Original Article

Interobserver reliability when using the Van Herick method to measure anterior chamber depth

Ahmed Javed, Mohamed Loutfi[1], Stephen Kaye[1], Mark Batterbury[1]

Department of Ophthalmology, Ninewells Hospital and Medical School, Dundee, DD2 1UB, [1]St. Paul's Eye Unit, Royal Liverpool University Hospital, Liverpool, L7 8XP, UK

Introduction: The Van Herick method is a quick and easy way to estimate anterior chamber depth, which allows grading of patients according to the likelihood of having primary acute closed-angle glaucoma. However, as the test is highly subjective, measurements and thus grading may vary between observers.

Aim: The aim of this study was to investigate the degree of variation of Van Herick scores among observers and to investigate agreement between temporal and nasal scores.

Materials and Methods: A total of 15 observers measured the temporal and nasal Van Herick scores from 18 patients, grouped into cohorts at outpatient glaucoma and corneal clinic. Analysis of data involved assigning a patient to a Van Herick grade based on the median score and then determining the mean standard deviation and percentage consistency for each grade.

Results: We found that Grades 1 and 4 had a high mean percentage consistency (80% and 84.6%, respectively) and a low mean standard deviation (0.45 and 0.26, respectively). Grades 2 and 3 had low mean percentage consistencies (57.5 and 5, respectively) and high mean standard deviations (0.71 and 0.89, respectively). The temporal and nasal scores showed good agreement ($\kappa = 0.61$ $P < 0.001$).

Conclusion: The Van Herick score has a good interobserver reliability for Grades 1 and 4; however, Grades 2 and 3 require further tests such as gonioscopy or ocular coherence tomography. Temporal and nasal scores demonstrated good agreement; therefore, if the nasal score cannot be measured due to nasal bridge size, the temporal can be used as an approximation.

Keywords: Acute closed-angle glaucoma, anterior chamber depth, Van Herick

Introduction

Glaucoma is an optic neuropathy which is defined by progressive degeneration of the retinal ganglion cell layer.[1] It is traditionally categorized into a chronic open-angle and primary acute closed-angle glaucoma (PACG). It affects more than 70 million people worldwide with around 20% of those having PACG.[1]

The acute form of glaucoma is largely theorized due to a narrower anterior chamber angle compared to the normal population. The anterior chamber angle is formed by the iris root and peripheral cornea. The trabecular meshwork and canal of Schlemm are found within this structure and allow the unobstructed flow of aqueous. In a normal eye, the iridocorneal angle is approximately 30°.[2] Risk factors for PACG are ethnic, such as the Inuit as well as having a family history and increased age.[3] The majority of patients with PACG occur because of a pupil block mechanism wherein the peripheral iris obstructs aqueous outflow by contact with the trabecular meshwork.[4] This leads to a rapid increase in intraocular pressure and patients seek medical attention complaining of severe headache and vomiting. Untreated PACG can result in irreversible blindness.[1] The acute nature of this disorder makes it one of the few ophthalmological emergencies.

Correspondence:
Dr. Ahmed Javed, The Countess of Chester Hospital, Chester, Cheshire, CH2 1UL, UK.
E-mail: ahmedjaved@nhs.net

The advent of YAG laser iridotomy as a preemptive measure means that PACG is a preventable condition in high-risk individuals.[5,6] Therefore, assessing the trabecular meshwork to ensure adequate aqueous outflow is essential to preventing such attacks.[5,6] Gonioscopy involves the use of a mirror or a prism to visualize the structure of the anterior chambers. Gonioscopy is the most commonly used tool to assess the anterior chamber angle; however, it requires significant skill and experience.[7] Computerized techniques to image the angle such as ocular coherence tomography (OCT) or ultrasound biomicroscopy is able to measure the iridocorneal angle; however, these are inefficient as they require technicians and are expensive.[8,9]

A commonly used optical technique is the Van Herick method. In a paper by Herick *et al.* in 1969, this method to measure the anterior chamber depth (ACD) was described.[10] This method involved using the slit lamp to project a light at 60° to the patient's temporal and nasal limbi, and compare ACD to corneal thickness to give a score which approximates an angle measurement. A Grade 4 is where the ACD is equal to or greater than the corneal thickness. In Grade 3, the ACD is ½ of corneal thickness; in Grade 2, it is ¼; and in Grade 1, it is <¼. Grade 0 means the angle is closed. Grade 3 and 4 suggest that the patient is safe as the anterior chamber angle is wide open. In Grade 2, the angle may close, and in Grade 1, the angle is likely to close.

The aim of this research project was to determine the interobserver reproducibility in an outpatient clinic. The secondary aim was to determine the agreement between temporal and nasal scores.

Materials and Methods

Patients were recruited from a corneal and glaucoma clinic at St. Paul's Eye Unit, Royal Liverpool University Hosptial, UK. Pseudophakic patients were excluded from the study. Each patient was examined by the primary clinician and several observers. The observers varied from glaucoma nurse specialists, optometrists, ophthalmology trainee doctors; the primary clinician was an ophthalmology consultant. The observer was given an information sheet detailing the Van Herick method as described in the original paper, then asked to examine one eye only and record his/her score; both temporal and nasal scores were recorded. The nasal side and temporal side were measured first and second alternatively; the right and left eye were also alternated between patients. The Van Herick method used was as described previously in the introduction.

Data analysis

To ascertain the ability of several observers to score a patient, the median score among the observers for each patient was used to classify a patient between Grades 0 and 4. The standard deviation between every observer for a given grade was then calculated. A mean of all the standard deviations for each grade was then calculated and plotted. To calculate percentage consistency for each grade, the mode Van Herick scores among each group of

observers for each patient was calculated as a percentage of the total number of scores. If two grades recurred equally, for example, two Grade 2 and two Grade 1 and one Grade 1; this was taken to have a 40% consistency.

To determine if a correlation existed between temporal and nasal scores, the data were summated and compared for agreement using the Cohen's kappa statistical analysis on SPSS (version 20.0; SPSS Science, Chicago, IL, USA).

Results

The study included 15 observers in total and 18 patients who were organized into four cohorts. The first cohort contained five observers and six patients, one patient were excluded from the results due to peripheral anterior synechiae (PAS). The second cohort had four observers and six patients. The third and fourth cohort had five observers and three patients. There was one patient with a Grade of 1; 4 with a Grade of 2; 5 with a Grade of 3; and 23 with a Grade of 4.

The average percentage agreement between observers for each grade is shown in Table 1. The average standard deviation among the observers for each grade is shown in Table 2. Grade 1 and 4 had the highest percentage of agreement at 80% and 84.6%, respectively; this was mirrored with these two groups also having the lowest mean standard deviations (0.45 and 0.26) [Figure 1]. The nasal and temporal measurements had good agreement $\kappa = 0.61$, $P < 0.001$.

Discussion

The Van Herick method, as first described in 1969, is a quick and easy way to estimate the ACD to assess the likelihood of developing PACG.[10] This is particularly important for new patients who receive mydriatics as this can precipitate PACG.

Since the acceptance of the Van Herick method as a regular part of the ophthalmic examination, the interobserver reliability has

Table 1: The average percentage consistency for each grade

Grade	0-≤1	1-≤2	2-≤3	3-≤4
Average percentage consistency	80	57.5	50	84.6
SD	N/A	20.2	7.1	19.3

SD: Standard deviation, N/A: Not available

Table 2: The mean standard deviation between observer cohorts for each grade

Grade	0-≤1	1-≤2	2-≤3	3-≤4
Average standard deviation between observer cohorts	0.45	0.71	0.89	0.26
SD	N/A	0.25	0.22	0.30

SD: Standard deviation, N/A: Not available

Figure 1: A scatter plot to show the average percentage of agreement for each grade
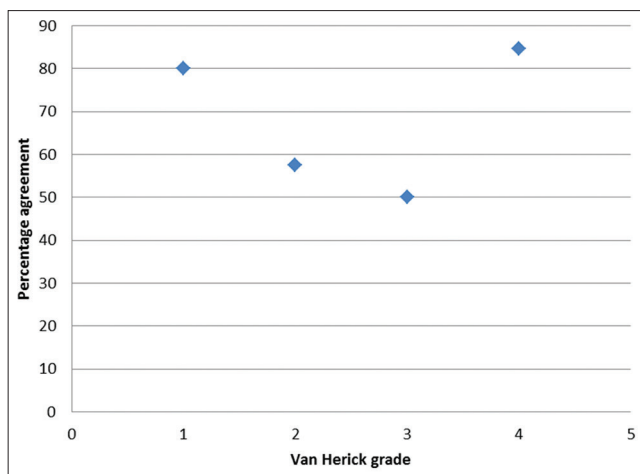

Figure 2: A scatter plot to show the mean standard deviation of agreement for each grade

only been measured by one study.[11] Other studies have shown that the Van Herick method has had poor results as a screening tool for PACG with nasal limbus measurements overestimating the angle size.[12,13] Foster *et al.* showed good reliability between two observers on a modified seven-point Van Herick score system, represented by a kappa score of 0.76.[11] A limitation of Foster's study was the use of only two observers; using a higher number of observers will allow for a more accurate measure of interobserver reliability.[11]

In this study, we used a total of 15 observers, included a glaucoma nurse specialist who regularly performs the Van Herick method, an optometrist, ophthalmology trainees, and consultants. The variety of observer grades is important as the method should not require a high level of skill and should be easy to teach and learn for anyone familiar with the basics of a slit lamp examination.

Before starting our study, it was observed that many clinicians only measured the temporal aspect and not the nasal to take an average. This was partly for practical reasons as we found in several patients, the nasal bridge prevented light from being projected at a 60° angle. In our Caucasian group, we found that often only a temporal score was possible. As the Van Herick test can be subjective, it was expected that there would be differences in scores between observers; however, it was important to ascertain if these differences were clinically significant. A clinician scoring Grade 3 when in reality the patient is Grade 4 or vice versa would not cause much harm. However, if a clinician scores a Grade 3 when in reality the angle should be a Grade 2 or even 1 is putting the patient at risk of an episode of PCAG.

From our results, we can see that Grade 4 has the lowest mean standard deviation among all the grades [Figure 2]. This is to be expected as clinicians should be able to say that the anterior chamber is wide open and the ratio of the corneal surface thickness and ACD is easily visible if it is 1:1 or greater. Grade 3 had the least consistency between observers and the highest mean standard deviation [Table 1]. This emphasizes the subjectivity of
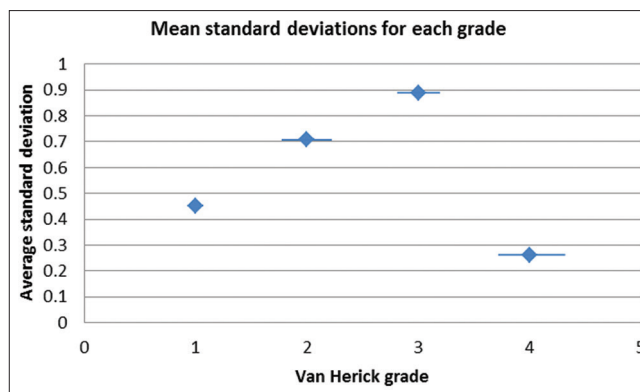
the test although Grade 3 is still a safe grade, patients who are considered to be Grade 3 could, in actual fact, be a Grade 2 which would necessitate further clinical examination. Grade 2 scores also showed poor consistency and had a high mean standard deviation of scores [Table 1]. Grades 2 and 3 are clearly difficult to differentiate using the Van Herick; however, it is an important grade as Grades 2 and below can be considered at risk of closing and Grade 3 can be considered open. The poor consistency in these two grades highlights that further clinical examination is necessary whether this is gonioscopy or OCT if a clinician suspects a Grade 3 or 2. Grade 1 showed a good consistency with a high percentage agreement and a low mean standard deviation. The fact that a Grade 1 shows a good consistency means that the observers were able to consistently detect an anterior chamber they felt was narrow. Although Grade 1 results were promising, there was only one patient in this category. This means that the standard deviation for the percentage consistency and mean standard deviation were not calculated. The high consistency between Grades 1 and 4 means that the test is useful if the patient is likely to have an attack or not likely to have an attack at all, and thus it is a useful screening device; however, in the middle grades, it should be remembered that the high variability of scores mean that further assessment using more advanced techniques need to be used to ensure patient safety.

The temporal and nasal scores showed good agreement with each other; however, this agreement was not considered to be strong agreement as the κ value was 0.61. Thus, both sides as far as possible should be measured as originally suggested. Taking an average of the two may not be useful. This is because if nasally the patient scored a 2 and temporally a 4, then the average grade is a 3, which means that the patient is considered to have an open angle, rather one should side with precaution and take the lowest score of the two.

In our cohort, we excluded patients who were pseudophakic as this opens the anterior chamber; however, we had one patient with a PAS. This patient was scored 0 in the temporal side by 3 of the observers; however, one scored him as a 3 and another as 2, nasally everyone scored him as a 4. We excluded this patient from the temporal to nasal statistical analysis. It is important to

note that, in patients with PAS, the score is highly dependent on which part of the limbus is examined as the anterior chamber angle may seem closed at one point yet at different millimeters above or below that point can be seen to be open. Variation of grades, such as this, means that patients with PAS require caution with the Van Herick test and should be considered for OCT or gonioscopy as routine.

We present several novel points, first, Grade 1 and 4 have a high interobserver reliability as both have low mean standard deviations, 0.45 and 0.26, respectively, as well as having a high mean percentage agreement of 80% and 84.6%, respectively. Grades 2 and 3 have poor interobserver reliability as the mean standard deviations are 0.71 and 0.89, respectively, and a low mean percentage consistency of 57.5 and 50, respectively. We also show good agreement between nasal and temporal scores ($\kappa = 0.61$) unless a patient has PAS. In conclusion, the Van Herick test is a useful screening tool if the Grade is 1 or 4, however, further investigation needs to take place if the result is a Grade 2 or 3.

Conclusion

The Van Herick score has a good interobserver reliability for Grades 1 and 4, however, Grades 2 and 3 require further tests such as gonioscopy or OCT. Temporal and nasal scores are in good agreement, therefore, if the nasal score cannot be measured due to the nasal bridge size, the temporal can be used as an approximation.

Financial support and sponsorship
Nil.

Conflicts of interest
There are no conflicts of interest.

References

1. Weinreb RN, Aung T, Medeiros FA. The pathophysiology and treatment of glaucoma: A review. JAMA 2014;311:1901-11.
2. Rüfer F, Schröder A, Klettner A, Frimpong-Boateng A, Roider JB, Erb C. Anterior chamber depth and iridocorneal angle in healthy white subjects: Effects of age, gender and refraction. Acta Ophthalmol 2010;88:885-90.
3. Nongpiur ME, Ku JY, Aung T. Angle closure glaucoma: A mechanistic review. Curr Opin Ophthalmol 2011;22:96-101.
4. Gazzard G, Friedman DS, Devereux JG, Chew P, Seah SK. A prospective ultrasound biomicroscopy evaluation of changes in anterior segment morphology after laser iridotomy in Asian eyes. Ophthalmology 2003;110:630-8.
5. He M, Friedman DS, Ge J, Huang W, Jin C, Cai X, *et al.* Laser peripheral iridotomy in eyes with narrow drainage angles: Ultrasound biomicroscopy outcomes. The Liwan eye study. Ophthalmology 2007;114:1513-9.
6. Kanski JJ. Clinical Ophthalmology: A Systematic Approach. 6th ed. Edinburgh, New York: Butterworth-Heinemann, Elsevier; 2007. p. vii, 931.
7. Park SB, Sung KR, Kang SY, Jo JW, Lee KS, Kook MS. Assessment of narrow angles by gonioscopy, van Herick method and anterior segment optical coherence tomography. Jpn J Ophthalmol 2011;55:343-50.
8. Wang D, Pekmezci M, Basham RP, He M, Seider MI, Lin SC. Comparison of different modes in optical coherence tomography and ultrasound biomicroscopy in anterior chamber angle assessment. J Glaucoma 2009;18:472-8.
9. Radhakrishnan S, Goldsmith J, Huang D, Westphal V, Dueker DK, Rollins AM, *et al.* Comparison of optical coherence tomography and ultrasound biomicroscopy for detection of narrow anterior chamber angles. Arch Ophthalmol 2005;123:1053-9.
10. 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;68:626-9.
11. Foster PJ, Devereux JG, Alsbirk PH, Lee PS, Uranchimeg D, Machin D, *et al.* 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;84:186-92.
12. Congdon NG, Youlin Q, Quigley H, Hung PT, Wang TH, Ho TC, *et al.* Biometry and primary angle-closure glaucoma among Chinese, white, and black populations. Ophthalmology 1997;104:1489-95.
13. 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;24:251-6.