tvst

https://doi.org/10.1167/tvst.8.5.16

Article

Remote Grading of the Anterior Chamber Angle Using Goniophotographs and Optical Coherence Tomography: Implications for Telemedicine or Virtual Clinics

Jack Phu[1,2], Henrietta Wang[1,2], Vincent Khou[1,2], Sophia Zhang[1,2], and Michael Kalloniatis[1,2]

[1] Centre for Eye Health, University of New South Wales, Kensington, NSW, Australia
[2] School of Optometry and Vision Science, University of New South Wales, Kensington, NSW, Australia

Correspondence: Jack Phu, Centre for Eye Health, University of New South Wales, Kensington, NSW, Australia. e-mail: jphu@cfeh.com.au

Received: 7 May 2019
Accepted: 4 August 2019
Published: 25 September 2019

Keywords: glaucoma; angle closure; narrow angles; gonioscopy; anterior segment optical coherence tomography

Citation: Phu J, Wang H, Khou V, Zhang S, Kalloniatis M. Remote grading of the anterior chamber angle using goniophotographs and optical coherence tomography: implications for telemedicine or virtual clinics. Trans Vis Sci Tech. 2019;8(5):16, https://doi.org/10.1167/tvst.8.5.16

Purpose: To evaluate the agreement and accuracy of grading goniophotographs and anterior segment optical coherence tomography (AS-OCT) results for assessment of the anterior chamber angle, and elicit factors driving concordance between perceived grade and ground truth.

Methods: Three clinicians evaluated the goniophotographs and AS-OCT results of 75 patients. Graders' impressions of the angle grade, trabecular pigmentation, and iris contour were compared with the ground truth gonioscopic examination result when physically performed by a senior optometrist. Percentage agreement and kappa statistics were calculated. Binary logistic regression was used to elicit factors for accurate grading.

Results: Exact angle matches and binary (open or closed) evaluations were above guessing rate for all graders. There was a systematic bias toward underestimating the angle structure across all graders, especially at the superior angle, by approximately 1 ordinal unit. Kappa statistics showed fair-moderate agreement for exact (0.387–0.520) and binary (0.347–0.520) angle evaluations. Agreement was unchanged when using a multimodal approach (0.373–0.523). Factors driving concordance were primarily related to the extremes of the anterior chamber angle configuration (shallow or deep structures, and iris contour). However, prediction models did not fully explain the levels of concordance with the ground truth (maximum R^2 amongst models 0.177).

Conclusions: Although moderate agreement between graders and ground truth could be obtained under binary evaluations, angle grades were generally underestimated. Factors affecting concordance were primarily the extremes of the ground truth angle and iris contour.

Translational Relevance: We highlight factors affecting accuracy of grading goniophotography and AS-OCT images of the anterior chamber angle.

Introduction

Glaucoma is one of the leading causes of irreversible blindness worldwide, and its rate of underdiagnosis highlights the need for improving case detection in the general community.[1–6] Poor accessibility to appropriate eye care required for accurate diagnosis has been cited as a reason for underdiagnosis.[7] Several collaborative care, referral refinement, and telemedicine or teleglaucoma pathways have been developed in different health care settings, designed to address issues such as the unequal distribution of health resources, provision of timely care, and provision of a channel for expert glaucoma care.[8–13]

One specific improvement to glaucoma care driven by these alternative collaborative care or virtual clinic pathways has been the ability to refine patient cohorts that necessitate more specialized eye care, which in the most appropriate cases are onward referred for timely treatment.[9–11] Data required to appropriately titrate glaucoma risk and triage patients for open angle glaucoma are highly conducive for these suggested clinical pathways, as results such as automated perimetry,[14,15] color fundus photography,[16–18] and optical coherence tomography (OCT)[19] are readily captured and interpreted in remote settings. However, a complete glaucoma assessment also necessitates assessment of the anterior chamber angle.[20,21] Though previous studies have examined concordance of gonioscopy and other anterior segment imaging techniques by physical examiners,[22,23] there remains a gap in the knowledge regarding concordance between clinicians for grading goniophotography images and anterior segment OCT (AS-OCT) under remote settings. Understanding the conduciveness of anterior segment imaging results for remote interpretation is critical for successful implementation of alternative health care pathways, such as telemedicine approaches, for case detection of angle closure spectrum disease. Although a recent large clinical trial has shown that relatively few high-risk patients eventually progress to angle closure glaucoma even in the presence of prophylactic laser iridotomy, angle closure glaucoma remains a significant cause of irreversible blindness in the glaucoma family of diseases and may be preventable if detected early enough.[24]

Therefore, the purpose of this study was to examine the concordance of assessing goniophotographs and AS-OCT by graders on a virtual system, in order to assess this aspect of the feasibility of remote virtual clinic or collaborative care approaches for angle closure disease. We also evaluated potential reasons or features of the anterior chamber angle that could affect agreement and accuracy, as strategies to exploit factors that may improve concordance may be desirable.

Methods

Patient Images for Analysis

The medical records of consecutive patients referred to the Centre for Eye Health for anterior chamber angle assessment[25] were examined for suitability for inclusion in this project. During the study period, data used for analysis were collected in a prospective fashion for the purpose of this study, and the cohort was consecutively examined to reduce potential spectrum bias.[26] The inclusion criteria included the following: patient age > 18 years; having provided written informed consent for their de-identified medical images to be used for research purposes; having undergone a complete anterior chamber angle and glaucoma assessment at the Centre for Eye Health; and having images of sufficient quality for remote viewing on a digital platform (see below section). Exclusion criteria included the following: having not provided informed consent; having an incomplete data set for analysis; and having images that were of insufficient quality for grading. The study adhered to the Declaration of Helsinki, and ethics approval was provided by the Human Research Ethics Committee of the University of New South Wales. In total, 75 subjects (mean age 57.1 years, SD 10.2 years; 30 males, 45 females) met the inclusion criteria for analysis during the study period.

Images for Analysis

Goniophotographs were captured using a standardized procedure for all patients. Following instillation of two drops of topical anesthetic (proparacaine hydrochloride 0.5%; Alcaine, Novartis Pharmaceuticals, NSW, Australia), the patient was lined up on a slit lamp (Haag-Streit BX900, Device Technologies, Belrose, Australia) and a goniolens (G4, Volk Optical, Mentor, OH; or Ocular Four Mirror Mini Gonio, Ocular Instruments, Bellevue, WA) with coupling gel (carbomer 980 0.2%; Viscotears Gel PF 0.6 mL, Bausch & Lomb Australia, Chatswood, Australia) was placed on the eye. The slit beam width was reduced to a maximum of 2 mm, and its height was reduced to a maximum of 5 mm. Instrument magnification was set to 15×. The light-emitting diode light source of the slit lamp was set at the lowest illumination setting and the neutral density filter (10%) was used; that is, the lowest light condition under which goniophotography could be reliably performed was used in order to attempt to replicate the conditions of the AS-OCT. The slit lamp aperture was set at 2. Room lighting was off throughout testing. The Canon 5D Mark IV (Canon, Tokyo, Japan) served as the attached camera on the slit lamp, and it was set to ISO 400, f-stop of f/0 and shutter speed of a 1/200 of a second. For consistency, the same slit lamp, camera system, room, and lighting set up was used for every subject within the study.

For capturing the angle photograph, the slit beam was oriented such that it was always approximately parallel to the quadrant: horizontal for the superior and inferior angles, and vertical for the nasal and temporal angles. Instead of using the usual corneal wedge technique, the parallel orientation of the slit beam in conjunction with the thin beam helped to reduce the amount of light entering the pupil whilst simultaneously providing a wider impression of the quadrant. During the examination, the clinician was able to tilt or manipulate the goniolens for full clinical documentation; however, for consistency for the purposes of the study, the primary gaze result was used as the ground truth to simplify the grading process. The photographs were therefore captured while the lens was in primary gaze (no tilt) and with no pressure on the eye. Each photo was saved individually as a JPEG file derived from the original RAW file. Each patient contributed a total of eight images for analysis, four (superior, inferior, nasal, and temporal quadrants) from each eye.

AS-OCT was performed using the Spectralis OCT (Heidelberg Engineering, Heidelberg, Germany) with the anterior segment module. The 1 anterior chamber angle (ACA) scan protocol was used, at an automatic real time (ART) level of at least 50. Scans were taken at the nasal and temporal meridians, as close to the horizontal midline as possible. Scanning was performed in a dark room, with the patient fixating upon a dark external target to minimize artificial pupil constriction. The resultant scan was directly exported as a JPEG file using the Heidelberg Eye Explorer software (Heidelberg Engineering). Each patient contributed four images for analysis (nasal and temporal from each eye).

Image Grading

Three independent, masked graders were tasked with grading the images. The graders were highly experienced optometrists staffing the Glaucoma Management Clinic within the Centre for Eye Health,[10] and regularly assess, diagnose, and manage patients with glaucoma. The 600 goniophotographs and 300 AS-OCT images were evaluated on a computer screen. The photographs were examined as per the exact orientation of image capture; however, the grader was free to rotate or enlarge the image as required. Contrast levels could also be adjusted as per the grader's preference.

For the goniophotographs, the grader recorded the following information: the deepest visible angle structure (no structures, Schwalbe's line, anterior trabecular meshwork, posterior trabecular meshwork, scleral spur, and ciliary body band), the amount of trabecular meshwork pigmentation if applicable (none, mild, moderate, or heavy), and the iris contour (flat, regular, or steep). These were then converted into a categorical scale. Though the distance between units were not necessarily linear, these were regarded as ordinal for the purpose of the study. For angle structures, a 0 to 4 grade was used, with no structures or Schwalbe's line representing 0 (as they are functionally similar for grading), and then each subsequent structure was 1 greater on the scale. As the main purpose of the study was to examine grading based on structure, as per current clinical guidelines for diagnosis of angle closure spectrum disease,[20,27,28] we provide further details regarding the assessment of the trabecular pigmentation and iris contour in the Supplementary Material. Other features of the gonioscopic examination, although noted clinically, were not a focus of the present study and were therefore not evaluated by the graders.

For the AS-OCT scans, the grader was asked to infer what structure they predict would be visible if gonioscopy was performed at the same meridian. The same grading scheme was used as per the gonioscopy grading above. The inference was made on the basis of clinical experience, where an imaginary tangent is drawn along the iris contour to the inferred angle structure. Examples of goniophotographs and AS-OCT scans are shown in Figure 1. Note that the grading of each image was performed independently of all other images, such that goniophotographs were not matched to the same patient's OCT result.

One clinician, masked to the graders' impressions and to the final diagnosis, performed the grading of the quality of the photography (amount of blur and obscuration of the photographs in an incremental scale of 25% intervals of the entirety of the image) and for the visibility of key angle structures on AS-OCT (Schwalbe's line, Schlemm's canal, scleral spur, and ciliary body band).

For both the goniophotographs and the AS-OCT results, the ground truth angle structures, pigmentation, and iris configuration used in the present study were derived from the physical examination gonioscopy results (not photography) using a goniolens conducted by a single senior optometrist working within the Centre for Eye Health. The ground truth result was agreed upon by remote review by an ophthalmologist, who could contact the examining clinician for additional clinical information, or could consult with another ophthalmologist for equivocal cases.[11] For consistency, the same single optometrist, who had previously demonstrated high gonioscopic agreement (number of angles with exact match 69/96 [71.9%] and binary match 95/96 [99.0%] when performing gonioscopy; of the discordant cases, most were within 1

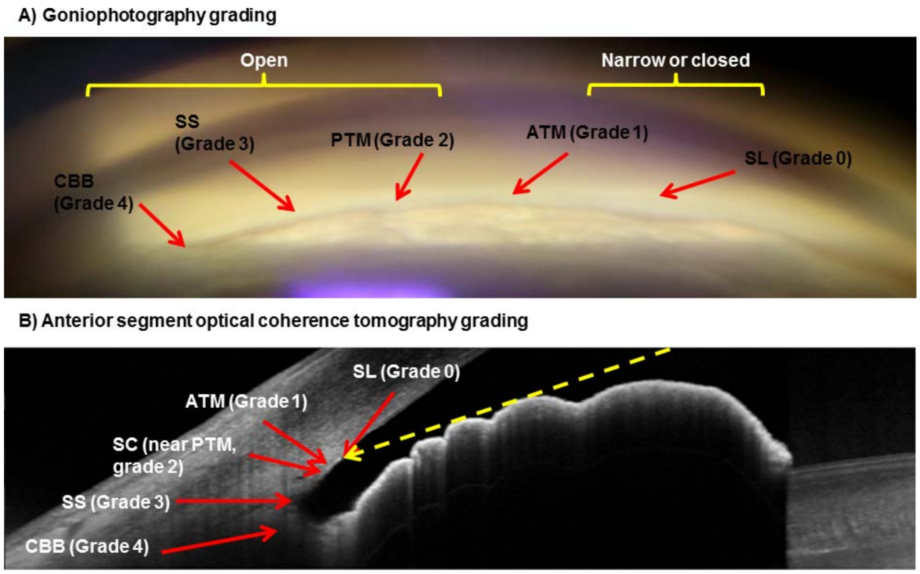

Figure 1. Examples of images used for grading in the present study (A, goniophotograph; B, AS-OCT scan). Note that these are angles from different eyes. The grades used for assessing the deepest visible angle structure are shown: CBB, ciliary body band; SS, scleral spur; PTM, posterior trabecular meshwork; ATM, anterior trabecular meshwork; SL, Schwalbe's line.

ordinal unit [20/27, 74.1%]) in terms of physical examination with a glaucoma specialist ophthalmologist within the Centre for Eye Health, performed the examinations and captured all images used in the study. The reason for having only one senior optometrist obtaining the ground truth gonioscopy result was because of the potential variability and generally only fair-to-moderate agreement in exact grading between optometrists and ophthalmologists,[29] and so this examining clinician needed to specifically demonstrate a high level of agreement with the glaucoma specialist prior to conducting this study. These records were extracted directly from the patient's file and compared with the graders' results. The ground truth consisted of the following angle grades on gonioscopy: 49 (8.2%) no structures visible, 36 (6.0%) anterior trabecular meshwork, 178 (29.7%) posterior trabecular meshwork, 134 (22.3%) scleral spur, and 203 (33.8%) ciliary body band. Consecutive and nontargeted (i.e., not a necessarily high-risk cohort, as the overall diagnostic yield for angle closure disease within this clinic was only around 36%[25]) subject recruitment was performed to minimize the potential effect of spectrum bias, and to reflect a more real clinical scenario within which the majority of the distribution of assessed angles are likely to be open. Thus, only 14.2% of the total sample had closed angles.

Multimodal Approach to Grading

To test if using both imaging modalities could be additive in terms of the resultant accuracy of grading, the graders were asked to regrade the images and were given both goniophotography and AS-OCT scans for the same patient. The grading and recording were as per the above, but limited to only angle structures and iris contour, and limited to the nasal and temporal quadrants, as these were mutually assessed using goniophotography and AS-OCT.

Factors Affecting Angle Evaluations Using Goniophotography and AS-OCT

We also used binary logistic regression analysis (SPSS Statistics version 25; IBM Corporation, New York, NY) to determine whether there were factors that could account for the agreement between each grader and ground truth using the goniophotographs. Agreement (coded in a binary fashion) was used as the dependent variable, and covariates for analysis are listed in Supplementary Table S1. Note that some factors were different for goniophotography (blur and obscuration of the goniophotographs) and AS-OCT (visibility of key anatomical structures) when performing the analysis.

The binary logistic regression analysis (the primary outcome was correct or incorrect grading) was performed for both exact structure matches and for binary grading, and separately for each grader. The models were assessed using Nagelkerke R^2, and a $P < 0.05$ was considered significant for individual covariates. Covariates listed as significant on the parameter estimates were then compared across all graders and conditions. Since the covariates were nonbinary, parameter estimates were also able to identify specific levels of the ordinal scale that contributed significantly to grader accuracy relative to the ground truth.

Statistical Analysis

Statistical analysis was carried out using GraphPad Prism version 8 (GraphPad, La Jolla, CA) and SPSS Statistics version 25 (IBM Corporation). Owing to the use of ordinal data, we were able to directly compare the responses between grader and ground truth, and generated difference plots between them (grader score – ground truth score). A positive difference indicated that the grader thought that the angle was more open than the ground truth, while a negative difference indicated that the grader thought the angle was narrower.

Agreement between each grader and ground truth was firstly examined using the number of exact matches (i.e., a difference of 0), and was expressed as a proportion of total comparisons. For angle structures, this meant that a proportion of 0.2 represented a guessing rate. Alongside exact matches of angle structure, exact matches were also examined in terms of whether the angle was classified as narrow/closed (ordinal grade 0 or 1) or open (ordinal grade 2 or greater). For this binary analysis, a proportion of 0.5 represented the guessing rate.

Intraobserver agreement using two different techniques (not repeatability, which we refer to as repeated grading by the same observer using specific techniques) was assessed using impressions of the nasal and temporal angles obtained for goniophotography and AS-OCT. This was performed as per the above method for proportion agreement for exact angle match and binary matches.

To complement the proportion agreement statistic, we also performed kappa statistic calculations. When exact matches were considered, results were arranged in a 5×5 matrix and compared using weighted Fleiss's kappa. When open versus narrow/closed classifications were considered, a 2×2 matrix was used to determine agreement using Cohen's kappa. However, we note that kappa statistics could be confounded by selective biases attributable to a conservative grading behavior that may be adopted by graders.

Note that sensitivity, specificity, and area under the receiver operating characteristic were not calculated in the present study. This was due to distribution of cases within the present cohort, where such a small sample of patients with angle closure disease requiring intervention would easily lead to a skewed result.

Results

Proportion Agreement of Angle Structure Evaluation Using Goniophotography

We firstly determined the proportion of times exact agreement occurred between each grader and the ground truth using goniophotography (Figs. 2A, 2B). There was no significant difference between graders ($F_{2,6} = 0.098$, $P = 0.9081$), and no difference between eyes ($F_{1,3} = 0.2127$, $P = 0.6761$). When right and left eye results were pooled within observer, there was a significant effect of angle direction (H(4) = 14.60, P = 0.0002), but multiple comparisons only showed a significant difference between inferior and superior directions (P = 0.0021). Proportions were significantly above guessing rate for all directions (inferior, $P < 0.0001$; superior, $P = 0.0235$; nasal, $P < 0.0001$; temporal, $P = 0.0005$).

When open versus closed judgements were considered, there was again no significant difference between graders ($F_{2,6}$ = 1.321, P = 0.3347) or eye ($F_{1,3} = 0.6517$, $P = 0.4786$) (Figs. 2C, 2D). There was a significant effect of angle direction (H(4) = 20.52, P = 0.0001), but multiple comparisons only showed a difference between superior and inferior directions (P = 0.0001), inferior and temporal (P = 0.0369), and superior and nasal (P = 0.0170). Proportions were significantly above guessing rate for all directions (inferior, $P < 0.0001$; superior, $P = 0.0283$; nasal, $P < 0.0001$; and temporal, $P = 0.0002$).

The ordinal scales were used to assess for systematic biases in disagreements across all three observers (Fig. 3). For all observers, where there were differences the angle was graded narrower than the ground truth (except temporally for grader 2). The superior angle tended to be underestimated by the greatest magnitude across all observers, and this was significant between superior and inferior angles for graders 1 (P = 0.0043) and 2 (P = 0.0002), and between superior and temporal angles for graders 1 ($P > 0.0001$) and 2 ($P = 0.0002$). Grader 1 also showed a significant difference between nasal and temporal estimates ($P = 0.0224$).

Agreement of Angle Structure Evaluation Using AS-OCT

For AS-OCT results, there was no effect of eye laterality ($F_{1,1} = 6.760$, $P = 0.2338$), nor was there an effect of grader ($F_{2,2} = 8.714$, $P = 0.1029$), similar to the goniophotography grading (Figs. 4A, 4B). The

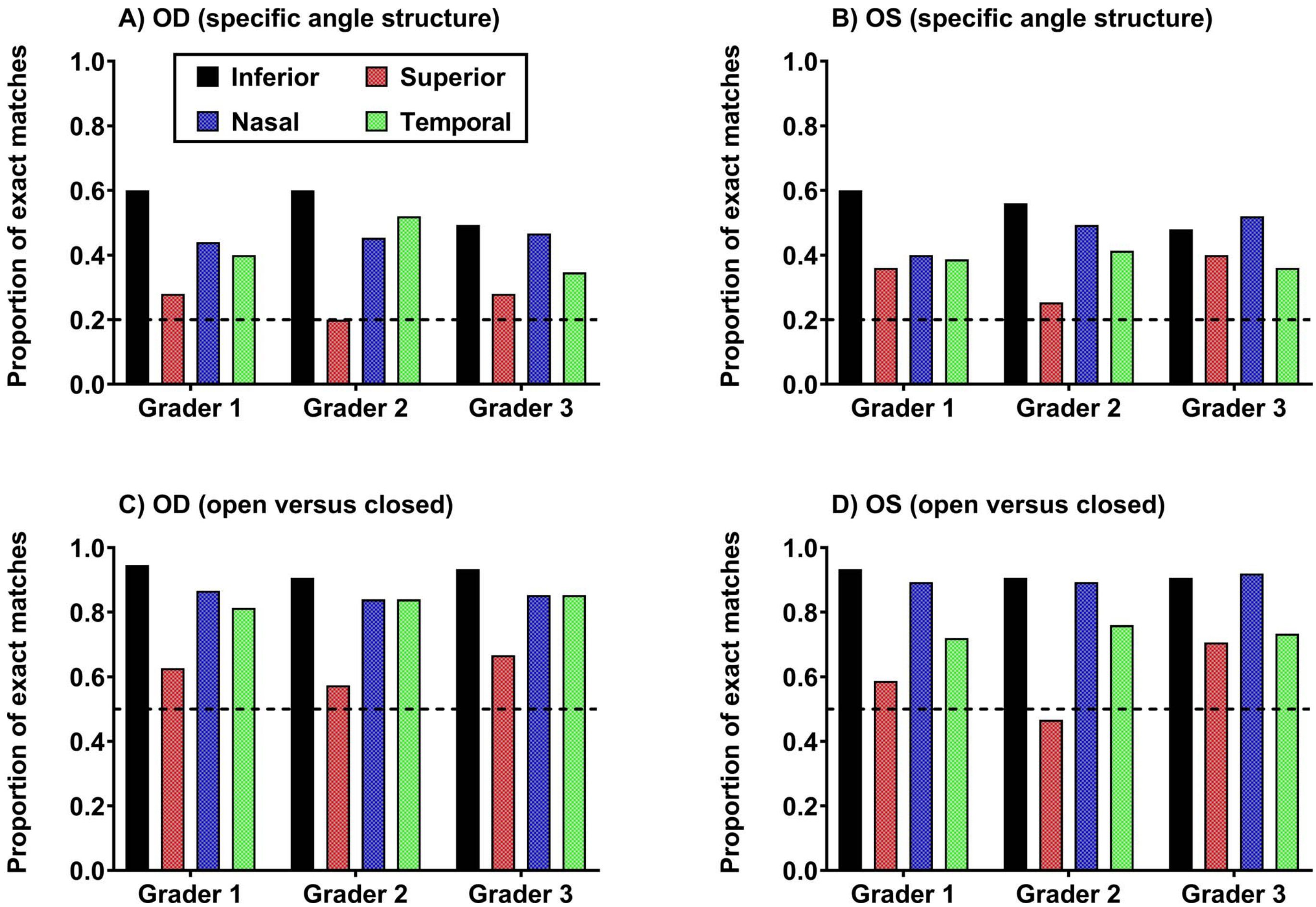

Figure 2. Proportion of exact matches for each grader for each angle direction (coded by color). Right and left eye results are shown separately. Specific angle structure matches are shown in (A) and (B), with the *horizontal dashed line* indicating the guessing rate of 0.2. Open versus closed matches are shown in (C) and (D), with the *horizontal dashed line* indicating the guessing rate of 0.5.

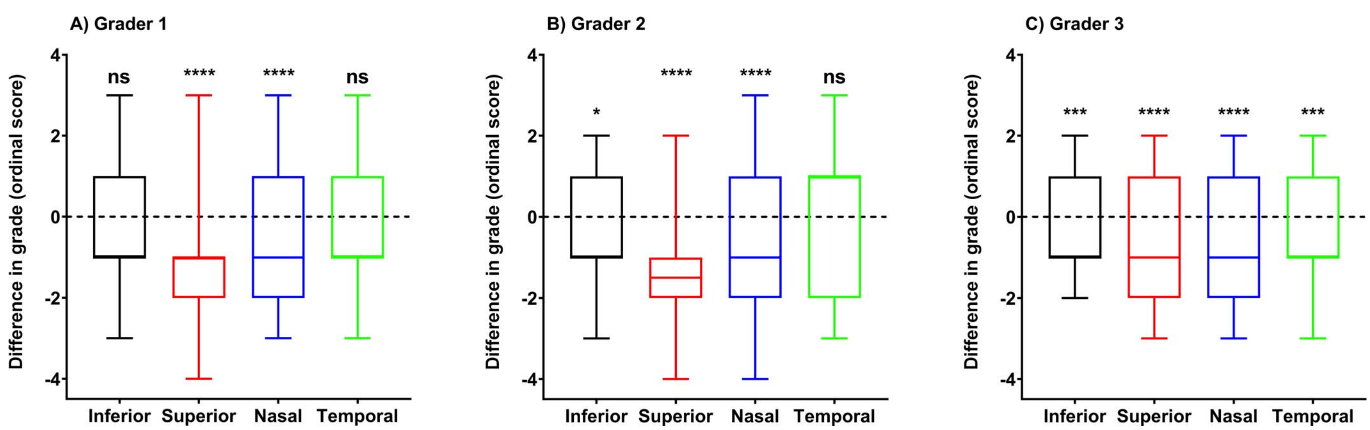

Figure 3. Box and whiskers plot (*horizontal lines* indicate median, the *boxes* indicate quartiles, and the *tails* indicate the range) showing difference in grade in ordinal score (grader – ground truth) for each angle direction. Each grader's results are shown separately, but right and left eyes were pooled together for each observer. The *black horizontal dashed line* indicates no difference in grade. The *asterisks* above indicate the level of significance for a one-sample *t*-test (difference from 0; *$P < 0.05$; ***$P < 0.001$; ****$P < 0.0001$) and ns, indicates $P > 0.05$.

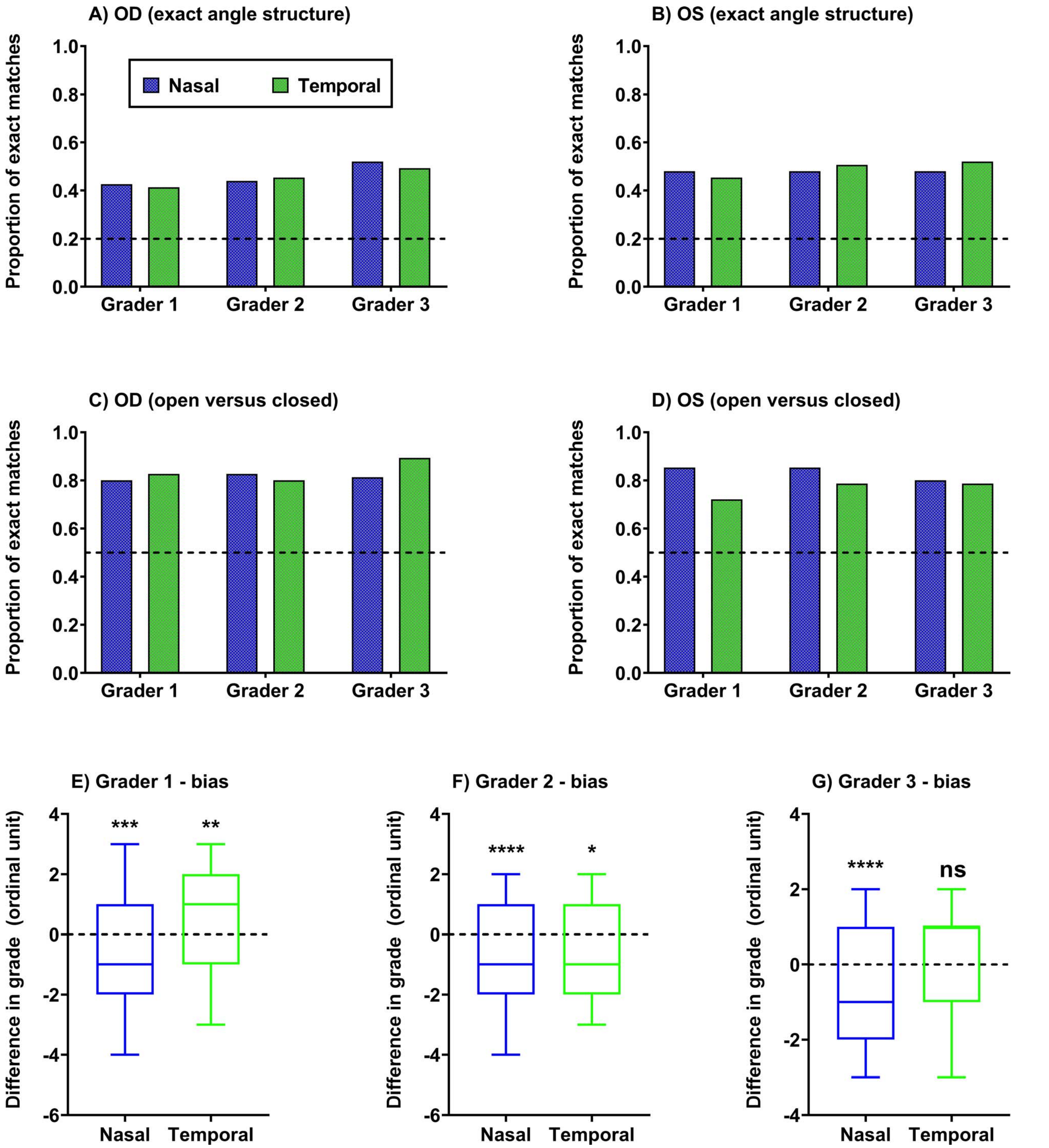

Figure 4. (A, B) Proportion of exact matches for each grader for each angle direction (coded by color). The *horizontal dashed line* indicates the guessing rate of 0.2. (C, D) Proportion of binary (open versus closed) matches for each grader for each angle direction (coded by color). The *horizontal dashed line* indicates the guessing rate of 0.5. For A–D, right and left eye results are shown separately. (E–G) Box and whiskers plot (*horizontal lines* indicate median, the *boxes* indicate quartiles, and the *tails* indicate the range) showing difference in exact grade in ordinal score (grader − ground truth) for nasal and temporal angle directions. Each grader's results are shown separately, but right and left eyes were pooled together for each observer. The *black horizontal dashed line* indicates no difference in grade. The *asterisks* above indicate the level of significance for a one-sample *t*-test (difference from 0; *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$; ****$P < 0.0001$) and ns, indicates $P > 0.05$.

rates of agreement were significantly greater than guessing rate and both nasal ($P = 0.0002$) and temporal ($P < 0.0001$) following pooling. When open versus closed judgements were considered, there was again no effect of laterality ($F_{1,1} = 0.2975$, $P = 0.6821$) or grader ($F_{2,2} = 0.2482$, $P = 0.8010$), and agreement levels were significantly greater than guessing across all conditions as well (nasal, $P < 0.0001$; temporal, $P < 0.0001$) (Figs. 4C, 4D). All graders had a tendency for grading the angle as narrower than the ground truth (Figs. 4E–4G).

Intraobserver Agreement Between Goniophotography and AS-OCT

Nasal and temporal angle directions had both examination results for goniophotography and imaging, and thus we were able to examine intraobserver agreement between the techniques for the same subject. Percentage agreement was similar across all three graders (1–3, respectively) for exact angle structure (31.7%, 38.3%, 40.7%) and for binary matching (77.3%, 79.3%, 84.5%).

These data were also analyzed to determine if one technique had a greater level of agreement with the ground truth within each observer, since each image was graded independently. There was no systematic difference between techniques for each grader for determination of exact angle structure ($P = 0.3485$–>0.9999) or for open versus closed judgements ($P = 0.3547$–0.5087), suggesting that graders had similar impressions of the angle with each technique even when assessed and used independently, and that intraobserver consistency was maintained.

Does a Multimodal Approach for Grading Improve Accuracy?

When both goniophotography and AS-OCT imaging results were available to the graders (nasal and temporal angles only), there was no change in exact grading ($H(3) = 2.667$, $P = 0.3611$) or binary accuracy ($H(3) = 2.000$, $P = 0.5278$) across each grader when right and left eye results were pooled together (Fig. 5). Grader 2 maintained a tendency to underestimate the angle width, while graders 1 and 3 had no systematic bias.

Agreement Using Kappa Statistics

Levels of agreement were assessed between grader and the ground truth, and also between each grader (Table). There was no significant difference between kappa values for exact structure matching or a binary choice between open and closed for goniophotography ($P = 0.6875$) or OCT ($P = 0.0625$). There was a tendency for higher agreement between graders in comparison to individual grader compared with the ground truth when comparing a binary decision ($P = 0.0312$), but this did not reach statistical significance for exact matches ($P = 0.0625$). There was a slight tendency for AS-OCT to have higher agreement compared with the corresponding condition assessed using goniophotography, but this was not consistent amongst observers and so therefore did not reach statistical significance (exact matches: $P = 0.5625$ open versus closed: $P = 0.0938$). When both imaging modalities were provided to the graders, there was a slight increase in agreement for graders 1 and 3, but a slight decrease for grader 2.

Using kappa statistics, intraobserver agreement—assessed by the consistency between goniophotography and OCT—for exact angle structure varied across all observers (graders 1: 0.234; 2: 0.360; and 3: 0.443). Agreement between techniques when a binary grade was considered was higher, but still only fair to moderate (graders 1: 0.300; 2: 0.396; and 3: 0.597). This suggests that each grader used goniophotography and OCT differently when evaluating the angle structures. Kappa values were similar across all pairwise grader comparisons when a multimodal approach was used.

Factors Affecting Agreement in Grading Images

Binary logistic regression was used to explore factors that may explain the levels of concordance with the ground truth with respect to grading angle structures. When considering grading using goniophotography, factors that commonly emerged as being significant for concordance were ground truth angles at the "extremes" of the angle spectrum (no structures visible or scleral spur/ciliary body), and to a lesser extent, the ability to visualize pigmentation (Supplementary Table S3). For grading using OCT, the iris contour was an important factor for determining concordance (Supplementary Table S4). Similar themes occurred when considering the combination of goniophotography and OCT (Supplementary Table S5). However, these covariates did not fully explain the agreement between grader and ground truth or between graders, suggesting individual and other factors that were not revealed in the study played a role in concordance.

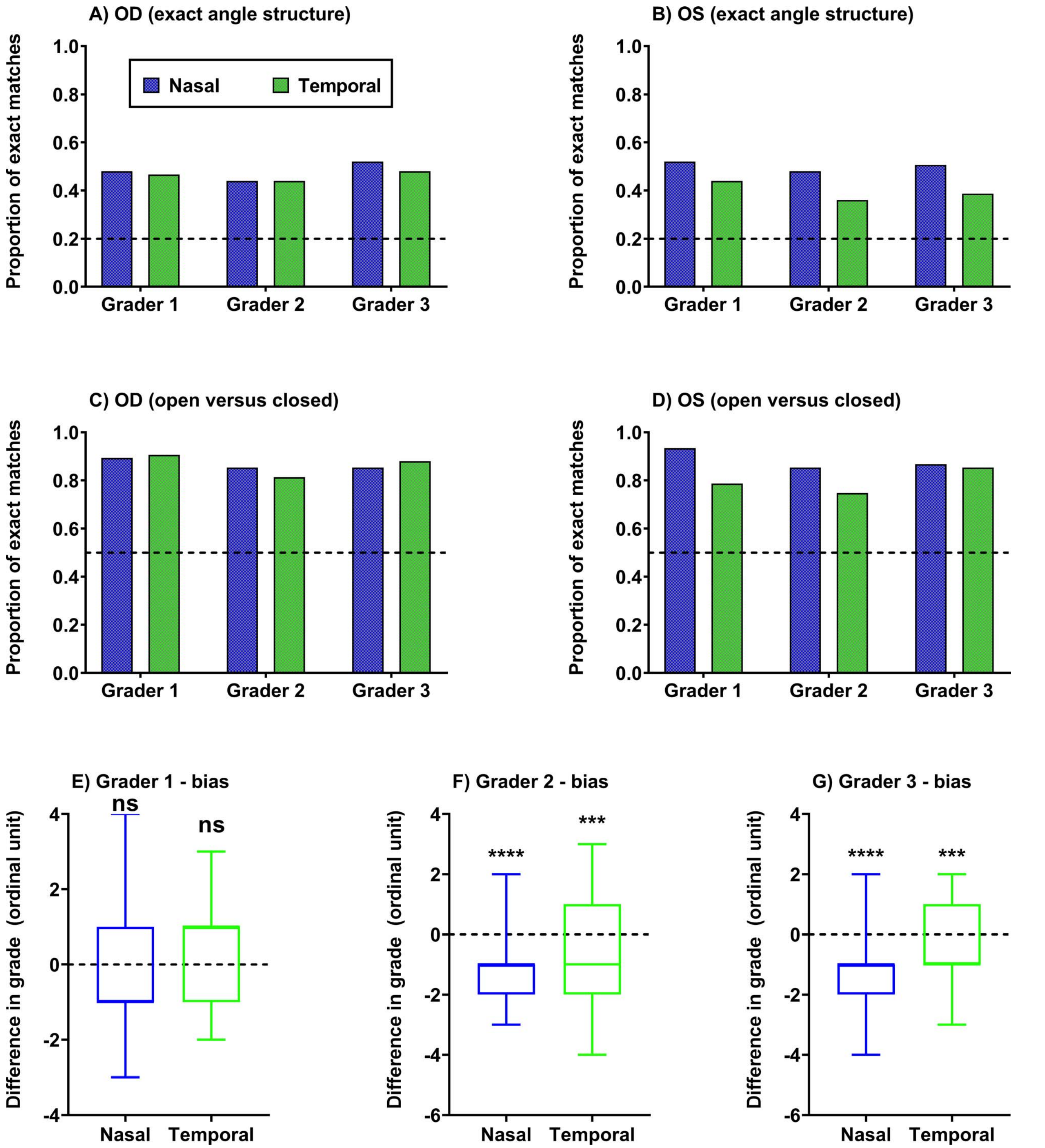

Figure 5. (A, B) Proportion of exact matches for each grader for each angle direction (coded by color) when a multimodal approach was used. The *horizontal dashed line* indicates the guessing rate of 0.2. (C, D) Proportion of binary (open versus closed) matches for each grader for each angle direction (coded by color). The *horizontal dashed line* indicates the guessing rate of 0.5. For A–D, right and left eye results are shown separately. (E–G) Box and whiskers plot (*horizontal lines* indicate median, the *boxes* indicate quartiles, and the *tails* indicate the range) showing difference in exact grade in ordinal score (grader – ground truth) for nasal and temporal angle directions.

Table. Agreement Between Graders and the Ground Truth Using Kappa Statistics When Assessing Angle Structures Seen on Goniophotography or AS-OCT. Conditions With Exact Angle Matches and Binary Grading (Open Versus Closed) Are Shown Separately

	Goniophotography		AS-OCT		Both Modalities	
	Exact Match	Open vs. Closed	Exact Match	Open vs. Closed	Exact Match	Open vs. Closed
Grader 1 vs. ground truth	0.387	0.347	0.399	0.413	0.457	0.476
Grader 2 vs. ground truth	0.432	0.358	0.428	0.475	0.413	0.373
Grader 3 vs. ground truth	0.451	0.431	0.520	0.520	0.523	0.451
Grader 1 vs. grader 2	0.764	0.748	0.621	0.660	0.578	0.502
Grader 1 vs. grader 3	0.415	0.451	0.491	0.573	0.461	0.449
Grader 2 vs. grader 3	0.511	0.565	0.628	0.717	0.523	0.550

Discussion

Aside from the necessary step of appropriately placing health care workers to deliver expert care within their respective domains, effective virtual clinic and collaborative care strategies also require the dissemination of clinical information in a modality conducive for reliable interpretation.[30] In the present study, we addressed the question of the application of goniophotography and AS-OCT images to facilitate a telemedicine approach for angle closure spectrum disease. In isolation, there are concerns regarding the accuracy of grading goniophotography and AS-OCT images; a multimodal approach tended to slightly but not significantly improve the agreement. Agreement levels were also slightly lower overall in comparison to the work of Murakami et al.,[31] who used the EyeCam instrument for capturing goniophotographs. The difference could be ascribed to a number of reasons, such as image capture technique (slit lamp versus supine position), gaze position (primary versus lens tilt), and the distribution of examined angles (where Murakami et al.[31] had a cohort with a greater number of closed angles). Our study provides an additional contribution to this discussion by demonstrating that the use of an ordinal scale revealed a conservative grading behavior across graders, which we define as the tendency to grade an angle as narrower than the ground truth, so as to not "miss" a narrow angle.[31] Using these data, we were also able to elicit factors that may pose barriers or may be facilitative of these techniques eventuating in virtual clinic strategies, and allow us to propose methods for improvement.

In comparison to previous studies that commonly report on binary decision making (i.e., angle open or closed),[31–33] the use of an ordinal scale allowed us to determine if systematic biases in angle grading were present, and the magnitude of this discordance. Our results demonstrated a systematic underestimation of angle grade compared with the ground truth, which was also reflected in the grader bias contributing to the overall low to moderate kappa values. Interestingly, this was also converse to the results of Murakami et al.,[31] who appeared to show that the EyeCam grader tended to identify fewer cases of closure compared with physical gonioscopy. The systematic bias was reduced when a multimodal approach was used, as seen by the higher kappa values, indicating less guessing behavior.

The uncertainty of grading was reflected in the overall low kappa values for exact matches and intraobserver agreement when comparing goniophotography and AS-OCT, and also appeared to vary by grader and by quadrant. Aside from criteria differences between individual graders, a recent study has also demonstrated regional variation in terms of agreement, particularly for the superior quadrant.[34] This was suggested to be related to the anatomical variation attributable to different techniques. Importantly, this raises concerns about the selection of the technique used for the ground truth of anterior chamber angle evaluation. The role of quantitative information that could arguably be more objective obtained using imaging techniques has been explored in the literature, with demonstrably high rates of accuracy for detection of angle closure.[35,36] Although these techniques offer an attractive option for mitigating the subjectivity inherent in gonioscopic evaluation and photography interpretation, current grading methods for determining cut-offs for treatable angle closure spectrum disease are still based upon physical examination by gonioscopy. In the absence of established normative distributions and consensus around cut-offs for significant progression, developing strategies for optimizing the gonioscopic evaluation remains relevant.

In the context of a telemedicine or virtual clinic modality, conservative criteria and underestimation (poor specificity) are likely more favorable compared with one in which angle width is overestimated (poor sensitivity). The conservative diagnosis may be related to the wary attitude of the three grading clinicians toward angle closure disease. This may appear to be at odds with previous reports of underdiagnosis of angle closure disease by optometrists in the literature[37,38]; however, given our present pre-screened, referred cohort, it is expected that the graders would examine these cases more carefully. The uncertainty of a two-dimensional static image, and similarly, single line scans using OCT, as the only source of information regarding the anterior chamber angle may have also driven more conservative behavior. Simultaneously, this may also be reflective of a conservative attitude toward gonioscopy as a technique. Gonioscopy is known to require significant clinical skill, and has been shown to pose a challenge to optometrists and ophthalmologists alike, depending on the health care setting.[39–41] Further specialized training and a feedback system may provide further benefits to improve grading accuracy and concordance.[40]

Further to this, we note that the focus of this paper has been in assessment of angle structures, primarily for the differentiation of angle closure spectrum disease. Telemedicine and virtual clinic approaches for the anterior chamber angle are also relevant for examination of secondary risk factors for glaucoma such as pigment dispersion, pseudoexfoliation, and angle recession. Methods specifically targeting identification and quantification of structures such as trabecular pigmentation and configuration would be useful under those situations.

We sought to identify factors contributing to agreement that could potentially be addressed to improve concordance. Broadly, our results suggest that the extremes of the angle appearance appeared to play the most consistent role in determining agreement, whereby patients with either completely closed angles or open angles were most likely correctly evaluated. This appears consistent with the high matches with binary judgements, and with the apparent irrelevance of distinguishing between functionally identical structures, such as the scleral spur and ciliary body band. When using gonioscopy in isolation, identifying the trabecular meshwork appeared to be a consistent feature necessary for accurate grading. The inability to visualize the posterior trabecular meshwork could therefore lead to more conservative grading behavior, with less confidence in visualizing deeper angle structures. Thus, techniques for identifying the location of the trabecular meshwork, such as the presence of pigmentation, blood reflux in Schlemm's canal, or hyporeflectivity of the canal on OCT, may be beneficial for accurate grading, similar to previous reports.[42] Again this highlights the limitation of solely using static goniophotography images in assessing the angle structures, and the potential advantages of a dynamic physical examination.

Interestingly, the ability to visualize certain angle structures did not emerge as a consistently significant factor driving agreement when using AS-OCT alone. Previous studies have suggested that difficulties in image acquisition could lead to the inability to visualize certain structures, which affect quantification of angle parameters.[43,44] The present results suggest that at least qualitative evaluation or the prediction of expected angle structure is driven instead by the iris contour, rather than visibility of the structures themselves. Image quality, also important in the gonioscopic examination,[45] also did not emerge as a significant, consistent factor for agreement. This was likely because there was a sufficient amount of visible angle to arrive at a decision in the majority of cases, and highlights the benefits of a broad photograph, rather than a restrictively narrow corneal wedge. Although this series of patient images represented a consecutive cohort, there could be an element of selection bias reducing the overall number of poor quality images, which may practically play a much larger role in routine clinical practice.

Unlike some other grading systems,[46–48] we did not employ the corneal wedge technique, as the use of a beam parallel to the examined angle provided a broader view of the angle for eliciting the deepest visible structure. The corneal wedge may provide an additional dimension of data, that is, the quantification of the angle width. There may be some subtle differences in the grading ability between these two gonioscopic views, and this may benefit from further study. Similarly, we only employed a limited number of OCT scans. More scans are likely beneficial to anterior chamber angle evaluation,[49] though this should also be weighed against the volume of data presented in remote clinical settings.

Based on these results, features specific to the imaging modality emerged as significant factors for concordance: trabecular pigmentation and iris contour for goniophotography and OCT, respectively. These occurred less commonly in the multimodal approach. This suggests that graders may rely more on features specific to a modality if their grading task is limited to that particular image alone. The complementary, but not overriding, relationship between goniophotography and AS-OCT may be related to differences in technique: contact versus noncontact, minimal versus almost no light. These differences have been suggested to account for differences in detection rate of angle closure between techniques.[50] To further temper these findings, all three iris configurations emerged as important factors in determining concordance, but it should be noted that the extremes and nonregular contour (i.e., flat or steep) were found to be much more significant. A practical interpretation of this result is therefore not the iris configuration as a whole, but rather the need to assess if the iris approach deviates from a regular insertion or contour. Overall, the underlying ground truth angle remained the most consistent, significant factor across graders and conditions, and the combination of results reinforces the challenge of borderline cases where view of the trabecular meshwork may be equivocal. This may also be especially relevant in patients with varying degrees of trabecular pigmentation.[51]

The present study utilized three highly trained optometrists who work within a specialized glaucoma clinic as graders, as the purpose was to evaluate concordance in an existing clinical model. Such graders have been demonstrated in the literature to be at least on part with a junior ophthalmologist in terms of glaucoma diagnosis.[52,53] The current gold standard for glaucoma assessment is by a glaucoma specialist ophthalmologist. In the future, this work can be extended to include this group of graders. It is important to also note that there is even a fair amount of variability even amongst trained ophthalmologists, and thus, the relatively low agreement found within this study should be tempered and considered relative to real-world practice.[54]

Our consecutive patient subset also consisted of only a small number of cases with an eventual diagnosis of narrow and occludable angles or worse requiring intervention. This is reflective of the overall low prevalence of glaucoma and angle closure disease within an optometric population in Australia, but as a result has potentially led to undersampling of the narrow angle group. Our factor analysis suggests that agreement statistics would improve should the sample within that subset increase, given that the extremes of angle appearance were associated with greater correct classification. Notably, the imaging methods in the present study reflected a current clinical protocol that was different to a large-scale evaluation of another form of anterior chamber angle photography, as discussed above.[31] Furthermore, a clinician performing a physical examination has access to an unimpeded physical view and the opportunity to focus upon specific areas of interest with various manipulations of technique. This would not be replicable using static images, and may have resulted in an additional factor for discordance. Image resolution and the fact that images are taken out of one eyepiece (without a stereoscopic view) are also potential factors reducing agreement.

Finally, we only had one clinician performing the physical examination with review of the results performed by an ophthalmologist, from which the ground truth had been derived.[38] Though having more physical examiners could enhance confidence in the fidelity of the ground truth, having one examiner reflects the process of a normal virtual clinic, as having multiple examiners would render a remote clinical arrangement redundant or less cost-effective.

Our graders exhibited generally moderate agreement levels and overall conservative behavior in remote viewing of anterior segment images. Based on our data, the recommendation appears to be to grade angles as open or closed/narrow, rather than by specific structure under conditions of remote viewing to optimize clinical relevance and agreement. More extreme angle structures were related to higher concordance, and thus this presents a challenge for identifying and suitably triaging borderline cases. Factors specific to an imaging modality appear to be less relevant when utilizing a multimodal approach, which in itself only slightly increases concordance in some cases, as graders tend to use each technique differently: these differences should be reconciled in specific targeted approaches to training to promote consistency. In combination with recent findings of the overall low prevalence of conversion to glaucoma within the angle closure spectrum disease family,[24] there appears to be a number of barriers to utilizing goniophotography and anterior segment imaging techniques for the purposes of remote evaluation and screening of the risk of angle closure spectrum disease, which need to be addressed in practical settings prior to widespread use.

Acknowledgments

The authors thank the staff at Centre for Eye Health for their assistance. The authors acknowledge the funding provided to the Centre for Eye Health by Guide Dogs NSW/ACT for the delivery of clinical services and salary support for JP, HW, SZ, and MK. VK is supported by an Australian Government Research Training Program PhD scholarship and a Guide Dogs NSW/ACT PhD scholarship. The funding bodies had no role in the conception or conduct of the study.

Disclosure: **J. Phu**, None; **H. Wang**, None; **V. Khou**, None; **S. Zhang**, None; **M. Kalloniatis**, None

References

1. Boodhna T, Crabb DP. Disease severity in newly diagnosed glaucoma patients with visual field loss: trends from more than a decade of data. *Ophthalmic Physiol Opt*. 2015;35:225–230.
2. Boodhna T, Saunders LJ, Crabb DP. Are rates of vision loss in patients in English glaucoma clinics slowing down over time? Trends from a decade of data. *Eye (Lond)*. 2015;29:1613–1619.
3. Quigley HA, Broman AT. The number of people with glaucoma worldwide in 2010 and 2020. *Br J Ophthalmol*. 2006;90:262–267.
4. Tham YC, Li X, Wong TY, Quigley HA, Aung T, Cheng CY. Global prevalence of glaucoma and projections of glaucoma burden through 2040: a systematic review and meta-analysis. *Ophthalmology*. 2014;121:2081–2090.
5. Chua J, Baskaran M, Ong PG, et al. Prevalence, risk factors, and visual features of undiagnosed glaucoma: the Singapore Epidemiology of Eye Diseases Study. *JAMA Ophthalmol*. 2015;133: 938–946.
6. Mitchell P, Smith W, Attebo K, Healey PR. Prevalence of open-angle glaucoma in Australia. The Blue Mountains Eye Study. *Ophthalmology*. 1996;103:1661–1669.
7. Shickle D, Todkill D, Chisholm C, et al. Addressing inequalities in eye health with subsidies and increased fees for General Ophthalmic Services in socio-economically deprived communities: a sensitivity analysis. *Public Health*. 2015; 129:131–137.
8. Baker H, Ratnarajan G, Harper RA, Edgar DF, Lawrenson JG. Effectiveness of UK optometric enhanced eye care services: a realist review of the literature. *Ophthalmic Physiol Opt*. 2016;36:545–557.
9. Bourne RR, French KA, Chang L, Borman AD, Hingorani M, Newsom WD. Can a community optometrist-based referral refinement scheme reduce false-positive glaucoma hospital referrals without compromising quality of care? The community and hospital allied network glaucoma evaluation scheme (CHANGES). *Eye (Lond)*. 2010;24:881–887.
10. Huang J, Hennessy MP, Kalloniatis M, Zangerl B. Implementing collaborative care for glaucoma patients and suspects in Australia. *Clin Exp Ophthalmol*. 2018;46:826–828.
11. Jamous KF, Kalloniatis M, Hennessy MP, Agar A, Hayen A, Zangerl B. Clinical model assisting with the collaborative care of glaucoma patients and suspects. *Clin Exp Ophthalmol*. 2015;43:308–319.
12. Keenan J, Shahid H, Bourne RR, White AJ, Martin KR. Cambridge community Optometry Glaucoma Scheme. *Clin Exp Ophthalmol*. 2015; 43:221–227.
13. Verma S, Arora S, Kassam F, Edwards MC, Damji KF. Northern Alberta remote teleglaucoma program: clinical outcomes and patient disposition. *Can J Ophthalmol*. 2014;49:135–140.
14. Lin AP, Katz LJ, Spaeth GL, et al. Agreement of visual field interpretation among glaucoma specialists and comprehensive ophthalmologists: comparison of time and methods. *Br J Ophthalmol*. 2011;95:828–831.
15. Tanna AP, Bandi JR, Budenz DL, et al. Interobserver agreement and intraobserver reproducibility of the subjective determination of glaucomatous visual field progression. *Ophthalmology*. 2011;118:60–65.
16. 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;239:342–350.
17. Kong YX, Coote MA, O'Neill EC, et al. Glaucomatous optic neuropathy evaluation project: a standardized internet system for assessing skills in optic disc examination. *Clin Exp Ophthalmol*. 2011;39:308–317.
18. Spalding JM, Litwak AB, Shufelt CL. Optic nerve evaluation among optometrists. *Optom Vis Sci*. 2000;77:446–452.

19. Strouthidis NG, Chandrasekharan G, Diamond JP, Murdoch IE. Teleglaucoma: ready to go? *Br J Ophthalmol*. 2014;98:1605–1611.
20. Prum BE Jr, Herndon LW Jr, Moroi SE, et al. Primary Angle Closure Preferred Practice Pattern((R)) Guidelines. *Ophthalmology*. 2016;123: P1–P40.
21. Prum BE Jr, Rosenberg LF, Gedde SJ, et al. Primary Open-Angle Glaucoma Preferred Practice Pattern((R)) Guidelines. *Ophthalmology*. 2016;123:P41–P111.
22. Hu CX, Mantravadi A, Zangalli C, et al. Comparing gonioscopy with visante and cirrus optical coherence tomography for anterior chamber angle assessment in glaucoma patients. *J Glaucoma*. 2016;25:177–183.
23. Campbell P, Redmond T, Agarwal R, Marshall LR, Evans BJ. Repeatability and comparison of clinical techniques for anterior chamber angle assessment. *Ophthalmic Physiol Opt*. 2015;35: 170–178.
24. He M, Jiang Y, Huang S, et al. Laser peripheral iridotomy for the prevention of angle closure: a single-centre, randomised controlled trial. *Lancet*. 2019;393:1609–1618.
25. Phu J, Hennessy MP, Spargo M, Dance S, Kalloniatis M. A collaborative care pathway for patients with suspected angle closure glaucoma spectrum disease [published online ahead of print May 23, 2019]. *Clin Exp Optom*. https://doi.org/10.1111/cxo.12923.
26. Rao HL, Kumbar T, Addepalli UK, et al. Effect of spectrum bias on the diagnostic accuracy of spectral-domain optical coherence tomography in glaucoma. *Invest Ophthalmol Vis Sci*. 2012;53: 1058–1065.
27. Asian Pacific Glaucoma Society (APGS). Asia Pacific Glaucoma Guidelines. Amsterdam, The Netherlands: Kugler Publications; 2016 [updated 2016]. Available at: http://www.icoph.org/dynamic/attachments/resources/asia_pacific_glaucoma_guidelines_2016_third_edition.pdf. Accessed October 9, 2018.
28. International Council of Ophthalmology. International Council of Ophthalmology Guidelines for Glaucoma Eye Care. 2016 [updated 2016]. Available at: http://www.icoph.org/downloads/ICOGlaucomaGuidelines.pdf. Accessed October 9, 2018.
29. Phu J, Wang H, Khuu SK, et al. Anterior chamber angle evaluation using gonioscopy: consistency and agreement between optometrists and ophthalmologists. *Optom Vis Sci*. 2019;96. Accepted for Publication. https://doi.org/10.1097/OPX.0000000000001432
30. Rathi S, Tsui E, Mehta N, Zahid S, Schuman JS. The current state of teleophthalmology in the United States. *Ophthalmology*. 2017;124:1729–1734.
31. Murakami Y, Wang D, Burkemper B, Lin SC, Varma R. A population-based assessment of the agreement between grading of goniophotographic images and gonioscopy in the Chinese-American Eye Study (CHES). *Invest Ophthalmol Vis Sci*. 2016;57:4512–4516.
32. Barkana Y, Dorairaj SK, Gerber Y, Liebmann JM, Ritch R. Agreement between gonioscopy and ultrasound biomicroscopy in detecting iridotrabecular apposition. *Arch Ophthalmol*. 2007;125: 1331–1335.
33. Narayanaswamy A, Sakata LM, He MG, et al. Diagnostic performance of anterior chamber angle measurements for detecting eyes with narrow angles: an anterior segment OCT study. *Arch Ophthalmol*. 2010;128:1321–1327.
34. Xu BY, Pardeshi AA, Burkemper B, et al. Differences in anterior chamber angle assessments between gonioscopy, EyeCam, and anterior segment OCT: the Chinese American Eye Study. *Transl Vis Sci Technol*. 2019;8:5.
35. Nongpiur ME, Haaland BA, Perera SA, et al. Development of a score and probability estimate for detecting angle closure based on anterior segment optical coherence tomography. *Am J Ophthalmol*. 2014;157:32–38.e1.
36. Nongpiur ME, Haaland BA, Friedman DS, et al. Classification algorithms based on anterior segment optical coherence tomography measurements for detection of angle closure. *Ophthalmology*. 2013;120:48–54.
37. Jindal A, Myint J, Edgar DF, Nolan WP, Lawrenson JG. Agreement among optometrists and ophthalmologists in estimating limbal anterior chamber depth using the van Herick method. *Ophthalmic Physiol Opt*. 2015;35:179–185.
38. Khan S, Clarke J, Kotecha A. Comparison of optometrist glaucoma referrals against published guidelines. *Ophthalmic Physiol Opt*. 2012;32:472–477.
39. Liu L. Australia and New Zealand survey of glaucoma practice patterns. *Clin Exp Ophthalmol*. 2008;36:19–25.
40. Zangerl B, Hayen A, Mitchell P, Jamous KF, Stapleton F, Kalloniatis M. Therapeutic endorsement enhances compliance with national glaucoma guidelines in Australian and New Zealand

optometrists. *Ophthalmic Physiol Opt*. 2015;35: 212–224.
41. Zebardast N, Solus JF, Quigley HA, Srikumaran D, Ramulu PY. Comparison of resident and glaucoma faculty practice patterns in the care of open-angle glaucoma. *BMC Ophthalmol*. 2015;15: 41.
42. Sakata LM, Lavanya R, Friedman DS, et al. Assessment of the scleral spur in anterior segment optical coherence tomography images. *Arch Ophthalmol*. 2008;126:181–185.
43. Seager FE, Wang J, Arora KS, Quigley HA. The effect of scleral spur identification methods on structural measurements by anterior segment optical coherence tomography. *J Glaucoma*. 2014;23:e29–e38.
44. Liu S, Li H, Dorairaj S, et al. Assessment of scleral spur visibility with anterior segment optical coherence tomography. *J Glaucoma*. 2010;19:132–135.
45. Lee B, Szirth BC, Fechtner RD, Khouri AS. Are disposable and standard gonioscopy lenses comparable? *J Glaucoma*. 2017;26:e157–e159.
46. Scheie HG. Width and pigmentation of the angle of the anterior chamber; a system of grading by gonioscopy. *AMA Arch Ophthalmol*. 1957;58: 510–512.
47. Shaffer RN. Primary glaucomas. Gonioscopy, ophthalmoscopy and perimetry. *Trans Am Acad Ophthalmol Otolaryngol*. 1960;64:112–127.
48. Spaeth GL. The normal development of the human anterior chamber angle: a new system of descriptive grading. *Trans Ophthalmol Soc U K*. 1971;91:709–739.
49. Blieden LS, Chuang AZ, Baker LA, et al. Optimal number of angle images for calculating anterior angle volume and iris volume measurements. *Invest Ophthalmol Vis Sci*. 2015;56:2842–2847.
50. Nolan WP, See JL, Chew PT, et al. Detection of primary angle closure using anterior segment optical coherence tomography in Asian eyes. *Ophthalmology*. 2007;114:33–39.
51. Kinori M, Hostovsky A, Skaat A, Schwartsman J, Melamed S. A novel method for quantifying the amount of trabecular meshwork pigment in glaucomatous and nonglaucomatous eyes. *J Glaucoma*. 2014;23:e13–e17.
52. Azuara-Blanco A, Burr J, Thomas R, Maclennan G, McPherson S. The accuracy of accredited glaucoma optometrists in the diagnosis and treatment recommendation for glaucoma. *Br J Ophthalmol*. 2007;91:1639–1643.
53. Marks JR, Harding AK, Harper RA, et al. Agreement between specially trained and accredited optometrists and glaucoma specialist consultant ophthalmologists in their management of glaucoma patients. *Eye (Lond)*. 2012;26:853–861.
54. Rigi M, Bell NP, Lee DA, et al. Agreement between gonioscopic examination and swept source fourier domain anterior segment optical coherence tomography imaging. *J Ophthalmol*. 2016;2016:1727039.