Intraobserver and interobserver reliability of the R/D score for evaluation of iris configuration by ultrasound biomicroscopy, in patients with pigment dispersion syndrome

MO Balidis^{1,2}, C Bunce¹, K Boboridis^{1,2},
J Salzman¹, RPL Wormald¹ and
MH Miller¹

Abstract

The work has been presented at The Association for Research in Vision and Ophthalmology (ARVO) Annual Meeting, Fort Lauderdale, Florida, May 1999.

Purpose To evaluate inter- and intraobserver variability of the R/D score in assessing the iris configuration in Pigment Dispersion Syndrome patients.

Methods Fifty-seven high-resolution ultrasound biomicroscopy images were obtained by a single ophthalmologist. All images were examined twice by each of three ophthalmologists, the second assessment being at least 2 weeks after the first. Each observer was masked to their colleagues' and their previous measurements. R/D scores were calculated at each examination.

Agreement between and amongst observers was assessed using Bland–Altman plots. In addition, the R/D scores were categorised and reassessed using the Kappa statistic. **Results** Intraobserver variability was small, the average differences between first and second scores of each observer being less than 0.01 units. Agreement within observers was 89% or higher, with Kappa values of 0.8 or higher, indicating almost perfect agreement. Interobserver variability was, however, greater. Although there was substantial agreement between two of the observers (87% agreement, first assessment; 80%, second assessment with respective kappa statistics of 0.78 and 0.66), they only

moderately agreed with the other observer (kappa statistics between 0.55 and 0.68).

Conclusions This study suggests that when using R/D scores to demonstrate changes in iris configuration, assessments should preferably be made by the same observer.

Eye (2002) 16, 722–726. doi:10.1038/sj.eye.6700116

Keywords: intraobserver reliability; interobserver reliability; pigment dispersion syndrome; iris configuration; ultrasound biomicroscopy; high-resolution ultrasound

Introduction

High frequency ultrasound biomicroscopy (UBM) has been a valuable tool in imaging anatomical structures of the anterior and middle segment of the eye. Since its original description in the mid 1980s by Drs Foster and Sherrard and Pavlin,¹ it has been used extensively in the diagnosis and management of conditions traditionally difficult to image with conventional ophthalmic ultrasounds.^{2–4}

For an imaging technique to be applied clinically for the diagnosis and follow-up of patients with glaucoma or other anterior segment pathology, the generated images need to be assessed in a reproducible method, independent of observer's experience. Measurements may be quantitative or qualitative—quantitative methods allowing greater accuracy but being more susceptible to variability.

¹Moorfields Eye Hospital
London, UK

²Department of
Ophthalmology
University of Thessaloniki
Medical School
Thessaloniki, Greece

Correspondence:
MO Balidis
Ermou 46
54623 Thessaloniki
Greece
E-mail: balidis@otenet.gr

Received: 7 August 2001
Accepted: 12 February
2002

This study was conducted to evaluate intra- and interobserver agreement in assessing iris configuration using ultrasound biomicroscopy and R/D scores.^{5,6}

Methods

A single examiner obtained 57 high-resolution ultrasound images of 57 eyes. The images were randomly selected from a library of 172 UBM scans of 30 PDS patients (59 eyes).

All subjects were recruited from the glaucoma clinic at Moorfields Eye Hospital, London, UK. Examinations were conducted using the commercial version of ultrasound biomicroscopy (Humphrey Inc UBM, San Leandro, CA, USA) using a 50 MHz frequency transducer, with the patient in supine position. High resolution ultrasound is capable of achieving lateral resolution of approximately 40 μm with tissue penetration of about 4–5 mm, producing images of 5 $\times$ 5 mm field. We scanned each eye in four different meridians in a radial fashion at the limbus at approximately 90, 180, 270 and 360 degrees.

Three clinicians used the UBM images to estimate the configuration of the iris, each being masked to their colleagues' measures. One of the observers (MOB) was also the examiner and was relatively more experienced in analysing UBM scans. Each clinician measured the iris profile twice in a period of at least 2 weeks, in order to assess the within-observer variability.

Our technique for measuring iris configuration makes use of readily identifiable anatomical features of the eye, as illustrated in Figure 1. The outermost point of iridolenticular contact: *A*, is joined by a straight line: *AB*, to the peripheral end of the iris pigment epithelium IPE: *B*. The point of greatest displacement of the iris (either posterior or anterior). *C* (usually found at the midperiphery), is joined to *A* by the straight line *AC*. Two angles are then measured. The first, termed the reference angle, *R*, is the angle between the tangent of the lens at point *A* to line *AB*. The second, termed the displacement angle, *D*, is the angle between the tangent of the lens at point *A* to the line *AC*. In an eye with a planar configuration, the ratio *R/D* will equal 1. The greater the ratio, the greater the concavity of the iris.

To be consistent with our previous work, we categorised R/D scores as:

- (i) Anterior Bowing (convex) if the R/D score was ≤ 0.9 ;
- (ii) Flat if the R/D score was >0.9 and <1.1 ;
- (iii) Posterior Bowing (concave) if the R/D score was ≥ 1.1 .

Figure 1 R/D score—method of measuring iris configuration. (a) Posterior bowing configuration in a PDS eye. (b) Same scan as in (a) with application of R/D score method of measurement. The outermost point of iridolenticular contact: *A*, is joined by a straight line: *AB*, to the peripheral end of the iris pigment epithelium IPE: *B*. The point of greatest displacement of the iris *C* (usually found at the midperiphery), is joined to *A* by the straight line *AC*. The third line is the tangent line of the lens at point *A*. The two angles created are then measured.

Statistical analysis

Data were entered into Excel spreadsheets and analysed in Stata (Release 5.0, Stata Corp., TX, USA). Bland Altman plots were constructed to assess whether there was any systematic bias between readings and whether the difference between readings varied in any way with the size of the measure. Since there was no evidence of heteroscedasticity, we summarised the

findings as the median and 95% limits of agreement. Kappa statistics were then used to assess agreement between the R/D scores.⁷⁻⁹

We categorised R/D scores as:

- (i) Anterior Bowing (convex) if the R/D score was ≤ 0.9 ;
- (ii) Flat if the R/D score was >0.9 and <1.1 .
- (iii) Posterior Bowing (concave) if the R/D score was ≥ 1.1 .

Results

Table 1 presents the number of missing measures for each observer. The observers had a variable degree of experience both in using image analysis programs and interpretation of ultrasound images (Observer 2 had highest scores in unclassified measures).

Figures 2 a-c are the Bland Altman plots for each intraobserver comparison under evaluation. It can be seen that in each case, points are clustered around the line of no difference, although there does appear slightly greater variability at extremes.

Table 2 summarises this information in terms of medians, interquartile ranges and maximum differences and illustrates how close are the observations made at different times by the same observer.

Tables 3 and 4 present percentile agreement and kappa statistics for assessment of intraobserver and interobserver variability, respectively, using the categorised R/D scores. A kappa of greater than 0.81 may be interpreted as near perfect agreement, between 0.61 and 0.8 interpreted as substantial, and values between 0.41 and 0.6 moderate. We thus see that there is near perfect agreement within observers, percentile agreement ranging between 89 and 91% and the kappa statistics between 0.82 and 0.85. Agreement between observers, however, ranges from moderate to substantial.

From the plots most of the observations lie fairly close to $y = 0$. There seems to be less agreement in higher (Observer 1, three values) and lower (Observer 2, two values) measurements. The intraobserver variability in 140 measurements (24 images were unclassified by the observers) was minimal, but also

Table 1 Unclassified UBM scans

Observers	1st Reading	2nd Reading
Observer 1	4	3
Observer 2	12	12
Observer 3	0	0

Figure 2 (a-c) Intraobserver variability. Bland Altman plots for each intraobserver comparison under evaluation. The plots present the difference in readings against the average of readings.

Table 2 Intraobserver variability

Observers	Median range	IQ range	Maximum difference
Observer 1	0.001	(−0.005, 0.005)	0.61
Observer 2	0	(−0.002, 0.03)	0.91
Observer 3	−0.012	(−0.07, 0.01)	0.49

Table 3 Intraobserver agreement using percentage and kappa statistics

Observers	% Agreement	Kappa statistics
Observer 1	91	0.85
Observer 2	89	0.82
Observer 3	89	0.82

Table 4 Interobserver agreement using percentage and kappa statistics

		Observer 2		Observer 3	
		% Agreement	Kappa statistics	% Agreement	Kappa statistics
Observer 1	1st Reading	80%	0.68	72%	0.55
	2nd Reading	73%	0.58	76%	0.6
	Reading				
Observer 2	1st Reading			87%	0.78
	2nd Reading			80%	0.66
	Reading				

the interobserver variability appears to be satisfactory especially between the two less experienced observers.

Discussion

Ultrasound biomicroscopy revealed to physicians the anatomical morphology of the anterior segment,¹⁰ which was impossible to be imaged before. Accurate and reproducible measurements of intraocular structures depend on the resolution of the scanner, the skills of the examiner and the standardisation of conditions (to avoid internal dynamic changes) during image acquisition. An experienced examiner can improve the image quality by altering the acquisition properties (ie focusing at the areas where the anatomical landmarks of interest are located, improving the contrast).

The above-mentioned parameters are important for the quality of the images, but even if the reproducibility is high, variability can occur during interpretation.

Various methods of measuring anterior segment

structures have been proposed and applied in UBM studies.^{11,12} Very few though have been assessed for their reproducibility.

From previous studies it was evident that the source of error was the identification of anatomical reference points and the use of complicated methods requiring subjective interpretation. Our method utilises only two reference points, easily identifiable and is based on angle and ratios calculations to avoid distance measurements, which can vary considerably with minimal movement of the cursor.

The within-observer agreement, as presented in our study was almost perfect, indicating that a single observer recognises anatomical landmarks and places the lines with minimum variability in repeated measurements. The between-observer agreement was substantial, pointing out that the level of subjective interpretation was still considerable. Possibly a more detailed description of the method, including examples, may help to avoid the disagreement in ratios.

An important aspect of reliability between observers is that there were only two (0.011%) major disagreements (both close to the limits of the categories), showing that by using a descriptive method, the interobserver agreement is high.

With the advantage of high lateral and axial resolution, detailed examination is feasible. Modern ocular scanning must be accompanied by standardised measuring and scoring techniques, to be applied in clinical practice.

References

- Pavlin CJ, Harasiewicz K, Sherar MD, Foster FS. Clinical use of ultrasound biomicroscopy. *Ophthalmology* 1991; **98**: 287–295.
- Liebmann JM, Ritch R. Ultrasound biomicroscopy of the anterior segment. *J Am Optom Assoc* 1996; **67**: 469–479.
- Pavlin CJ. Ultrasound biomicroscopy in pigment dispersion syndrome [letter; comment]. *Ophthalmology* 1994; **101**: 1475–1477.
- Sokol J, Stegman Z, Liebmann JM, Ritch R. Location of the iris insertion in pigment dispersion syndrome. *Ophthalmology* 1996; **103**: 289–293.
- Spaeth GL, Azuara-Blanco A, Araujo SV, Augsburger JJ. Intraobserver and interobserver agreement in evaluating the anterior chamber angle configuration by ultrasound biomicroscopy. *J Glaucoma* 1997; **6**: 13–17.
- Tello C, Liebmann J, Potash SD, Cohen H, Ritch R. Measurement of ultrasound biomicroscopy images: intraobserver and interobserver reliability. *Invest Ophthalmol Vis Sci* 1994; **35**: 3549–3552.
- Koch GG, Landis JR, Freeman JL, Freeman DHJ, Lehen RC. A general methodology for the analysis of experiments with repeated measurement of categorical data. *Biometrics* 1977; **33**: 133–158.
- Landis JR, Koch GG. An application of hierarchical kappa-type statistics in the assessment of majority

- agreement among multiple observers. *Biometrics* 1977; **33**: 363–374.
- 9 Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; **33**: 159–174.
 - 10 Pavlin CJ, Macken P, Trope G, Feldman F, Harasiewicz K, Foster FS. Ultrasound biomicroscopic features of pigmentary glaucoma. *Can J Ophthalmol* 1994; **29**: 187–192.
 - 11 Spaeth GL, Aruajo S, Azuara A. Comparison of the configuration of the human anterior chamber angle, as determined by the Spaeth gonioscopic grading system and ultrasound biomicroscopy. *Trans Am Ophthalmol Soc* 1995; **93**: 337–347.
 - 12 Urbak SF. Ultrasound biomicroscopy. I. Precision of measurements. *Acta Ophthalmol Scand* 1998; **76**: 447–455.