Original Research

The development and evaluation of the new Ocular Surface Disease Index-6

Heiko Pult[a,b,c,*], James S. Wolffsohn[b]

[a] Dr. Heiko Pult - Optometry and Vision Research, Weinheim, Germany
[b] Ophthalmic Research Group, Life and Health Sciences, Aston University, Birmingham, United Kingdom
[c] School of Biomedical & Life Sciences, Cardiff University, Cardiff, United Kingdom

ARTICLE INFO

ABSTRACT

Aim: To evaluate if the OSDI can be shortened and the score calculation simplified without significantly impacting the outcome of the questionnaire.

Methodology: Study#1: 264 participants completed the OSDI questionnaire (174 females; mean age: 34.4 ± 12.3 yrs) and the results were analyzed to detect those questions of each subscale that are the most discriminative by multiple regression and RASCH analyses, resulting in the OSDI-6. Study#2: OSDI-6 was compared to the OSDI and the 5-item Dry Eye Questionnaire (DEQ-5) in 120 patients (73 females; mean age: 51 ± 20.2 yrs) to evaluate predictive ability of the OSDI-6. Repeatability was analyzed in 50 participants.

Results: Study#1: The mean OSDI score was 13.1 ± 11.5. The most discriminant questions were questions 1, 4, 7, 9, 10 and 11. Infits and outfits of the OSDI-6 were between 1.26 and 0.78 (STRATA = 3). The OSDI-6 was significantly correlated to the OSDI ($r = 0.898$, $p < 0.001$). Area under the curve (AUC) of the OSDI-6 to predict OSDI was 0.967 $p < 0.001$. Study#2: The mean OSDI score was 20.3 ± 16.6, the mean DEQ-5 score 7.9 ± 4.6 and the mean OSDI-6 score 10.3 ± 8.6. AUC of the OSDI-6 to predict OSDI was 0.901 ($p < 0.001$) and 0.803 ($p < 0.001$) to predict DEQ-5. The OSDI-6-V2 was significantly correlated to the OSDI and its sub-scales ($r > 0.842$, $p < 0.001$). Repeatability of the OSDI was 0.72 ± 0.11 (Kappa; $p < 0.001$), the DEQ-5 was 0.75 ± 0.06 ($p < 0.001$) and the OSDI-6 was 0.80 ± 0.05 ($p < 0.001$).

Conclusions: The OSDI-6 seems to be a repeatable questionnaire and a good alternative to the full original OSDI for use in clinical practice as well as research.

1. Introduction

Dry eye is one of the most frequent causes of patient visits to eye care practitioners [1]. The prevalence of dry eye diagnosed based on both symptoms and signs ranges in different studies from 9% to 30%, and based on symptoms alone from 7% to 52% [2]. Dry eye is a symptomatic disease characterized by a vicious cycle of tear film instability and hyperosmolarity, which leads to increased ocular surface inflammation, damage and neurosensory abnormalities [3]. Moderate to severe dry eye is associated with significant ocular discomfort, limitations in performing daily activities, reduced vitality, poor general health and often depression [3]. Although the relationship between symptoms and signs of dry eye varies across individuals and types of dry eye [4,5], the ability to accurately quantify ocular surface symptoms is an important screening tool to assist dry eye evaluation and diagnosis. It is also critical for monitoring the progression of the condition and response to treatments. Especially in clinical practice, symptoms are typically captured through the patient case history [6,7].

Consequently, symptom measurements are very similar to clinical signs of dry eye. Therefore, the use of validated symptom questionnaires is recommended.

The Ocular Surface Disease Index (OSDI) [8] may be the most widely used questionnaire for dry eye clinical trials and clinical practice [1]. The OSDI measures frequency of experiencing associated symptoms of dry eye, environmental triggers and vision-related quality of life. The OSDI includes 12 questions grouped by poor symptoms and visual disturbance (light sensitivity, eyes feel gritty, painful or sore eyes, blurred vision or poor vision); visual function/tasks (problems when reading, driving at night, working on a computer or watching TV); and environmental questions (problems in windy conditions, places/areas with low humidity or areas that are air conditioned). The final OSDI score is calculated by a formula (OSDI score = sum of scores x 25/ number of answered questions), accounting for the fact that questions 6 to 12 are optional.

In clinical practice, patients can struggle to understand differences between some questions, such as between blurred and poor vision or

differences between low humidity and air conditioning. Practitioners are often under time pressure, and therefore appreciate simple, fast and exact tools. As the OSDI can be used as a first screening tool in patient care, clinicians could benefit from a shorter version of the OSDI with a quick and easy scoring system.

The aim of this study was therefore to evaluate whether the OSDI can be shortened and the score calculation simplified without significantly impacting the outcome of the questionnaire.

2. Methodology

This project consisted of two studies: In the first study, the completed OSDI questionnaires of sequential clinic patients who were non-contact lens wearers were reviewed to detect those questions of each subscale which are the most discriminative. This resulted in the first version of the OSDI-6 questionnaire (OSDI-6-V1). In the second study, the OSDI-6-V1 was compared to the original OSDI and the five-item Dry Eye Questionnaire (DEQ-5) [9]. As this was an international study, an English version of the questionnaire was used at the UK site and a German version [10] was used at the German site. Participants gave written informed consent before participating in the study. Ethical approval was given by Aston University. All procedures were conducted in accordance with the Declaration of Helsinki (1983).

Participants were excluded from the study if they were younger than 18 years of age, had worn contact lenses in the past 3 months, were taking eye drops for an ocular condition other than dry eye, had a history of any eye surgery in the past year or reported other significant ocular pathology.

2.1. Study #1

A total of 264 participants completed the OSDI questionnaires provided by the two sites (Aston University, School of Optometry and Vision Sciences, Birmingham, UK and Horst Riede GmbH, Weinheim, Germany).

2.1.1. Statistical analyses

Data were tested for normality using the Shapiro Wilk test. As the residuals of the regression were normally distributed, the detection of the most predictive questions from each of the OSDI subscales was conducted using multiple, linear regression analyses (backwards; p-out = 0.05). Furthermore, infits and outfits of the single questionnaire were analyzed by RASCH analyses. Infit and outfit values should not be outside a range of 0.7–1.3 [11]. The resulting six questions (two from each subscale) were combined to form the new OSDI-6.

The ability of the OSDI-6-V1 to predict the OSDI diagnoses (cut-off value 13[1]) was analyzed by receiver operating characteristic curve (ROC). The optimal cut-off threshold was determined by the point whereby the Youden's index was maximized (Youden's index = sensitivity + specificity −1). Correlations between the OSDI and OSDI-6-V1 were analyzed by Spearman's correlation. The number of participants was defined by following the protocol of Dougherty et al. [11], who analyzed 172 participants. The post-hoc power calculation resulted in a power of 1-β = 1.0. The data were analyzed using SPSS 25.0 (SPSS Inc., Chicago, USA) and Winstep Version 3.63.0 (Winsteps.com, Beaverton, Oregon, USA) and BiAS 11.04 (Dr. H. Ackermann, Goethe University, Frankfurt, Germany).

2.2. Study #2

Based on the regression analyses formula of the contribution of the OSDI-6-V1 items to the overall OSDI score, the questions were weighted. To simplify scoring, this was done by counting the scores of the questions 1, 3 and 5 and adding a double weighting of the scores of questions 2, 4 and 6 (resulting in a total score between 0 and 36). Alternatively, the OSDI-6-V1 could be calculated by simply summating the score of each question, without weighting (resulting in a total score between 0 and 24). However, this needed to be analyzed in terms of predictive ability and repeatability and was tested in Study #2 in a new patient cohort.

The questions of the original OSDI refer to the past week, which may be too short of a time period. For example, the DEQ-5 refers to symptoms experienced over the past month. Hence, the OSDI-6-V1 was reworded to also ask about symptoms of the past month in Study #2. Furthermore, the wording of the score descriptors was amended based on the wording of the DEQ-5 to read "constantly," "mostly," "often," "sometimes," and "never," instead of "all of the time," "most of the time," "half of the time," "some of the time" and "none of the time."

This reworded questionnaire was named OSDI-6-V2. Each of 120 participants (73 females; mean age: 51 ± 20.2 years) completed the three questionnaires, OSDI-6-V2, OSDI and DEQ-5. Of these, 50 randomly selected participants (23 females; mean age: 55.8 ± 17.7 years) completed them twice, separated by one day. The order of completion was randomized, and patients were not allowed to re-read previously completed questionnaires.

2.2.1. Statistical analyses

Data were tested for normality using the Shapiro Wilk test. All data were normal distributed, except from the OSDI (p = 0.045) and the OSDI-6-V2 (weighted version; p = 0.032). The ability of the OSDI-6-V2 (weighed and unweighted scoring) to predict the OSDI diagnoses (dry eye ≥ 13[1]) and the DEQ-5 (dry eye ≥ 6[1]) was analyzed by an ROC. Correlation between OSDI and OSDI-6-V2 were analyzed by Spearman's correlation. A prior power calculation resulted in the need for at least 82 participants (1-β = 0.80; ρ = 0.3). Repeatability of OSDI-6-V2 (weighted and un-weighted), OSDI and DEQ-5 were evaluated by Cohens's Kappa statistics, Bland-Altman analyses and the Intra Class Coefficient (ICC). A prior power calculation for this resulted in the need for 41 participants (1-β = 0.80; dz = 0.4).

3. Results

3.1. Study #1

The mean OSDI score of this cohort of population was 13.1 ± 11.5 (174 females, 90 males; mean age: 34.4 ± 12.3 years). Based on the regression analyses (using backwards selection), the most predictive two questions of each of the three subcategories of the OSDI were "eyes that are sensitive to light" (OSDI question number 1), "blurred vision" (OSDI question number 4), "driving at night" (OSDI question number 7), "watching TV" (OSDI question number 9), "windy conditions" (OSDI question number 10) and "places or areas with low humidity" (OSDI question number 11). The equations derived from the regression analyses were as follows:

OSDI - sub-category 1 (questions 1–5) = 1.12 x question 1 + 1.92 x question 4 + 0.43.

OSDI - sub-category 2 (questions 6–9) = 0.86 x question 7 + 2.05 x question 9 + 0.54.

OSDI - sub-category 3 (questions 10-11) = 0.99 x question 10 + 1.92 x question 11 + 0.13.

The resulting 6 questions (Q) were combined to form the OSDI-6-V1 (Table 1).

Table 1
Questions of the OSDI-6-V1 questionnaire.

Question 1	Eyes that are sensitive to light
Question 2	Blurred vision
Question 3	Driving at night
Question 4	Watching TV
Question 5	Windy conditions
Question 6	Places or areas with low humidity

The formula for determining the overall OSDI-6-V1 score was derived from linear weighting of each item based on the multivariate regression coefficients. The regression coefficients of the equation of the multiple linear regression analyses were rounded. Based on this, the OSDI-6-V1 total score was calculated with the following formula:

$$OSDI\text{-}6\text{-}V1 = 1 \times Q1 + 2 \times Q2 + 1 \times Q3 + 2 \times Q4 + 1 \times Q5 + 2 \times Q6$$

The OSDI-6-V1 was significantly correlated to the OSDI ($r = 0.898$, $p < 0.001$). Area under the curve (AUC) of the OSDI-6-V1 to predict participants with an OSDI score greater than the cut-off of 13 was 0.967 (95% confidence interval of AUC = 0.942; 0.992 and SD = 0.0127; $p < 0.001$; sensitivity = 0.905, specificity = 0.813; OSDI-6-V1 cut-off value of 3; positive likelihood ratio (LR +) = 4.84; negative LR (LR-) = 0.12).

RASCH analyses resulted in the same questions being included as determined by the regression analyses; infits and outfits were between 1.26 and 0.78 by a STRATA of 3 for the OSDI-6-V1 and between 1.68 and 0.57 by a STRATA of 4 for the OSDI.

3.2. Study #2

Ten of the patients did not answer at least one of the mandatory questions 1 to 5 of the OSDI, nine patients did not answer at least one of the questions of the OSDI-6-V2 and six patients did not answer at least one question of the DEQ-5. As the first six questions of the OSDI are mandatory and as the OSDI-6-V1 questionnaire does not allow questions to be skipped – due to the simplified scoring model - those questionnaires were not included in the statistical analyses. This resulted in 110 questionnaires (68 females; mean age: 52 ± 20.4 years) analyzed with 43 of them (19 females; mean age: 53.7 ± 18.2 years) included in the repeatability study.

The mean OSDI score was 20.3 ± 16.6, the mean DEQ-5 score was 7.9 ± 4.6, the mean OSDI-6-V2 (weighted) score was 23.3 ± 19.0 and the mean OSDI-6-V2 (un-weighted) score was 10.3 ± 8.6. ROC showed that the OSDI-6-V2 weighted and OSID-6-V2 un-weighted scores both were significant discriminators of the full OSDI or DEQ-5 diagnostic cut-offs (Table 2). However, the OSDI-6-V2 un-weighted did slightly better than the OSDI-6-V2 weighted with a 1.3% larger AUC (de Long paired test; $p = 0.007$) when predicting the OSDI diagnoses and 0.3% larger AUC (de Long paired test; $p = 0.012$) when predicting the DEQ-5. The ability of the OSDI-6-V2 to predict OSDI diagnoses was better than that of the DEQ-5 (Table 3). AUCs of the OSDI, OSDI-6-V2 weighted and OSDI-6-V2 un-weighted were similar when predicting the DEQ-5 diagnostic cut-offs.

The OSDI-6-V2 was significantly correlated with the OSDI ($r = 0.893$, $p < 0.001$) as well as each pair of OSDI-6-V2 questions to the related sub-scale of the original OSDI questionnaire (subscale 1: $r = 0.866$; subscale 2: $r = 0.842$; subscale 3: $r = 0.947$; all $p < 0.001$).

Bland-Altman analyses showed that all questionnaires were repeatable (Table 3, Figs. 1–2). Based on linearly weighted Kappa analyses, the original OSDI was moderately repeatable [12] (Kappa

Table 3
Bland-Altman analyses of the questionnaires showing p-value, negative and positive 95% limit of agreement (LoA), 95% coefficient of repeatability (CR) and intraclass correlation coefficient (ICC).

Questionnaire	p-value	95% LoA		95% CR	ICC
OSDI	0.691	− 10.6	10.0	10.3	0.971
OSDI-6-V2 (weighted)	0.328	− 4.5	5.2	4.8	0.952
OSDI-6-V2 (un-weighted)	0.475	− 2.9	3.3	3.1	0.960
DEQ-5	0.575	− 4.2	4.6	4.4	0.894

0.720 ± 0.110, $p < 0.001$), as were the DEQ-5 (0.753 ± 0.059, $p < 0.001$) and the OSDI-6-V2 (weighted) (0.745 ± 0.055, $p < 0.001$). However, the OSDI-6-V2 (un-weighted) showed a strong repeatability [12] (0.800 ± 0.053, $p < 0.001$). As the OSDI-6-V2 unweighted was significantly better in discriminating between subjective dry eye and normal than the OSDI-6-V2 un-weighted, and its repeatability was also better than of the OSDI-6-V2 weighted, the OSDI-6-V2 unweighted was ultimately used as the final questionnaire, named the Ocular Surface Disease Index 6 (OSDI-6) (Table 4).

4. Discussion

This study analyzed whether a simplified, shorter version of the full OSDI may be comparable to the OSDI in terms of predictive ability of dry eye symptoms and repeatability. The OSDI is composed of three subcategories reflecting dry eye symptoms, tasks and environmental impacts, and this concept was maintained in the OSDI-6. However, while the OSDI has a different numbers of questions per subcategory, the OSDI-6 contains just two questions per subcategory.

In the first part of the study, the most discriminative questions to be used in the OSID-6-V1 were analyzed by two different approaches; multiple regression analyses and RASCH analyses. Interestingly, both concepts resulted in the same set of questions. The RASCH analyses showed that the number of questions could significantly be reduced, which was also implicated by the infit and outfit results of a study of Simpson et al. [13] However, this is different from the results of Dougherty et al. [11], although in their study the five-category response structure of the OSDI was not used; the authors' combined "half of the time" and "most of the time" category thresholds of the OSDI five-category response structure were shown to be disordered in that study.

As the OSDI is one of the most frequently used questionnaires for dry eye symptomology assessment, based on a five-category response, this was used unchanged in our study as well as in Simpson et al. [13] Interestingly, many of the questions that were excluded in the OSDI-6 were those that had been noted to be more difficult for patients to conceptualize. Experiences in patient care show that many patients do not really understand the difference between "eyes that feel gritty" and "painful or sore eyes" or are able to differentiate between "blurred vision" and "poor vision." "Watching TV" or "computer work" also appear to be similar as well as "places or areas with low humidity (very dry)" or "areas that are air conditioned." The ability of the OSDI-6-V1
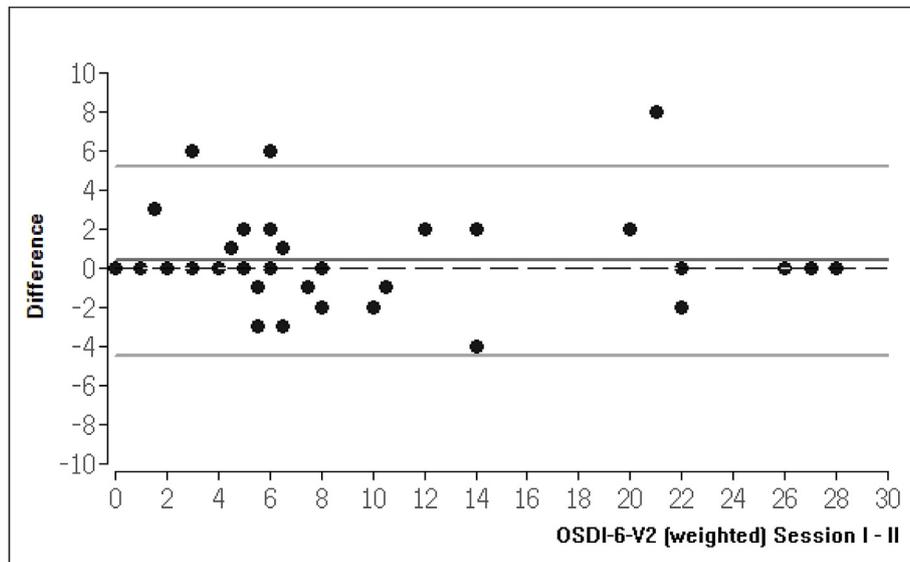
Table 2
Area under the ROC curve (AUC), 95% confidence interval (CI) of AUC, standard deviation (SD) of AUC, p-value (p), cut-off value and sensitivity and specificity and positive and negative likelihood ratios (LR +, LR-) of the different questionnaires. An alternative cut-off value for the OSDI-6-V2 weighted version is shown in square brackets.

Grouping Variable	Source of the Curve	AUC	95% CI AUC	SD AUC	P	Cut-Off Value	Sensitivity	Specificity	LR +	LR-
OSDI	OSDI-6-V2 weighted	0.888	0.828; 0.949	0.031	< 0.001	3	0.968	0.556	2.180	0.058
						(5)	(0.750)	(0.800)	(3.750)	(0.313)
	OSDI-6-V2 un-weighted	0.901	0.846; 0.957	0.028	< 0.001	4	0.703	0.867	5.286	0.342
	DEQ-5	0.824	0.747; 0.901	0.039	< 0.001	6	0.781	0.711	2.702	0.308
DEQ-5	OSDI-6-V2 weighted	0.827	0.752; 0.902	0.038	< 0.001	5	0.701	0.727	2.568	0.411
	OSDI-6-V2 un-weighted	0.830	0.757; 0.904	0.037	< 0.001	4	0.672	0.818	3.692	0.401
	OSDI	0.825	0.749; 0.901	0.039	< 0.001	13	0.788	0.721	2.824	0.294

Fig. 1. Bland-Altmann plot showing the agreement between session I and II of the OSDI-6-V2 (weighted; n = 43; identical values are not shown).

to predict OSDI outcome was very high in the first part of this study.

In the second part of the study, wording of questions and scores were updated based on some of the wording of the DEQ-5. Again, this needed to be tested against the OSDI and resulted once more in a high predictive ability of the OSDI-6-V2. In this study, patients were asked to complete a set of dry eye questionnaires, including the OSDI, OSDI-6-V2 and DEQ-5. Analyzing the predictive ability of each questionnaire against the others resulted in relatively similar, high predictive abilities, based on ROC analyses. In Study #1, the multiple regression analyses resulted in different coefficients that could be applied to weighting the OSDI-6-V1 questions. However, Study #2 did not confirm this more complex approach as being better in terms of predictive ability. Furthermore, repeatability of the OSDI-6-V2 improved when not using those coefficients.

Consequently, the OSDI-6-V2 with unweighted questions was chosen as the final simplified version of the OSDI questionnaire, named the OSDI-6. This final OSDI-6 (Table 3) is half the length of the original, but results in a very similar score both overall as well as for each of the three subscales of the original OSDI.

Nevertheless, there is a limitation of the new OSDI-6. In the OSDI,

patients are able to skip questions 6 to 12, if they feel they are not relevant. This is addressed by a special formula to calculate the final OSDI score. On one hand, this makes sense: for example, if the patient is not driving at night, how could they answer this question? On the other hand, this could result in patients only answering the first five questions of the OSDI, and therefore not fulfilling the concept of the three sub-categories. Furthermore, when completing the OSDI or the OSDI-6 in an interview, why not highlight that "driving at night" does not only mean being the driver but also a passenger, or if the patient is not watching TV asking for comparable situations? This would simplify the calculation of the questionnaire's score. In this study, patients were asked to self-complete the questionnaires. Even though the first five questions of the OSDI and those of the DEQ-5 are mandatory, some of the patients did not complete them. Interviewing the patient to complete a questionnaire would be a way around this, but can artificially decrease the result [14]. An alternative is an app that will not allow completion until all questions are responded to. One third (34%) of the patients did not complete all questions of the OSDI, whereas only 9% did not complete all questions of the OSDI-6. The most common question of the OSDI-6 that was not answered by patients was "driving at night." This could be
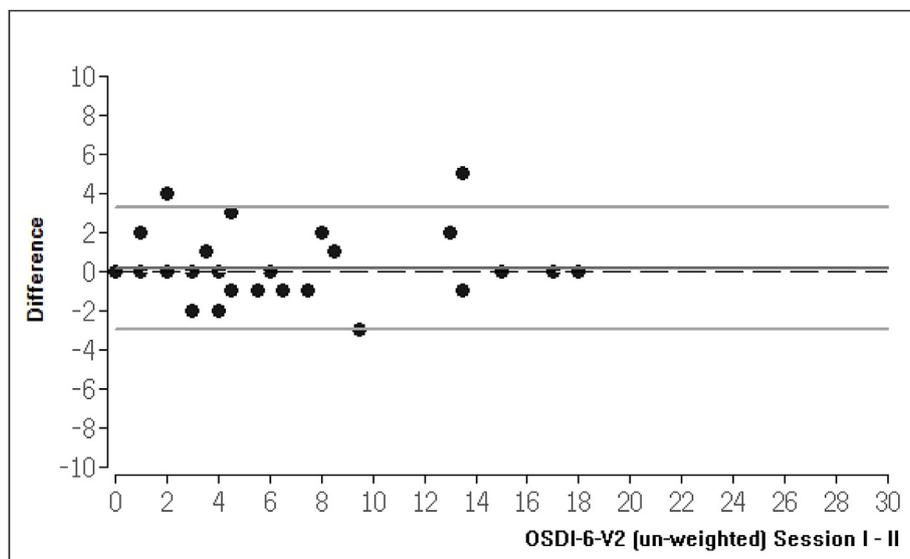

Fig. 2. Bland-Altmann plot showing the agreement between session I and II of the OSDI-6-V2 (un-weighted; n = 43; identical values are not shown).

Table 4
The proposed OSDI-6. The response of each question is scored resulting in a score between 0 and 24.

	Constantly	Mostly	Often	Sometimes	Never
Have you experienced any of the following *during a typical day of the last month?*					
1. Eyes that are sensitive to light?	4	3	2	1	0
2. Blurred vision?	4	3	2	1	0
Have problems with your eyes limited you in performing any of the following *during a typical day of the last month?*					
3. Driving or being driven at night?	4	3	2	1	0
4. Watching TV, or a similar task?	4	3	2	1	0
Have your eyes felt uncomfortable in any of the following situations *during a typical day of the last month?*					
5. Windy conditions?	4	3	2	1	0
6. Places or areas with low humidity?	4	3	2	1	0

clarified by amending the question to "driving or being driven at night." Based on intra-class coefficients (ICC), all the questionnaires showed excellent repeatability [15], although there is some reservation about the validity of this metric [16]. Kappa analyses showed moderate repeatability of all of the questionnaires, while the OSDI-6 un-weighted was slightly better with a strong repeatability [12].

5. Conclusions

The OSDI-6 showed a strong repeatability, better than that of the OSDI and DEQ-5. The OSDI-6 seems to be a good alternative questionnaire to the OSDI in clinical practice. It can be easily and quickly completed, giving a calculated score predictive of the longer OSDI as well as the DEQ-5.

Disclosure

The authors have no conflicts to disclose.

References

[1] Wolffsohn JS, Arita R, Chalmers R, Djalilian A, Dogru M, Dumbleton K, Gupta PK, Karpecki P, Lazreg S, Pult H, Sullivan BD, Tomlinson A, Tong L, Villani E, Yoon KC, Jones L, Craig JP. TFOS DEWS II diagnostic methodology report. Ocul Surf 2017;15:539–74.
[2] Stapleton F, Alves M, Bunya VY, Jalbert I, Lekhanont K, Malet F, Na KS, Schaumberg D, Uchino M, Vehof J, Viso E, Vitale S, Jones L. TFOS DEWS II epidemiology report. Ocul Surf 2017;15:334–65.
[3] Craig JP, Nelson JD, Azar DT, Belmonte C, Bron AJ, Chauhan SK, et al. The TFOS dry eye workshop II: executive summary. Ocul Surf 2017;4:802–12.
[4] Begley CG, Chalmers RL, Abetz L, Venkataraman K, Mertzanis P, Caffery BA, Snyder C, Edrington T, Nelson D, Simpson T. The relationship between habitual patient-reported symptoms and clinical signs among patients with dry eye of varying severity. Investig Ophthalmol Vis Sci 2003;44:4753–61.
[5] Begley C, Chalmers RL, Abetz L, Venkataraman K, Mertzanis P, Caffery BA, Snyder C, Edrington T, Nelson D, Simpson T. The relationship between habitual patient-reported symptoms and clinical signs among patients with dry eye of varying severity. Investig Ophthalmol Vis Sci 2003;44:4753–61.
[6] Nichols KK, Smith JA. Association of clinical diagnostic tests and dry eye surveys: the NEI visual function questionnaire-25 and the ocular surface disease index. Cornea November 2000;19(6). Supplement 2:S109 2000.
[7] Survey DRK. Of preferred tests for diagnostis of the tear film and dry eye. Cornea 2000;19:483–6.
[8] 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.
[9] Chalmers RL, Begley CG, Caffery B. Validation of the 5-Item Dry Eye Questionnaire (DEQ-5): discrimination across self-assessed severity and aqueous tear deficient dry eye diagnoses. Contact Lens Anterior Eye 2010;33:55–60.
[10] Michel M. Evaluation of contact lens dry eye using validated questionnaires and objective tests [Diploma Thesis]. Jena University of Applied Sciences; 2008.
[11] Dougherty BE, Nichols JJ, Nichols KK. Rasch analysis of the ocular surface disease index (OSDI). Investig Ophthalmol Vis Sci 2011;52:8630–5.
[12] McHugh ML. Interrater reliability: the kappa statistic. Biochem Med 2012;22:276–82.
[13] Simpson TL, Situ P, Jones LW, Fonn D. Dry eye symptoms assessed by four questionnaires. Optom Vis Sci 2008;85. https://doi.org/10.1097/OPX.0b013e318181ae36. E692-E9.
[14] Wolffsohn JS, Cochrane AL, Watt NA. Implementation methods for vision related quality of life questionnaires. Br J Ophthalmol 2000;84:1035–40.
[15] Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med 2016;15:155–63.
[16] McAlinden C, Khadka J, Pesudovs K. Statistical methods for conducting agreement (comparison of clinical tests) and precision (repeatability or reproducibility) studies in optometry and ophthalmology. Ophthalmic Physiol Opt 2011;31:330–8.