Open Access

Original Article

Agreement between Subjective and Objective Parameters Using Meibography for Assessment of Dry Eye Syndrome

Ijlal Taimoor, Syed Abid Hassan Naqvi, Taimoor Ashraf Khan, Saquib Naeem, Waseem Yousaf, Muhammad Usman Ghani

Department of Ophthalmology, Armed Forces Institute of Ophthalmology/National University of Medical Sciences (NUMS) Rawalpindi Pakistan

ABSTRACT

Objective: To determine the agreement between subjective and objective parameters using meibography.
Study Design: Cross-sectional observational study.
Place and Duration of Study: Armed Forces Institute of Ophthalmology (AFIO), Rawalpindi Pakistan from August 2021 till February 2022.
Methodology: 100 patients were included in the study. Non-contact infra-red meibography was performed on each patient. Subjective assessment was done using the four-grade and five-grade scales while objective assessment was done by applying computerized grading.
Results: A total of one hundred (n=100) patients participated in the study. Out of total, 60% patients were male and 40% were females, with a mean age 42.25±14.9 years. There was statistically significant difference was seen Session I OII vs Session II OII as *p*-value <0.001 in Grade-4, Session-I OIII vs Session-II OIII as *p*-value <0.001 in Grade-5. There was not statistically significant difference seen in Computerized Grading (100-grade scale) as *p*-value >0.05. Correlations between observers was best with the 5-grade scale followed by the computerized 100-grade scale and 4-grade scale.
Conclusions: We found that the reliability of the 5-grade scale was poorer to that of the 4-grade scales. We also concluded that computerized grading offers a better intra and inter-observer assessment.

Keywords: Computerized grading system, Dry eye disease, Meibography, Objective grading system, Subjective grading system.

How to Cite This Article: *Taimoor I, Naqvi SAH, Khan TA, Naeem S, Yousaf W, Ghani MU. Agreement between Subjective and Objective Parameters Using Meibography for Assessment of Dry Eye Syndrome. Pak Armed Forces Med J 2023; 73(Suppl-2): S374-377 DOI: https://doi.org/10.51253/pafmj.v73iSUPPL-2.9858*

INTRODUCTION

The tear film helps to keep the human eye moist. The three layers that make up the tear film are the outside lipid layer, which is created by the meibomian glands, the middle aqueous layer, which is maintained by the lacrimal gland, and the inner mucous layer, which is produced by the conjunctival goblet cells. Meibum, which is secreted by the meibomian glands, aids the outer layer of the tear film in preventing the aqueous layer from evaporating and prevents the eye from drying out. Additionally, it functions as a surfactant, allowing the tear film to spread.[1,2]

Meibomian gland dysfunction (MGD) causes poor quality and amount of meibum secretion, affecting the composition and purity of the tear film's outer (lipid) layer. As a result, there is an increase in tear film instability, inflammation, tear hyperosmolarity, and ocular surface injury. This interplay of all these adverse outcomes leads to a complex syndrome called Dry Eye Syndrome varying in severity over a vast spectrum of symptoms and signs.[3,4]

Correspondence: Dr Ijlal Taimoor, Department of Ophthalmology, Armed Forces Institute of Ophthalmology, Rawalpindi Pakistan
Received: 27 Jan 2023; revision received: 26 Jul 2023; accepted: 27 Jul 2023

Meibography is a specialized imaging technique that was first reported by Tapie *et al.* in 1977 who visualized meibomian glands on Infrared (IR) photography with the help of infrared light.[5] It was officially named as Meibography by Maters *et al.* in 19916. Maters *et al.* also significantly improved the traditional technique by using video IR meibography which greatly accelerated the imaging of meibomian glands.[7] Meibography allows us to visualize the morphology of the meibomian glands in vivo. Before the advent of Meibography, meibomian glands could only be visualized ex vivo histologically. Since then, various types of Meibographic techniques have been developed and can be classified either as contact or non-contact. Contact meibography provides imaging using a specialized camera that transilluminates the eyelid after the skin is everted with the help of a light probe. This method relies on operator expertise, is time consuming and causes discomfort to the patient.[5,8] These problems led to the development of non-contact meibography which was first described in literature by Arita *et al.* in 2008.[9] This technique images a digitally everted eyelid with the help of a slit lamp having an IR filter that is coupled to a video camera. The fact that it does not use a light probe as compared to contact meibography

results in it being much better tolerated by the patients, faster and simple to use. A greater surface area of the eyelid is able to be imaged. Although other newer techniques such as Laser confocal meibography,[10] and Optical Coherence Tomographic Meibography (OCTM),[11] have the advantages of visualizing microscopic structures and microenvironment of the meibomian glands along with the ability to perform a volumetric analysis of the morphology of the meibomian glands respectively, non-contact IR meibography remains the mainstay and most popular technique.

Normal Meibomian glands are typically visualized as hypo-illuminescent grape like clusters while the underlying tarsus and ducts appear as hyperilluminescent.[8] In comparison, abnormal meibomian glands show dilated ducts and enlarged, tortuous glands. However, imaging the meibomian glands alone by itself is not enough if those images cannot be interpreted. As such several grading systems have been proposed which would help in documenting the progression and treatment response in MGD. Meiboscore and Meibograde systems were first described by Arita *et al.* and Call *et al.* respectively.[9,12] Similar variations of the 4-grade scale have also been described.[13,14] Recently, a 5-grade scale and a computerized 100-grade scale have been proposed.[15,16]

The objective of our study was to compare whether a 5-grade scale provides a better intra and inter-observer repeatability as compared to existing 4-grade scales and whether objective grading was superior to subjective grading.

METHODOLOGY

This was a cross-sectional study conducted at Armed Forces Institute of Ophthalmology (AFIO), Rawalpindi, Pakistan which is a specialized tertiary centre for people with eye disease. The study was conducted from August 2021 to February 2022. Non-probability convenient sampling technique was used. Ethical approval was granted by the Ethics Review Committee (ERC) at AFIO, (ERC dated: 22 December 2020). A sample size of 75 was calculated by Open Epi Online software keeping reference prevalence of Dry eye disease to be 5%.[1] We included all the patients fulfilling the inclusion exclusion criteria during the study period.

Inclusion Criteria: We included patients diagnosed with Meibomian Gland Dysfunction.

Exclusion Criteria: We excluded patients presenting with chronic diseases such as diabetes and hypertension, any history of ophthalmic surgery, any recent ocular infections which could change the curvature of the ocular surface, any seasonal allergies which had ocular manifestations, use of any topical eye drops, any current or previous use of contact lenses and any recent ocular trauma. Furthermore, images that were out of focus, improperly illuminated and those that did not capture the whole eyelid were discarded as they would have led to biased results.

Data was collected by a single investigator. All the participants signed a written informed consent proforma for enrolment into the study. A single investigator performed Non-contact infra-red meibography on each patient with the help of portable non-contact meibograph (PCNM). Hundred images were selected randomly (1 from each patient). Meibography images were classified according to three different scales: a 4-grade scale, a 5-grade scale, and a computerized 100-grade scale by three different observers (O1, O2, O3). The 2nd session was repeated the following day. Observers were masked against sessions as well as each other to eliminate any bias. Subjective assessment was done using the four-grade and five-grade scales while objective assessment was done by applying computerized grading and measuring the area of MG loss using Image J software.

RESULTS

A total of one hundred (n=100) patients participated in the study. Out of total, 60% patients were male and 40% were females. The ages of participants ranged from a minimum 18 years to a maximum 66 years with a mean age 42.25±14.9 years. Observed Meibomian gland (MG) loss ranged from 0-4 on the 5-grade scale and from 0-3 on the 4-grade scale. There was statistically significant difference seen in Session-I OII vs Session-II OII as p-value <0.001 in grade-4, Session-I OIII vs Session-II OIII as p-value <0.001 in grade-5. There was not statistically significant diffe-rence seen in Computerized Grading (100-grade scale) as p-value >0.05 shown in Table-I.

Subjective grading Intra-observer agreement: 95% confidence interval (CI) for O1, O2 and O3 are all similar. CI was poorer for all 3 observers for the 5-grade scale when compared to the 4-grade scale (Table-II).

Correlation between the observers was better than the 4-grade scale but worse than the 5-grade scale. The detail of Correlation shown in Table-III.

DISCUSSION

While Meibography techniques are well established and have been well documented, no officially

Table-I: Standard Deviation and Mean classified according to various grading systems

	4-Grade Scale		5-Grade Scale		Computerized Grading (100-grade scale)	
	Mean±SD	p-value	Mean±SD	p-value	Mean±SD	p-value
Session-I OI	1.79±0.31	0.789	2.06±0.24	0.134	0.29±0.13	0.746
Session-II OI	1.80±0.24		2.11±0.23		0.30±0.28	
Session-I OII	1.90±0.81	<0.001	2.08±0.17	0.712	0.34±0.14	0.159
Session-II OII	1.00±0.89		2.09±0.21		0.31±0.16	
Session-I OIII	1.82±0.91	1.000	2.30±1.06	<0.001	0.31±0.12	0.656
Session-II OIII	1.80±0.28		2.2±1.11		0.32±0.19	

Table-II: 95% confidence interval (CI) (n=100)

Confidence Interval	4-Grade Scale	5-Grade Scale	Computerized Grading (100-grade scale)
O1: S1-S2	1.73–1.85	2.06–2.11	0.29–0.29
O2: S1-S2	1.75–1.85	2.06–2.15	0.28–0.29
O3: S1-S2	1.74–1.85	2.05–2.13	0.28–0.29
S2: O1-O2	1.74–1.85	2.05–2.14	0.28–0.29
S2: O1-O3	1.76–1.87	2.05–2.12	0.28–0.29
S2: O2-O3	1.74–1.86	2.04–2.10	0.28–0.29

Table-III: Correlation between sessions and observers according to various grading systems (n=100)

Correlation	4-Grade Scale		5-Grade Scale		Computerized grading (100-grade scale)	
	r-value	*p*-value	r-value	*p*-value	r-value	*p*-value
OI: SI-SII	-0.124	0.219	0.228	0.023	0.053	0.471
O2: S1-S2	0.205	0.041	-0.248	0.124	0.006	0.956
O3: S1-S2	0.460	< 0.001	0.120	2.33	0.138	0.172
S2: O1-O2	0.009	0.936	0.113	0.64	0.027	0.792
S2: O1-O3	0.944	<0.001	0.326	<0.001	0.097	0.338
S2: O2-O3	0.287	0.004	0.541	< 0.001	0.028	0.76

recognized grading scales exist for interpreting these images. Pflugfelder *et al*. proposed a 4-grade scale based on gland dropout where "Grade-0=No gland dropout, Grade-1=33% gland dropout, Grade-2=34-66% gland dropout and Grade-3=>66% gland dropout".[14] Similarly, Nichols *et al.* proposed a 4-grade scale based on loss of partial meibomian glands where "Grades 0-3 corresponded to no partial glands, <25% partial glands, 25–75% partial glands and >75% partial glands respectively".[13] Arita's 4-grade "meiboscore" method is based on percentage of meibomian gland loss area where grades 0–3 correspond to no loss of meibomian glands, loss less than 1/3rd of total meibomian gland area, loss between 1/3rd and 2/3rd of total meibomian gland area and loss >2/3rd of total meibomian gland area respectively.[9] However, recently Pult *et al.* proposed a newer 5-grade scale which has been postulated as being more sensitive regarding treatment efficacy and progression of severity of MGD. Furthermore, Pult *et al*. also reported on a computerized 100-grade scale by comparing it to subjective grading.[17] The mean age of patients was 42±14.9 years. 60% were males in our study population. Significant differences were found between sessions (p<0.001) and observers in our study.

Pult *et al.* in their study found no significant differences between sessions and observers which directly contradicted the findings of our study.[16] For the subjective grading scales, they also reported 95% CI of O1 and O3 as being better than O2 while our study reported 95% CI as similar for O1 and O2 compared to O3 for the subjective grading scales. However, limits of agreement in both the studies were same for the five-grade and the four grade scales. The major difference in both studies was that 95% CI was found to be superior for 5-grade scale versus the 4-grade scale by Pult *et al.* This phenomenon was not observed in our results as 95% CI was poorer amongst all observers for the 5-grade scale in comparison to the 4-grade scale. This contradicts Pult *et al.* major observation of the 5-grade scale having a much better intra-observer repeatability as opposed to the 4-grade scale. Our study also identified that the CI sessions correlated well for all grading systems for observer 3 while also correlating well for the five-grade and four-grade systems for observer.[2] Similar results were seen by Pult *et al.* We also did not find inter-observer agreement to be superior for the 5-grade scale when compared to the 4-grade scale which was reported by the German study. The only results in our study that seemed to back the findings of Pult *et al.* were those of computerized grading. We observed that the 95% CI were best for computerized grading for both intra and inter observer agreements. Similarly, intra-observer correlations were best for the computerized grading as well. However, surprisingly the inter-observer correlation was best for the 5-grade scale followed by the computerized grading and the 4-grade scale.

All in all, while our study does support the findings of the published study by Pult *et al.* in terms of computerized grading being the best scale to evaluate MGD, we did not find the 5-grade scale to be a better scale as compared to the 4-grade scale as suggested by Paul *et al.* The reasons given by Pult et al. were fourfold: 1) better intra-observer repeatability was obser-

ved for the 5-grade scale, 2) five-grade scale showed a better inter-observer agreement, 3) the 5-grade scale gave more consistent increments and 4) the kappa statistics supported a higher reliability for the 5-grade scale. Our results directly contradicted reasons 1 and 2 and showed the opposite. Our results did support reason 3, as inter-observer correlation was highest for five-grade scale and is largely due to the ease of classifying the images due to smaller increments. Unfortunately, we are unable to verify the fourth reason as our study did not calculate the kappa statistic.

LIMITATIONS OF STUDY

The Limitation of our study is the absence of kappa statistic. The simple kappa statistic represents the true measure of agreement while the weighted kappa statistic considers the degree of disagreement between the various sessions of the observers. The absence of these means that we are unable to comment on the reliability of the various grading scales. Furthermore, we believe that proper training of the observers are of the utmost importance to eliminate any possible bias in the interpretation of the images. Furthermore,, we didn't document the operative difficulty and patient discomfort during the procedure. Previous data shows an easy learning curve and limited patient inconvenience during meibography study.[18] However, we believe Artificial Intelligence based on its technology will eliminate the need for even having an observer interpreting the images using a scale. As such it will eliminate observer bias in future.

CONCLUSION

While we have seen the evolution of Meibographic Techniques over the decades, scaling grades for interpretation of those images have not evolved at the same pace. Several have been proposed. We found the reliability of the 5-grade scale to be poorer to that of the 4-grade scales which is contrary to the findings reported in an earlier study. However, we also think that computerized grading offers a better intra and inter-observer assessment as reported by several studies.

Conflict of Interest: None.

Author's Contribution

Following authors have made substantial contributions to the manuscript as under:

IT & SAHN: Data acquisition, data analysis, drafting the manuscript, critical review, approval of the final version to be published.

TAK & SN: Conception, study design, approval of the final version to be published.

WY & MUG: Data interpretation, critical review, approval of the final version to be published.

Authors agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

REFERENCES

1. Salmon J. Kanski's Clinical Ophthalmology E-Book: A Systematic Approach: Elsevier Health Sciences; 2019, [Internet] avaiable at: https://shop.elsevier.com/books/kanskis-clinical-ophthalmology/salmon/978-0-7020-7711-1
2. McCulley JP, Shine WE. Meibomian gland function and the tear lipid layer. The ocular surface 2003; 1(3): 97-106. https://doi: 10.1016/s1542-0124(12)70138-6.
3. Chhadva P, Goldhardt R, Galor A. Meibomian gland disease: the role of gland dysfunction in dry eye disease. Ophthalmol 2017; 124(11): S20-S6. https://doi: 10.1016/j.ophtha.2017.05.031.
4. Kim HM, Eom Y, Song JS. The relationship between morphology and function of the meibomian glands. Eye & Contact Lens 2018; 44(1):1-5. https://doi: 10.1097/ICL.0000000000000336.
5. Tapie R. Etude biomicroscopique des glandes de meibomius. Ann Oculistique 1977; 210: 637-648.
6. Mathers WD, Shields WJ, Sachdev MS, Petroll. Meibomian gland dysfunction in chronic blepharitis. Cornea 1991; 10(4): 277-285. https://doi: 10.1097/00003226-199107000-00001.
7. Mathers WD, Daley T, Verdick R. Video imaging of the meibomian gland. Archives of ophthalmology 1994; 112(4): 448-449. https://doi: 10.1001/archopht.1994.01090160022008.
8. Jester JV, Rife L, Nii D, Luttrull JK. In vivo biomicroscopy and photography of meibomian glands in a rabbit model of meibomian gland dysfunction. Investigative ophthalmology & visual science 1982; 22(5): 660-667.
9. Arita R, Itoh K, Inoue K. Noncontact infrared meibography to document age-related changes of the meibomian glands in a normal population. Ophthalmol 2008; 115(5): 911-915.
10. Kobayashi A, Yoshita T, Sugiyama K. In vivo findings of the bulbar/palpebral conjunctiva and presumed meibomian glands by laser scanning confocal microscopy. Cornea 2005; 24(8): 985-988. https://doi: 10.1097/01.ico.0000160976.88824.2b.
11. Bizheva K, Lee P, Sorbara L, Hutchings N, Simpson T. In vivo volumetric imaging of the human upper eyelid with ultrahigh-resolution optical coherence tomography. J Biomed Opt. 2010; 15(4): 040508. https://doi: 10.1117/1.3475957.
12. Call C, Wise R, Hansen M. In vivo examination of meibomian gland morphology in patients with facial nerve palsy infrared meibography. Ophthalmic Plast Reconstr Surg 2012; 28(6): 396-400.
13. Nichols JJ, Berntsen DA, Mitchell GL, Nichols KK. An assessment of grading scales for meibography images. Cornea 2005; 24(4): 382-388. https://doi: 10.1097/01.ico.0000148291. 38076.59.
14. Pflugfelder SC, Tseng S, Sanabria O, Kell H, Evaluation of subjective assessments and objective diag-nostic tests for diagnosing tear-film disorders known to cause ocular irritation. Cornea 1998; 17(1): 38-56. https://doi: 10.1097 /00003226-199801000-00007.
15. Pult H, Riede-Pult BH, Nichols JJ. Relation between upper and lower lids' meibomian gland morphology, tear film, and dry eye. Optometry and Vision Science 2012; 89(3): E310-E5. https://doi: 10.1097/OPX.0b013e318244e487.
16. Pult H, Riede-Pult BH. An assement of subjective and objective grading of meibography images. Investigative Ophthalmology & Visual Science 2012; 53(14): 588-.
17. Pult H, Nichols JJ. A review of meibography. Optometry and Vision Science 2012; 89(5): E760-E9. https://doi: 10.1097/OPX.0b013e3182512ac1.
18. O'Dell L, Halleran C, Schwartz S. An assessment of subjective meibography image grading between observers and the impact formal gland interpretation training on inter-observer agreement of grading scores. Inv Ophth Vis Sci 2020; 6: 486