Agreement among optometrists and ophthalmologists in estimating limbal anterior chamber depth using the van Herick method

Running title: Inter-observer agreement in estimating limbal anterior chamber depth

Anish Jindal[1, 2], Joy Myint[3], David F Edgar[1], Winifred P Nolan[2] and John G Lawrenson[1]

1. Division of Optometry and Visual Science, City University London, London UK
2. Moorfields Eye Hospital NHS Foundation Trust, London UK
3. Postgraduate Medicine, Life and Medical Sciences, University of Hertfordshire, Hatfield UK

Correspondence: John G Lawrenson Email address: J.G.Lawrenson@city.ac.uk

Key words: van Herick test, angle closure, inter-observer agreement

Disclosure: The authors report no conflicts of interest and have no proprietary interest in any of the materials mentioned in this article.

Acknowledgment: Part of this study was funded by an unrestricted educational grant from Pfizer Ophthalmology to Dr Myint. The authors would like to thank thank the reviewers of this study for their helpful suggestions.

Abstract

Purpose: to evaluate the inter-observer agreement for measuring limbal anterior chamber depth (LACD) using the van Herick test in community optometrists, glaucoma specialist optometrists and ophthalmologists

Methods: the study was divided into two phases. In the first phase, a random sample of 100 UK community optometrists were given an opportunity to select and grade eight digital slit-lamp images of anterior chamber angles using the original van Herick 4 point grading scale. The images were included in a clinical decision making study using computerized virtual case vignettes. In the second phase, hospital-based glaucoma specialist optometrists and glaucoma sub-specialist ophthalmologists graded the LACD of the right eye using a 7 point % grading scale in 57 consecutively presenting patients with suspect glaucoma. Inter-observer agreement was assessed using linearly weighted kappa (κ_w).

Results: inter-observer agreement for community optometrists was moderate, with a mean κ_w for grading photographic images of 0.50 (95% confidence interval (CI) 0.43-0.57). Overall, ninety two percent of observations were within one grade of the actual grade, although grading of narrow angles was associated with a 13% false negative error rate (based on a ≤ grade 2 threshold). For phase 2 of the study, pairwise comparisons between optometrists and ophthalmologists showed that agreement was moderate to substantial (mean κ_w =0.54-0.65) with a false negative rate of 1.9% (based on a ≤25% threshold). Grading accuracy of specialist optometrists and ophthalmologists were equivalent.

Conclusions: In summary, the present study found that community optometrists showed moderate inter-observer agreement for grading LACD. Glaucoma specialist optometrists showed moderate to substantial agreement with weighted kappa values that were equivalent to sub-specialist ophthalmologists. The augmented 7 point % grading scale is intuitive and potentially offers greater accuracy for grading narrow angles than the traditional 4 point scale for grading LACD.

Introduction

In European-derived populations, it has been estimated that approximately 0.4% of those over 40 years old have primary angle closure glaucoma (PACG),[1] corresponding to 130,000 individuals in the UK. Using the Department of Health's Hospital Episode Statistics a marked reduction in the number of incident cases of acute angle closure has occurred over the last 15 years, which has been attributed to increased rates of phakoemulsification surgery for cataract and higher rates of prophylactic Nd:YAG laser peripheral iridotomy (LPI).[2] However, in many parts of the world PACG is at least as common as primary open angle glaucoma (POAG) and PACG remains a major cause of bilateral blindness worldwide.[3] The natural history of PACG involves three conceptual stages: initial contact between the peripheral iris and trabecular meshwork, followed by adhesions (synechiae) between the iris and meshwork with or without raised IOP (this stage is termed primary angle closure (PAC)); the final stage of the disease involves the development of glaucomatous optic neuropathy. Not every case of PACG progresses through all three stages, for example initial iridotrabecular contact can be followed by raised IOP and subsequently glaucomatous optic neuropathy without the presence of synechiae[4]. In 'high risk' Asian populations approximately half of affected individuals are asymptomatic which often results in late presentation with advanced visual field loss.[5] Consequently, targeted screening has been suggested in these populations as prophylactic treatment of eyes with anatomically narrow angles can potentially reduce the risk of developing glaucomatous optic neuropathy.[6]

Eyes at risk of angle closure tend to have smaller anterior chamber dimensions with shorter axial and peripheral anterior chamber depths.[7,8] Although gonioscopy, performed in a darkened room, remains the 'reference standard' for assessing angle configuration and depth[9,10], the technique is unsuitable for screening the general population.[6] Gonioscopy is not always well tolerated and the interpretation of the appearance of the angle is highly dependent on the skill and experience of the practitioner. A variety of alternative non-invasive methods are available for evaluating the configuration of the drainage angle (see Dabasia et al 2013 and 2014 for recent reviews).[11,12] A simple method for estimating limbal anterior chamber depth (LACD) using the slit lamp biomicroscope was originally described by van Herick et al in 1969[13] and is based on an estimation of the ratio of the chamber depth at the limbus and the peripheral corneal thickness expressed as a fraction (<1/4, 1/4, 1/4 to 1/2, and ≥1). The original four point grading scheme (which has been extended to five by including 0 (closed)) was found to give comparable results to those obtained with a five point grading scheme for estimating the angular width of the peripheral

anterior chamber by gonioscopy.[13] Application of the van Herick test to an Indian population gave sensitivities and specificities of 61.9% and 89.3% respectively.[14] However, expressing the LACD in more refined percentage increments of the peripheral corneal thickness increased the sensitivity and specificity to 91% and 93%,[7] suggesting that the test could be used by primary eye care practitioners to identify those at risk of angle closure. In the UK, the use of the van Herick test is specified in professional guidance that advises community optometrists on the assessment and referral of glaucoma suspects.[15,16] The 2009 NICE glaucoma guideline[9] also recommended determining LACD as an alternative to gonioscopy if clinical circumstances rule out gonioscopy (e.g. when people with physical or learning disabilities are unable to participate in the examination by gonioscopy).

Previous studies of inter-observer reliability using both the traditional[14]and augmented[7] grading scales have shown substantial agreement (weighted kappa values: 0.74 and 0.76). However, these studies were performed using glaucoma subspecialist ophthalmologists. There were two primary aims of this research. First, to evaluate the inter-observer agreement in a random sample of UK community optometrists for the measurement of LACD from slit lamp images using the original 4 point van Herick grading system. Second, to evaluate the inter-observer agreement for the measurement of LACD, using the revised 7 point percentage grading scale, between glaucoma specialist optometrists and glaucoma sub-specialist ophthalmologists working in the same eye unit.

Methods

Ethical approval for the study was given by the School of Health Sciences Research and Ethical Committee (City University London) and London-Surrey Borders NRES Committee (ethics approval no 114480). All subjects gave informed consent and the research was carried out in accordance with the tenets of the Declaration of Helsinki.

The study was divided into two phases:

Phase 1. LACD estimation by community optometrists

A randomized sample of 250 community optometrists practising in England and Scotland was selected from the UK General Optical Council's (GOC) register of optometrists and invited to participate in an interactive vignette-based clinical decision making study. The vignettes were in the form of a standard 'virtual record card' which could be interrogated to provide answers to questions that could be asked when taking a clinical history and obtain results of tests and investigations that could form part of a routine eye examination. One hundred optometrists agreed to participate in the study and were remunerated for their time to complete four case scenarios. Assessment of the anterior chamber angle using the van Herick test was included as an available clinical investigation. If the test was selected, a digital slit-lamp image of both eyes would appear with a pull-down menu to grade the angle using the traditional 4-point grading scheme (Figure 1).

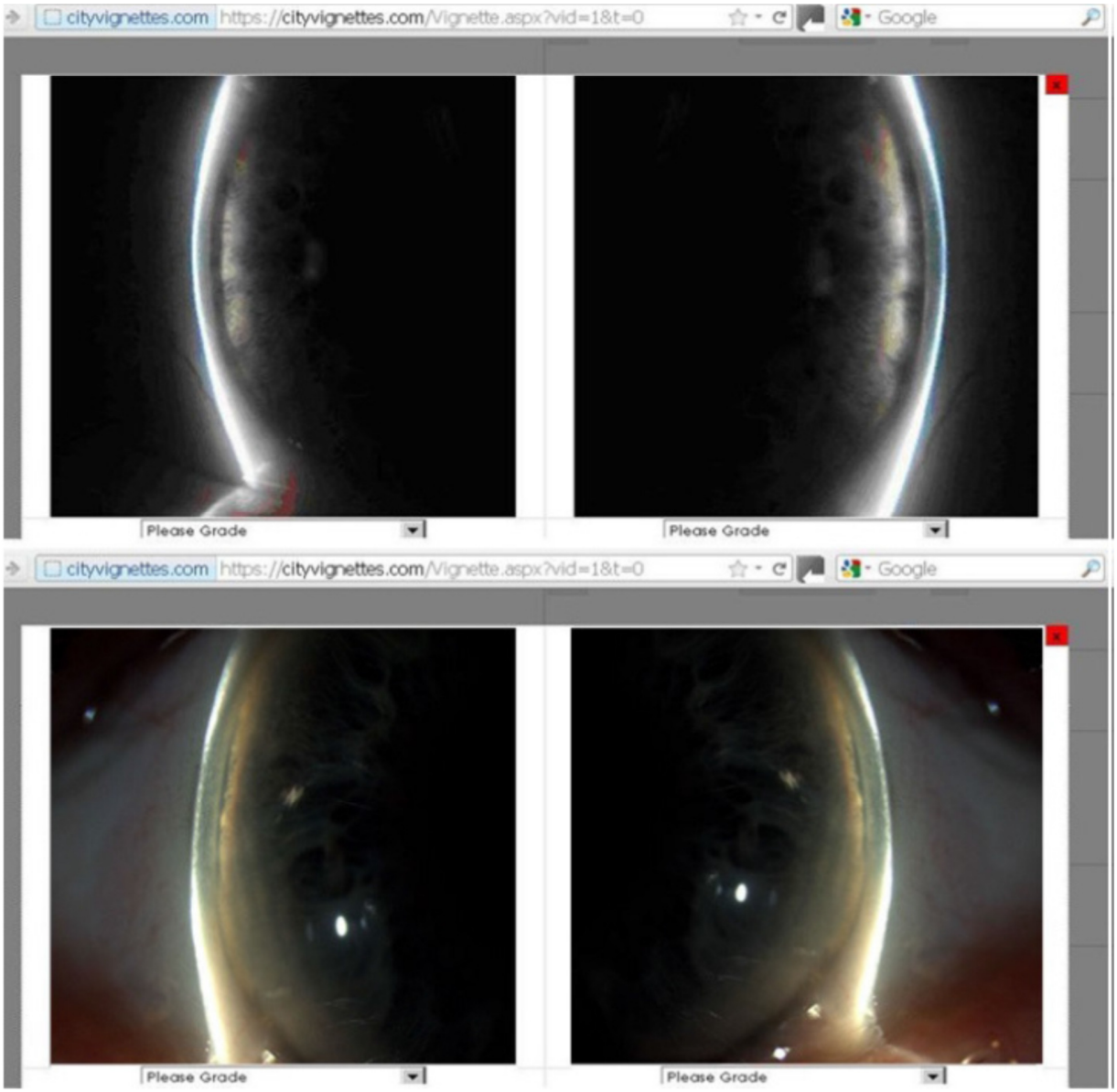

Figure 1. Example of digital vignette images for estimating LACD

Phase 2. LACD estimation by specialist optometrists and ophthalmologists

Four observers (two hospital-based optometrists with a specialist interest in glaucoma and two ophthalmologists with a sub-specialty interest in glaucoma) independently estimated the LACD in a series of patients referred as glaucoma suspects into the glaucoma service at Moorfields Eye Hospital (MEH), London. Although all observers were familiar with the van Herick technique, prior to the start of

the study, each clinican was briefed regarding the standardized LACD measurement procedure and participated in two practice sessions to perform the technique and become familiar with the data recording sheets. A total of 57 consecutive adult glaucoma suspects were invited to participate in the agreement study. Subjects were excluded only if they could not be examined using a table mounted slit lamp due to physical limitations or had gross nystagmus. After providing informed consent, each subject was assigned a unique reference number and demographic data were recorded (including age, gender and ethnicity). The measurement of LACD was carried out using the same slit lamp (Model 900 BM, Haag-Streit, Bern, Switzerland, http://www.haagstreituk.com) under constant dim room lighting conditions. The illumination column of the slit-lamp was offset at 60 degrees temporal to the microscope and a narrow slit beam was positioned at the most peripheral point of the cornea, just inside the limbus. Each observer viewed the central portion of the beam using 16 x magnification and estimated the size of the gap between the reflection of the beam on the iris and the corneal section as a percentage of the total thickness of the cornea. The result was recorded using the augmented seven point percentage grading scale (0%, 5%, 15%, 25%, 40%, 75%, ≥100%)[7] or designated as ungradable. Each subject had their LACD measured at the temporal limbus of the right eye by each observer in turn and the result was recorded on separate data recording sheets. If the right eye could not be graded due to anterior segment abnormalities that obscured the limbus (e.g. corneal scarring, peripheral corneal degeneration or pterygium) the left eye was assessed. If the left eye was also ungradable, this was recorded as both eyes ungradable. Each observer was masked to the findings of the other observers. At the end of the clinical session, each observer inserted their individual data collection sheets into an envelope which was sealed, labeled with the date and observer number and passed on to an independent data coordinator. All data were transferred from the individual observer data collection sheets into an Excel spreadsheet by the independent data co-coordinator. Data analysis was performed by a member of the research team who was masked as to the identity and profession of the observer.

Statistical Analysis

The level of agreement between observers was calculated using the kappa statistic. For the community optometrists, kappa was calculated by comparing the grade recorded by each participant using linear weights against the actual grade determined by direct linear measurement of each image on the computer screen. This was carried out independently by 3 members of the research term and a consensus grade agreed. The weighted kappa (κ_w) attaches greater significance to larger differences between ratings than smaller differences. For phase 2 of the study weighted kappa was calculated for all pairwise combinations of LACD grade estimates for each observer. Landis and Koch[17] have proposed the following as standards for strength of agreement for the kappa statistic;<0=poor, 0.00-0.20=slight, 0.21-0.40=fair, 0.41-0.60=moderate, 0.61-0.80=substantial and 0.81-1.0=almost perfect. All statistical analyses were performed using Medcalc v12.50 (Medcalc Software. Mariakeerke, Belgium, http://www.medcalc.org).

Results

Phase 1

The demographic characteristics of the participating community optometrists are shown in Table 1. The cohort comprised optometrists with a range of clinical experience, although the majority (64%) had not undertaken any further postgraduate training or accreditation. Forty five percent of optometrists worked in independent practices and 55% in multiples

A total of eight van Herick images showing open and narrow angles were available for LACD grading across the four vignettes. If all images had been graded this would have resulted in 800 independent grading observations. However, since selection of the van Herick test was optional, a total of 41% of images (N=329) were graded (Table 2). The mean κ_w with respect to the actual grade was 0.50 (95% confidence interval (CI) 0.43-0.57), which corresponds to 'moderate' agreement.

Although overall, 92% of reported grades showed complete concordance or were within one grade of the actual grade (Table 3), grading errors may still have been critical in terms of the commonly used ≤ grade 2 threshold for referral for suspected PAC[11,13]. For example, 12.7% of reported grades for the grade 1 images were documented as >2, which may have potentially resulted in a failure to refer leading to missed cases of angle closure. Although the majority of graders showed good agreement with the

actual grade there were some very large discrepancies, with 6 observations differing by 3 grades, which suggests that some optometrists may have misinterpreted the directionality of the scale.

Demographic category	**%**
Geographical location	
• England	73
• Scotland	27
Gender	
• Male	37
• Female	63
Year of first registration with the GOC	
• 1960-9	1
• 1970-79	6
• 1980-89	25
• 1990-99	22
• 2000-09	44
• 2010+	2
Further post-graduate training	
• Higher degree/professional diploma/local accreditation	36
• None	64
Mode of practice	
• Independent	45
• Multiple	55

Table 1. Demographics of the community optometrist cohort

Image	Observations (N)	van Herick Grade
1	32	1
2	31	1
3	38	4
4	39	4
5	53	3
6	53	3
7	41	3
8	42	4

Table 2. van Herick grade and number of reported observations for each image

Actual van Herick Grade	Observations (N)	% of reported grades differing from actual grade						
		-3	**-2**	**-1**	**0**	**1**	**2**	**3**
Grade 1 (2 images)	63	NA	NA	NA	52	35	10	3
Grade 3 (3 images)	147	NA	6	30	44	20	NA	NA
Grade 4 (3 images)	119	3	4	29	64	NA	NA	NA

Table 3. Percentage differences between actual van Herick grade and observer grades for the LACD represented in the photographic images

Phase 2

A total of 57 patients were recruited into this phase of the study. The mean age was 59.3 years (range: 20-86 years). Gender distribution was 47% male and 53% female. Self-reported ethnicity comprised:

56% white, 21% South Asian, 16% African and 7% other. All angles were gradable and the distribution of the median LACD grades for all four observers is shown in Figure 2. Pair wise comparisons between observers showed that inter-observer agreement was moderate to substantial (mean κ_w =0.54-0.65). Table 4 shows the linearly weighted kappa for each observer pairing with 95% confidence intervals. Eighty three percent of pairwise observations by glaucoma specialist optometrists and ophthalmologists showed either complete concordance or differed by one grade (based on the revised 7-point percentage scale).

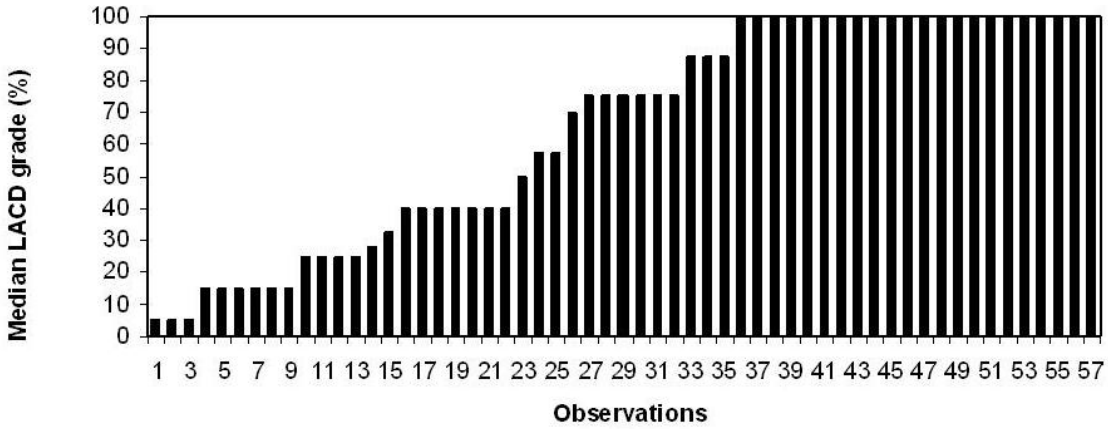

Figure 2. Distribution of the median LACD arranged in rank order for all four observers for each of the 57 subjects using a 7-point % grading scale (0%, 5%, 15%, 25%, 40%, 75%, ≥100%).

Observers	**Weighted Kappa**	**Standard error**	**95% CI**
Optom 1 vs Optom 2	0.63	0.06	0.51-0.75
Optom 1 vs Ophthal 1	0.62	0.06	0.50-0.74
Optom 1 vs Ophthal 2	0.54	0.07	0.41-0.68
Optom 2 vs Ophthal 1	0.64	0.07	0.51-0.77
Optom 2 vs Ophthal 2	0.65	0.07	0.50-0.79
Ophthal 1 vs Ophthal 2	0.60	0.06	0.49-0.72

Table 4. Weighted Kappa for the comparison between optometrists and ophthalmologists

A frequency plot of the differences between graders for all 342 pairwise comparisons is shown in Figure 3. Based on the median grade, 13 of the 57 subjects would have met the ≤25% screening threshold criterion for referral for suspect PAC. In only one of these cases of narrow angles a difference in grading would have led to a failure to refer on the basis of a ≤25% threshold. However, in approximately 8% of cases the grading discrepancy may have led to a false positive referral.

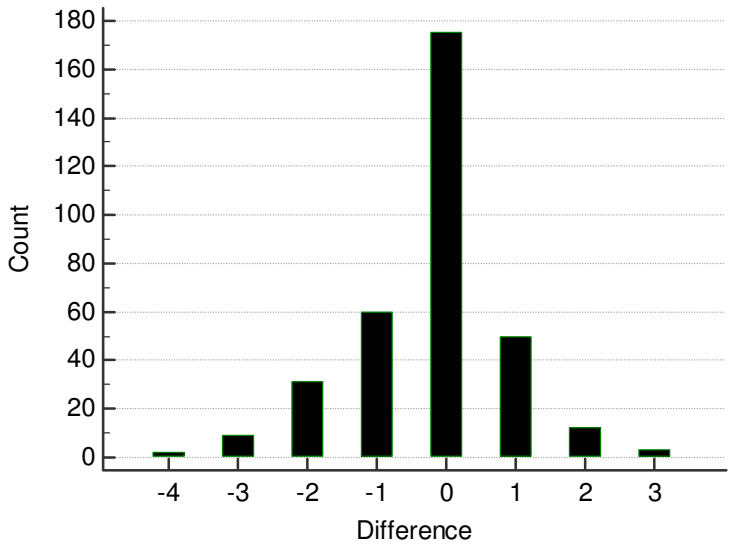

Figure 3. Frequency plot of differences between grades for all 342 pairwise comparisons for the LACD percentage grading scale converted onto a 7-point integer scale (1=0%, 2=5%, 3=15%, 4=25%, 5=40%, 6=75%, and 7=≥100%).

Discussion

To our knowledge this is the first published study of inter-observer reliability for estimating LACD by non-ophthalmologists. Current professional guidance for community optometrists in the UK refers to the detection of eyes at risk of angle closure. For example, guidance jointly produced by the College of Optometrists and Royal College of Ophthalmologists on the referral of glaucoma suspects[15] states that an optometrist should refer patients for further investigation if they identify 'a narrow anterior drainage angle on van Herick testing consistent with a significant risk of acute angle closure within the foreseeable future'.[15] Furthermore, the current College of Optometrists guidance for examining patients at risk from glaucoma states that 'Assessment of the anterior eye and angle (e.g. by slit lamp van Herick technique) is advisable for all patients suspected of having glaucoma'.[16] The current study assessed inter-rater reliability for the estimation of LACD in a randomized sample of community optometrists and also in a group of optometrists and ophthalmologists with a sub-specialist interest in glaucoma. Differences in the nature of the grading task between the two phases of the study and variability in the number of raters and sample sizes precluded a direct comparison of inter-rater agreement between phases 1 and 2.

Phase 1 of the study compared the agreement between community optometrists in LACD grading using digital photographs of anterior chamber angles using the traditional four point grading scheme.[13] Moderate inter-observer reproducibility (κ_w =0.50) was demonstrated for LACD estimation, with 92% of observations showing either complete concordance or being within one grade of the actual grade. However, there were some large discrepancies which imply that there may have been some confusion regarding the directionality of the scale.

Based on the average absolute deviation of the grade estimates for each image there was no suggestion that grading accuracy varied between open and narrow angles. Grading accuracy becomes much more significant when grading narrow angles. Approximately 13% of observed grades for the grade 1 images were graded as 3 or 4, which if mirrored in actual optometric practice may have led to a decision not to refer a patient when referral for investigation for potential angle closure was indicated. Interestingly, the weighted kappa value for van Herick grading in the current study is compares favorably with those obtained in previous studies of inter-optometrist variability in optic disc assessment, for example, κ_w for the assessment of vertical cup: disc ratio has been reported as 0.39 and 0.40.[18,19]

Improvement in the diagnostic performance of the van Herick test has been achieved by extending the original scale to a seven point scale percentage scale, which includes additional grades in the more clinically significant range of narrow angles.[7] Bailey and colleagues[20] have discussed the problems associated with measurement scales that are too coarse and advocated the use of finer scales to enhance the sensitivity of clinical measurements. As well as potentially being more accurate, the reporting of LACD as a series of fixed percentage estimates of corneal thickness is more intuitive than the originally proposed ordinal scale. Further training in the use of the extended scale together with standardization of the LACD assessment technique could potentially improve test performance and accuracy.

The augmented grading scheme was used for phase 2 of the study, which compared the performance of glaucoma specialist optometrists and ophthalmologists in the assessment of LACD in glaucoma suspects. Inter-observer agreement between optometrists and ophthalmologists was found to be equivalent with κ_w values in the range 0.54-0.65. These levels of agreement were lower than previously reported for other glaucoma sub-specialist observers.[7,14] However, concordance between observers was excellent in those patients with narrow angles, with only one of the 52 pairwise comparisons (2%) in patients with a median LACD ≤25% differing by more than one grade.

There are a number of inherent limitations in the current study, which may impact on the generalizability of the results. To investigate inter-rater variability amongst community optometrists we adopted a pragmatic experimental design, by asking participants to grade high quality photographic images. Other sources of error may be introduced when performing the test on real patients e.g. not positioning the beam as close as possible to the limbus or not maintaining a 60 degree angle between the slit-lamp illumination and observation systems. It is therefore likely that agreement between optometrists for measurement of LACD obtained under clinical examination conditions will be lower. Also, the cohort who opted to grade the images may not be representative, since these subjects may be more confident of their ability to grade angles than the average UK community optometrist. Consequently, it is possible that the performance of this group overestimates that of community UK optometry as a whole. Another limitation was the small number of narrow angles available in both

phases of the study for grading, since accuracy in grading narrow angles is clinically more meaningful. However, there was no evidence that the variance of grading estimates differed between open and narrow angles. A further limitation was the small sample size for assessing the agreement between specialist optometrists and ophthalmologists, although all participating clinicians were typical of those trained and accredited in the UK. In terms of statistical power, the sample of subjects for phase 2 was sufficiently large to detect a statistically significant kappa value ≥ 0.4 at a significance level of 0.05 with 80% power.[21]

In summary, the present study found that community optometrists showed moderate inter-observer agreement for grading LACD from photographic images using the traditional 4 point grading scale, although there were a small number of large discrepancies which suggests some unfamiliarity with the grading scale. Glaucoma specialist optometrists, using the augmented 7 point % grading scale, showed moderate to substantial agreement with κ_w values that were equivalent to that found between sub-specialist ophthalmologists. The augmented 7 point grading scale is intuitive and potentially provides greater accuracy when grading narrow anterior chamber angles. Further training and targeted continuing education on the use of the van Herick technique, together with standardization of the LACD measurement could potentially improve test performance and accuracy.

References

1. Day AC, Baio G, Gazzard G, Bunce C, Azuara-Blanco A, Munoz B, Friedman DS, Foster PJ. The prevalence of primary angle closure glaucoma in European derived populations: a systematic review. *Br J Ophthalmol* 2012;96(9):1162-7.

2. Day AC, Foster PJ. Increases in rates of both laser peripheral iridotomy and phacoemulsification have accompanied a fall in acute angle closure rates in the UK. *Br J Ophthalmol* 2011;95(9):1339-40.

3. Johnson GJ, Foster PJ. Can we prevent angle-closure glaucoma? *Eye* 2005;19(10):1119-24.

4. Foster PJ, Buhrmann R, Quigley HA, Johnson GJ. The definition and classification of glaucoma in prevalence surveys. *Br J Ophthalmol* 2002;86(2):238-42. 5. Ang LP, Aung T, Chua WH, Yip LW, Chew PT. Visual field loss from primary angle-closure glaucoma: a comparative study of symptomatic and asymptomatic disease. *Ophthalmology* 2004;111(9):1636-40.

6. Sakata LM, Aung T. Can we screen for angle closure? *Am J Ophthalmol* 2011;151(4):567-8.

7. Foster PJ, Devereux JG, Alsbirk PH, Lee PS, Uranchimeg D, Machin D, Johnson GJ, Baasanhu J. Detection of gonioscopically occludable angles and primary angle closure glaucoma by estimation of limbal chamber depth in Asians: modified grading scheme. *Br J Ophthalmol* 2000;84(2):186-92.

8. Devereux JG, Foster PJ, Baasanhu J, Uranchimeg D, Lee PS, Erdenbeleig T, Machin D, Johnson GJ, Alsbirk PH. Anterior chamber depth measurement as a screening tool for primary angle-closure glaucoma in an East Asian population. *Arch Ophthalmol* 2000;118(2):257-63.

9. National Institute for Health and Clinical Excellence (2009) Glaucoma: Diagnosis and Management of Chronic Open Angle Glaucoma and Ocular Hypertension. Available online at:
http://www.nice.org.uk/guidance/CG85, accessed 25/01/15

10.Dabasia PL, Edgar DF, Lawrenson JG. Methods of measurement of the anterior chamber angle. Part 1: Angle closure glaucoma and gonioscopy. *Optom Pract* 2013; 14(3): 107 – 114

11. Dabasia PL, Edgar DF, Lawrenson JG. Methods of measurement of the anterior chamber angle. Part 2: Screening for angle closure and angle closure glaucoma. *Optom Pract* 2013;14(4): 147 – 154

12. Dabasia PL, Edgar DF, Lawrenson JG. Methods of measurement of the anterior chamber angle. Part 3: Screening for angle closure and angle closure glaucoma. *Optom Pract* 2014; 15(1): 11 – 18

13. van Herick W, Shaffer RN, Schwartz A. Estimation of width of angle of anterior chamber. Incidence and significance of the narrow angle. *Am J Ophthalmol* 1969; 68(4): 626–9

14. Thomas R, George T, Braganza A, Muliyil J. The flashlight test and van Herick's test are poor predictors for occludable angles. *Aust N Z J Ophthalmol* 1996;24(3): 251–6

15. College of Optometrists (CoO) (2010). Guidance on the Referral of Glaucoma Suspects by Community Optometrists. Available online at http://www.college-optometrists.org/en/utilities/document-summary.cfm/docid/B7251E0C-2436-455A-B15F1E43B6594206, accessed 25/01/15

16. College of Optometrists (CoO). (2015) Guidance for professional practice.
http://guidance.college-optometrists.org/guidance-contents/knowledge-skills-and-performance-domain/examining-patients-a/#open:111,114. Accessed 26/01/1517. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33(1):159-74.

18. Abrams LS, Scott IU, Spaeth GL, Quigley HA, Varma R. Agreement among optometrists, ophthalmologists, and residents in evaluating the optic disc for glaucoma. *Ophthalmology* 1994;101(10):1662-7.

19. Harper R, Radi N, Reeves BC, Fenerty C, Spencer AF, Batterbury M. Agreement between ophthalmologists and optometrists in optic disc assessment: training implications for glaucoma co-management. *Graefes Arch Clin Exp Ophthalmol* 2001;239(5):342-50.

20. Bailey IL, Bullimore MA, Raasch TW, Taylor HR. Clinical grading and the effects of scaling. *Invest Ophthalmol Vis Sci* 1991;32(2):422-32

21. Sim J, Wright CC. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Phys Ther* 2005 Mar;85(3):257-68.