Repeatability and Reproducibility of Pterygium Grading Based on Clinical Translucence Appearance

Mohd Radzi Hilmi¹, ², *, Khairidzan Mohd Kamal ², Mohd Zulfazal Che Azemin ¹, Zulhilmi Abdul Razak ² and Mohd Hafidz Ithnin ¹

¹Department of Optometry and Visual Sciences, Kulliyyah of Allied Health Sciences, Inter-national Islamic University Malaysia (IIUM), Kuantan, Pahang, Malaysia
²Department of Ophthalmology, Kulliyyah of Medicine, International Islamic University Malaysia (IIUM), Kuantan, Pahang, Malaysia

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

Received date: 16 September, 2021 | Accepted date: 1 October, 2021 | Published date: 5 October, 2021

Citation: Hilmi MR, Kamal KM, 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/JORVC2100102

Copyright: © 2021 Hilmi MR, 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

Background: The goal of this study was to evaluate the repeatability and reproducibility of pterygium grading based on clinical translucence appearances and its reliability in clinical setting.

Methods: A total of 93 primary pterygium eyes from 93 participants who visited a tertiary ophthalmic clinic were recruited in this study. Two (2) ophthalmologists and two (2) optometrists evaluated and graded the 93 primary pterygium images in randomized fashion. Graders were instructed to utilise the clinical translucence appearance of pterygium to grade them into type I, II and III. Repeatability testing was done by a single expert by comparing grading of each image on two separate sessions, with one month interval between sessions. Reproducibility was tested by comparing the grading obtained by both experts and optometrists.

Results: Paired and independent T-test results showed no significance difference between graders for both experts and optometrists’ group (all \( P > 0.05 \)). Intra-grader and inter-grader reliability estimates were high with intraclass correlation ranging from 0.97 to 0.98 for both groups.

Conclusions: Pterygium clinical grading based on its translucence appearance is reliable and repeatable in clinical setting, easily to be graded, interpreted, and recommended for clinicians with different levels of experience.

Keywords: pterygium; translucency; fleshiness; clinical grading; reliability

Abbreviations: UV: Ultraviolet, IIUM: International Islamic University Malaysia, IREC: research ethical committee, SPSS: Predictive analytics software, ICC: intraclass correlation coefficients, LoA: limits of agreement.

Introduction

Pterygium is an abnormal benign lesion which originates from the bulbar conjunctiva and progresses towards the central cornea [1-3]. It is an established fact that prevalence of pterygium is closely related to ultraviolet (UV) ray exposures [4-6]. Pterygium is also known as one of the established factors which contribute to induced corneal astigmatism [7-12]. However, to date, it is difficult to predict its progression due to its non-specific tissue growth in nature.
Clinically, there are several grading’s been employed in assessing pterygium. Pterygium is assessed based on several approaches such as via its morphology [13], extension or length [7,8,14,15], its size [16-21] and based on its encroachment relative to the corneal size [7,8,15,22]. Although these approaches are clinically relevant, the main lacking in these gradings are lack of consistency in measurement as they were subjectively measured. Hence, individual variation is an issue which could lead to subjective bias. In 1997, Donald Tan and his team [13] had proposed a classification of pterygium which is based on its clinical translucence appearance (Figure 1).

They [13] proposed a clinical grading which was based on three (3) types or grades known as atrophy, intermediate and fleshy. The framework of this theory was the loss of translucency of pterygium tissue is related to increased fleshiness that could represent the fibrous vascular component of pterygium. The authors [13] divided these classifications into three (3) grades as described below:

- Grade-I (atrophy) - denoted as pterygium which episcleral vessels underlying the body of pterygium were unobscured and clearly distinguished (Figure 1a).
- Grade-II (intermediate) - denoted as pterygium which episcleral vessel details were indistinctly seen or partially obscured (Figure 1b).
- Grade-III (fleshy) - denoted as pterygium which episcleral vessels underlying the body of pterygium were totally obscured (Figure 1c).

To the best of our literature search, no study has been done in proving the repeatability and reliability of Tan’s classification of pterygium. Hence, this study aims to establish the reliability and repeatability of the subjectively measured Tan’s clinical grading of pterygium utilizing two (2) ophthalmologists and two (2) optometrists with different years of experience in the field. This is important to ensure that individual variations in subjective clinical grading are kept minimum.

**Methods**

93 primary pterygium eyes from 93 patients were recruited in this study. This study applied digitized pterygium images and other data obtained during the baseline visit. Image acquisition, image analyses and its clinical grading were identical as in our previous work [23-26]. A group of 93 primary pterygium patients from 93 individuals who visit an eye-specialist were selected in order to display a wide range of severity of pterygium patients. All participants in this study were selected based on specific criteria as previously described [23-26]. Assessment, diagnosis and classification of primary pterygium were performed by a single ophthalmologist (KMK). The study was conducted according to recommendation of the tenets of the Declaration of Helsinki and approved by the International Islamic University Malaysia (IIUM) research ethical committee (IREC) (IIUM/310/G13/4-4-125). Written and informed consent was obtained from all participants prior to any procedures performed. Image acquisition and description of clinical grading were similar as previously described [23-26].

![Figure 1](image1.jpg)

**Figure 1.** Reference images of Tan’s pterygium clinical grading. (a) Grade I. (b) Grade II. (c) Grade III.
For repeatability testing, each grader (MRH, KMK, ZAR and MHI) was given two (2) sets of images consisting of 93 pterygium images in each set were randomized using randomization software [27]. Each set of images was tested in two (2) different sessions. The time interval between the first and second session was set at least a month apart. The results between sessions were compared to determine its individual repeatability. For reproducibility testing, each grader (MRH, KMK, ZAR and MHI) graded each of 93 images based on Tan’s classification of pterygium [13]. The set of images for both assessors were also randomized as described above. The redness appearance was graded by both corneal experts (KMK and ZAR) and experienced (MRH) and inexperienced optometrist (MHI) were based on the clinical scales and reference images as in previous work [13]. Figure 1 below showed the illustration of the reference images.

All 93 images were presented using PowerPoint™ presentation (Microsoft Corporation, Redmond, WA, USA) and displayed on a monitor with 17-inch 1280 X 1024 resolution. This is important to ensure the quantification and evaluation processes conform to normal human colour perception. Statistical analyses were performed using IBM SPSS (Predictive analytics software) (Version 19, SPSS Inc., Chicago, IL, USA). Paired T-test was employed to descriptively evaluate the difference between two sessions, while independent T-test was employed to evaluate difference between graders (both experts and between experienced-inexperienced clinicians). The intra-grader repeatability and inter-grader reproducibility were assessed using a combination of intraclass correlation coefficients (ICC) and limits of agreement (LoA) respectively.

Results

The analysis included 93 participants, with 50.5% (n = 47) were men. The mean and standard deviation (mean ± SD) of pterygium redness based on subjective clinical grading by experts for grader 1 (KMK) was 1.73 ± 0.61 for session 1 (baseline) and 1.70 ± 0.63 for session 2. Paired T-test results showed no significance difference between session 1 and 2 (P = 0.132). For grader 2 (ZAR), the mean ± SD of pterygium redness was 1.72 ± 0.66 for baseline and 1.73 ± 0.53 for session 2 (P = 0.158). Independent T-test results showed no significance difference between both graders (P = 0.596).

With regards to experienced-inexperienced grader 3 (MRH) and 4 (MHI), the mean and standard deviation (mean ± SD) of pterygium redness for grader 3 (MRH) was 1.75 ± 0.65 for baseline and 1.73 ± 0.55 for session 2 (P = 0.193). While for grader 4 (MHI), the mean ± SD of pterygium redness was 1.74 ± 0.66 for baseline and 1.73 ± 0.75 for session 2 (P = 0.216), with independent T-test results showed no significance difference between both graders (P = 0.623). Descriptive results were summarized in Table 1 below. Normality testing was evaluated using Shapiro-Wilk test, and the results showed normal data distribution.

In comparison between the corneal experts, the intra-observer agreement between initial and second grading 1 month apart was 0.98 (95% CI, 0.971 - 0.987; P < 0.001). Bland-Altman analysis showed narrow limits of agreement (LoA) of 0.67 which indicates good intra-observer agreement of the redness grading. Inter-observer agreement was 0.97 (95% CI, 0.964 - 0.984; P < 0.001) with Bland-Altman analysis showing a narrow LoA of 0.76. In comparison between experienced and inexperienced optometrists, the intra-observer agreement between initial and second grading 1 month apart was 0.97 (95% CI, 0.965 - 0.975; P < 0.001) with Bland-Altman analysis showed narrow limits of agreement (LoA) of 0.72. Inter-observer agreement was found 0.96 (95% CI, 0.964 - 0.978; P < 0.001) with LoA of 0.75. Table 2 below summarizes the intra and inter-grader reliability estimates based on expert subjective grading and between experienced and inexperienced optometrists which indicates excellent reliability, with ICC of 0.97 to 0.98 and 0.96 to 0.97 respectively.

Discussion

It is an established fact that subjective grading of pterygium is difficult and prone to subjective bias. This is due to the nature of pterygium evaluation itself which requires manual measurement of pterygium based on its horizontal length or its encroachment. However, lack of studies has been done in evaluating pterygium based on its translucency known as Tan’s classification of pterygium. Hence, this study aims to evaluate the agreement of Tan’s classification of pterygium
based on a subjectively graded method using two different clinicians with different experiences. It is important as standardization is the key in assessment, diagnosing and monitoring pterygium patients and in any ocular pathology. By having this, the evaluation would be less subjectively biased and less prone to individual variation [28-30]. Hence, error between observers can be minimized and improve its reliability.

The current study results showed that the intra and inter-grader reliability testing in subjectively measured the pterygium redness were excellent with all ICCs were more than 0.90. Bland-Altman analysis revealed narrow LoA for all tests. Narrow LoA of less than 1.0 indicates good agreement between observers [31]. Previous works [32,33] had employed approximately similar techniques in assessing agreement of clinical grading of the anterior eye. The differences were only in the description of the scales and its scales interval which made it more detailed and specific for each condition. Redness grading is important as it may be applied to describe different severity of primary pterygium [23] and also indirectly indicates presence of inflammatory factors which further promotes progression of a disease due to angiogenesis factor from microvessel density perspective [34-38].

There are few ways in assessing repeatability and reliability of a clinical grading such as by employing more images, hence less number of grader and vice versa. Both CCLRU proposed by Terry et al.[32] and the EFRON clinical grading proposed by Efron [33] employed the latter approach by labouring multiple and vast number of graders in evaluating clinical grading of the corneal complication. The drawback of this approach is it is time-consuming and standardization of time and location setting of the grading was also difficult to be made.

Other previous works [23-26,28,39] employed the less labour-intensive approach by employing more images in a standardized form and clinical setting using visual technology as demonstrated in this study. This approach proved to be better in terms of standardization as all images were presented in a uniform manner with standardization of illumination and brightness. However, it is worth noting that a possible confounding factor for this approach is how normal the colour vision status of the grader is. However, this can be easily solved by assessing the normality of colour vision of each grader as suggested in previous work [40].

**Conclusion**

Pterygium clinical grading based on its translucence appearance is reliable and repeatable in clinical settings and easily to be graded and interpreted.

**Disclosure**

The author and authors report no conflicts of interest in this work.

**References**


