Spektrum Augenheilkd.
https://doi.org/10.1007/s00717-025-00593-8

spektrum der augenheilkunde

Comparison of Schirmer test values between patients with aqueous deficient dry eye and healthy participants

Hugo Pena-Verdeal · Jacobo Garcia-Queiruga · Belen Sabucedo-Villamarin · Carlos García-Resúa · Maria J. Giraldez · Eva Yebra-Pimentel

Received: 29 July 2024 / Accepted: 4 February 2025

Summary
Purpose The present study aimed to compare the results of the Schirmer test over time between the eyes of healthy participants and those with aqueous-deficient dry eye (ADDE) and to calculate the diagnostic cut-off values.
Methods An observational cross-sectional study was carried out with 20 healthy volunteers and 20 patients with ADDE who fulfilled the inclusion/exclusion criteria (40 eyes/group). A digital photograph was taken of each strip in both eyes simultaneously every minute during a Schirmer test procedure until the test ended (minutes 1, 2, 3, 4, and 5). Results from the strips were measured by a masked observer using ImageJ software and compared between the two groups.
Results There was a difference in the Schirmer test results for the various time points in all analyses between the two groups (Greenhouse–Geisser, both $p \leq 0.001$; Sidak, all $p \leq 0.001$). In the pairwise analysis of each time point, there was a difference in Schirmer values between the groups (unpaired t test, all $p \leq 0.001$). There was no inter-eye difference in Schirmer absolute values between time points for both groups (Friedman, both $p \geq 0.090$; Wilcoxon, all $p \geq 0.062$), except for the first vs. the fourth and fifth time points in the ADDE participants (Wilcoxon, both $p \leq 0.030$). In the pairwise analysis of each time point, there was a difference in inter-eye absolute values between the groups (unpaired t test, all $p \leq 0.010$), except for the fourth vs. the final time point (Mann–Whitney, both $p \geq 0.414$).
Conclusion A complete report of the progression of Schirmer test values over time and between the eyes of healthy participants and ADDE patients is presented here, along with the cut-off criteria for the different time points studied.

Keywords Dry eye diagnosis · Schirmer test · Aqueous deficient dry eye · Tear volume

Vergleich von Ergebnissen im Schirmer-Test zwischen Patienten mit trockenem Auge durch zu geringen Tränenfilm und gesunden Probanden

Zusammenfassung
Ziel Ziel der vorliegenden Studie war es, die Ergebnisse im Schirmer-Test zwischen den Augen von gesunden Probanden und durch zu geringen Tränenfilm bedingten trockenen Augen über die Zeit zu vergleichen und die diagnostischen Grenzwerte zu ermitteln.
Methoden Dazu wurde eine Querschnitts-Beobachtungsstudie mit 20 gesunden Probanden und 20 Patienten mit durch zu geringen Tränenfilm bedingten trockenen Augen durchgeführt, welche die Ein-/Ausschlusskriterien erfüllten (40 Augen/Gruppe). Bei beiden Augen wurde gleichzeitig eine digitale Fotografie von jedem Streifen in jeder Minute während des Schirmer-Tests gemacht, bis der Test zu Ende war (Minute 1, 2, 3, 4 und 5). Die Ergebnisse der Streifen wurden mittels verdeckter Beobachtung unter Verwendung der gemeinfreien Software ImageJ gemessen und zwischen den beiden Gruppen verglichen.

H. Pena-Verdeal, MSc PhD · J. Garcia-Queiruga, OD MSc PhD (✉) · B. Sabucedo-Villamarin, OD MSc · C. García-Resúa, OD MSc PhD · M. J. Giraldez, MSc PhD · E. Yebra-Pimentel, OD PhD
Departamento de Física Aplicada (Área de Optometría), Facultade de Óptica e Optometría, Universidade de Santiago de Compostela, Santiago de Compostela, Galicia, Spain
jacobogarcia.queiruga@usc.es

H. Pena-Verdeal, MSc PhD · J. Garcia-Queiruga, OD MSc PhD · C. García-Resúa, OD MSc PhD · M. J. Giraldez, MSc PhD · E. Yebra-Pimentel, OD PhD
Instituto de Investigación Sanitaria (IDIS), Travesía da Choupana S/N, Santiago de Compostela, Spain

Published online: 10 March 2025

Ergebnisse Zwischen den beiden Gruppen bestand ein Unterschied bei den Ergebnissen des Schirmer-Tests für die verschiedenen Zeitpunkte in allen Auswertungen (Greenhouse-Geisser: beide $p \leq 0{,}001$; Sidak: alle $p \leq 0{,}001$). In der paarweise erfolgenden Auswertung jedes einzelnen Zeitpunkts gab es einen Unterschied bei den Ergebnissen des Schirmer-Tests zwischen den Gruppen (ungepaarter t-Test: alle $p \leq 0{,}001$). Kein Unterschied zwischen den Augen bestand bei den absoluten Ergebnissen des Schirmer-Tests für die verschiedenen Zeitpunkte in beiden Gruppen (Friedman: beide $p \geq 0{,}090$; Wilcoxon: alle $p \geq 0{,}062$), außer für den ersten vs. den vierten und fünften Zeitpunkt bei den Teilnehmern mit durch zu geringen Tränenfilm bedingten trockenen Augen (Wilcoxon: beide $p \leq 0{,}030$). In der paarweise erfolgenden Auswertung für jeden Zeitpunkt fand sich ein Unterschied zwischen den Augen für die absoluten Werte im Vergleich der Gruppen (ungepaarter t-Test: alle $p \leq 0{,}010$), außer für den vierten vs. den letzten Zeitpunkt (Mann-Whitney: beide $p \geq 0{,}414$).
Schlussfolgerung In der vorliegenden Arbeit wird eine vollständige Darstellung des Verlaufs der Ergebnisse im Schirmer-Test über die Zeit und im Vergleich zwischen den Augen von gesunden Probanden und Patienten mit durch zu geringen Tränenfilm bedingten trockenen Augen gegeben, auch werden die Grenzwertkriterien für die verschiedenen untersuchten Zeitpunkte genannt.

Schlüsselwörter Diagnostik bei trockenem Auge · Schirmer-Test · Trockenes Auge mit reduzierter Tränenfilmproduktion · Tränenvolumen

Introduction

Dry eye disease (DED) was redefined by the Tear Film and Ocular Surface Society (TFOS) at the Dry Eye Workshop II (DEWS II) as a multifactorial disease of the ocular surface characterized by loss of tear film homeostasis accompanied by ocular symptoms [1, 2]. The TFOS DEWS II classified DED into two main types: aqueous-deficient dry eye (ADDE) and evaporative dry eye (EDE; [2]). The tear aqueous component is important for ocular surface health and its loss of homeostasis may be at the same time a key pathogenic mechanism and a diagnostic sign in the ADDE subtype [3].

Secreted tears are distributed over the ocular surface during the blinking process, departing via evaporation and drainage. Tear volume and production are key components of tear dynamics since they are directly related to the final total aqueous component [3–7]; a decreased tear volume on those variables is assumed to be a diagnostic indicator of ADDE. The Schirmer test, phenol red test, and the tear meniscus height (TMH) have been used as diagnostics tools for identifying individuals with ADDE [8–10]. The Schirmer test is an easy-to-use method; nevertheless, its 5-min duration is a disadvantage for most clinicians, generating discomfort due to the invasiveness of the strip compared to other tests [11, 12]. Simplifying clinical procedures by reducing the performance time may have benefits for practitioners and patients alike.

On the other hand, tear film parameters usually remain stable over time and between eyes in healthy individuals. However, in ocular surface diseases such as DED, disruptions to tear film homeostasis lead to increasing changes in the diagnostic values over time or between eyes [13]. This variability offers valuable insights into the clinical settings. Lower disparities may indicate temporary effects within compensatory mechanisms [14, 15]. Studying inter-eye variations is becoming a significant field in scientific research, potentially streamlining diagnostic processes and reducing associated costs. While prior research has explored inter-eye differences in primary diagnostic indicators such as osmolarity, the potential of the Schirmer test for the diagnosis of DED or even for the detection of its subtype has not been investigated to date [1, 14–19].

The present study aimed to (a) analyze and compare the values of the Schirmer test over time and between eyes of healthy volunteers and individuals with ADDE as well as to (b) calculate the diagnostic cut-off criteria of both study parameters for each time point.

Material and methods

In accordance with the recommendations of the TFOS DEWS II Diagnostic Methodology subcommittee, first a series of tests were conducted to differentiate between participants with DED and healthy individuals: assessment of DED symptoms, fluorescein tear film break-up time (FBUT), and corneal and conjunctival staining [1, 2, 20]. The diagnostic thresholds used to identify DED were Ocular Surface Disease Index (OSDI) ≥13, FBUT <10 s, and corneal staining (Oxford grade) >1 [1, 2, 20]. Additionally, TMH was measured to rule in or to rule out the presence of ADDE, with a diagnostic cut-off set at TMH ≥0.20 mm [2, 20, 21]. Participants were included and classified as having ADDE if they failed to fulfil all the inclusion criteria in both eyes and were assigned to the healthy group if they passed the cut-off inclusion criteria in all the tests in both eyes.

Once volunteers were recruited and grouped based on the inclusion criteria, qualified participants were scheduled for a single session where the Schirmer test was performed. Throughout the study, laboratory conditions of temperature, light, and humidity were kept constant (temperature 20–23 °C, humidity 50–60%).

Sample size and inclusion criteria

We used PS: Power and Sample Size Calculation, Version 3.1.2 (William D. Dupont and Walton D.

Plummer), for the sample size calculation. The standard deviation (SD) reported in the literature for the Schirmer test was assumed to be 3.93 [1, 22]. To have 80% power for a significance level of $\alpha = 0.05$ (type I error associated) with a confidence level of 95% to detect minimal clinical difference of 5 mm between healthy and pathological participants, the minimum number of participants required in each group was 11; moreover, the sample size calculated was in concordance with the sample size proposed by the TFOS for study designs where Schirmer values were assessed [1]. Adjusting for a 10% loss to follow-up, a minimal sample size of 13 participants per group was calculated. Finally, the study design was adjusted to include 20 participants per group in order to reduce the margin of error from 2.14 to 1.72 mm, enhancing the precision of the results. This decision was made to improve the scientific validity and robustness of the findings, ensuring a closer representation of the population's true characteristics.

Two similar-sized groups of 20 qualified volunteer participants were recruited from the optometry clinic of the center: one group of healthy participants and another of individuals with ADDE, according to the study's inclusion criteria [23]. Participants were recruited for each group based on the TFOS DEWS II Diagnostic Methodology subcommittee recommendations. Characteristics of the sample are reported in Table 1. No participant was under any type of topical and systemic medications or used artificial tears at the time of the study. Informed consent was obtained from all participants included in the study. The study protocol adhered to the tenets of the Declaration of Helsinki and was approved by the institutional ethics committee of the university.

Dry eye test battery procedures for inclusion

The symptomatic status of the patients was assessed using the OSDI questionnaire, a 12-item self-administered questionnaire designed for rapid assessment of ocular surface symptoms related to chronic dry eye disease, severity, and their effects on the patient [24, 25]. Total OSDI questionnaire scores were calculated according to published guidelines [24, 25].

The TMH was monitored, recorded, and semiautomatically quantified following a previous protocol [19, 21, 23]. Subjects were positioned by the slit-lamp and instructed to look at a target located to maintain primary eye gaze while the lower tear meniscus was observed through a Topcon SL-D4 (Topcon Corporation, Japan) biomicroscope with natural blinks. The videos were recorded by the Topcon DC-4 digital camera (Topcon Corporation) attached to the illumination system. Then, videos were stored on a computer connected to the slit-lamp microscope. To avoid reflex tearing, a short light beam (3 mm wide and 5 mm height) with moderate illumination was used to prevent the light from shining directly into the pupil during measurements [19, 21, 23]. The central meniscus was captured at the 6 o'clock position; in all cases, the observation and illumination systems were always set at 0° and without tilt of the illumination column. Tear meniscus images extracted from recorded videos were then measured by a second masked observer with computer-assistance image analysis using a Java-based open-source image processing software (ImageJ software v1.53i; National Institutes of Health, Bethesda, MD; http://imagej.nih.gov/ij/; [19, 21, 23, 26]).

After 5–10 min of the meniscus recording, a 2-µL volume of non-preserved 2% sodium fluorescein was instilled into the conjunctival sac with a micro-pipette [16, 27, 28]. Individuals were instructed to blink three or four times naturally, without squeezing, to evenly distribute the fluorescein over the cornea [16, 27, 28]. Within 30 s of instillation, the eye was examined with the SL-D4 slit-lamp biomicroscope set at ×16 magnification using a cobalt blue light and a Wratten 12 yellow filter [28]. After asking individuals to blink three times, the fluorescent tear film was video recorded on the DC-4 camera attached to the slit-lamp while the recording was stored on a computer. Participants were requested to keep the experimental eye open for as long as possible [29, 30]. The process was repeated three times. A fourth video of the whole anterior ocular surface was then videotaped to evaluate ocular surface staining. The upper eyelid was lifted slightly to grade the whole surface when the situation required it. A second masked observer evaluated the videos. A masked observer quantified the FBUT (defined as the time from the last blink to the first dark spot/line) in frames from the video recordings obtained using VirtualDub Software v1.10.4 (GNU General Public License, https://www.virtualdub.org/); frame results were then converted into seconds (15 frames = 1 s). The FBUT was an average as only the two most similar measurements were used to reduce variability [16, 31]. Corneal and conjunctival staining were categorized following the Oxford grading scheme that

Table 1 Descriptive statistics of healthy participants and patients with ADDE for the different inclusion criteria tests[a]

	Age (years) Mean ± SD	OSDI (score) median (IQR)	FBUT (s) Mean ± SD	Ocular surface staining (Oxford Scheme) Median (IQR)	TMH (mm) Mean ± SD
Healthy	24.0 ± 9.58	10.42 (8.33–12.00)	17.0 ± 9.40	0 (0–0)	0.22 ± 0.04
ADDE	35.3 ± 15.13	25.00 (15.23–35.09)	5.3 ± 2.68	1 (1–2)	0.10 ± 0.02

OSDI Ocular Surface Disease Index, *FBUT* fluorescein tear break-up time, *TMH* tear meniscus height, *IQR* interquartile range, *SD* standard deviation, *ADDE* aqueous-deficient dry eye
[a]Corneal staining data were classified based on the Oxford Scale

corresponds to 0–5 different levels of ocular surface damage [16, 32].

Schirmer test procedure

Participants were asked to be seated and instructed to look slightly up and keep both eyes open. The Schirmer test was performed by folding a Tear Strips Whatman 41 paper (Contactcare Ophthalmics and Diagnostics, 5 × 35 mm) at the notch, and hooking the folded end over the temporal one-third of the lower lid margin in both eyes by two investigators simultaneously. During the test, participants were instructed to fixate their sight on a straight target, refrain from blinking, and avoid horizontal eye movements that might cause corneal irritation and consequently tear hypersecretion. From the start until the end of the test, a digital picture from each strip without a light flash or other external illumination was taken every minute by two examiners at the same time to establish the different time point references (minutes 1, 2, 3, 4, and 5) in both eyes. Once the 5-min period ended, the strips were removed. To mask the data, measurements were associated with an alphanumeric code for randomized analysis by another observer not involved in the measurement process.

Another masked observer not involved in any steps of the clinical sessions measured the images by using a Java-based open-source image processing software (ImageJ software, [26]). The images of the database were presented to the masked observer in a randomized order. Since the length of the strip is a known parameter (35 mm) and the beginning of the strip where the notch was placed is difficult to capture, the following two parameters were used to obtain the results of the test: the non-wetted area and 1 mm of the strip as reference for the pixel conversion. The length of those parameters was marked with the *straight* tool, which allows the user to set a line with a free size and position. The line was chosen near the middle of the strip, perpendicular to the border, from one limit of the parameter to another. Then, the length was calculated using the command *Analyse* > *Measure* with which the software gives the parameter length.

Statistical analysis

We used SPSS statistical software v. 25.0 for Windows (SPSS Inc., Chicago, IL, USA) for data analyses. Significance was set at a $p \leq 0.05$ for all the analyses. Before the analysis, the normal distribution of the Schirmer results was checked using the Kolmogorov–Smirnov test; all time points showed a normal distribution (all $p > 0.05$). Hence, parametric tests were used. An ANOVA for repeated measures was performed, with Mauchly's W test used to assess the sphericity assumption (the assumption that the variances of the differences between all possible pairs of conditions are equal; [33, 34]). In cases where sphericity was violated ($p \leq 0.05$), the Greenhouse–Geisser or Huynh–Feldt correction was applied, adjusting the degrees of freedom based on the departure from sphericity (represented by epsilon, ε) to control for potential type I error inflation [33, 35]. Once general differences were identified, a Sidak test was used to detect significant pairwise differences [33]. Additionally, an unpaired t test was used at each time point to compare differences between results obtained for healthy volunteers and participants with ADDE [33].

Schirmer inter-eye differences were calculated as the absolute difference between values obtained from both eyes (|OD-OS|) of participants [14, 18]. Since the absolute inter-eye difference in the Schirmer test showed a non-normal distribution (Shapiro–Wilk test: $p < 0.05$), differences in this parameter between time points were assessed using the Friedman test, while the Wilcoxon test was used to detect significant pairwise differences; to avoid type I errors arising from multiple comparisons in the |OD-OS| analysis, statistical significance for the Wilcoxon test was divided by the number of comparisons performed to give a value of $p \leq 0.005$ [33]. Additionally, the Mann–Whitney U test was used at each time point to compare differences between results obtained in healthy volunteers and participants with ADDE.

The optimal cut-off value of the studied parameters to distinguish between healthy volunteers and participants with ADDE was estimated with the receiver operating characteristics (ROC) procedure [8, 36, 37]. Each theoretical cut-off value (from the lowest to the highest value observed in the study population) was used to estimate the sensitivity and specificity of the test. Results were then graphed with the sensitivity as a function of (1-specificity). area under the curve (AUC), and 95% confidence intervals (CIs) were also provided. This method of representing the relationship between the putative cut-off values and the effectiveness of the test provides a convenient way for selecting the threshold that finally provides the best combination between sensitivity and specificity (the optimal cut-off value is usually chosen as the hinge point of the curve).

Results

Schirmer values between time points

A statistical difference in the wetted length of the Schirmer test value between all time points was found for both groups, i.e., healthy volunteers (Mauchly's W: $p < 0.001$, $\varepsilon = 0.444$; Greenhouse–Geisser, $p < 0.001$, Table 2) and ADDE participants (Mauchly's W: $p < 0.001$, $\varepsilon = 0.383$; Greenhouse–Geisser, $p < 0.001$, Table 2); paired analysis also showed that there was a statistical difference in the Schirmer value for any pairwise analysis between time points (Sidak test, all $p < 0.001$). When Schirmer test results for each time point were compared between groups, there was

Table 2 Descriptive statistics and analysis of differences between Schirmer test results obtained for each time point for each group

		Healthy ($n=40$) Mean ± SD	ADDE ($n=40$) Mean ± SD	Healthy vs. ADDE (unpaired *t* test)
				p
Schirmer value (mm)	Minute 1	16.84 ± 8.30	3.13 ± 2.26	<0.001
	Minute 2	23.17 ± 8.54	4.80 ± 2.66	<0.001
	Minute 3	26.96 ± 7.99	6.56 ± 3.14	<0.001
	Minute 4	29.08 ± 7.55	8.35 ± 3.98	<0.001
	Minute 5	30.81 ± 7.15	9.75 ± 4.47	<0.001
Time point differences (Greenhouse–Geisser)	*p*	<0.001	<0.001	–

All values in mm
SD standard deviation, *ADDE* aqueous-deficient dry eye

Table 3 ROC procedure and cut-off criteria of Schirmer test value (mm) obtained for each time point studied

		AUC	SD	*p*	95% CI		Youden's J statistic	Cut-off criteria		
					Lower	Upper		Value	Sensitivity (in %)	Specificity (in %)
Schirmer value (mm)	Minute 1	0.959	0.019	<0.001	0.921	0.997	0.825	7.68	85	97.5
	Minute 2	0.980	0.012	<0.001	0.957	1.000	0.875	10.85	87.5	100
	Minute 3	0.988	0.008	<0.001	0.973	1.000	0.825	13.13	92.5	95
	Minute 4	0.978	0.014	<0.001	0.951	1.000	0.825	13.53	90	95
	Minute 5	0.976	0.013	<0.001	0.950	1.000	0.825	14.81	87.5	97.5

AUC area under the curve, *SD* standard deviation, *ROC* receiver operating characteristics

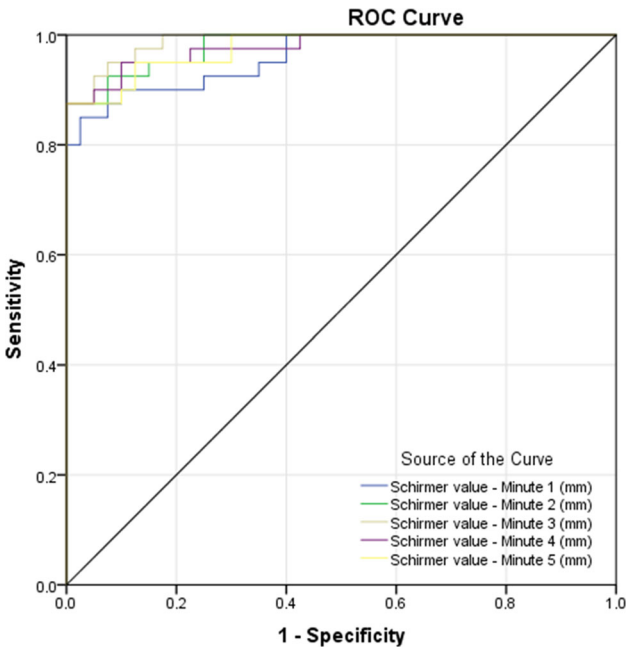

Fig. 1 Receiver operating characteristic (*ROC*) curve showing the relationship between sensitivity and specificity of the Schirmer test value (healthy vs. aqueous-deficient dry eye [ADDE]) according to theoretical thresholds (systematically chosen cut-off points); for each of the values observed in the study population (from the lowest to the highest in either the healthy or the ADDE group), the sensitivity and sensibility indexes have been calculated and reported in the graph; $n=80$

a statistical difference in all the unpaired analyses between time points (unpaired *t* test, all $p \leq 0.001$, Table 2). According to the ROC procedure, the optimal cut-off value is usually chosen as the hinge point of the curve between sensitivity and specificity for each of the observed values in the study population. The ROC procedures showed that the Schirmer value has a diagnostic capability to differentiate between participants' categories (Table 3). By calculating the Youden's index, we found the cut-off value for the Schirmer value on each time point to distinguish healthy volunteers from participants with ADDE (all $p<0.001$). The analysis is graphically represented in Fig. 1.

Absolute inter-eye differences in Schirmer test results between time points

For both participant groups, healthy and ADDE, no difference was found in the wetted length of the Schirmer |OD-OS| value between all time points (Friedman, both $p \geq 0.090$, Table 4); paired analysis also showed that there were no differences in the Schirmer test results |OD-OS| on any pairwise analysis between time points (Wilcoxon test, all $p \geq 0.062$) for both groups, except for the first time point in the ADDE participants versus the fourth and the fifth time points (Wilcoxon test, both $p \leq 0.030$). When Schirmer test |OD-OS| results on each time point were compared between groups, there was no statistical difference in the fourth and final time points (Mann–Whitney *U*, all $p \geq 0.414$, Table 4) while statistical differences were found between groups in the other three initial time points (Mann–Whitney *U*, both $p \leq 0.020$, Table 4). The ROC procedures showed that the Schirmer |OD-OS| value has a diagnostic capability to differentiate between participants' categories on the first-, second- and third-minute time points (all $p \leq 0.020$, Table 5). By calculating Youden's index,

Table 4 Descriptive statistics and analysis of differences between Schirmer |OD-OS| results obtained for each time point for each group

		Healthy ($n=20$) Mean ± SD	ADDE ($n=20$) Mean ± SD	Healthy vs. ADDE (unpaired t test)
				p
Schirmer |OD-OS| value (mm)	Minute 1	7.28 ± 5.36	2.23 ± 1.17	0.004
	Minute 2	8.21 ± 6.29	2.64 ± 1.84	0.004
	Minute 3	7.95 ± 6.42	3.11 ± 2.17	0.020
	Minute 4	7.06 ± 7.17	3.79 ± 2.23	0.429
	Minute 5	6.74 ± 7.79	3.55 ± 2.46	0.414
Time point differences (Friedman)	*P*	0.090	0.092	–

All values in mm
SD standard deviation, */OD-OS/* absolute inter-eye difference, *ADDE* aqueous-deficient dry eye

Table 5 ROC procedure and cut-off criteria of Schirmer |OD-OS| value (mm) obtained for each time point studied

		AUC	SD	p	95% CI		Youden's J statistic	Cut-off criteria		
					Lower	Upper		Value	Sensitivity (in %)	Specificity (in %)
Schirmer |OD-OS| value (mm)	Minute 1	0.766	0.080	0.004	0.609	0.924	0.550	5.28	55	100
	Minute 2	0.763	0.089	0.005	0.588	0.937	0.600	4.94	70	90
	Minute 3	0.715	0.090	0.020	0.538	0.892	0.500	8.43	50	100
	Minute 4	0.575	0.095	0.417	0.388	0.762	0.300	8.82	30	100
	Minute 5	0.423	0.097	0.402	0.233	0.612	0.825	11.66	25	100

AUC area under the curve, *SD* standard deviation, *ROC* receiver operating characteristics, */OD-OS/* absolute inter-eye difference

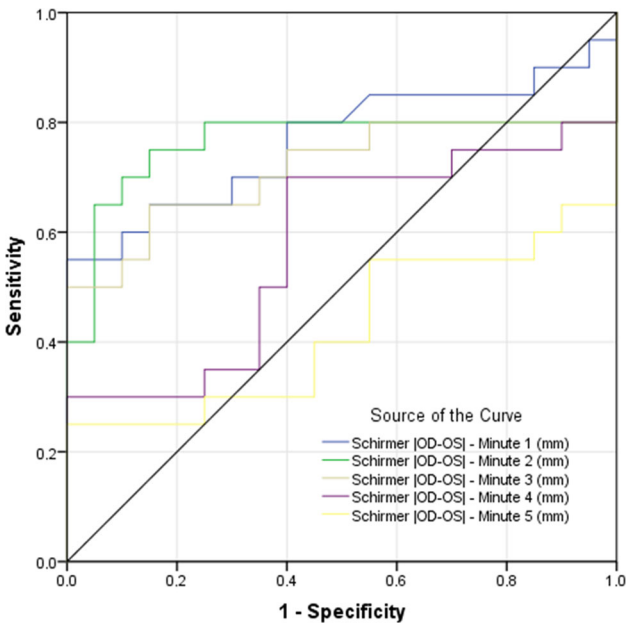

Fig. 2 Receiver operating characteristic (*ROC*) curve showing the relationship between sensitivity and specificity of the Schirmer |OD-OS| value (healthy vs. aqueous-deficient dry eye [ADDE]) according to theoretical thresholds (systematically chosen cut-off points); for each of the values observed in the study population (from the lowest to the highest in either the healthy or the ADDE group), the sensitivity and sensibility indexes have been calculated and reported in the graph; $n=40$. |*OD-OS*| absolute inter-eye difference

we found the cut-off value for the Schirmer |OD-OS| value on each time point to distinguish healthy volunteers from participants with ADDE. The analysis is graphically represented in Fig. 2.

Discussion

The main disadvantage of the Schirmer test, a 5-min examination, is its prolonged duration; it is common for individuals to express an unpleasant experience caused by the discomfort generated by the strip during the procedure. Consequently, some clinical research is needed to seek the possibility of minimizing the duration of the test, as well as the need to perform the test in both eyes of each participant. The present study aimed to analyze and compare the values of the Schirmer test over time and between the eyes of healthy volunteers and participants with ADDE.

According to the DEWS II classification report, there are two main DED types: ADDE and evaporative dry eye [2]. It is important to differentiate between the two DED types due to the subsequent management that will be different for each. While the EDE subtype is usually related to lipid layer alterations (therefore, diagnosed based on meibomian gland malfunctions), the ADDE subtype is related to the aqueous phase of the tear (discriminated based on a test that evaluates total tear volume). Tears are secreted and subsequently distributed over the ocular surface during the process of blinking, departing from the ocular surface via evaporation and drainage. Tear meniscus evaluation offers a noninvasive indication of the total volume since it is related to the total tear volume [4–7]. It has been estimated that the tear meniscus holds 75–90% of the total tear volume [4]; therefore, the evaluation of its characteristics has been revealed as a useful tool [38–42]. The TMH has been proposed in the DED diagnostic test battery of the DEWS II as a criterion to detect the presence and estimate the severity of the ADDE subtype [20].

One strength of the present study is the use of strict criteria to differentiate between healthy volunteers and patients with DED based on the diagnostic test battery of the DEWS II (higher symptomatology, lower stability, and damage presence) and to ensure focus on individuals with ADDE (lower tear volume).

There was a statistically significant difference in the Schirmer values between all time points of the test in both groups. When the results were compared between groups point by point, a difference was found even from the first time point. These results may suggest that the 5-min test could be shortened by establishing adequate cut-off criteria to discriminate ADDE individuals; previous authors made a similar suggestion [12, 43]. In the present study, these criteria were calculated and reported.

It is important to note that in the first minute of the test, the wetting of the filter strip was conducted at an accelerated rate, but in the next 4 min, it seemed to decrease significantly (Table 2). Previous reports have hypothesized that this phenomenon may occur because of the presence of a series of processes [9–12, 44]. First, there is a natural initial accumulation of tears in the fornix previous to the strip insertion; this is probably why the wetting of the filter strip occurred at an accelerated rate for the first minute and then slowed down afterward [9, 10, 12]. Second, the sensory receptor adaptation input of the central nervous system becomes reduced over time to prevent sensory overload and therefore the sensation is perceived as less intense; this may mean that the rapid wetting of the first minute could occur mainly by reflex tearing [12, 44]. Third, the evaporation of the aqueous phase from the strips increases over time, since a higher wetting surface allows for a physically higher evaporation rate; there is another inner influence on this variable, for instance, comparisons between different testing sites or between different regions of the world could be affected by environmental differences [9, 12]. And finally, it has been hypothesized that there is a negative hydrostatic pressure within the tear meniscus over the wetting part of the filter paper strip that slows the wetting process over time because of this pressure [11, 12]. The fact that no differences were found in the present analyses between healthy volunteers and participants with ADDE regarding the first time point of the test could be a significant limitation to reducing the clinical performance time of the test to only 1 min. However, the present study also emphasizes that the specific cut-off values identified during later time points offer a more reliable and practical approach for diagnosing ADDE. It indicates that accurate differentiation between patients with ADDE and healthy participants can be achieved as early as the second minute of the Schirmer test. Using this early time point, the test duration can be shortened without compromising the diagnostic accuracy, providing a more comfortable and efficient method for distinguishing patients with ADDE from healthy individuals.

In the second part of the analyses, it was found that there is no variation in the absolute Schirmer inter-eye difference between time points, neither in patients with ADDE nor in healthy participants; besides, there is no difference in this value during the first, the fourth, or the final time point between both groups. In this sense, previous reports also found no significant difference in test scores between eyes in the Schirmer test [12] and in other DED diagnostics tests [16, 17]. The variability or increasing variation in parameters is a statistical characteristic of the cause of heteroscedasticity in DED patients. The heteroscedasticity is considered a clinical indication of the loss of tear film homeostasis that occurs with dry eye [16, 18, 45]; indeed, inter-eye variability is a hallmark and a clinical target of DED. It has been suggested that the higher value of the two eyes obtained on a test should be used in clinical practice during dry eye management, while the lower value seems to reflect the transient effects of compensatory mechanisms [1, 14, 15, 18]. The present study results open the door to the possibility of reducing the test to only one eye when both eyes show a similar severity status, especially if the test is performed for its full 5-min period. It is proposed in the literature that the Schirmer test measures the reflex tearing in response to irritation of the conjunctival surface by the inserted Schirmer strip [11]; this may be the origin of the individual's discomfort. Performing the test with only one eye, as well as establishing a shorter period for the test, may reduce the unpleasant experience for the participants and improve the time spent in a clinic session. These cut-off values may also support the option of conducting the Schirmer test on only one eye under specific conditions, especially when both eyes show similar severity. The present study found no significant differences between the eyes during the test, which suggests that performing the Schirmer test on just one eye may be a valid approach in certain cases. However, this should only be considered when it is clear that both eyes exhibit similar severity, as determined by the inclusion criteria and diagnostic classification used in the present study, which required both eyes to meet specific conditions. This approach cannot be verified in other patient populations. While this adjustment could reduce discomfort for the patient, streamline the testing process, and save valuable clinical time, it is important to note that this remains a proposal that requires further investigation. This approach offers a useful starting point for future research aimed at optimizing testing protocols without compromising diagnostic accuracy. By shortening the test duration and focusing on one eye, clinicians can improve the overall patient experience while still effectively diagnosing ADDE.

It is important to note that the Schirmer test itself has some limitations. On the traditional Schirmer test,

the reflex tearing occurs throughout the strip, therefore the reflex tear is not reflective of the basal tear produced by individuals in their everyday life [46]. In addition, participants are requested to refrain from blinking for as long as possible during the tests; when the eyes are open, tear evaporates to the environment and, therefore, the Schirmer wetted length is influenced by both the lack of production and the evaporation rate, which is one of the limitations of the traditional Schirmer test when trying to study ADDE. Although in the present study, patients with ADDE were selected under specific criteria for the dry eye cohort, every individual has a different evaporation rate, and therefore this may confound results. In future studies, a similar study design but with the eyes closed could eliminate the evaporation of the tear on the ocular surface.

Conclusion

In summary, the present study provides a complete report of the progression of Schirmer test values over time and between eyes of healthy volunteers and patients with ADDE; it also established the cut-off criteria for the different time points for both Schirmer and Schirmer |OD-OS| values. The findings suggest that reducing the Schirmer test to 2 min and performing it on only one eye when both eyes exhibit similar severity could be a practical approach in the population studied. However, given the limited sample size, these findings should be considered to be preliminary. Future research with larger samples is necessary to validate the reported cut-off values and establish robust diagnostic criteria applicable across different populations and time points. Sample size is too small to draw this general conclusion, and the authors encourage reserachers to conduct a study with a larger sample.

Funding Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature.

Declarations

Conflict of interest H. Pena-Verdeal, J. Garcia-Queiruga, B. Sabucedo-Villamarin, C. García-Resúa, M.J. Giraldez and E. Yebra-Pimentel declare that they have no competing interests.

Ethical standards Informed consent was obtained from all patients being included in the study. The study protocol adhered to the tenets of the Declaration of Helsinki and was approved by the institutional Ethics Committee of the university.

References

1. Wolffsohn JS, Arita R, Chalmers R, Djalilian A, Dogru M, Dumbleton K, Gupta PK, Karpecki P, Lazreg S, Pult H, et al. TFOS DEWS II Diagnostic Methodology report. Ocul Surf. 2017;15:539–74. https://doi.org/10.1016/j.jtos.2017.05.001.
2. Craig JP, Nichols KK, Akpek EK, Caffery B, Dua HS, Joo CK, Liu Z, Nelson JD, Nichols JJ, Tsubota K, et al. TFOS DEWS II Definition and Classification Report. Ocul Surf. 2017;15:276–83. https://doi.org/10.1016/j.jtos.2017.05.008.
3. Stahl U, Willcox M, Stapleton F. Osmolality and tear film dynamics. Clin Exp Optom : J Aust Optom Assoc. 2012;95:3–11. https://doi.org/10.1111/j.1444-0938.2011.00634.x.
4. Holly FJ. Physical chemistry of the normal and disordered tear film. Trans Ophthalmol Soc U K. 1985;104((Pt4):374–80.
5. Li M, Du C, Zhu D, Shen M, Cui L, Wang J. Daytime variations of tear osmolarity and tear meniscus volume. Eye Contact Lens. 2012;38:282–7. https://doi.org/10.1097/ICL.0b013e31825fed57.
6. Bron AJ, Tomlinson A, Foulks GN, Pepose JS, Baudouin C, Geerling G, Nichols KK, Lemp MA. Rethinking dry eye disease: a perspective on clinical implications. Ocul Surf. 2014;12:1–31. https://doi.org/10.1016/j.jtos.2014.02.002.
7. Yokoi N, Bron AJ, Tiffany JM, Maruyama K, Komuro A, Kinoshita S. Relationship between tear volume and tear meniscus curvature. Arch Ophthalmol. 2004;122:1265–9. https://doi.org/10.1001/archopht.122.9.1265.
8. de Monchy I, Gendron G, Miceli C, Pogorzalek N, Mariette X, Labetoulle M. Combination of the Schirmer I and phenol red thread tests as a rescue strategy for diagnosis of ocular dryness associated with Sjogren's syndrome. Investig Ophthalmol Vis Sci. 2011;52:5167–73. https://doi.org/10.1167/iovs.10-6671.
9. Kim YH, Graham AD, Li W, Radke CJ, Lin MC. Human Lacrimal Production Rate and Wetted Length of Modified Schirmer's Tear Test Strips. Transl Vis Sci Technol. 2019;8:40. https://doi.org/10.1167/tvst.8.3.40.
10. Li S, Kim YH, Li W, Lin MC, Radke CJ. Human Lacrimal Production Rates from Modified Schirmer-Tear Test. Optom Vis Sci. 2018;95:343–8. https://doi.org/10.1097/OPX.0000000000001196.
11. Yokoi N, Kinoshita S, Bron AJ, Tiffany JM, Sugita J, Inatomi T. Tear meniscus changes during cotton thread and Schirmer testing. Investig Ophthalmol Vis Sci. 2000;41:3748–53.
12. Karampatakis V, Karamitsos A, Skriapa A, Pastiadis G. Comparison between normal values of 2- and 5-minute Schirmer test without anesthesia. Cornea. 2010;29:497–501. https://doi.org/10.1097/ICO.0b013e3181c2964c.
13. Bron AJ, de Paiva CS, Chauhan SK, Bonini S, Gabison EE, Jain S, Knop E, Markoulli M, Ogawa Y, Perez V, et al. TFOS DEWS II pathophysiology report. Ocul Surf. 2017;15:438–510. https://doi.org/10.1016/j.jtos.2017.05.011.

14. Lemp MA, Bron AJ, Baudouin C, Benitez Del Castillo JM, Geffen D, Tauber J, Foulks GN, Pepose JS, Sullivan BD. Tear osmolarity in the diagnosis and management of dry eye disease. Am J Ophthalmol. 2011;151(e791):792–8. https://doi.org/10.1016/j.ajo.2010.10.032.
15. Jacobi C, Jacobi A, Kruse FE, Cursiefen C. Tear film osmolarity measurements in dry eye disease using electrical impedance technology. Cornea. 2011;30:1289–92. https://doi.org/10.1097/ICO.0b013e31821de383.
16. Pena-Verdeal H, Garcia-Queiruga J, Sabucedo-Villamarin B, Giraldez MJ, Garcia-Resua C, Yebra-Pimentel E. Capability of the Inter-Eye Differences in Osmolarity, Break-Up Time and Corneal Staining on the Diagnostic of Dry Eye. Ocul Immunol Inflamm. 2023; https://doi.org/10.1080/09273948.2023.2285449.
17. Nilsen C, Graae Jensen P, Gundersen M, Utheim OA, Gjerdrum B, Gundersen KG, Jahanlu D, Potvin R. The Significance of Inter-Eye Osmolarity Difference in Dry Eye Diagnostics. Clin Ophthalmol. 2023;17:829–35. https://doi.org/10.2147/OPTH.S402556.
18. Pena-Verdeal H, Garcia-Resua C, Vazquez-Sanchez C, Garcia-Queiruga J, Giraldez MJ, Yebra-Pimentel E. Inter-eye osmolarity differences in patients with symptomatic and non-symptomatic dry eyes. ABO. 2020;83:103–8. https://doi.org/10.5935/0004-2749.20200024.
19. Pena-Verdeal H, Garcia-Queiruga J, Sabucedo-Villamarin B, Garcia-Resua C, Giraldez MJ, Yebra-Pimentel E. A Comprehensive Study on Tear Meniscus Height Inter-Eye Differences in Aqueous Deficient Dry Eye Diagnosis. J Clin Med. 2024;https://doi.org/10.3390/jcm13030659.
20. Wolffsohn JS, Wang MTM, Vidal-Rohr M, Menduni F, Dhallu S, Ipek T, Acar D, Recchioni A, France A, Kingsnorth A, et al. Demographic and lifestyle risk factors of dry eye disease subtypes: A cross-sectional study. Ocul Surf. 2021;21:58–63. https://doi.org/10.1016/j.jtos.2021.05.001.
21. Sabucedo-Villamarin B, Pena-Verdeal H, Garcia-Queiruga J, Giraldez MJ, Garcia-Resua C, Yebra-Pimentel E. Categorization of the Aqueous Deficient Dry Eye by a Cut-Off Criterion of TMH Measured with Tearscope. Life. 2022; https://doi.org/10.3390/life12122007.
22. Nichols KK, Mitchell GL, Zadnik K. The repeatability of clinical measurements of dry eye. Cornea. 2004;23:272–85.
23. Pena-Verdeal H, Garcia-Resua C, Barreira N, Giraldez MJ, Yebra-Pimentel E. Interobserver variability of an open-source software for tear meniscus height measurement. Contact Lens Anterior Eye : J Br Contact Lens Assoc. 2016;39:249–56. https://doi.org/10.1016/j.clae.2016.02.009.
24. Miller KL, Walt JG, Mink DR, Satram-Hoang S, Wilson SE, Perry HD, Asbell PA, Pflugfelder SC. Minimal clinically important difference for the ocular surface disease index. Arch Ophthalmol. 2010;128:94–101. https://doi.org/10.1001/archophthalmol.2009.356.
25. Schiffman RM, Christianson MD, Jacobsen G, Hirsch JD, Reis BL. Reliability and validity of the Ocular Surface Disease Index. Arch Ophthalmol. 2000;118:615–21.
26. Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. Nat Methods. 2012;9:671–5.
27. Abelson MB, Ousler GW 3rd, 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;506:1121–5.
28. Johnson ME, Murphy PJ. The Effect of instilled fluorescein solution volume on the values and repeatability of TBUT measurements. Cornea. 2005;24:811–7.
29. Liu H, Begley CG, Chalmers R, Wilson G, Srinivas SP, Wilkinson JA. Temporal progression and spatial repeatability of tear breakup. Optom Vis Sci. 2006;83:723–30. https://doi.org/10.1097/01.opx.0000237546.88464.6d.
30. Begley CG, Himebaugh N, Renner D, Liu H, Chalmers R, Simpson T, Varikooty J. Tear breakup dynamics: a technique for quantifying tear film instability. Optom Vis Sci. 2006;83:15–21. https://doi.org/10.1097/01.opx.0000195569.36185.fd.
31. Cho P, Brown 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;70:30–8.
32. Bron AJ, Evans VE, Smith JA. Grading of corneal and conjunctival staining in the context of other dry eye tests. Cornea. 2003;22:640–50.
33. Armstrong RA, Davies LN, Dunne MC, Gilmartin B. Statistical guidelines for clinical studies of human vision. Ophthalmic Physiol Opt. 2011;31:123–36. https://doi.org/10.1111/j.1475-1313.2010.00815.x.
34. Armstrong RA. Recommendations for analysis of repeated-measures designs: testing and correcting for sphericity and use of manova and mixed model analysis. Ophthalmic Physiol Opt. 2017;37:585–93. https://doi.org/10.1111/opo.12399.
35. Armstrong RA. When to use the Bonferroni correction. Ophthalmic Physiol Opt. 2014;34:502–8. https://doi.org/10.1111/opo.12131.
36. Hajian-Tilaki K. Receiver Operating Characteristic (ROC) Curve Analysis for Medical Diagnostic Test Evaluation. Caspian J Intern Med. 2013;4:627–35.
37. Nahm FS. Receiver operating characteristic curve: overview and practical use for clinicians. Korean J Anesthesiol. 2022;75:25–36. https://doi.org/10.4097/kja.21209.
38. Patel S, Wallace I. Tear meniscus height, lower punctum lacrimale, and the tear lipid layer in normal aging. Optom Vis Sci. 2006;83:731–9.
39. Johnson ME, Murphy PJ. The agreement and repeatability of tear meniscus height measurement methods. Optom Vis Sci. 2005;82:1030–7.
40. Santodomingo-Rubido J, Wolffsohn JS, Gilmartin B. Comparison between graticule and image capture assessment of lower tear film meniscus height. Contact Lens Anterior Eye : J Br Contact Lens Assoc. 2006;29:169–73. https://doi.org/10.1016/j.clae.2006.05.004.
41. Garcia-Resua C, Santodomingo-Rubido J, Lira M, Giraldez MJ, Vilar EY. Clinical assessment of the lower tear meniscus height. Ophthalmic Physiol Opt : J Br Coll Ophthalmic Opt. 2009;29:487–96. https://doi.org/10.1111/j.1475-1313.2009.00634.x.
42. Pult H, Purslow C, Murphy PJ. The relationship between clinical signs and dry eye symptoms. Eye (Lond). 2011;25:502–10. https://doi.org/10.1038/eye.2010.228.
43. Bawazeer AM, Hodge WG. One-minute schirmer test with anesthesia. Cornea. 2003;22:285–7. https://doi.org/10.1097/00003226-200305000-00001.
44. Clinch TE, Benedetto DA, Felberg NT, Laibson PR. Schirmer's test. A closer look. Arch Ophthalmol. 1983;101:1383–6. https://doi.org/10.1001/archopht.1983.01040020385009.
45. Sullivan B. Challenges in using signs and symptoms to evaluate new biomarkers of dry eye disease. Ocul Surf. 2014;12:2–9. https://doi.org/10.1016/j.jtos.2013.10.002.
46. Li N, Deng XG, He MF. Comparison of the Schirmer I test with and without topical anesthesia for diagnosing dry eye. Int J Ophthalmol. 2012;5:478–81. https://doi.org/10.3980/j.issn.2222-3959.2012.04.14.