M, Kampik A. Hyperosmolarity of the tear film in dry eye syndrome. *Research projects in dry eye syndrome. Volume 45: Karger Publishers; 2010. p 129-138.*
23. Mathers W. Evaporation from the ocular surface. *Exp Eye Res.* 2004;78(3):389-394.
24. Tung CI, Perin AF, Gumus K, Pflugfelder SC. Tear meniscus dimensions in tear dysfunction and their correlation with clinical parameters. *Am J Ophthalmol.* 2014;157(2):301-310. e1.
25. Schirmer O. Studien zur physiologie und pathologie der tränenabsonderung und tränenabfuhr. *Graefes Arch Clin Exp Ophthalmol.* 1903;56(2):197-291.
26. De Roeth A. Lacrimation in normal eyes. *AMA Arch Ophthalmol.* 1953;49(2):185-189.
27. Van Bijsterveld O. Diagnostic tests in the sicca syndrome. *Arch Ophthalmol.* 1969;82(1):10-14.
28. Danjo Y. Diagnostic usefulness and cutoff value of Schirmer's I test in the Japanese diagnostic criteria of dry eye. *Graefes Arch Clin Exp Ophthalmol.* 1997;35(12):761-766.
29. Lee JH, Hyun PM. The reproducibility of the Schirmer test. *Korean J Ophthalmol.* 1988;2(1):5-8.
30. Lucca JA, Nunez JN, Farris RL. A comparison of diagnostic tests for keratoconjunctivitis sicca: Lactoplate, Schirmer, and tear osmolarity. *Eye Cont Lens* 1990;16(2):109-112.
31. Paschides C, Kitsios G, Karakostas K, Psillas C, Moutsopoulos H. Evaluation of tear break-up time, Schirmer's-I test and rose bengal staining as confirmatory tests for keratoconjunctivitis sicca. *Clin Exp Rheumatol.* 1988;7(2):155-157.
32. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiol.* 1982;143(1):29-36.
33. Pult H, Purslow C, Murphy PJ. The relationship between clinical signs and dry eye symptoms. *Eye* 2011;25(4):502-510.

34. Su TY, Ho WT, Lu CY, Chang SW, Chiang HK. Correlations among ocular surface temperature difference value, the tear meniscus height, Schirmer's test and fluorescein tear film break up time. *Br J Ophthalmol.* 2015;99:482-487.
35. Wang Y, Xu J, Sun X, Chu R, Zhuang H, He JC. Dynamic wavefront aberrations and visual acuity in normal and dry eyes. *Clin Exp Optom.* 2009;92(3):267-273.
36. Hong J, Sun X, Wei A, Cui X, Li Y, Qian T, Wang W, Xu J. Assessment of tear film stability in dry eye with a newly developed keratograph. *Cornea* 2013;32(5):716-721.