

RESEARCH ARTICLE

The Berkeley Dry Eye Flow Chart: A fast, functional screening instrument for contact lens-induced dryness

Andrew D. Graham, Erika L. Lundgrin, Meng C. Lin*

Clinical Research Center, School of Optometry, University of California, Berkeley, California, United States of America

* mclin@berkeley.edu



Abstract

Purpose

In this article, we introduce a novel flow chart-based screening tool for the categorization of contact lens-induced dryness (CLIDE) and its impact on daily visual activities: the Berkeley Dry Eye Flow Chart (DEFC).

Methods

One hundred thirty (130) experienced soft contact lens wearers discontinued lens wear for 24 hrs, passed a baseline screening and eye health examination, completed the Ocular Surface Disease Index (OSDI) then were dispensed fresh pairs of their habitual lenses. After 6 hrs of wear, subjects were administered a battery of symptom questionnaires, and underwent non-invasive tear breakup time (NITBUT) measurement, grading of distortion in reflected topographer mires, grading of lens surface wettability, and a fluorescein examination of the ocular surface. Subjects returned after at least 48 hrs and repeated all assessments after 6 hrs of wear of a second fresh pair of habitual lenses.

Results

The repeatability of the DEFC between visits was within 1%, and Limits of Agreement and Coefficient of Repeatability were comparable to those of the other CLIDE assessments. Higher DEFC score was significantly related to shorter pre-lens NITBUT, higher OSDI score, and higher Visual Analog Scale (VAS) ratings of average and end-of-day severity and frequency of dryness (all $p < 0.001$). For CLIDE as diagnosed based on DEFC score, the highest sensitivities and specificities were achieved by the OSDI and VAS ratings; pre-lens NITBUT exhibited good sensitivity but poor specificity. The optimum pre-lens NITBUT diagnostic threshold was found to be ≤ 2.0 sec for debilitating CLIDE, and the OSDI threshold was ≥ 11.4 .

OPEN ACCESS

Citation: Graham AD, Lundgrin EL, Lin MC (2018) The Berkeley Dry Eye Flow Chart: A fast, functional screening instrument for contact lens-induced dryness. PLoS ONE 13(1): e0190752. <https://doi.org/10.1371/journal.pone.0190752>

Editor: Andrew D. Pucker, University of Alabama at Birmingham, UNITED STATES

Received: April 12, 2017

Accepted: December 6, 2017

Published: January 24, 2018

Copyright: © 2018 Graham et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data, as well as the Dry Eye Flow Chart (DEFC) are available from Open Science Framework at the following URL: <https://osf.io/dfwnq/>. Users are welcome to download and utilize this information, provided that proper attribution is granted.

Funding: This work was supported by the Morton D. Sarver Research Fund, Roberta Smith Research Fund and Clinical Research Center Unrestricted Fund. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Conclusions

The DEFC provides a means of quickly categorizing CLIDE patients based on severity and frequency of symptoms, and on the degree to which symptoms impact daily life. The DEFC has several potential advantages as a CLIDE screening and monitoring tool, has good repeatability, and is significantly related to commonly employed clinical assessments for CLIDE.

Introduction

Contact-lens-induced dryness (CLIDE) is a condition with a multifaceted etiology and widely varying presentation, with no consensus among researchers or clinicians as to how it should be diagnosed [1–4]. The International Workshop on Contact Lens Discomfort (CLD) classified dryness during contact lens wear as one form of CLD, reserving the term “contact lens-related dry eye” for a pre-existing dry eye (DE) condition which (rather confusingly) “may or may not be exaggerated” during lens wear [2]. Most studies to date have not attempted to rigorously define CLIDE but rather have simply examined dryness symptoms that occur during lens wear or have compared “symptomatic” and “asymptomatic” groups of subjects defined in various ways [5–10]. Less commonly, CLIDE has been described as the presence of dryness symptoms only during lens wear, or that resolve when lenses are removed [11–14].

CLIDE symptoms have typically been assessed using questionnaire instruments developed for DE and not specifically designed or validated for contact lens wear. These include the Ocular Surface Disease Index (OSDI) [15,16], the Dry Eye Questionnaire (DEQ) [13], the Standard Patient Evaluation of Eye Dryness (SPEED) [17], the Impact of Dry Eye on Everyday Life (IDEEL) questionnaire [18], the Subjective Evaluation of Symptoms of Dryness (SESoD) [19], and McMonnies’ Dry Eye Index (DEI) [20–22]. The OSDI has been extensively validated for DE patients, and although not designed or specifically validated for CLIDE [4], has been widely used in its study. McMonnies’ DEI also was designed and validated for DE patients, but was later validated in a study of 70 contact lens wearers [6].

Of the few published questionnaire instruments that were developed specifically for CLIDE (as distinct from DE in general), the 8-item short form of the Contact Lens Dry Eye Questionnaire (CLDEQ-8) is the most widely employed [13,23,24]. The CLDEQ-8 includes 5-point ordinal rating scale items for the frequency and end-of-wearing-period intensity of discomfort, dryness, and changeable, blurry vision during lens wear, as well as for frequency of desire to close the eyes and of contact lens removal due to symptoms, all of which are then summed for a total CLIDE score. A subset of CLDEQ-8 questions regarding the frequency and end-of-wearing-period intensity of symptoms has been employed by Young [25,26] to assign a categorization of CLIDE-positive, CLIDE-negative or marginal.

The purpose of this study is to describe a new flow chart-based questionnaire for the categorization of CLIDE: the Berkeley Dry Eye Flow Chart (DEFC; Fig 1) [27,28]. The DEFC categorizes individuals into one of five CLIDE types based on the presence/absence of dryness sensation during lens wear over the previous week, the perception of discomfort caused by dryness symptoms when present, and the frequency with which dryness symptoms interfere with daily activities such as reading or using the computer (interference may include having to interrupt the activity to use eye drops, having to remove contact lenses, or having to stop the activity altogether). The DEFC was developed by U.C. Berkeley Clinical Research Center staff, including optometrists, clinical researchers, and biostatisticians, through discussions with

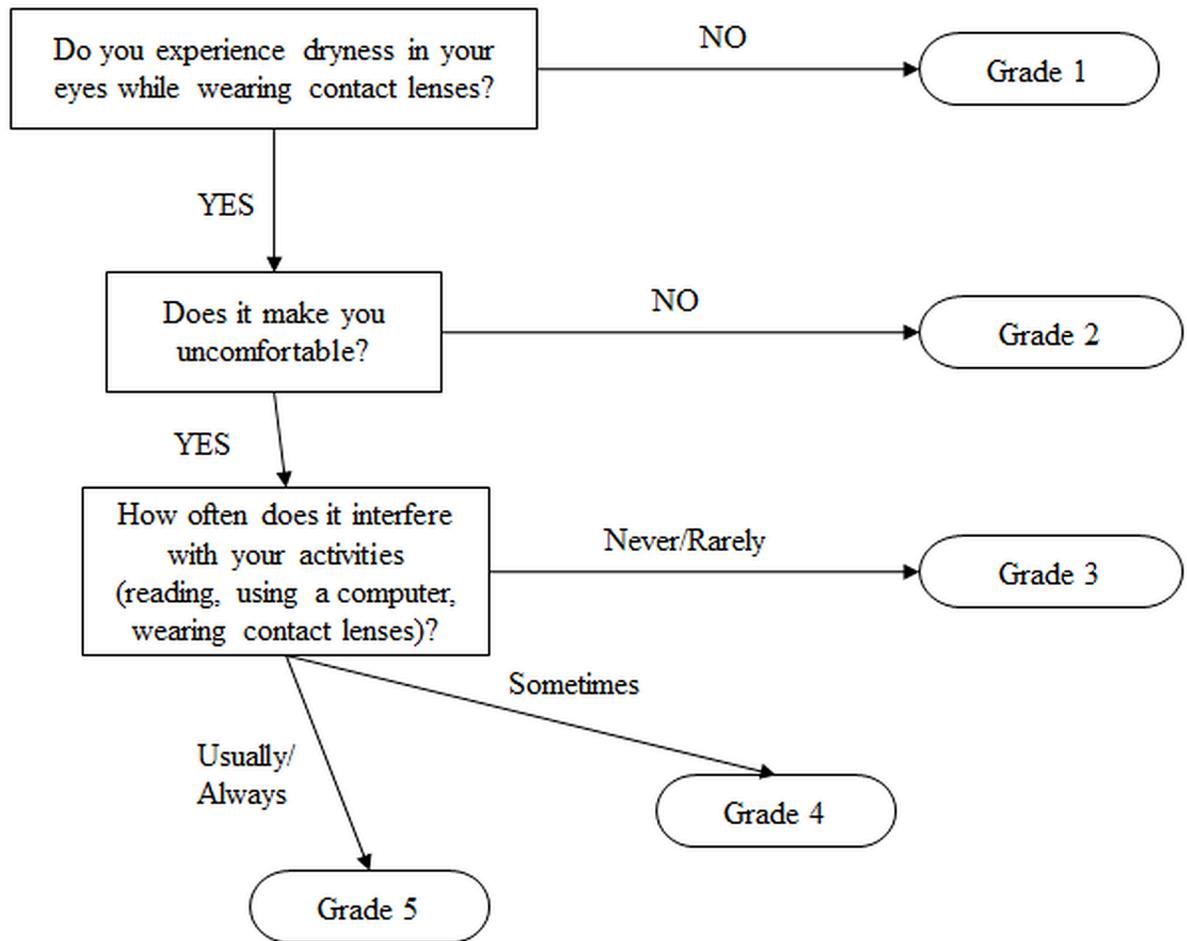


Fig 1. The Berkeley Dry Eye Flow Chart (DEFC).

<https://doi.org/10.1371/journal.pone.0190752.g001>

colleagues in clinical practice who felt that numerical grading of symptoms or Likert-style checkbox questionnaires were insufficient to capture how a symptomatic patient really felt. While an internal mental state is a latent trait, it may be made manifest (and measured) in a change of expression or behavior related to the symptoms. From a clinical perspective, the extent to which symptoms interfere with daily activities is a good indicator of the severity of the problem, and suggestive of how it should be further treated. From a research perspective, it was determined that the instrument should be short and easy to complete, thus reducing the potential for misclassification bias. It was finally determined that the most promising design would be a flow-chart questionnaire, with binary items related to the presence or absence of symptoms, their severity and frequency, and their impact on daily life.

There are several potential advantages inherent in the design of the DEFC. It is very simple to read and understand, reducing the potential for later misdiagnosis due to incorrectly completed questionnaires from patients who are fatigued or confused by a longer or more difficult instrument, and reducing the potential for bias in clinical studies due to over-selecting for the capabilities required to successfully complete longer, more difficult questionnaires [29]. Because the DEFC score incorporates aspects of symptom presence, sensation, severity, frequency and especially impact on everyday activities, different types of functional CLIDE can be identified by setting the DEFC score at appropriate diagnostic thresholds. For example, a

researcher may wish to assign potential clinical trial subjects to either a control group with no CLIDE symptoms (DEFC score < 2) or a CLIDE group with symptoms spanning the entire range encountered in the target population (DEFC score ≥ 2); a clinician may wish to screen patients (n.b., with further clinical assessment before initiating actual treatment) whose symptoms require no or only mild intervention such as occasional application of artificial tears (DEFC score < 4) from those whose symptoms are severe enough to be debilitating and require more thorough investigation and possibly more aggressive treatment (DEFC score ≥ 4). Finally, the DEFC score provides a fast and easy way to gather a basic, functional, actionable reading of CLIDE status, which can be logistically important in facilitating large numbers of potential subjects being quickly pre-qualified in a clinical trial setting or classified for epidemiologic study, or large numbers of patients being efficiently pre-screened in an institutional setting (e.g., an on-site screening at a senior care facility) [30,31].

Our goals in this paper are: (1) to establish the repeatability of the the DEFC and to compare its performance with that of other known subjective and clinical measures of CLIDE; (2) to determine whether several clinical tests and subjective symptom measures of CLIDE are related to the DEFC score, and to determine their effect sizes and significance of association; (3) to determine the diagnostic performance (sensitivity and specificity, optimum threshold) of these subjective and clinical measures for a DEFC-based diagnosis of CLIDE. We will present this analysis for the two different examples of DEFC-based CLIDE diagnosis described above: No Symptoms vs. Any Symptoms, and No/Mild Symptoms vs. Debilitating Symptoms. These findings will serve to illuminate whether and to what degree clinical outcomes and subjective symptom measures are reflected in different types of CLIDE, based on the functional classification offered by the DEFC. With its characteristics and performance documented, the DEFC could prove to be a useful addition to the investigational and screening tools available for CLIDE, and to its diagnosis and management in the clinical setting.

Materials and methods

Subjects

Subjects were recruited from the University of California, Berkeley campus and surrounding community through posted flyers and direct referrals. Potential subjects were soft contact lens wearers 18 years of age or older who had had eye examinations within the previous 2 years. Eligible subjects were free of any ocular surface pathology or health conditions with ocular manifestations, and were not taking any prescription or over-the-counter medications that could affect the ocular surface or tear film.

Study protocol and procedures

Subjects who passed an initial eligibility screening by telephone and elected to participate were instructed to report to the U.C. Berkeley Clinical Research Center (CRC) for a baseline examination and final eligibility determination, wearing their habitual lenses. Upon arrival at the baseline visit, subjects completed a background questionnaire to collect demographic data as well as medical and contact lens histories, and to complete the Ocular Surface Disease Index (OSDI). Subjects then completed the DEFC followed by 100-point rating scale questionnaires on the severity and frequency of all-day and end-of-day dryness symptoms, presented as Visual Analog Scales (VAS) [32,33]. For dryness severity, a rating of 0 = No Dryness and a rating of 100 = Severe Dryness; for dryness frequency, a rating of 0 = Never and a rating of 100 = All the Time. VAS ratings were given by the subject making a vertical mark through a 10 cm horizontal line with only the end points labelled as described, to indicate where his or her

eyes fell on the continuum of symptoms, the location of which was automatically determined by scanning software (Teleform, Hewlett Packard, Palo Alto, CA, USA).

Eligible subjects were then scheduled for 2 follow-up visits, at least 48 hrs apart. Subjects were dispensed 2 fresh pairs of their habitual lenses, and were instructed to insert a fresh pair of lenses at least 6 hrs prior to each follow-up visit. At these visits, subjects first completed all dryness symptom questionnaire instruments, following which, a corneal topographer (E300, Medmont International Pty. Ltd., Melbourne, Victoria, AUS) was used to measure non-invasive tear breakup time (NITBUT). Breakup time was recorded at the first sign of a break or distortion in the reflected mires. Three measurements per eye were taken and averaged, with at least 30 seconds between measurements to allow blinking to refresh the tear film. Following the NITBUT measurement, in a mild form of “stress test”, subjects were instructed to blink and then hold the eyes open, and at 10 sec the severity of distortion or haze in the reflected topographer mires was graded on a 0–4 scale (0 = No haze or distortion; 4 = Extreme haze or distortion) in the central, nasal, temporal and inferior zones. A slit lamp (SL 120, Carl Zeiss Meditec AG, Oberkochen, Baden-Württemberg, GER) with white light was then used to grade *in vivo* contact lens surface wettability on a 0–4 scale (0 = Non-wetting; 4 = Complete wetting). Following removal of the lenses and instillation of 2 μ l of 0.35% sodium fluorescein to the bulbar conjunctiva, cobalt blue light and a Wratten #12 yellow barrier filter were employed to grade corneal and conjunctival staining according to the Brien Holden Vision Institute grading scales [34]. Conjunctival staining was graded separately in the nasal, temporal, superior and inferior quadrants. Corneal staining was graded separately in the 4 quadrants and in the central 5mm zone, in terms of the type, extent and depth of staining.

This study adhered to the tenets of the Declaration of Helsinki, was approved by institutional review board (U.C. Berkeley Committee for the Protection of Human Subjects) and was conducted in compliance with Health Insurance Portability and Accountability Act (HIPAA) guidelines for data safety and subject anonymity. All subjects provided written informed consent after a thorough description of the procedures, potential benefits, and risks involved in the study.

Statistical methods

Repeatability of the DEFC instrument was assessed by comparing scores at the 2 visits, taken under as close to identical conditions as possible. The 2 visits were approximately 48 hours apart, both after at least 24 hours discontinuation of lens wear, insertion of a fresh pair of lenses, followed by 6 hours of wear. We examined the distribution of the differences (Visit 2 – Visit 1) in DEFC score, the mean difference, Limits of Agreement (LoA) [35], Difference vs. Means (DVM) plots [35], the % of scores within 1 point on the 5-point DEFC scale from visit to visit, and conducted a variance component analysis to estimate the within-subject variability in DEFC scores from which we calculated the Coefficient of Repeatability (CR) [36]. Strictly speaking, the true repeatability of an instrument requires measuring the identical phenomenon on multiple occasions under identical conditions with the same observer, which is not possible for subjective symptoms which can vary from day to day for some subjects, even under nearly identical conditions. We therefore conducted the same repeatability analysis on our other CLIDE measures (VAS ratings of symptoms, NITBUT, etc.) and converted those results to DEFC-equivalent scales, in order to show that the repeatability of the DEFC instrument, in the presence of inherently labile symptoms, is comparable to that of other commonly employed CLIDE assessments.

After repeatability assessment and a thorough exploratory analysis, a multivariable linear mixed effects modeling approach was taken to determine which clinical, laboratory and subjective measures were significantly related to DEFC score, accounting for potential within-

subject correlations engendered by the repeated measures design. Because DEFC score is a subject-level variable, for eye-level variables such as tear breakup time or lens surface wettability, we examined grades for the worse eye (e.g., faster breakup time, poorer wettability), the better eye, the average of the two eyes and the sum of the two eyes as potential explanatory variables for DEFC score. For corneal and conjunctival staining, we also examined the grades zone-by-zone and summed over all zones.

After identifying which variables were significantly related to DEFC score, we employed Receiver Operating Characteristic (ROC) analysis [37] to estimate the sensitivities and specificities of these diagnostics to a DEFC-based determination of CLIDE status, for the two different thresholds (DEFC score ≥ 2 and ≥ 4) described above. In the ROC analysis we calculated the sensitivity and specificity using every possible value of each clinical and subjective measure as a diagnostic threshold, within the ranges we observed. We developed an algorithm to implement our adaptation of the method proposed by Emir, et al. [38] for estimating sensitivity and specificity in the presence of repeated measures. The optimum threshold was considered to be the value that resulted in sensitivity and specificity at the smallest Euclidean distance to the (0,1) point on the ROC plot (i.e., the point geometrically closest to 100% sensitivity and 100% specificity).

Results

Subject characteristics

One hundred thirty (130) subjects completed the study. Subjects ranged in age from 18 to 61 years, with a mean (SD) of 26.6 (9.7) years. A majority of subjects (n = 89, 68%) were female. Sixty Asian subjects (46%) and 70 non-Asian subjects (54%) were analyzed. The Asian group consisted of subjects of Chinese, Japanese, Korean, Southeast Asian or Pacific Islander descent. Non-Asian subjects included Caucasians, African-Americans, and Latinos. Subjects were all experienced soft contact lens wearers (38% silicone hydrogel, 62% hydrogel), with mean (SD) wearing times of 5.7 (1.7) days/week and 11.8 (3.4) hours/day. Further detail on the demographic makeup and baseline characteristics of our subjects is provided in Table 1.

Repeatability of the DEFC

The mean difference in DEFC scores (Visit 2 – Visit 1) was 0.05 units (1.00%) on the 5-point DEFC scale, and was symmetrical about this difference, ranging from +3 to -3 units (Fig 2).

Table 1. Baseline subject characteristics.

	Asian (n = 60)		Non-Asian (n = 70)	
	Female (n = 44)	Male (n = 16)	Female (n = 45)	Male (n = 25)
Age (yrs)	23.6 (6.9)	22.9 (5.4)	29.9 (11.6)	28.4 (10.2)
Students	34 (77.3%)	11 (68.8%)	23 (51.1%)	19 (76.0%)
Computer (hrs/day)	5.5 (3.7)	4.7 (2.6)	5.0 (2.6)	5.5 (3.0)
CL Wear Hx (yrs)	5.8 (6.5)	6.9 (12.5)	6.9 (7.6)	4.0 (4.8)
CL Wear (hrs/day)	11.6 (3.0)	12.3 (3.4)	12.1 (3.3)	11.4 (4.3)
CL Wear (days/wk)	5.9 (1.4)	5.1 (2.0)	5.6 (2.0)	5.5 (1.6)
Eye Drops Use	33 (75.0%)	8 (50.0%)	27 (60.0%)	16 (64.0%)

The n = 130 subjects are shown stratified on gender and on Asian vs. non-Asian descent. The mean (SD) in each stratum is shown for subject age, computer use, and contact lens (CL) wearing history and daily average wearing pattern. The number (n) and % of subjects in each stratum who are students and who use eye drops are also shown.

<https://doi.org/10.1371/journal.pone.0190752.t001>

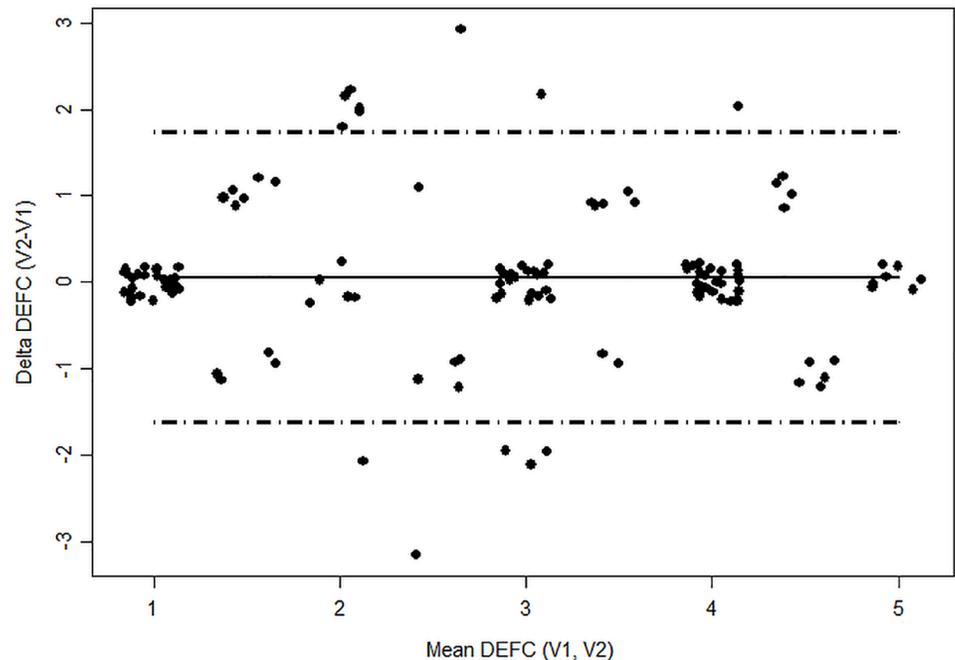


Fig 2. Difference-vs.-mean plot of DEFC scores. The mean difference in DEFC scores (Visit 2 – Visit 1) was 0.05 units (1.00%) on the 5-point DEFC scale. There was no pattern of dependence of the inter-visit difference on the size of the mean. The discrete ordinal DEFC scores are jittered to reveal coincident data.

<https://doi.org/10.1371/journal.pone.0190752.g002>

The between-visit differences in other CLIDE measures ranged from 0.02% to 4.60% (Table 2) and were also symmetrical. Only 2 subjects differed by 3 DEFC units between visits, with 117 subjects (90.00%) having a difference of either 0 (n = 84) or 1 (n = 33) unit. Scaling the other CLIDE measures to DEFC-equivalent units for comparison purposes, we found percentages of between-visit differences of ≤ 1 unit ranging from 55.38 to 80.77%. LoA for DEFC scores were [-1.62, 1.73], or [-32.40%, 34.60%]. Other CLIDE measures had lower LoA ranging from -35.50 to -58.40%, and upper LoA ranging from 37.00 to 55.80%. The CR for DEFC score was 0.42, or 8.47%, meaning that the true absolute difference between repeated DEFC scores under these conditions lies within less than half a unit (0.42) on the 5-point DEFC scale, with 95% probability. Overall, the DEFC showed good repeatability, comparable to or better than that of other CLIDE measures.

DEFC score: Relationships to other CLIDE assessments

Subjects reported DEFC scores spanning the range of CLIDE symptoms (1–5), with a median score of 3, and a mean (SD) score of 2.8 (1.4), combined over the 2 visits. Table 3 shows the distribution of DEFC scores at each visit, along with the mean (SD) of each CLIDE assessment, stratified on DEFC score.

We found there to be significantly ($p < 0.001$) shorter pre-lens NITBUT (i.e., greater tear film instability) with higher DEFC score (i.e., worse symptoms). We examined a number of variables derived from pre-lens NITBUT on both the raw and natural log scales, and found that the best model fits were achieved using the worse pre-lens NITBUT of the two eyes, with three individual measurements above 10 sec truncated at that time to better approximate normality and avoid undue leverage effects. The pre-lens NITBUT ranged from approximately 4.0 sec on average among subjects with no CLIDE symptoms (DEFC score = 1) to 1.9 sec on

Table 2. Repeatability of the CLIDE diagnostic instruments.

INSTRUMENT	Limits of Agreement						Coefficient of Repeatability		n (%) Diffs w/in 1 DEFC Level Equiv.
	Raw Scale			% of Max			Raw Scale	% of Max	
	Mean Diff	Lower	Upper	Mean Diff	Lower	Upper			
DEFC	0.05	-1.62	1.73	1.00	-32.40	34.60	0.42	8.47	117 (90.00)
NITBUT	-0.46	-5.84	4.92	4.60	-58.40	49.20	0.14	1.38	81 (62.79)
Distortion	-0.01	-8.93	8.92	0.06	-55.80	55.80	8.72	54.50	72 (55.38)
Wettability	0.03	-1.42	1.48	0.75	-35.50	37.00	0.12	3.03	105 (80.77)
Avg. Severity VAS	-0.30	-50.56	49.96	0.30	-50.56	49.96	4.74	4.74	95 (73.08)
Avg. Frequency VAS	-1.08	-44.26	42.09	1.08	-44.26	42.09	30.45	30.45	92 (70.77)
EOD Severity VAS	1.24	-46.65	49.13	1.24	-46.65	49.13	0.70	0.70	94 (73.44)
EOD Frequency VAS	-0.02	-42.77	42.72	0.02	-42.77	42.72	0.31	0.31	98 (76.56)

The mean inter-visit (Visit 2 –Visit 1) difference in DEFC score (1.00%) was within the normal variation encountered with other CLIDE measures, which ranged from 0.02% to 4.60%. The coefficient of repeatability estimates that 95% of inter-visit differences will lie within less than half a DEFC level (0.42 on the 5-point scale). Ninety percent of inter-visit differences in DEFC score were within 1 level. The repeatability of the DEFC is comparable to that of the other CLIDE assessments.

EOD = End Of Day. VAS = Visual Analog Scale. DEFC Level Equiv. = 20% of the observed range of the variable (i.e., because 1 DEFC level is 20% of its own 5-point scale).

<https://doi.org/10.1371/journal.pone.0190752.t002>

average among subjects whose CLIDE symptoms are severe enough to frequently interfere with daily activities (DEFC score = 5). Table 4 shows the sensitivity and specificity of pre-lens NITBUT for our first DEFC-based diagnosis of CLIDE (No Symptoms vs. Any Symptoms) to be 65% and 50%, respectively, at a threshold of 2.5 sec. For our second DEFC-based diagnosis

Table 3. DEFC score vs. other CLIDE assessments.

	DEFC SCORE				
	1	2	3	4	5
n (%):	72 (27.7)	31 (11.9)	59 (22.7)	74 (28.5)	24 (9.2)
Mean (SD):					
Pre-Lens NITBUT	3.97 (2.85)	3.15 (1.79)	2.73 (2.14)	2.38 (2.69)	1.86 (1.72)
Pre-Lens Distortion	9.60 (3.79)	9.87 (3.51)	9.17 (3.83)	9.74 (3.85)	9.83 (2.79)
Pre-Lens Wettability	2.95 (0.69)	2.77 (0.72)	2.92 (0.64)	2.82 (0.60)	2.97 (0.71)
OSDI	7.96 (8.55)	10.69 (14.06)	11.33 (10.30)	19.23 (12.78)	25.80 (17.24)
Avg. Severity VAS	18.91 (26.36)	21.67 (23.58)	35.89 (27.98)	50.55 (21.44)	48.00 (22.94)
Avg. Frequency VAS	13.91 (15.39)	25.54 (29.43)	40.37 (28.43)	56.10 (21.13)	60.12 (25.66)
EOD Severity VAS	25.32 (24.69)	37.16 (29.89)	52.32 (26.11)	70.66 (19.75)	68.37 (25.88)
EOD Frequency VAS	25.28 (24.68)	39.67 (33.65)	55.47 (28.11)	73.70 (19.29)	70.87 (28.55)

Shown are the sample sizes at each level of the DEFC (Visit 1 and Visit 2 combined; n = 260), along with the mean (SD) of each CLIDE assessment that was compared with the DEFC. Pre-lens NITBUT is shorter the higher the DEFC score, and all subjective ratings of symptoms are higher the higher the DEFC score.

EOD = End-Of-Day.

<https://doi.org/10.1371/journal.pone.0190752.t003>

Table 4. Sensitivity and specificity of CLIDE assessments for two DEFC-based diagnoses.

Diagnosis 1: No Symptoms vs. Any Symptoms					
	Sensitivity (%)	95% CI	Specificity (%)	95% CI	Threshold
Pre-Lens NITBUT	65.4	[59.9, 70.9]	50.0	[40.4, 59.6]	≤ 2.5
Pre-Lens Distortion	50.3	[44.6, 56.1]	52.9	[43.3, 62.5]	≥ 12
Pre-Lens Wettability	48.6	[42.8, 54.4]	58.7	[49.2, 68.1]	≤ 2.5
OSDI	55.9	[50.2, 61.7]	80.8	[73.2, 88.3]	≥ 11.4
Avg. Severity VAS	65.7	[60.2, 71.2]	76.0	[67.7, 84.2]	≥ 21
Avg. Frequency VAS	72.3	[67.2, 77.6]	76.0	[67.7, 84.2]	≥ 21
EOD Severity VAS	71.3	[66.1, 76.6]	76.9	[68.8, 85.0]	≥ 45
EOD Frequency VAS	70.3	[65.0, 75.6]	81.7	[74.3, 89.2]	≥ 48
Diagnosis 2: No/Mild Symptoms vs. Debilitating Symptoms					
	Sensitivity (%)	95% CI	Specificity (%)	95% CI	Threshold
Pre-Lens NITBUT	73.2	[66.2, 80.2]	48.1	[41.7, 54.5]	≤ 2.0
Pre-Lens Distortion	52.3	[44.4, 60.2]	52.3	[46.0, 58.7]	≥ 12
Pre-Lens Wettability	50.3	[42.4, 58.2]	55.7	[49.4, 62.0]	≤ 2.5
OSDI	71.2	[64.1, 78.4]	70.0	[64.2, 75.9]	≥ 11.4
Avg. Severity VAS	82.4	[76.3, 88.4]	68.4	[62.4, 74.3]	≥ 24
Avg. Frequency VAS	74.5	[67.6, 81.4]	75.5	[70.1, 81.0]	≥ 42
EOD Severity VAS	73.2	[66.2, 80.2]	73.0	[67.3, 78.6]	≥ 63
EOD Frequency VAS	81.7	[75.6, 87.8]	67.5	[61.5, 73.5]	≥ 54

Sensitivities tended to be slightly higher for DEFC-based Diagnosis 2. Specificities for Diagnoses 1 and 2 tended to be about the same, or slightly lower for Diagnosis 2 (but not significantly so). With the exception of OSDI, thresholds were higher (i.e., symptoms ratings higher, NITBUT shorter) for identifying subjects with debilitating CLIDE symptoms.

EOD = End-Of-Day.

<https://doi.org/10.1371/journal.pone.0190752.t004>

of CLIDE (No/Mild Symptoms vs. Debilitating Symptoms), the sensitivity and specificity of pre-lens NITBUT were 73% and 48%, respectively, at a threshold of 2.0 sec.

We did not find a significantly worse grade of distortion of the reflected topographer mires at 10 sec post-blink with higher DEFC score (Table 3). We examined a number of variables derived from grading of distortion, including the individual grades in each quadrant, and the sum distortion grade over all quadrants. We found no significant association between any distortion grade variable and DEFC score, with the sum score across quadrants from the worse eye ($p = 0.540$) having the best model fit, although all models were similar. With no significant association, grade of distortion was clearly not useful for distinguishing subjects with no symptoms from those with any CLIDE symptoms, having a sensitivity and a specificity of 0.50 and 0.52, respectively (Table 4); nor was grade of distortion useful for distinguishing subjects with no/mild symptoms from those with debilitating symptoms requiring intervention, with a sensitivity and a specificity of 0.52 and 0.52, respectively.

Clinician grading of in vivo contact lens wettability did not show a trend related to DEFC score (Table 3), and indeed in the mixed effects models was not a significant factor (e.g., $p = 0.607$ for the worse eye, that is, the eye with the less wettable lens). As expected, grade of lens wettability was not useful for distinguishing subjects with no symptoms from those with any CLIDE symptoms, having a sensitivity and a specificity of 0.48 and 0.58, respectively (Table 4); nor was grade of contact lens wettability useful for distinguishing subjects with no/mild symptoms from those with debilitating symptoms requiring intervention, with a sensitivity and a specificity of 0.50 and 0.55, respectively.

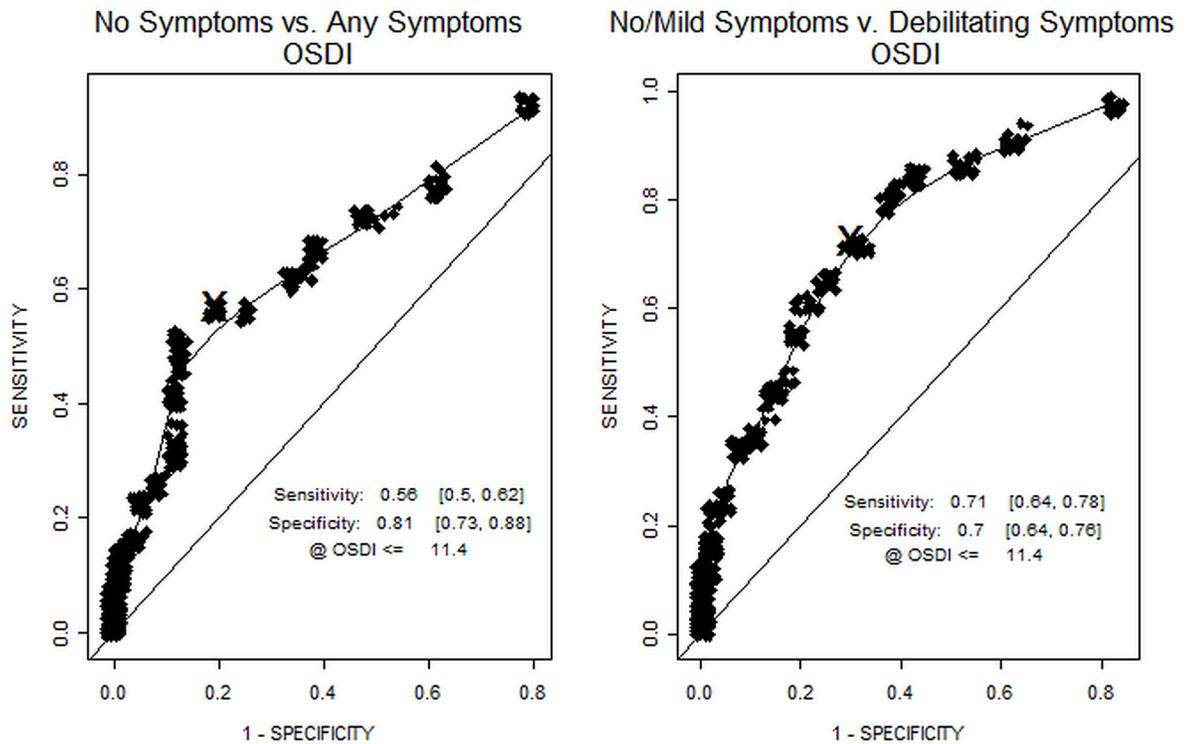


Fig 3. ROC curves for the OSDI. For the first DEFC-based CLIDE diagnosis, the OSDI showed a low sensitivity but a relatively high specificity. For the second DEFC-based diagnosis, the sensitivity of the OSDI improved with a small concomitant reduction in specificity.

<https://doi.org/10.1371/journal.pone.0190752.g003>

The OSDI score was significantly related to the DEFC score ($p < 0.001$), and showed a consistent increase across the levels of the DEFC (Table 3) from a mean (SD) of 8.0 (8.6) among subjects with no CLIDE symptoms (DEFC score = 1) to 25.8 (17.2) among subjects with the most severe symptoms (DEFC score = 5). As shown in Fig 3, for the first DEFC-based CLIDE diagnosis (No Symptoms vs. Any Symptoms), the OSDI exhibited a low sensitivity of 0.55 but a relatively high specificity of 0.80 (Table 4). For the second DEFC-based diagnosis (No/Mild Symptoms vs. Debilitating Symptoms), the OSDI showed a sensitivity and specificity of 0.71 and 0.70, respectively.

There were significant associations between DEFC score and subjective VAS ratings of symptoms, including on-average severity and frequency of dryness while wearing lenses, and end-of-day (EOD) severity and frequency of dryness (all $p < 0.001$). A clear trend of higher VAS ratings with higher DEFC score was found (Table 3), and EOD ratings tended to be slightly higher than on-average ratings (Fig 4). On-average dryness severity ranged from approximately 19 on the 100-point VAS scale for those with no symptoms (DEFC score = 1) to 48 for those with the most debilitating symptoms (DEFC score = 5), while EOD dryness severity ranged from approximately 25 to 68, on average, across those same DEFC categories. Similarly, on-average dryness frequency ranged in mean from approximately 14 to 60 across the DEFC score spectrum, while EOD dryness frequency means ranged from approximately 25 to 71. Sensitivities and specificities for our two DEFC-based diagnoses of CLIDE are shown in Table 4. In general, the subjective VAS ratings had higher sensitivities and specificities than did any of the clinical measurements, and were similar to those of the OSDI.

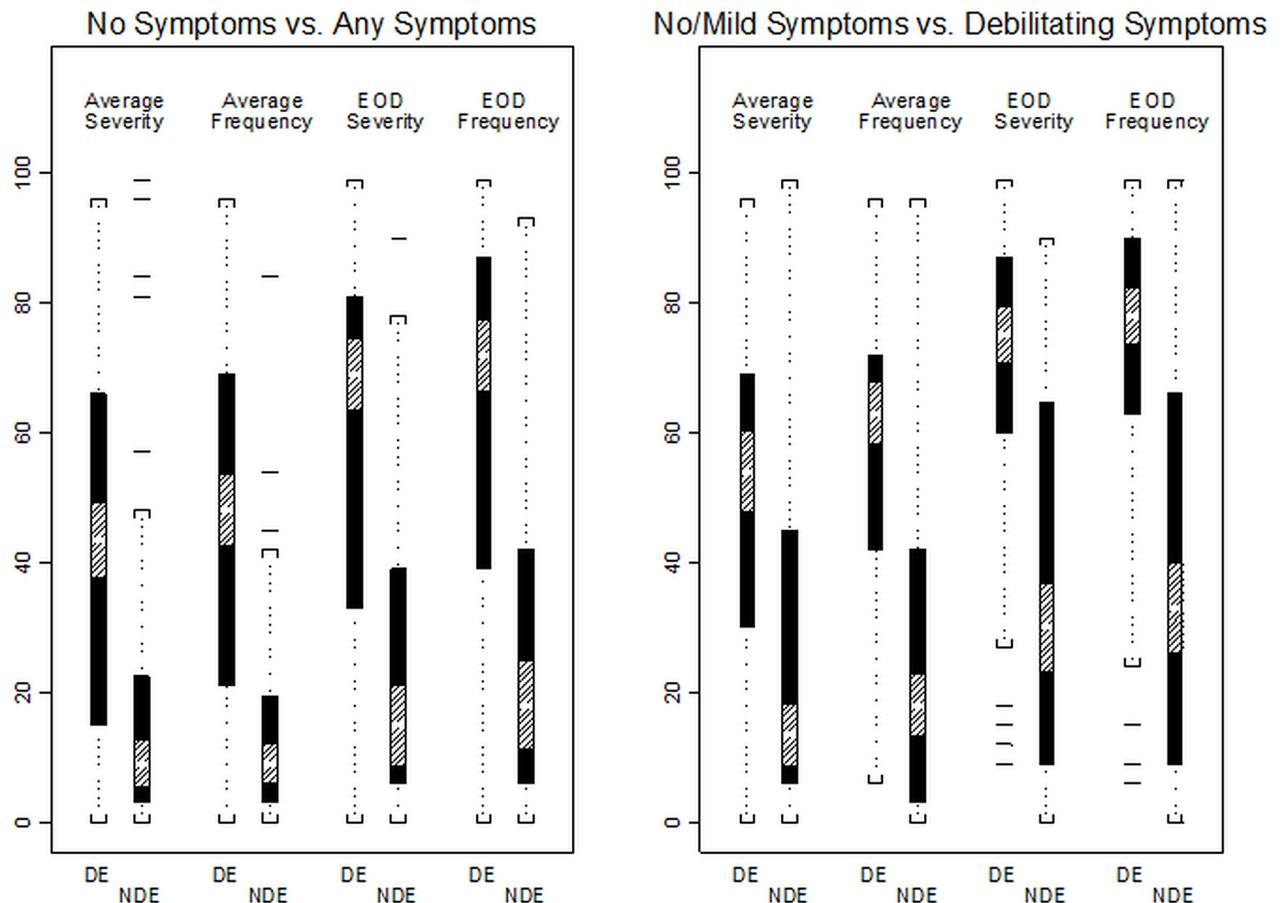


Fig 4. A trend of higher VAS ratings with higher DEFC score. EOD ratings tended to be slightly higher than on-average ratings. The subjective VAS ratings had higher sensitivities and specificities than did any of the clinical measurements, and were similar to those of the OSDI. *DE* = Contact Lens-Induced Dry Eye (DEFC score ≥ 2); *NDE* = Non-DE (DEFC score = 1).

<https://doi.org/10.1371/journal.pone.0190752.g004>

Discussion

In this article, we have introduced the Berkeley Dry Eye Flow Chart, which has the potential advantages of being simple to understand, fast to administer to large numbers of subjects or patients, and provides a simple 1–5 score that directly reflects not only the presence of dryness symptoms and the level of discomfort they induce, but the functional impact of these symptoms on daily visual activities such as reading or using a computer. We have shown that the repeatability of the DEFC is comparable to or better than that of other common clinical and subjective measures of CLIDE. We have demonstrated a direct relationship between higher DEFC score and shorter pre-lens tear breakup time, higher VAS ratings of severity and frequency of symptoms, both on-average and at end-of-day, and higher OSDI score. Finally, we have demonstrated how the DEFC can be used to categorize subjects into CLIDE and non-CLIDE groups based on different diagnostic criteria, and estimated the sensitivity and specificity of each subjective and clinical CLIDE assessment to our two different DEFC-based diagnoses, along with the optimum diagnostic threshold for each assessment.

The DEFC was significantly related to, and showed clinically significant trends in the expected direction with, subjective ratings of symptoms using the VAS and the OSDI, and with pre-lens tear film instability measured by NITBUT. These are among the most commonly

employed assessments for both CLIDE and DE in general. In contrast, grades of topographer mires distortion and *in vivo* contact lens wettability were not significantly related to DEFC score, and furthermore did not suggest a linear trend of worse distortion or poorer wettability with higher DEFC score. These assessments, however, are not standard tests for CLIDE or DE in clinical practice. We included them in this study because some previous work in our lab [39] and others' [10,40] suggested that these might prove to be useful assessments related to severity of symptoms. This, however, turned out not to be the case in our study population. Thus the validation of the DEFC is based on the more commonly accepted assessments of tear film instability measurement and subjective ratings of symptoms.

The presence of a contact lens partitions the tear film into pre- and post-lens segments [41], and the pre-lens tear film is typically less stable than the pre-corneal tear film without lenses [42]. It remains equivocal whether pre-lens tear film stability differs between symptomatic and asymptomatic contact lens wearers. Estimates of average pre-lens NITBUT from asymptomatic hydrogel lens wearers range from 2.8 sec to 10.8 sec depending on lens type as well as on lens care solution formulation [43–46]. Pre-lens NITBUT has been found across a similar range in symptomatic subjects, and while symptomatic and asymptomatic subjects have been found to be significantly different in some studies [47], they have not in others [48,49]. Ours is the first study to show a distinct trend in pre-lens NITBUT *in vivo* across the range of CLIDE severity, from approximately 4 sec on average among subjects with no CLIDE symptoms to less than 2 sec among subjects with the most debilitating CLIDE. In addition, this study is the first to establish diagnostic thresholds (≤ 2.5 sec for any CLIDE symptoms, ≤ 2.0 sec for debilitating symptoms) for pre-lens, *in vivo* NITBUT (albeit at sensitivities and specificities of 65 and 50, respectively, for any CLIDE symptoms, and 73 and 48, respectively, for debilitating symptoms). It should be noted that the sensitivities and specificities, and the diagnostic thresholds obtained by ROC analysis in this study must be interpreted in the context of our study population: its demographic makeup, contact lens history, ocular characteristics, and so forth. For example, a much older population with substantially different refractive demands could very well elicit different diagnostic performance and different optimum diagnostic thresholds.

Corneal and conjunctival staining were also assessed at each exam. Although visible ocular surface damage can be indicative of severe, prolonged DE or CLIDE [50], in our study there were no significant relationships between DEFC score and any of our ocular surface staining variables. This is because, in our sample of young, healthy, experienced contact lens-wearing subjects, who wore fresh pairs of their habitual contact lenses for 6 hours only prior to assessment, there were very few cases of even moderate staining—certainly not sufficient to test any hypothetical statistical association.

The ROC analysis showed that even the best assessments for CLIDE do not operate with anything close to perfect sensitivity and specificity. In other words, even for the best assessments there will be false positives and false negatives. The sensitivities and specificities we found in this study are broadly in line with the relatively few examples published to date for CLIDE diagnosis, and with the somewhat more commonly reported diagnosis of DE in general. It is impossible to compare the sensitivities and specificities we found directly to other published examples because differences in study design, population characteristics, and in the criteria for determining a positive diagnosis all result in widely varying estimates across studies. For example, the sensitivities of NITBUT were 65% and 73% for our two DEFC-based diagnoses, respectively, compared with published examples ranging from 55% to 90%; the specificities for NITBUT that we found were 50% and 48%, while published examples ranged from 21% to 85% [51–55]. Studies employing various types of questionnaires have estimated sensitivities of 60% to 86%, compared with approximately 66% to 82% for our VAS questionnaire items, and specificities of 46% to 94%, compared with our range of approximately 68% to

82% [51,55–58]. A great many other potential diagnostic instruments (primarily for DE) have been investigated, including the Schirmer test, fluorescein tear breakup time, tear osmolarity, imprint cytology, vital staining, tear ferning, and a variety of subjective symptom instruments, leading to published estimates of sensitivity ranging from as low as 39% to as high as 98%, and estimates of specificity ranging from 10% to 100% [15,20,51,59–66]. The primary difference between studies is the set of criteria by which DE or CLIDE was defined, which renders direct comparisons of diagnostic performance uninformative. At the very least, the DEFC provides another means of defining CLIDE or DE, with certain practical advantages over other methods, and against which the most common diagnostic tools perform with sensitivities and specificities comparable to previous definitions of CLIDE or DE.

In the current study, sensitivities tended to be slightly higher for our second DEFC-based diagnoses (No/Mild Symptoms vs. Debilitating Symptoms), in which subjects with mild symptoms that do not cause discomfort sufficient to interfere with activities are not distinguished from subjects with no symptoms whatsoever. This shows that CLIDE assessments tend to be better at identifying subjects with debilitating symptoms than they are at distinguishing symptom-free subjects from those with any level of CLIDE symptoms from severe to mild. Specificities for both DEFC-based diagnoses of CLIDE tended to be about the same, or slightly lower for the second diagnosis (but not significantly so). With the exception of OSDI, thresholds were higher (i.e., symptoms ratings higher, NITBUT shorter) for subjects with debilitating CLIDE symptoms.

There were some differences particular to the OSDI in diagnosing CLIDE as based on our DEFC categorization. For the first DEFC-based diagnosis of CLIDE (No Symptoms vs. Any Symptoms), the OSDI had low sensitivity (56%), meaning that the OSDI score did not identify a number of subjects with $DEFC \geq 2$ as being CLIDE-positive. On the other hand, it had high specificity (81%), meaning that it is better at correctly identifying subjects with $DEFC = 1$. OSDI scores above the ROC threshold of 11.4 are probably characteristic mostly of those with CLIDE symptoms severe enough to cause significant discomfort [15] and possibly affect visual functioning in daily activities. Thus a number of subjects with only mild symptoms not affecting daily activities had OSDI scores that, while somewhat elevated over the truly CLIDE-negative subjects, did not exceed the ROC threshold and therefore were not diagnosed as CLIDE-positive. This is supported by the higher sensitivity (71%) of the OSDI for the second DEFC-based diagnosis of CLIDE (No/Mild Symptoms vs. Debilitating Symptoms), while still maintaining a specificity (70%) that is not significantly lower than for the first diagnosis. It is interesting to note that the ROC analysis-generated threshold for diagnosing CLIDE using the OSDI in our study population (≥ 11.4) is very close to the commonly accepted clinical threshold OSDI score (≥ 13) for diagnosing DE [67,68].

Our results illuminate a common difficulty in identifying optimal diagnostic criteria for multifaceted conditions such as CLIDE, arising from a common misconception. There are a number of published reports covering various clinical, laboratory and subjective assessments shown to be related to a CLIDE outcome. Often these variables are shown to be significant explanatory variables in multivariable association models of CLIDE, whether the outcome be presence/absence, risk, or severity of symptoms. However, a significant association with CLIDE does not necessarily make such an assessment a good diagnostic indicator [38]. It would seem natural to assume, from perusing the literature, that if certain questionnaire items or clinical assessments are consistently shown to be significantly related to CLIDE outcomes, they should therefore be routinely administered to ascertain a patient's CLIDE status; however, without directly examining the diagnostic performance of such assessments (i.e., sensitivity and specificity; determination of threshold values), there is no way of knowing whether or to what extent such assessments are useful as diagnostic indicators for CLIDE.

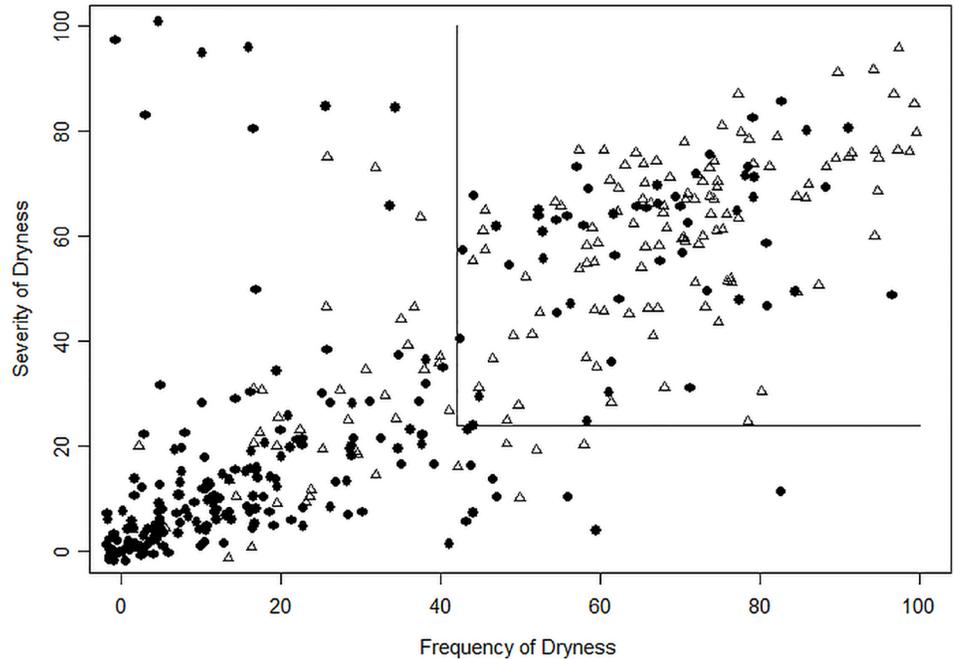


Fig 5. Potential pitfalls in using assessments that are significantly associated with CLIDE outcomes as diagnostic instruments. There is clear separation between CLIDE-negative (filled circles) and CLIDE-positive (open triangles) subjects; however, there are subjects with VAS ratings above the diagnostic thresholds (No/Mild Symptoms vs. Debilitating Symptoms thresholds shown) who do not have CLIDE (false-positives), as well as those with VAS ratings below the thresholds who do have CLIDE (false-negatives).

<https://doi.org/10.1371/journal.pone.0190752.g005>

Fig 5 depicts the potential pitfalls in using assessments that are significantly associated with CLIDE outcomes as diagnostic tools: while there is clear separation between CLIDE-negative (filled circles in the figure) and CLIDE-positive (open triangles) subjects, there are nevertheless a number of subjects with symptom severity and/or frequency above the optimum thresholds (as determined by ROC analysis) who did not self-classify as having CLIDE, as well as those with symptom ratings below the thresholds who do self-classify as CLIDE sufferers. These “false positives” and “false negatives” contribute to the diagnostic performance—the sensitivities and specificities—of the VAS rating scale symptom assessments.

It is important to understand that the ordinal categorical nature of the DEFC gives the instrument inherently good repeatability but relatively poor discrimination ability, or resolution. In other words, it would require a relatively large change in a patient’s or subject’s symptoms before there was a change in DEFC score, compared with, say, the 100-point VAS rating scale which has relatively poor repeatability but can record differences or changes in symptoms with much finer resolution. The uses for which the DEFC is designed make this appropriate: a clinician or researcher would not use the DEFC to determine whether there were fine between-subject differences in symptoms (e.g., that would show up on the VAS rating scale), but rather, would use the DEFC to determine whether differences in symptoms are sufficient to represent qualitatively different levels in terms of a patient’s or subject’s quality of life (e.g., dryness symptoms that do or do not cause significant ocular discomfort; symptoms that do or do not interfere with visual activities in daily life). While both types of instruments have a place in the clinical and research arenas, these distinctions highlight the need to employ the subjective response instrument that is most appropriate to the specific clinical or research application.

In this study, we compared individual assessments of CLIDE used in current research and clinical practice to our new metric, the Berkeley DEFC score. In practice, most clinicians would consider a number of different assessments before identifying a subject as having CLIDE or some other DE-related condition [50]. It is possible that the best diagnosis of CLIDE consists of a suite of clinical and subjective measures, each with its own diagnostic threshold, that collectively would provide the highest sensitivity and specificity for a test of CLIDE. Although there has been some limited work done in this area [51,52], the collective performance of commonly employed suites of diagnostic tools, the identification of the best set of assessments for diagnosis of CLIDE or DE in different populations, and determination of the optimum diagnostic thresholds for those assessments have not been well-studied to date. Work is ongoing in this area.

Conclusions

This article has introduced the Berkeley Dry Eye Flow Chart, discussed its potential advantages as a CLIDE screening tool, demonstrated its association with several known CLIDE assessments, and established its repeatability. In addition, we have demonstrated that the 5-level ordinal categorization of CLIDE based on the DEFC corresponds directly to trends in VAS ratings of symptoms, OSDI scores, and pre-lens NITBUT, and we have established the optimum diagnostic thresholds for these commonly employed assessments in our population by ROC analysis. We believe that the DEFC can provide a fast, accurate, functional categorization of CLIDE that takes into account the extent to which symptoms affect visual activities in daily life, and that it could prove a useful addition to the set of tools available to researchers and clinicians working with symptomatic contact lens wearers.

Author Contributions

Conceptualization: Andrew D. Graham, Meng C. Lin.

Data curation: Andrew D. Graham, Erika L. Lundgrin.

Formal analysis: Andrew D. Graham.

Funding acquisition: Meng C. Lin.

Investigation: Erika L. Lundgrin, Meng C. Lin.

Methodology: Andrew D. Graham, Meng C. Lin.

Project administration: Erika L. Lundgrin, Meng C. Lin.

Resources: Meng C. Lin.

Software: Andrew D. Graham.

Supervision: Meng C. Lin.

Validation: Andrew D. Graham.

Visualization: Andrew D. Graham.

Writing – original draft: Andrew D. Graham, Meng C. Lin.

Writing – review & editing: Andrew D. Graham, Meng C. Lin.

References

1. Lemp MA, for the Definition and Classification Subcommittee of the International Dry Eye Workshop. The Definition and Classification of Dry Eye Disease: Report of the Definition and Classification

- Subcommittee of the International Dry Eye Workshop. *The Ocular Surface*. 2007; 5(2):75–92. PMID: [17508116](#)
2. Nichols JJ, Willcox MDP, Bron AJ, Belmonte C, Cioloio JB, Craig JP, et al. The TFOS International Workshop on Contact Lens Discomfort: Executive Summary. *Invest Ophthalmol Vis Sci*. 2013; 54:TFOS7–TFOS13. <https://doi.org/10.1167/iovs.13-13212> PMID: [24058135](#)
 3. Sindt CW, Longmuir RA. Contact Lens Strategies for the Patient with Dry Eye. *The Ocular Surface*. 2007; 5(4):294–307. PMID: [17938839](#)
 4. Jalbert I, Golebiowski B, Stapleton F. Measuring Contact Lens Discomfort. *Current Ophthalmol Reports*. 2015; 3(2):106–110.
 5. Nichols JJ, Ziegler C, Mitchell GL, Nichols KK. Self-reported dry eye disease across refractive modalities. *Invest Ophthalmol Vis Sci*. 2005; 46(6):1911–1914. <https://doi.org/10.1167/iovs.04-1294> PMID: [15914603](#)
 6. Michel M, Sickenberger W, Pult H. The effectiveness of questionnaires in the determination of Contact Lens Induced Dry Eye. *Ophthalmic Physiol Opt*. 2009; 29(5):479–486. <https://doi.org/10.1111/j.1475-1313.2009.00658.x> PMID: [19689545](#)
 7. Chen J, Simpson TL. A role of corneal mechanical adaptation in contact lens-related dry eye symptoms. *Invest Ophthalmol Vis Sci*. 2011; 52(3):1200–1205. <https://doi.org/10.1167/iovs.10-5349> PMID: [20926812](#)
 8. Willen CM, McGwin G, Liu B, Owsley C, Rosenstiel C. Efficacy of cyclosporine 0.05% ophthalmic emulsion in contact lens wearers with dry eyes. *Eye Contact Lens*. 2008; 34(1):43–45. PMID: [18180683](#)
 9. Yeniad B, Beginoglu M, Bilgin LK. Lid-wiper epitheliopathy in contact lens users and patients with dry eye. *Eye Contact Lens*. 2010; 36(3):140–143. PMID: [20351557](#)
 10. Best N, Drury L, Wolffsohn JS. Predicting success with silicone-hydrogel contact lenses in new wearers. *Contact Lens Anterior Eye*. 2013; 36(5):232–237. <https://doi.org/10.1016/j.clae.2013.02.013> PMID: [23522993](#)
 11. Chalmers R, Long B, Dillehay S, Begley C. Improving contact lens-related dryness symptoms with silicone hydrogel lenses. *Optom Vis Sci* 2008; 85:778–784. PMID: [18677242](#)
 12. Djalilian AR, Mali JO, Holland EJ. The use of argon laser punctal stenosis in patients with contact lens-induced dry eyes. *Eye Contact Lens*. 2010; 36(3):144–148. PMID: [20461885](#)
 13. Begley CG, Chalmers RL, Mitchell GL, Nichols KK, Caffery B, Simpson TL, DuToit R, Portello J, Davis L. Characterization of ocular symptoms from optometric practices in North America. *Cornea*. 2001; 20(6):610–618. PMID: [11473162](#)
 14. Begley CG, Caffery B, Nichols KK, Chalmers RL. Responses of contact lens wearers to a dry eye survey. *Optom Vis Sci*. 2000; 77(1):40–46. PMID: [10654857](#)
 15. Schiffman RM, Christianson MD, Jacobsen G, Hirsch JD, Reis BL. Reliability and validity of the Ocular Surface Disease Index. *Arch. Ophthalmol*. 2000; 118(5):615–621. PMID: [10815152](#)
 16. Dougherty BE, Nichols JJ, Nichols KK. Rasch analysis of the Ocular Surface Disease Index (OSDI). *Invest Ophthalmol Vis Sci*. 2011; 52(12):8630–8635. <https://doi.org/10.1167/iovs.11-8027> PMID: [21948646](#)
 17. Ngo W, Situ P, Keir N, Korb D, Blackie C, Simpson TL. Psychometric properties and validation of the Standard Patient Evaluation of Dryness questionnaire. *Cornea*. 2013; 32(9):1204–1210. PMID: [23846405](#)
 18. Abetz L, Rajagopalan K, Mertzanis P, Begley C, Barnes R, Chalmers R, for the Impact of Dry Eye on Everyday Life (IDEEL) Study Group. Development and validation of the impact of dry eye on everyday life (IDEEL) questionnaire, a patient-reported outcomes (PRO) measure for the assessment of the burden of dry eye on patients. *Health and Quality of Life Outcomes*. 2011; 9:111. <https://doi.org/10.1186/1477-7525-9-111> PMID: [22152125](#)
 19. Simpson TL, Situ P, Jones LW, Fonn D. Dry eye symptoms assessed by four questionnaires. *Optom Vis Sci*. 2008; 85(8):692–699. PMID: [18677225](#)
 20. McMonnies CW. Key questions in a dry eye history. *J Am Optom Assoc*. 1986; 57:512–517. PMID: [3489027](#)
 21. McMonnies CW, Ho A. Patient history in screening for dry eye conditions. *J Am Optom Assoc*. 1987; 58:296–301. PMID: [3584792](#)
 22. Tang F, Wang J, Tang Z, Kang M, Deng Q, Yu J. Accuracy of McMonnies questionnaire as a screening tool for Chinese ophthalmic outpatients. *PLoS One*. 2016; 11(4):e0153047. <https://doi.org/10.1371/journal.pone.0153047> PMID: [27073922](#)

23. Nichols JJ, Mitchel GL, Nichols KK, Chalmers R, Begley C. The performance of the contact lens dry eye questionnaire as a screening survey for contact lens-related dry eye. *Cornea*. 2002; 21(5):469–475. PMID: [12072721](#)
24. Chalmers RL, Begley CG, Moody K, Hickson-Curran SB. Contact Lens Dry Eye Questionnaire-8 (CLDEQ-8) and opinion of contact lens performance. *Optom Vis Sci*. 2012; 89(10):1435–1442. PMID: [22960615](#)
25. Young G, Chalmers R, Napier L, Hunt C, Kern J. Characterizing contact lens-related dryness symptoms in a cross-section of UK soft lens wearers. *Contact Lens Anterior Eye*. 2011; 34(2):64–70. <https://doi.org/10.1016/j.clae.2010.08.005> PMID: [20832350](#)
26. Young G, Chalmers R, Napier L, Kern J, Hunt C, Dumbleton K. Soft contact lens-related dryness with and without clinical signs. *Optom Vis Sci*. 2012; 89(8):1125–1132. <https://doi.org/10.1097/OPX.0b013e3182640af8> PMID: [22820475](#)
27. Tran N, Graham AD, Lin MC. Ethnic Differences in Dry Eye Symptoms: Effects of Corneal Staining and Length of Contact Lens Wear. *Contact Lens Ant Eye*. 2013; 36(6):281–288.
28. Yeh TN, Graham AD, Lin MC. Relationships among Tear Film Stability, Tear Osmolarity, Corneal Staining History, and Dryness Symptoms. *Optom Vis Sci*. 2015; 92:e264–272. <https://doi.org/10.1097/OPX.0000000000000649> PMID: [26154693](#)
29. Choi BCK, Pak AWP. A catalog of biases in questionnaires. *Prev Chronic Dis*. 2005; 2(1). Available at: <http://www.cdc.gov/pccd/issues/2005/>.
30. West SK, Munoz B, Rubin GS, Schein OD, Bandeen-Roche K, Zeger S, et al. Function and visual impairment in a population-based study of older adults. The SEE Project. *Invest Ophthalmol Vis Sci*. 1997; 38:72–82. PMID: [9008632](#)
31. Bandeen-Roche K, Muñoz B, Tielsch JM, West SK, Schein OD. Self-reported assessment of dry eye in a population-based setting. *Invest Ophthalmol Vis Sci*. 1997; 38(12):2469–2475. PMID: [9375564](#)
32. Price DD, McGrath PA, Rafii A, Buckingham B. The validation of visual analogue scales as ratio scale measures for chronic and experimental pain. *Pain*. 1983; 17:45–56. PMID: [6226917](#)
33. Papas EB, Schultz BL. Repeatability and comparison of visual analogue and numerical rating scales in the assessment of visual quality. *Ophthalmic Physiol Opt*. 1997; 17:492–498. PMID: [9666923](#)
34. Terry RL, Schnider CM, Holden BA, Cornish R, Grant T, Sweeney D, et al. CCLRU standards for success of daily and extended wear contact lenses. *Optom Vis Sci*. 1993; 70(3):234–243. PMID: [8483586](#)
35. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet*. February 8, 1986:307–310.
36. Vaz S, Falkmer T, Passmore AE, Parsons R, Andreou P. The case for using the repeatability coefficient when calculating test-retest reliability. *PLoS One*. 2013; 8(9):e73990. <https://doi.org/10.1371/journal.pone.0073990> PMID: [24040139](#)
37. Metz CE. Basic principles of ROC analysis. *Seminars in Nuclear Med*. 1978; 8(4):283–298.
38. Emir B, Wieand S, Jung SH, Ying Z. Comparison of diagnostic markers with repeated measurements: a non-parametric ROC curve approach. *Statist Med*. 2000; 19:511–523.
39. Truong TN, Graham AD, Lin MC. Factors in Contact Lens Symptomatology: Evidence from a Multi Study Database. *Optom Vis Sci*. 2014; 91(2):133–141. PMID: [24317134](#)
40. Khanal S, Simmons PA, Pearce EI, Day M, Tomlinson A. Effect of artificial tears on tear stress test. *Optom Vis Sci*. 2008; 85(8):732–739. <https://doi.org/10.1097/OPX.0b013e318181ae1c> PMID: [18677237](#)
41. Nichols JJ, King-Smith PE. Thickness of the pre- and post-contact lens tear film measured in vivo by interferometry. *Invest Ophthalmol Vis Sci*. 2003; 44(1):68–77. PMID: [12506057](#)
42. Sweeney DF, Millar TJ, Raju SR. Tear film stability: a review. *Exp Eye Res*. 2013; 117:28–38. <https://doi.org/10.1016/j.exer.2013.08.010> PMID: [23973716](#)
43. Young G, Efron N. Characteristics of the pre-lens tear film during hydrogel contact lens wear. *Ophthalmic Physiol Opt*. 1991; 11:53–58.
44. Elliott M, Fandrich H, Simpson T, Fonn D. Analysis of the repeatability of tear break-up time measurement techniques on asymptomatic subjects before, during and after contact lens wear. *Contact Lens Ant Eye*. 1998; 21(4):98–103.
45. Faber E, Golding TR, Lowe R, Brennan NA. Effect of hydrogel lens wear on tear film stability. *Optom Vis Sci*. 1991; 68(5):380–384. PMID: [1852401](#)
46. Lau JK, So HT, Chiang TK, Cho P. Effect of multipurpose solutions on in vivo surface wettability of a silicone hydrogel lens. *Contact Lens Ant Eye*. 2016; 39(6):461–465

47. Fonn D, Situ P, Simpson T. Hydrogel lens dehydration and subjective comfort and dryness ratings in symptomatic and asymptomatic contact lens wearers. *Optom Vis Sci*. 1999; 76(10):700–704. PMID: [10524785](#)
48. Golding TR, Efron N, Brenna NA. Soft lens lubricants and prelens tear film stability. *Optom Vis Sci*. 1990; 67(6):461–465. PMID: [2381689](#)
49. Bruce AS, Golding TR, Au SWM, Rowhani H. Mechanisms of dryness in soft lens wear—role of BUT and deposits. *Clin Exp Optom*. 1995; 78(5):168–175.
50. Hessen M, Akpek EK. Dry eye: an inflammatory disease. *J Ophthalmic Vis Res*. 2014; 9(2):240–250. PMID: [25279127](#)
51. Alves M, Reinach PS, Paula JS, Vellasco e Cruz AA, Bachellet L, Faustino J, Aranha FP, Vigorito A, Antonio de Souza C, Rocha EM. Comparison of diagnostic tests in distinct well-defined conditions related to dry eye disease. *PLoS One*. 2014; 9(5):e97921. <https://doi.org/10.1371/journal.pone.0097921> PMID: [24848115](#)
52. Mengher LS, Pandher KS, Bron AJ. Non-invasive tear film break-up time: sensitivity and specificity. *Acta Ophthalmol*. 1986; 64(4):441–444.
53. Chui WS, Cho P, Brown B. Soft contact lens wear in Hong Kong Chinese: predicting success. *Ophthal Physiol Opt*. 2000; 20(6):480–486.
54. Zhang X, Chen Q, Chen W, Cui L, Ma H, Lu F. Tear dynamics and corneal confocal microscopy of subjects with mild self-reported office dry eye. *Ophthalmology*. 2011; 118(5):902–907. <https://doi.org/10.1016/j.ophtha.2010.08.033> PMID: [21146227](#)
55. Best N. (2013). The predictive ability of clinical tests for contact lens induced dry eye. (Doctoral Thesis). Aston University, Birmingham, England. [https://research.aston.ac.uk/portal/en/theses/the-predictive-ability-of-clinical-tests-for-contact-lens-induced-dry-eye\(faf401e5-63ca-443f-8d5c-8359df6cef79\).html](https://research.aston.ac.uk/portal/en/theses/the-predictive-ability-of-clinical-tests-for-contact-lens-induced-dry-eye(faf401e5-63ca-443f-8d5c-8359df6cef79).html)
56. Gulati A, Sullivan R, Buring JE, Sullivan DA, Dana R, Schaumberg DA. Validation and repeatability of a short questionnaire for dry eye syndrome. *Am J Ophthalmol*. 2006; 142(1):125–131. <https://doi.org/10.1016/j.ajo.2006.02.038> PMID: [16815260](#)
57. Kaido M, Uchino M, Yokoi N, Uchino Y, Dogru M, Kawashima M, Komuro A, Sonomura Y, Kato H, Kinoshita S, Tsubota K. Dry-eye screening by using a functional visual acuity measurement system: the Osaka study. *Invest Ophthalmol Vis Sci*. 2014; 55(5):3275–3281. <https://doi.org/10.1167/iovs.13-13000> PMID: [24801509](#)
58. Oden NL, Lilienfeld DE, Lemp MA, Nelson JD, Ederer F. Sensitivity and specificity of a screening questionnaire for dry eye. In: *Lacrimal Gland, Tear Film, and Dry Eye Syndromes 2* 1998 (pp. 807–820). Springer US.
59. Khanal S, Tomlinson A, McFadyen A, Diaper C, Ramesh K. Dry eye diagnosis. *Invest Ophthalmol Vis Sci*. 2008; 49:1407–1414. <https://doi.org/10.1167/iovs.07-0635> PMID: [18385057](#)
60. Tan LL, Morgan PB. Diagnostic ability of conventional dry eye tests and their correlation with ocular surface temperature. *J Clin Exp Ophthalmol*. 2016; 7(4):587.
61. Farris RL, Gilbard JP, Stuchell RN, Mandel ID. Diagnostic tests in keratoconjunctivitis sicca. *CLAO J*. 1983; 9(1):23–28. PMID: [6839430](#)
62. Albach KA, Lauer M, Stolze HH. Zur Diagnose der Keratoconjunctivitis sicca bei rheumatoider Arthritis. Die Wertigkeit verschiedener Tests. *Der Ophthalmologe*. 1994; 91:229–234.
63. Sullivan DA, Stern ME, Tsubota K, Dartt DA, Sullivan RM, Bromberg BB (Eds.). *Lacrimal Gland, Tear Film, and Dry Eye Syndromes 3: Basic Science and Clinical Relevance Part B*. *Adv Exp Med Biol*, Vol. 506. Springer US, 2002.
64. Versura P, Frigato M, Cellini M, Mulè R, Malavolta N, Campos EC. Diagnostic performance of tear function tests in Sjögren's syndrome patients. *Eye*. 2007; 21:229–237. <https://doi.org/10.1038/sj.eye.6702204> PMID: [16397619](#)
65. Lemp MA, Bron AJ, Baudouin C, Denitez 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; 151:792–798. <https://doi.org/10.1016/j.ajo.2010.10.032> PMID: [21310379](#)
66. Yeo S, Aung HT, Tong L. The Association of Dry Eye Symptoms with Socioeconomic Factors and quality of Life. *J Clin Res Ophthalmol*. 2014; 1(1):006–013.
67. Amparo F, Schaumberg DA, Dana R. Comparison of two questionnaires for dry eye symptom assessment. *Ophthalmol*. 2015; 122(7):1498–1503.
68. 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(1):94–101. <https://doi.org/10.1001/archophthalmol.2009.356> PMID: [20065224](#)