

# Reliability analysis of subjectively graded real-image pterygium based on its translucence appearance between young and experienced clinicians

Khairidzan Mohd Kamal<sup>2</sup>, Mohd Radzi Hilmi<sup>1,\*</sup>, Mohd Zulfaezal Che Azemin<sup>1</sup>, Mohd Hafidz Ithnin<sup>1</sup>

<sup>1</sup>Department of Optometry and Visual Sciences, Kulliyah of Allied Health Sciences, Inter-national Islamic University Malaysia (IIUM), Kuantan, Pahang, Malaysia

<sup>2</sup>Department of Ophthalmology, Kulliyah of Medicine, International Islamic University Malaysia (IIUM), Kuantan, Pahang, Malaysia

\*Corresponding author: Mohd Radzi Hilmi, Department of Optometry and Visual Science, Kulliyah of Allied Health Sciences, IIUM, Jalan Sultan Ahmad Shah, 25200 Bandar Indera Mahkota, Kuantan, Pahang, Malaysia

Received date: 16 September, 2021 | Accepted date: 2 October, 2021 | Published date: 11 October, 2021

Citation: Kamal KM, Hilmi MR, Azemin MZC, Razak ZA, Ithnin MH (2021) Repeatability and Reproducibility of Pterygium Grading Based on Clinical Translucence Appearance. J Ophthalmic Res Vis Care 1(1). doi <https://doi.org/10.54289/JORVC2100101>

Copyright: © 2021 Kamal KM, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

## Abstract

**Introduction:** Pterygium is commonly subjectively evaluated via anterior segment assessments during slit-lamp examination. Thus, this assessment prones individual variations as it requires subjective grading and adequate experience to ensure consistency of diagnosis and management.

**Purpose:** This study aimed to compare the reliability of subjectively graded real-image pterygium based on its translucence appearance between experienced clinicians.

**Design of study:** Prospective randomized study.

**Materials and methods:** Thirty (30) primary pterygium images from 30 pterygium patients were captured in a standardized magnification, illumination and formatting setting as previously de-scribed. All images were projected using PowerPoint presentation™ on a liquid crystal display (LCD) monitor with standard resolution. Two clinicians with different levels of experiences act as a grader and grade all images based on reference images provided. For reliability testing, intra-grader assessment was repeated twice with different sequences at least a month apart between each session. Both clinicians were given a set of 30 randomized pterygium images for all sessions. Reliability testing was evaluated using paired T-test and independent T-test.

**Results:** Descriptive analysis revealed observer 1 obtained mean grade of 2.19 (SD = 0.670) and 2.23 (SD = 0.713) for session 1 and 2 respectively. Observer 2 obtained 2.04 (SD = 0.853) and 2.08 (SD = 0.894) for session 1 and 2 respectively. Paired T-test showed the difference for both observers in both sessions were not statistically significant (P = 0.776 and P = 0.583) respectively. Reproducibility testing using Independent T-test results showed the difference between observers was not statistically significant (P = 0.275). Subjectively graded pterygium clinical grading based on its translucence appearance was repeatable and reproducible.

**Conclusion:** These findings could serve as a basis for future work on to evaluate performance of pterygium clinical grading based on its morphology with different levels of experience and larger number of samples.



**Keywords:** terygium, Cornea; Conjunctiva; Grading Scale; Subjective Grading

**Abbreviations:** UV: Ultraviolet, IIUM: International Islamic University Malaysia, IREC: International research ethical committee, HD: high definition, CSO: Costruzione Strumenti Oftalmici, JPEG: Joint Photographic Experts Group, SPSS: Predictive analytics soft-ware.

## Introduction

Pterygium is one of the most common ocular surface disorder especially in hot climate regions. Pterygium can be defined as defined as a fibrovascular encroachment of fibrous tissue of conjunctiva which originates from bulbar conjunctiva and progresses towards the central cornea [1]. It is an established fact that pterygium has higher prevalence in tropical countries which is also known as ‘Pterygium belt’ which is located between 30° north and south of the equator. Ultraviolet (UV) exposure has been dubbed as the main risk factor for its growth. It is well-known fact that pterygium causes significant reduction in visual acuity [2,3] and contrast sensitivity [4,5] due to changes on corneal curvature [6-8] as it progresses.

Pterygium has been evaluated in various approaches, with its length in comparison with the corneal size is the common one. However, the impact of pterygium on cornea somehow not necessarily depending on its length. Thus, this suggest new approach is needed. In 1997, Donald Tan and his co-workers [9] had proposed a grading which classifies pterygium based on its translucency into three groups: Grade I - atrophy, Grade II - intermediate and Grade III - fleshy. This grading has been widely used since its inception [1-5]. However, there is a lack of literature that addresses the reproducibility of the clinical grading between clinicians. Therefore, this study aims to determine the reproducibility of subjectively graded pterygium based on its translucency appearance using real images of pterygium.

## Materials and Methods

A cross-sectional study was conducted to evaluate the reliability of subjectively graded real-image pterygium based on its translucence appearance between two clinicians with different levels of experience. Sixty (60) primary pterygium images from 60 eyes were captured in a standardized magnification, orientation and illumination [10-12] by a single operator. The inclusion criteria include established

diagnosis of primary pterygium, both genders were included with age ranges from 20 to 70 years and free from any history of ocular trauma, ocular surgery and history of contact lens wear. Patients with significant ocular surface diseases such as recurrent pterygium, corneal opacity or irregularity due to diseases other than pterygium [1,3,13,14]. Participants were recruited from a University eye-specialist ophthalmic centre in order to obtain images which would display a wide range of severity of pterygium patients. Ethical approval was obtained by the International Islamic University Malaysia (IIUM) research ethical committee (IREC) (IIUM/310/G13/4/4-125) and this study conformed to the tenets of the Declaration of Helsinki. All participants were given adequate information regarding methods and risks of this study. Written consent was obtained prior to acquisition of images.

All images were acquired using a standardized protocol at 10 - 16X magnification, under diffuse white illumination and oriented at 30 - 35° to the corneal surface [1,3]. A high definition (HD) digital camera was used to capture the images through a SL-990 slit-lamp biomicroscope (Costruzione Strumenti Oftalmici (CSO), Firenze, Italy). All images were stored and processed using an image processing software, Phoenix™ version 1.2 (Costruzione Strumenti Oftalmici (CSO), Firenze, Italy) in the form of Joint Photographic Experts Group (JPEG) files.

All primary pterygium images were carefully examined based on Donald Tan’s grading by placed into a PowerPoint presentation (Microsoft Corporation, Redmond, WA, USA) and displayed on a monitor with 17- inch 1280 X 1024 resolution [3]. All images were randomly sorted using randomization software [15]. The images were matched and graded based on reference images as described in previous study [9] as shown in Figure 1 by two clinicians with different levels of experiences. The young clinician was defined as a clinician with less than 5 years of exposure in ocular surface field, while the experienced clinician was defined as a person

with more than 5 years' experience in respective fields. The experienced clinician is noted as observer 1 while the young clinician was noted as observer 2.

Statistical analyses were performed using IBM SPSS (Predictive analytics software) (Version 24, SPSS Inc., Chicago, IL, USA). Descriptive statistics were used to determine the mean and standard deviation (SD) of image graded by both observers. Reliability analyses were evaluated

via Paired and Independent T-test. Paired T-test was employed to assess the difference in measurement between two sessions. For intra-observer reliability testing, the process was repeated twice with different sequences at least a month apart between each session. The differences between two (2) observers were evaluated using Independent T-test. Both grading sessions were done simultaneously, but on different occasions. The alpha significance level was set at  $P < 0.05$ .

**Table 1.** Descriptive analysis results (n = 60)

| Variables  | Session 1 (mean ± SD) | Session 2 (mean ± SD) |
|------------|-----------------------|-----------------------|
| Observer 1 | 2.19 ± 0.670          | 2.23 ± 0.713          |
| Observer 2 | 2.04 ± 0.853          | 2.08 ± 0.894          |

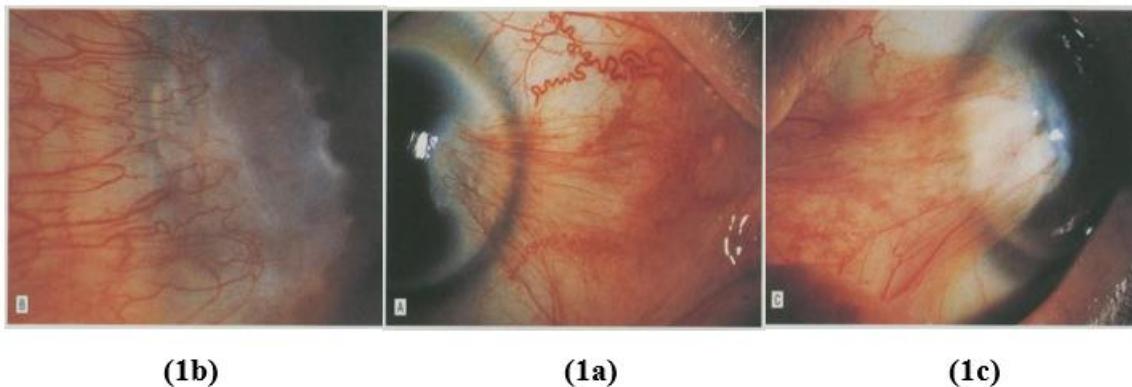
SD: standard deviation

**Table 2.** Repeatability and reproducibility assessment of Tan's pterygium clinical grading (n = 60)

| Variables  | Session 1 (mean ± SD) | Session 2 (mean ± SD) | P value | (95 % CI)       |
|------------|-----------------------|-----------------------|---------|-----------------|
| Observer 1 | 2.19 ± 0.670          | 2.23 ± 0.713          | 0.776*  | (-0.214, 0.190) |
| Observer 2 | 2.04 ± 0.853          | 2.08 ± 0.894          | 0.583*  | (-0.321, 0.111) |

\*: Based on Paired T-test,  $P = 0.05$  was set as the level of significance

#: Based on Independent T-test,  $P = 0.05$  was set as the level of significance



**Figure 1.** Reference images of Tan's pterygium clinical grading, (a) Grade I. (b) Grade II. (c) Grade III.

## Results

Based on thirty (30) primary pterygium, descriptive analysis revealed observer 1 obtained a mean grade of 2.19 (SD = 0.670) for session 1, and mean grade of 2.23 (SD = 0.713) for session 2. Observer 2 obtained 2.04 (SD = 0.853) for session 1, and mean grade of 2.08 (SD = 0.894) for session 2. Descriptive analysis results were summarized in **Table 1**. Paired T-test results showed the difference for both observers were not statistically significant for observer 1 ( $P = 0.776$ , 95

% CI: -0.214, 0.190) and observer 2 ( $P = 0.583$ , 95 % CI: -0.321, 0.111). Reproducibility testing using Independent T-test results showed the difference between observers was not statistically significant ( $P = 0.275$ , 95 % CI: -0.256, 0.364). Both repeatability and reproducibility findings were summarized in **Table 2**.

## Discussion

The purpose of this study was to evaluate the perceptual relationship between two clinicians with different levels of



experience in estimating the translucence appearance of pterygium based on reference images by Tan's clinical grading [9]. Previous works have shown that reproducibility for subjective clinical grading is frequently limited [16-19], due to variations between individuals. Study by Fieguth and Simpson (2002) has proved that individual variations can extend to approximately 50 %. In this study, the difference in the level of experiences could be translated based on the standard deviation which is wider in less-experienced clinicians compared to experienced clinicians.

The difference in inconsistency of measurement could happen due to various reasons such as lack of standardization of reference images, variations in methodology, technology and experience of individuals involved in assessing its reliability. With incorporations of reference image as the benchmark of a clinical grading and a standardized protocol, inter-observer performance improved, hence reducing individual variations and bias. Recent work [3] has demonstrated by inclusion of reference image in subjectively grading pterygium redness, the clinical grading performance is promising. Previous works [3,20,21] had demonstrated that subjective grading is permissible and reliable provided a standard protocol was employed. This study proves that using a static image as a reference image with an approximate experience between observers, the repeatability and reproducibility are very good. However, it is worth to re-evaluate the reliability of this grading on a bigger scale in terms of more pterygium images, with inclusion of variations in experience/training factor involved. Proper training could further reduce individual variations.

## Conclusion

Different levels of clinical exposure could provide differences in reliability of the grading. However, it is worth noting that subjective grading is reproducible and repeatable.

## References

1. Hilmi MR, Khairidzan MK, Azemin MZC, Azami MH, Ariffin AE. (2019) Corneo-ptyerygium Total Area Measurements Utilizing Image Analysis Method. *J Optom*, 12(4): 272-277.

2. Norazmar NA, Hilmi MR, Khairidzan MK, Azemin MZC, Maruziki NN, et al. (2019) Variation in Oculovisual Function Between Unilateral Primary Pterygium and Normal Adults Utilizing Computerized M&S Smart System Ii. *Int J Allied Health Sci*. 3(2):643-648.
3. Mohd Radzi H, Mohd Zulfaezal CA, Khairidzan MK, Mohd Izzuddin MT, Norfazrina AG, et al. (2017) Prediction of changes in visual acuity and contrast sensitivity function by tissue redness after pterygium surgery. *Curr Eye* 42(6):852-856.
4. Hilmi MR, Khairidzan MK, Azemin MZC, Azami MH, Ariffin AE. (2018) Measurement of Contrast Sensitivity Using the M&S Smart System II Compared with the Standard Pelli-Robson Chart in Patients with Primary Pterygium. *Makara J Health Res* 22(3):167-171.
5. Oh JY, Wee WR. (2010) The effect of pterygium surgery on contrast sensitivity and corneal topographic changes. *Clinical Ophthalmology*. Vol. 4: 315-319.
6. Hilmi MR, Nasir MS, Khairidzan MK, Azemin MZC, Maruziki NN, Norazmar NA, Musa NH. (2019). Topographic Changes on Anterior Corneal Curvature Between Unilateral Primary Pterygium and Normal Adults Utilizing Shape Factor and Mean Toric Keratometry Corneal Indices, *International Journal of Allied Health Sciences*, 3(2):734 - 740.
7. Hilmi MR, Musa NH, Khairidzan MK, Azemin MZC, Maruziki NN, et al. (2019) Changes in Apical Corneal Curvature in Unilateral Primary Pterygium and Normal Adults Using Simulated-K And Corneal Irregularity Measurement. *Int J Allied Health Sci*. 3(2): 588-594.
8. Hilmi MR, Maruziki NN, Khairidzan MK, Azemin MZC, Norazmar NA, Nasir MS, Musa NH. (2019). Topographic Changes as Predictor for Determining Anterior Corneal Curvature Stabilization Point Subsequent to Pterygium Excision Using Controlled Partial Avulsion Fibrin Glue Technique. *Int J Allied Health Sci*. August 3(2): 734-740.



9. Tan DT, Chee SP, Dear KGB, Lim ASM. (1997) Effect of pterygium morphology on pterygium recurrence in a controlled trial comparing conjunctival autografting with bare sclera excision. *Arch Ophthalmol.* 115(10): 1235-1240.
10. Azemin MZC, Hilmi MR, Kamal MK. (2014) Supervised pterygium fibrovascular redness grading using generalized regression neural network. In Fujita H (Ed). *New Trends in Software Methodologies, Tools and Techniques* (pp. 650–656). Amsterdam: IOS Press.
11. Che Azemin MZ, Hilmi MR, Mohd Tamrin MI, Mohd Kamal K. (2014). Fibrovascular redness grading using Gaussian process regression with radial basis function kernel. In *Biomedical Engineering and Sciences (IECBES), 2014 IEEE* pp. 113–116. IEEE.
12. Che Azemin MZ, Mohd Tamrin MI, Hilmi MR, Mohd Kamal K. (2015). GLCM texture analysis on different color space for pterygium grading. *ARPN J Eng Appl* 10: 6410-6413.
13. Che Azemin MZ, Gaffur NA, Hilm MR, Mohd Tamrin MI, Mohd Kamal K. (2016) Benchmarked pterygium images for human and machine graders. *J Engineering and Applied Sci.* 11(11): 2378-2382.
14. Che Azemin MZ, Mohd Tamrin MI, Hilmi MR, Mohd Kamal K. (2016) Inter-grader reliability of a supervised pterygium redness grading system. *Adv Sci Letters.* 22(10): 2885-2888.
15. Urbaniak GC, Plous S. (2013) *Research Randomizer.*
16. Efron N, Morgan PB, Katsara SS (2001) Validation of grading scales for contact lens complications. *Ophthalmic Physiol Opt.* 21(1):17-29.
17. Efron N, Morgan PB, Jagpal R. (2002) Validation of computer morphs for grading contact lens complications. *Ophthalmic Physiol Opt.* 22(4):341-349.
18. Fieguth P, Simpson T. (2002) Automated measurement of bulbar redness. *Invest Ophthalmol Vis Sci.* 43(2):340-347.
19. Murphy PJ, Lau JS, Sim MM, Woods RL. (2007) How red is a white eye? Clinical grading of normal conjunctival hyperaemia. *Eye* 21(5):633-638.
20. McMonnies CW, Chapman-Davies A. (1987) Assessment of conjunctival hyperemia in contact lens wearers. *Am J Optom Physiol Opt.* 64(4):246-250.
21. Efron N. (2000) Grading Scales. *Optician.* 219:44-45.