CLINICAL AND EXPERIMENTAL
OPTOMETRY

RESEARCH PAPER

The development of measurement tools for prosthetic eye research

Clin Exp Optom 2012 DOI:10.1111/j.1444-0938.2012.00754.x

Keith Raymond Pine* MBA
Brian Sloan† FRANZCO
Robert John Jacobs* PhD
* Department of Optometry and Vision Science, New Zealand National Eye Centre, The University of Auckland, Auckland, New Zealand
† Department of Ophthalmology, New Zealand National Eye Centre, The University of Auckland, Auckland, New Zealand
E-mail: r.jacobs@auckland.ac.nz

Submitted: 4 October 2011
Revised: 31 January 2012
Accepted for publication: 7 February 2012

Background: The aim was to develop tools to measure the condition of ocular prostheses and the socket's response to prosthetic eyewear.
Methods: A novel staining technique for displaying deposits on prosthetic eyes was developed. Equal interval perceptual grading scales for measuring inferior palpebral conjunctival inflammation, and anterior and posterior stained surface deposits on prosthetic eyes were developed from 800 photographs of 43 volunteers. The photographs for each scale were chosen by the authors. A group of four ophthalmologists, three optometrists and three senior students was consulted about selection criteria and asked to position the photographs along a 1.5 m rule to determine equal intervals. Photographs judged not to represent exactly equal perceptual intervals were exchanged with others from the original pool. The final scales (a five-photograph scale for inflammation and two 11 photograph scales for deposits) were assessed for inter-rater reliability and test-retest reliability by groups of senior optometry students.
Results: Standard deviations for inter-rater reliability tests were 0.52 scale units for the inflammation scale, 0.99 for the anterior surface deposits scale and 1.03 for the posterior surface deposits scale. The standard deviation of the test-retest differences for inflammation was 0.6 scale units and for both anterior and posterior surface deposits it was 0.71.
Conclusions: A novel technique for displaying and measuring the intensity and extent of deposit formation on prosthetic eye surfaces has been described. The two equal interval perceptual grading scales that have been developed to quantify the extent of deposit formation together with the equal interval perceptual scale for grading severity of palpebral conjunctival inflammation will for the first time allow the effects of prosthetic eye wear to be evaluated. Further research to validate the scale for palpebral conjunctival inflammation in a clinical setting is recommended. The technique for staining deposits on prosthetic eyes is recommended for clinical practice.

Key words: conjunctiva, deposits, grading scales, inflammation, prosthetic eyes

The literature on prosthetic eyes is not well developed. Available information is focused on early issues surrounding eye removal, such as surgical procedures, the effects on patients' visual perception and the emotional impact of eye loss. A small number of researchers have investigated the anophthalmic socket's response to prosthetic eye wear but aside from a link between giant papillary conjunctivitis and prolonged prosthetic eye use[1] wearing behaviour has not been found to have any effect on conjunctival cytological features[2,3] or the flora of the socket.[4–6] Two recent studies[2,3] have investigated potential links between prosthetic eye removal and cleaning regimes and inflammation in anophthalmic sockets. Chang

Camera	Canon 1000D
Lens	Macro EF-S 60mm f/2.8 USM
Flash	Cannon Macro Ring Light MR-14X
Camera setting	Manual
Exposure time	1/125 second
Aperture size	F/32
Focus	Automatic
Picture style	Faithful
White balance	Flash
Sensitivity	ISO 400
Flash setting	Manual exposure
Flash output	1/16
Distance from sensor plane to the prosthesis for photography	22 cm–27 cm

Table 1. Camera specifications and photographic settings

Water	70–75%
Ethyl alcohol	18–20%
Food red 105 (Rose Bengal)	4%
D sorbitol	3%
Sodium carboxymethyl cellulose (CMC-Na)	2%
Butyl p-hydroxybenzoate	<1%
Flavouring	<1%
Sodium salicylate	<0.1%

Table 2. GC Corporation plaque-disclosing gel ingredients

and colleagues[2] used an independent ophthalmic pathologist to estimate inflammation on a 0–3 scale, while Kim and colleagues[3] used a four-point verbally descriptive scale for bulbar conjunctival inflammation and the four-level verbally descriptive scale for tarsal conjunctival inflammation of Saini, Rajwanshi and Dhar.[7] Both investigations failed to find any significant link between inflammation and care regime; however, it is possible that these scales were too coarse for small changes in inflammation to be noticed. Bailey and colleagues[8] recommended using finer than four-point grading scales and Chong, Simpson and Fong[9] showed that scales using reference photographs had better repeatability than verbally descriptive scales.

No finely spaced photographic grading scales designed specifically for the investigation of anophthalmic sockets appear to have been developed, although many scales exist for other purposes.[10–12]

Hurst, Mitchell and Douthwaite[11] created a photographic grading system for contact lens deposits, but to investigate prosthetic eyes an alternative method for displaying and measuring deposits needs to be developed. The nature and dynamics of deposition on prosthetic eyes is very different from deposition on contact lenses. Deposits revealed by the staining solution used in this study form over all prosthetic eye surfaces except perhaps for the inter-palpebral zone. Because the body of a prosthetic eye is opaque, only very thick deposits are visible unless they are stained. Contact lens deposits on the other hand exist in the inter-palpebral zone and can easily be seen because the material on which they form is transparent.

This study describes a technique for displaying deposits on prosthetic eyes. It aimed to develop and confirm reliability of three photographic grading scales to aid prosthetic eye research: one for grading conjunctival inflammation in anophthalmic sockets and one each for stained deposits on the anterior and posterior surfaces of prosthetic eyes. The study included consultation with experienced ophthalmologists and optometrists and used perceptual and physical attributes when developing the scales similar to that described and recommended by Schulze, Jones and Simpson.[12]

METHODS

Data collection

Every two to four weeks over a three-month period, 43 volunteers had their anophthalmic sockets photographed with lower lids everted using a cotton bud (Table 1). The deposits on their prosthetic eyes were stained with a solution of 5 g of GC Corporation's plaque disclosing gel[13] (Table 2) dissolved in 30 ml of OcuPure saline solution.[14] The prosthetic eyes were immersed in the solution at 20°C (68°F) for a period of 2 min. After removing and blotting with tissue paper, the prosthetic eyes were photographed front and back against a black background, which included standard grey and colour scales to ensure the consistency of the photographic settings throughout the project. All the photographs of sockets and prosthetic eyes (800 in total) were printed 12 cm × 8 cm with a colour laser printer in a single session using a single batch of satin-finish photographic paper. The study had ethics approval from the University of Auckland Human Participants Ethics committee and the research

adhered to the tenets of the Declaration of Helsinki.

Development of grading scales and judgement criteria

Thirty photographs depicting a full range of severities for each condition of conjunctival inflammation, anterior surface deposits and posterior surface deposits were chosen by consensus among the authors. A provisional five-photograph inflammation scale and two provisional 11 photograph deposits scales were developed. The larger numbers of photographs in the two deposits scales were needed to incorporate both the pink and blue characteristics of the stained deposits within the same scale. Deposits that appeared lightly stained were pink, whereas those that were more heavily stained appeared blue.

The provisional scales were presented to a consultation group, comprised of four ophthalmologists, three optometrists and three senior optometry students. The consultation group was formally interviewed about the provisional scales and the judgement criteria they were likely to use in estimating severity of inflammation and deposit formation. For the conjunctival inflammation scale, the consultation group said that they focused primarily on the appearance of the 10 mm wide band across the inferior palpebral conjunctiva adjacent to the lid margin. They considered vasodilation of conjunctival blood vessels, apparent roughness of the conjunctival surface (for example, papillae) and the visible presence of any oedema. For the deposits scales, the consultation group noted the extent of the stained areas and the intensity of the pink and blue stains.

The members of the consultation group were then asked independently to place the photographs from the provisional scales along a 1.5 m rule in positions that represented their judgments of the relative severity of the factor being assessed, beginning with the least severe at 0 m and the most severe at 1.5 m. The photographs were assembled on a white table top and the lighting conditions (nominal correlated colour temperature 6,500 K, nominal colour rendering index greater than 90, illuminance greater than 500 lux) were kept identical for each session. This placement process identified images that were not consistent with an equal interval scale. These photographs were removed by the experimenters and replaced with others from the pool judged more likely to be placed at equal intervals. The members of the consultation group then repeated the placement task. The final outcome of this iterative process was the set of photographs shown in Figures 1, 2 and 3 that represent the best possible uniform scales.

Inter-rater reliability of the scales

SCALE FOR CONJUNCTIVAL INFLAMMATION

Four to six photographs of the anophthalmic sockets of 43 patients were taken at different times to provide images that covered a wide range of severity of inflammation. The resulting 218 photographs were individually coded and placed at the centre of a neutral grey Microsoft PowerPoint slide, which contained the grading scale to be tested. The slides were divided into four batches and presented in random order to 40 final year optometry students, who had at least three years of clinical training under supervision. A rest period with a short comical video followed the completion of each batch to reduce fatigue. The students were asked to grade the severity of conjunctival inflammation shown in each photograph to the nearest 0.1 unit by interpolation within the photographic scale as proposed by Bailey and colleagues.[8] For each photograph, the difference between individual severity assessments and the average of all assessments were plotted to determine the bias and the estimated limits of agreement.[15]

SCALES FOR ANTERIOR AND POSTERIOR SURFACE-STAINED DEPOSITS ON PROSTHETIC EYES

The process described was also followed for the two scales for deposits except that 156 photographs of stained anterior and 156 photographs of posterior surface deposits were used to create the slides for evaluating each scale. The scale for deposits on the anterior surface was assessed by 25 final year optometry students, while 20 assessed the scale for posterior surface deposits.

Test-retest reliability of the scales

SCALE FOR CONJUNCTIVAL INFLAMMATION

The 40 optometry students repeated the grading exercise after a half hour break. The differences between each individual grader's test and retest scores were plotted against the average of his/her test-retest score. Agreement of repeated measures was evaluated using correlation coefficients of concordance calculated using the formulae of Lin.[16,17]

SCALES FOR ANTERIOR AND POSTERIOR SURFACE-STAINED DEPOSITS ON PROSTHETIC EYES

The process for determining the test-retest reliability of the scales for deposits was the same as used for the inflammation scale except that only 18 final year students participated.

RESULTS

Inter-rater reliability of the scales

In Figures 4, 5 and 6, the differences between the individual grader's assessments of severity for each photograph are shown as a function of the average of all the assessments for that photograph. The 95 per cent confidence limits of the differences are marked on the plots. The standard deviation of the differences for conjunctival inflammation (Figure 4) was 0.52. For deposits on the anterior surface (Figure 5), it was 0.99 and for posterior surface deposits (Figure 6) it was 1.03.

Test-retest reliability of the scales

In Figures 7, 8 and 9, the differences between each grader's test and retest scores for each scale is shown as a function of the average of the individual grader's test-retest scores. The 95 per cent confidence limits of the differences are marked

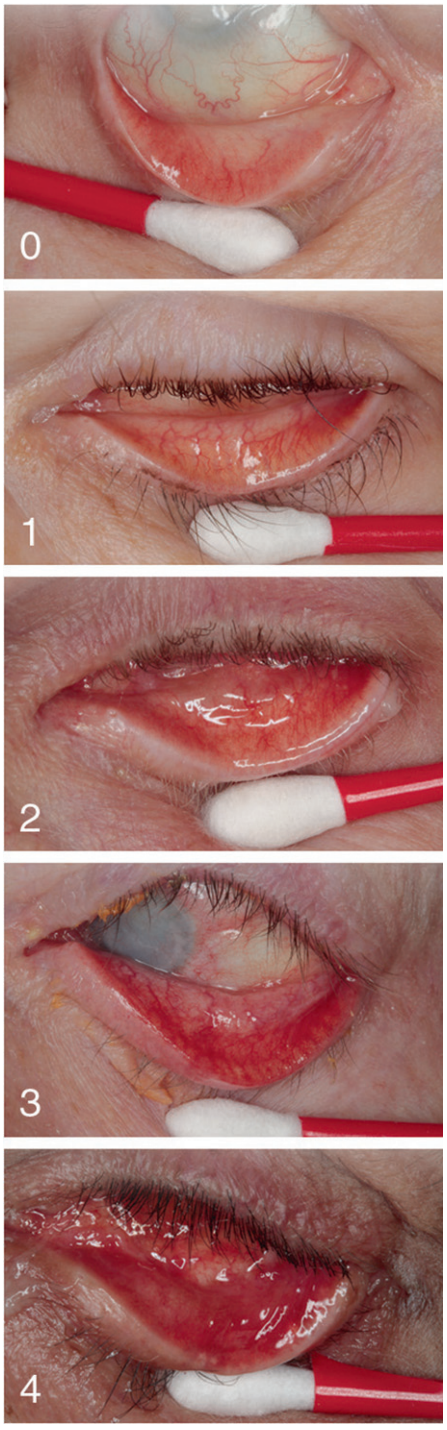

Figure 1. Grading scale for lower palpebral conjunctival inflammation in anophthalmic sockets

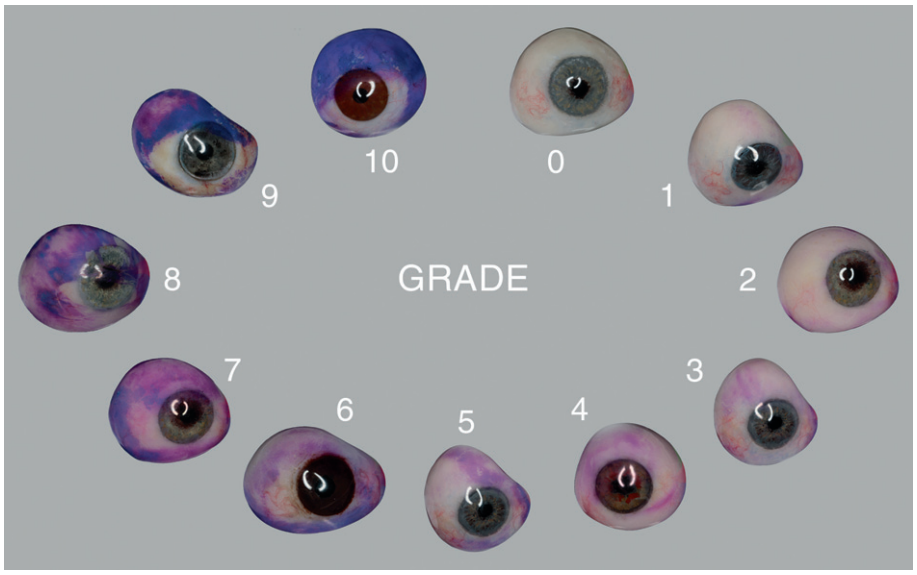

Figure 2. Grading scale for stained deposits on the anterior surface of prosthetic eyes

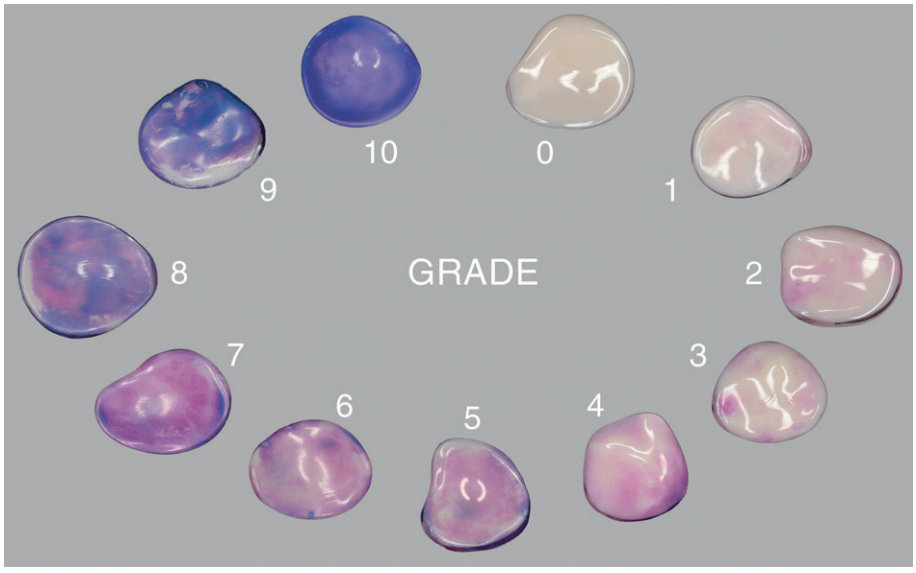

Figure 3. Grading scale for stained deposits on the posterior surface of prosthetic eyes

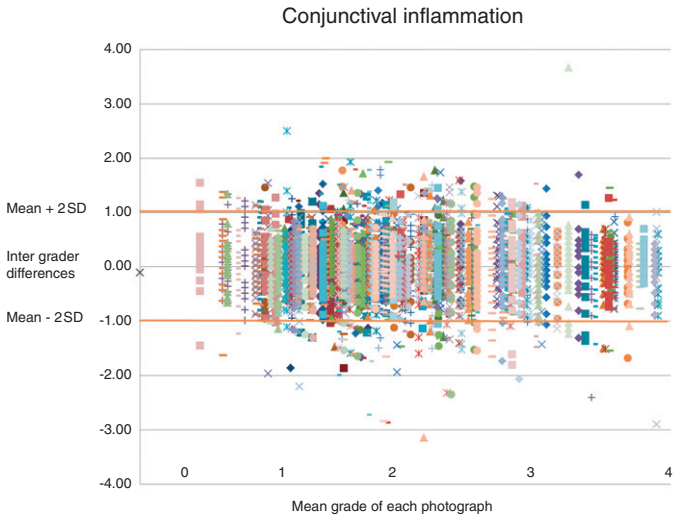

Figure 4. Inter-rater differences for the grading of inferior palpebral conjunctival inflammation. For each photograph, the difference between individual assessments and the average of all assessments was plotted. The mean of the differences was -0.01 and the standard deviation of the differences was 0.52.

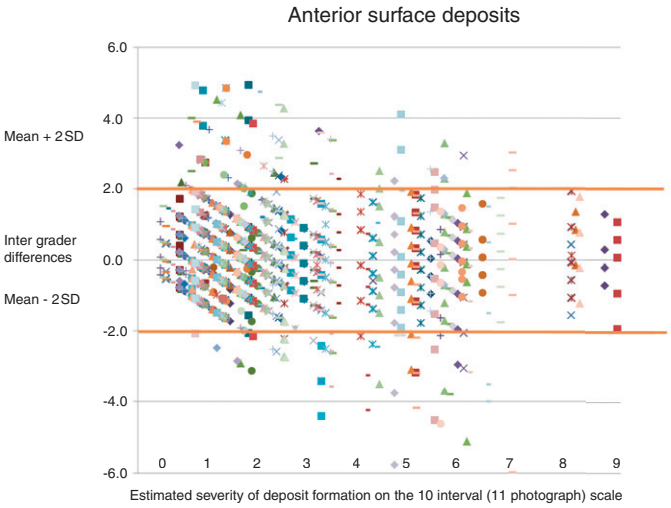

Figure 5. Inter-rater differences for the grading of deposits on the anterior surface of prosthetic eyes. For each photograph, the difference between individual assessments and the average of all assessments was plotted. The mean of the differences was 0.0 and the standard deviation of the differences was 0.99.

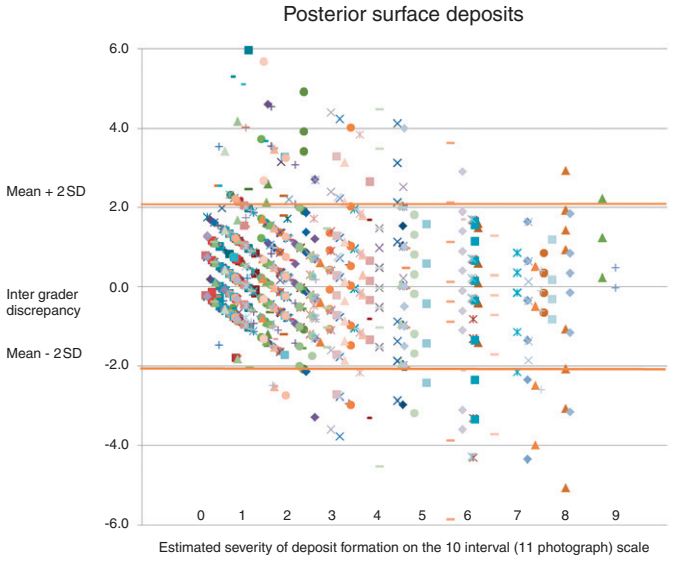

Figure 6. Inter-rater differences for the grading of deposit severity on the posterior surfaces of artificial eyes. For each photograph, the difference between individual assessments and the average of all assessments was plotted. The mean of the differences was 0.0 and the standard deviation of the differences was 1.03.

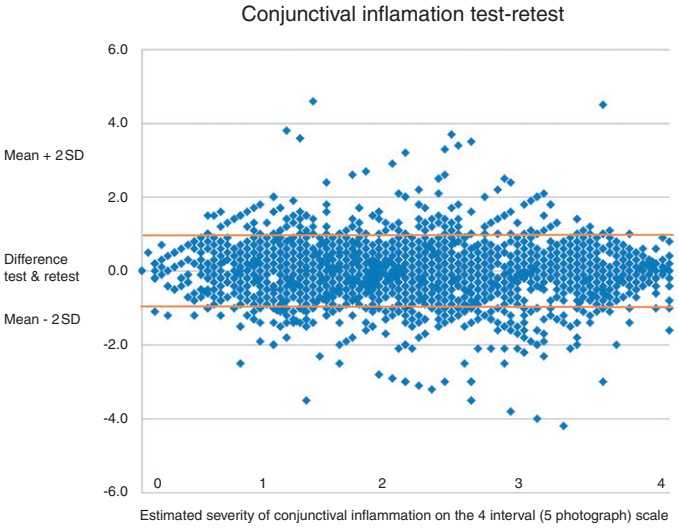

Figure 7. Test-retest reliability of the conjunctival inflammation grading scale. The differences between each grader's test and retest scores were plotted against the average of their test-retest scores. The mean of the differences was -0.03 and the standard deviation of the differences was 0.57.

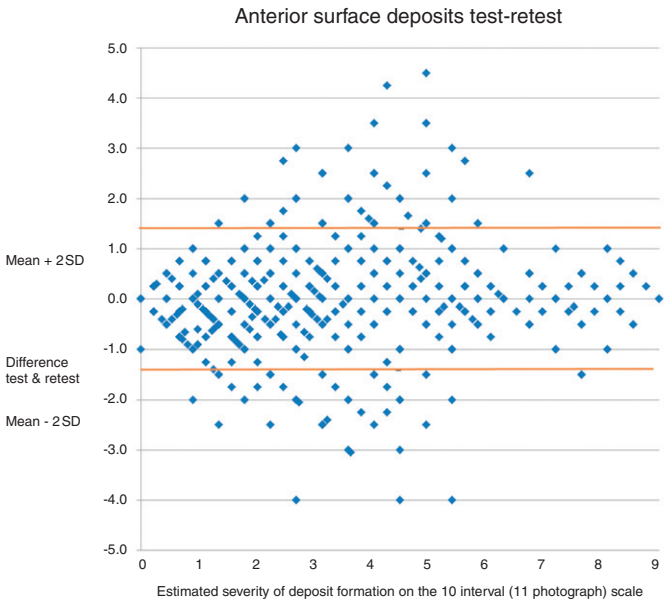

Figure 8. Test-retest reliability of the grading scale for deposits on the anterior surface of the prosthesis. The differences between each grader's test and retest scores were plotted against the average of their test-retest scores. The mean of the differences was 0.003 and the standard deviation of the differences was 0.71.

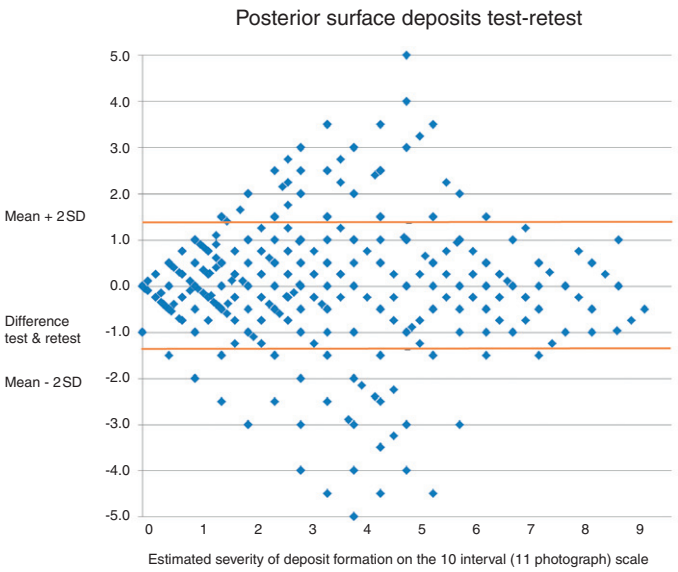

Figure 9. Test-retest reliability of the posterior deposits grading scale. The differences between each grader's test and retest scores were plotted against the average of their test-retest scores. The mean of the differences was -0.07 and the standard deviation of the differences was 0.71.

on the plots. The standard deviation of the test-retest differences for conjunctival inflammation (Figure 7) was 0.57 (concordance correlation coefficient = 0.808, 95 per cent confidence interval 0.800–0.816). For anterior surface deposits (Figure 8), the standard deviation of test-retest differences was 0.71 (concordance correlation coefficient = 0.786, 95 per cent confidence interval 0.771–0.801) and for posterior surface deposits (Figure 9), it was 0.71 (concordance correlation coefficient = 0.777, 95 per cent confidence interval 0.761–0.791).

DISCUSSION

Deposits form at the interface between the prosthetic eye surface and the conjunctiva and their effect on mucous discharge, socket comfort and socket health cannot be fully investigated unless tools to measure changes and outcomes have been created. The technique for staining deposits and the equal interval grading scales for measuring their intensity and extent developed here address this need and are key to further investigations into prosthetic eye wear.

The scale for measuring conjunctival inflammation is likely to be useful for clinical practice as it allows clinicians to monitor the level of conjunctival inflammation in anophthalmic sockets. While the scales for measuring deposits on prosthetic eyes may have limited use in clinical practice, the staining method is a valuable means of detecting and recording blemishes on the prosthesis. The plaque-disclosing gel used in this study is commonly used by dentists to demonstrate proper brushing and inter-dental cleaning techniques and could be used in the same way to educate prosthetic eye wearers about effective cleaning techniques. The grading scales for the deposits and the staining technique may have further applications in areas of research involving biofilm colonisation of synthetic materials.

In line with Efron, Morgan and Jagpa[18] and because of the novelty of the measuring tools required, only experienced clinicians or senior students were recruited to develop and confirm reliability of the three grading scales in this study.

Efron, Morgan and Jagpa's[18] suggestion to avoid 'display bias' between test and retest sessions was observed by having the sessions at the same venue. Difficulties in scheduling the same observers into the same venue prevented the second session being run on a later date. It is possible that despite the short breaks between batches of tests and the tailoring of the pace of presentations to the needs of the group, observer fatigue may have influenced the results. The effect of fatigue would likely be to increase the variability of results and reduce the reliability of the scales. The half-hour rest period between test and retest sessions may have allowed graders to remember their previous scores but this is doubtful because of the large number of photographs and the randomness of the order of presentation.

The size of the steps in the scales was somewhat arbitrary, although based on the authors' clinical experience. To allow

Reliability	Efron	Annunziaton	CCLRU	Vistakon	This study
Mean	0.55	0.58	0.53	0.59	0.40
Upper 95 per cent confidence limit	0.82	0.81	0.84	0.84	0.51
Lower 95 per cent confidence limit	0.28	0.36	0.34	0.34	0.29

Table 3. Comparison of grading reliability data from four different grading systems for bulbar conjunctival redness compared in the study of Efron, Morgan and Katsara[10] and the palpebral conjunctival inflammation scale developed in this study. The reliability data for this study were recalculated using the mean data of each observer to align with the method used in Efron, Morgan and Katsara.[10]

the scales to have the opportunity of detecting finer gradations of change, we followed the recommendation of Bailey and colleagues[8] and asked our observers to grade to the nearest 0.1 unit on the photographic scales. Ultimately, this study has demonstrated that the scales' steps appear to have the correct order of magnitude in terms of inter-rater and test-retest reliability. Strong correlations between the test and retest scores were also observed in this study.

To assess whether the grading scale differences created for this study are likely to perform comparably with scales created for grading contact lens complications, we compared our results with those provided by Efron, Morgan and Katsara.[10] They compared the reliability data of four different grading systems for bulbar conjunctival redness. While the palpebral conjunctival inflammation scale in this study is not limited to redness, the reliability data fall within the ranges calculated for these other grading systems. The comparison is shown in Table 3.

The ability to confidently detect change occurs with a grade change of 1.02 scale units for the conjunctival inflammation scale, 1.94 scale units for the anterior deposits scale and 2.04 units for the posterior deposits scale. The close agreement between the graders' abilities to detect change (with 95 per cent confidence) and each single step in the inflammation scale suggests that this scale may be useful in clinical practice, where practitioners wish to measure differences in severity of conjunctival inflammation that are clinically and statistically significant. The close agreement between graders' abilities to detect change (with 95 per cent confidence) and every second step in the two scales for deposits provides researchers with a tool for measuring deposits that measures statistically significant differences.

CONCLUSIONS

A novel and valuable technique for displaying and measuring the intensity and extent of deposit formation on prosthetic eye surfaces has been described. The two equal interval perceptual grading scales that have been developed to quantify the extent of deposit formation together with the equal interval perceptual scale for grading severity of palpebral conjunctival inflammation will for the first time allow the effects of prosthetic eye wear to be evaluated. The palpebral conjunctival inflammation scale and the technique for staining deposit formation on prosthetic eyes are recommended for use in clinical practice.

GRANTS AND FINANCIAL ASSISTANCE

This study was supported by The University of Auckland Post-graduate Research Student Support fund. Xuan Tai and Stella Shin, Department of Optometry and Vision Science at the University of Auckland were Summer Research Scholarship recipients.

REFERENCES

1. Srinivasan BD, Jakobiec FA, Iwamoto T, DeVoe AG. Giant papillary conjunctivitis with ocular prostheses. *Arch Ophthalmol* 1979; 97: 892–895.
2. Chang WJ, Tse DT, Rosa RH, Huang A, Johnson TE, Schiffman J. Conjunctival cytology features of giant papillary conjunctivitis associated with ocular prostheses. *Ophthal Plast Reconstr Surg* 2005; 21: 39–45.
3. Kim JH, Lee MJ, Choung H, Kim NJ, Hwang SW, Sung MS, Khwarg SI. Conjunctival cytologic features in anophthalmic patients wearing an ocular prosthesis. *Ophthal Plast Reconstr Surg* 2008; 24: 290–295.
4. Vasquez RJ, Linberg JV. The anophthalmic socket and the prosthetic eye: a clinical and bacteriologic study. *Ophthal Plast Reconstr Surg* 1989; 5: 277–280.
5. Christensen JN, Fahmy JA. The bacterial flora of the conjunctival anophthalmic socket in glass prosthesis-carriers. *Acta Ophthalmol* 1974; 52: 801–809.
6. Nath K, Krishn G, Gogi R, Kumar D. Bacterial and fungal flora of the sockets. *Indian J Ophthalmol* 1978; 26: 5–8.
7. Saini JS, Rajwanshi A, Dhar S. Clinicopathological correlation of hard contact lens related changes in tarsal conjunctiva by impression cytology. *Acta Ophthalmol* 1990; 68: 65–70.
8. Bailey IL, Bullimore MA, Raasch TW, Taylor HR. Clinical grading and the effects of scaling. *Investigat Ophthalmol Vis Sci* 1991; 32: 422–432.
9. Chong E, Simpson T, Fonn D. The repeatability of discrete and continuous anterior segment grading scales. *Optom Vis Sci* 2000; 77: 244–251.
10. Efron N, Morgan PB, Katsara SS. Validation of grading scales for contact lens complications. *Ophthalmic Physiol Opt* 2001; 1: 17–29.
11. Hurst MA, Mitchell SE, Douthwaite WA. Contact lens opacity grading system (CLOGS). *J BCLA* 1994; 17: 19–24.
12. Schulze MM, Jones DA, Simpson TL. The development of validated bulbar redness grading scales. *Optom Vis Sci* 2007; 84: 976–983.
13. GC Corporation. Plaque Disclosing Gel Material safety data sheet No Z-0003(E) June 2005, revision May 2006 GC Corporation, Tokyo. Japan. Available from: www.halas.com.au/pdfs/MSDS/GC/PlaqueDisclosingGel.pdf. Accessed September 2011
14. Lens Plus® Ocupure™ Saline Solution. Available from: www.home.intekom.com/pharm/allergan/lensplus.html. Accessed September 2011
15. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986; 1 (8476): 307–310.
16. Lin LI. A concordance correlation coefficient to evaluate reproducibility. *Biometrics* 1989; 45: 255–268.
17. Lin LI. A note on the concordance correlation coefficient. *Biometrics* 2000; 56: 324–325.
18. Efron N, Morgan PB, Jagpa R. Validation of computer morphs for grading contact lens complications. *Ophthalmic Physiol Opt* 2002; 22: 341–349.

Corresponding author:
Dr Robert J Jacobs
Department of Optometry and Vision Science
New Zealand National Eye Centre
The University of Auckland
Private Bag 92019
Auckland, 1142
NEW ZEALAND
E-mail: r.jacobs@auckland.ac.nz