Purpose. To validate a new computer-aided diagnosis (CAD) imaging program for the assessment of nuclear lens opacity.

Methods. Slit-lamp lens photographs from the Singapore Malay Eye Study (SiMES) were graded using both the CAD imaging program and manual assessment method by a trained grader using the Wisconsin Cataract Grading System. Cataract was separately assessed clinically during the study using Lens Opacities Classification System III (LOCS III). The repeatability of CAD and Wisconsin grading methods were assessed using 160 paired images. The agreement between the CAD and Wisconsin grading methods, and the correlations of CAD with Wisconsin and LOCS III were assessed using the SiMES sample (5547 eyes from 2951 subjects).

Results. In assessing the repeatability, the coefficient of variation (CoV) was 8.10% (95% confidence interval [CI], 7.21–8.99), and the intraclass correlation coefficient (ICC) was 0.96 (95% CI, 0.93–0.96) for the CAD method. There was high agreement between the CAD and Wisconsin methods, with a mean difference (CAD minus Wisconsin) of −0.02 (95% limit of agreement, −0.91 and 0.87) and an ICC of 0.81 (95% CI, 0.80–0.82). CAD parameters were also significantly correlated with LOCS III grading (all P < 0.001).

Conclusions. This new CAD imaging program assesses nuclear lens opacity with results comparable to the manual grading using the Wisconsin System. This study shows that an automated, precise, and quantitative assessment of nuclear cataract is possible. (Invest Ophthalmol Vis Sci. 2011;52:1314–1319) DOI:10.1167/iovs.10-5427

Validity of a New Computer-Aided Diagnosis Imaging Program to Quantify Nuclear Cataract from Slit-lamp Photographs

Carol Yim-lui Cheung,1 Huiqi Li,2 Ecosse L. Lamoureux,1,3 Paul Mitchell,4 Jie Jin Wang,3,4 Ava Grace Tan,4 Lily K. Jobari,1 Jiang Liu,2 Joo Hwee Lim,2 Tin Aung,1,5 and Tien Yin Wong1,5

Over the last three decades, epidemiologic studies have provided substantial data on the prevalence, incidence, progression, and risk factors for age-related cataract.1–11 In these studies cataract was graded either by ophthalmologists at the slit-lamp or by technicians (graders) who assessed the presence and severity of cataract from lens photographs by comparing to standard photographs. It has been acknowledged, however, that such assessments are inefficient, time-consuming, and imprecise.12–14 These drawbacks, particularly the need for a relatively long training period for the graders and the often substantial, inter-observer variation, limit the comparabilities of grading results across different studies, and have significantly limited further clinical research on cataract, particularly in genetic studies15–18 and in the evaluation of preventive therapies,19–22 in which accurate phenotyping of lens opacity is essential.

As an alternative to manual grading procedures, an objective, precise, and reproducible automatic method for assessment of lens opacities is appealing. We developed a new computer-aided diagnosis (CAD) imaging program for the assessment of nuclear lens opacity23–25 based on digital slit-lamp photographs, and aligned this program to the Wisconsin Cataract Grading System.1 In this article we describe the repeatability of this newly developed CAD nuclear cataract grading program on different images taken from the same eyes, and examine the agreement in grading between the CAD and the manual Wisconsin System in a large population-based sample.

Methods

Study Population

We used data from the Singapore Malay Eye Study (SiMES), a population-based cross-sectional survey of eye diseases in Malay adults aged 40 to 80 years to evaluate the validity of the CAD nuclear cataract imaging program. The methodology and objectives of the study population have been reported in detail elsewhere.20 In brief, subjects residing in 15 districts in Southwestern Singapore were selected using an age-stratified (by 10-year age group) random sampling method, from a computer-generated list provided by the Singapore Ministry of Home Affairs. Of 4168 eligible participants, 3280 (78.7%) participated in the study conducted from August 2004 through June 2006. Participants underwent a standardized interview, systemic and ocular examination, and laboratory investigations. Written informed consent was obtained from each participant and the study was conducted in adherence with the Declaration of Helsinki. Ethics approval was obtained from the Singapore Eye Research Institute Institutional Review Board.

Wisconsin Cataract Grading System

Slit-lamp lens photography was performed with a digital slit-lamp camera (Topcon model DC-1, with FD-21 flash attachment; Topcon, Investigative Ophthalmology & Visual Science. March 2011, Vol. 52, No. 3 Copyright 2011 The Association for Research in Vision and Ophthalmology, Inc.
The approach used has been previously reported. In brief, the contours of the lens and the nucleus are first detected using a modified active shape model (ASM) from the slit-lamp image. Twenty-one images of the lens and nucleus are automatically displayed. A single observer (HL) reviewed all images and classified them as failing if the segmentation was classified as failing and manual intervention was then performed to improve the lens and nuclear margin recognition and location.

**CAD Nuclear Cataract Imaging Program**

For consistency with the Blue Mountains Eye Study (BMES), cataract was assessed by the same grader (AGT) using the Wisconsin Cataract Grading System in SiMES, thus closely aligning the CAD nuclear cataract imaging program to the Wisconsin protocol. Details of the CAD approach used has been previously reported. In brief, the contours of the lens and the nucleus are first detected using a modified active shape model (ASM) from the slit-lamp image. Twenty-one imaging features (e.g., intensity of nucleus, color of nucleus, intensity of sucus) are then extracted and normalized to identify the intensity of nuclear opacity. For the latter, support vector machine (SVM) regression was used for nuclear opacity grading prediction by using a training set of 100 slit-lamp photographs taken from 100 randomly selected eyes. These were derived from 68 SiMES participants who had Wisconsin decimal grades for their lens nuclear opacity that ranged from 0.4 to 5.0.

Figure 1 shows the user interface of the CAD program. After loading the lens image into the program, both the predicted grade and detection of lens structure are automatically displayed. A batch processing mode is also available in the system for processing large numbers of images automatically. A single observer (HL) reviewed all lens detections after the automated grading processing by the CAD program on all SiMES images. When the overlap between the automated detection and the contour of real lens structure is estimated < 95%, the segmentation is classified as failing and manual intervention is then performed to improve the lens and nuclear margin recognition and location.

**Repeatability of the CAD and Wisconsin Grades**

Two consecutive slit-lamp lens images were captured from a subset of 160 eyes during the study. Each subject was instructed to sit back between the two shots while the slit-lamp beam was realigned. These image pairs were derived from 157 SiMES participants with Wisconsin grade ranging from 0.7 to 5.0, and used to assess the repeatability of nuclear opacity measurement on different images of the same eyes by the CAD and Wisconsin grading methods.

**Comparison with Clinical Cataract Assessment**

Cataract was separately defined by trained ophthalmologists during the study using Lens Opacities Classification System III (LOCS III) with a slit-lamp microscope (Haag-Streit model BQ-900; Haag-Streit, Koeniz, Switzerland) comparing to standard photographic slides for nuclear opalescence (NO), nuclear color (NC), cortical and posterior subcapsular cataract. In this study, a clinical diagnosis of nuclear cataract was defined as LOCS III NO scores of 4 or higher, or LOCS III NC scores of 4 or higher.

**Statistical Analysis**

All statistical analyses were performed using commercially available statistical software (SPSS, version 17.0; SPSS Statistics, Chicago, IL). Both eyes (if available) from each participant were selected for this study, after excluding pseudophakic and aphakic eyes. The grading score was analyzed on a continuous scale. We performed four sets of analyses. First, to estimate the repeatability of nuclear opacity measurements of the two assessment methods, we calculated the coefficient of variation (CoV) and intraclass correlation coefficient (ICC) for repeated slit-lamp images. Second, to assess agreement in the nuclear opacity scores between the CAD and Wisconsin grading methods, we constructed Bland-Altman plots and calculated the ICC between the two methods. Third, to determine correlations with the clinical cataract assessment (LOCS III) for CAD and the Wisconsin grades, Kendall’s tau coefficient was calculated and the area under receiver operating characteristic curves (AUCs) were used to assess the ability of the CAD and Wisconsin grading scores to differentiate nuclear cataract, as defined by clinical LOCS III NO and NC scores. The highest average of sensitivity and specificity were also calculated. Finally, to
demonstrate correlation between CAD and clinical cataract grading. Pearson correlation coefficient was calculated for selected CAD imaging features (mean intensity of nucleus, mean hue in HSV [hue, saturation, and value] color space of nucleus, mean saturation in HSV color space of nucleus and intensity of sulcus) and LOCS III NO and NC scores.

RESULTS

Repeatability of CAD and Wisconsin Grading Methods

Table 1 shows intra-session repeatability estimates for the CAD and Wisconsin nuclear opacity grading methods on two consecutive slit-lamp lens photographs from the subset of 160 eyes. The CoV was 8.10% for CAD and 9.23% for Wisconsin; and the ICC was 0.96 for the CAD and 0.90 for the Wisconsin grading methods, respectively.

Agreement between CAD and Wisconsin Grading

We initially analyzed all slit-lamp lens photographs of 5850 eyes from 3038 participants (92.6% of 3280 participants). Apart from the 100 images for the training set, 33/5750 (0.6%) images failed the CAD grading because of poor quality. CAD grading was automatically obtained for 5426/5750 (94.4%) images, and 291/5750 (5.0%) images needed manual intervention in the CAD grading.

We excluded images from the training set (n = 100), those that failed the CAD grading quality assessment (n = 33), those that were marked as ungradable in the Wisconsin grading (n = 174, in which of 13 also failed in CAD grading), and those with missing LOCS III information (n = 9). This left 5547 eyes from 2951 participants for the final analysis (90.0% of 3280 participants). The mean (standard deviation [SD]) age of this sample was 57.9 (10.8) years, including 1420 (48.1%) men and 1531 (51.9%) women.

Figure 2 shows the distributions of nuclear opacity grading scores using the CAD and Wisconsin grading methods in the SiMES population. The score range was 0.30 to 5.00 and interquartile range (25th to 75th percentile) was 1.80 to 2.80 by the CAD method. The range was 0.30 to 5.00 and interquartile range was 1.80 to 2.70 by the Wisconsin method. The mean (SD) and median scores were 2.34 (0.77) and 2.20 by the CAD method, and 2.36 (0.72) and 2.30 by the Wisconsin method. The grading score agreement between the CAD and Wisconsin methods is shown in Figure 3. The mean difference was −0.02 with 95% limits of agreement between −0.91 and 0.87. Table 2 further demonstrates the agreement in cataract severity categories between the CAD and Wisconsin grading methods, with a weighted kappa value of 0.925 (SE, 0.009). For the eyes with severe cataract, the CAD grading method tended to classify more eyes at a higher grade than the Wisconsin grading method, with 120/920 (13%) eyes classified as category 4 by the Wisconsin and classified as category 5 by the CAD.

Comparison with Clinical Cataract Assessment using the LOCS III System

The correlation with clinical LOCS III NO score was relatively better for CAD grading than that for Wisconsin grading (Kendall’s tau coefficient, 0.508 vs. 0.470, all P < 0.001). Table 3 demonstrates the ability of CAD and Wisconsin grading scores to differentiate clinical nuclear cataract, as defined by the clinical LOCS III grading. The AUC was significantly larger (P < 0.001) and sensitivity value was higher for CAD grading than that for Wisconsin grading.

Table 4 further shows that quantitative CAD imaging features (mean intensity of nucleus, mean hue in HSV color space of nucleus, mean saturation in HSV color space of nucleus, and intensity of sulcus) were significantly correlated with clinical LOCS III NO and NC scores (all P < 0.001).

DISCUSSION

Cataract is the leading cause of blindness and visual impairment worldwide, and the burden of cataract surgery is a...
major public health issue. An objective, reproducible, and efficient method to measure lens opacities is needed in cataract research to provide time effective grading, that will improve our ability to study the causes and risk factors for cataract, including genetic risk factors, and to evaluate new preventive and therapeutic measures. In this study, we tested a newly developed CAD imaging program aligned to the Wisconsin Cataract Grading System to grade the presence and severity of nuclear cataract from digital slit-lamp photographs. We showed that this imaging program is highly repeatable, has good agreement with the manual Wisconsin grading system, is sensitive to differentiate clinical cataract severity, and is correlated with clinical cataract assessment.

Our results demonstrated that the CAD grading score was highly repeatable in two consecutive slit-lamp lens photograph series from the same eyes, which was in fact even better than the manually-operated Wisconsin grading method, indicating that this CAD program is less sensitive to subtle variations in photography. Furthermore, we showed that the agreement of grading scores between the CAD program and Wisconsin system was very good when applied to the whole population, indicating that the CAD lens assessment is consistent with that of a highly trained photographic grader. We also found significant correlation between CAD imaging features and clinical nuclear cataract grading using the LOCS III system, suggesting that CAD assessment accurately reflects clinical impression of cataract severity.

There are few objective methods for the assessment of nuclear opacity. Some of these methods are limited by measuring lens density (pixel intensity) only. For example, Klein et al. developed a semiautomated computer grading of the black-and-white digitized images to initially assess progression of nuclear sclerosis. However, Klein et al. found that this program identified less progression than by human grading, possibly because color information was not assessed in the computer grading.

Other studies investigating the objective assessment of lens opacity have also used Scheimpflug imaging systems with computerized analysis of lens opacities. The main advantage of Scheimpflug imaging is that it can provide an evenly focused representation of a section of the lens and an objective measurement of lens density. Different study groups have shown that nucleus and overall lens density can be objectively measured using Scheimpflug imaging with image processing techniques. While these studies have shown that Scheimpflug imaging is highly correlated with LOCS III grading and can provide a reliable assessment of nuclear cataract, the Scheimpflug instrumentation is very expensive and the lens color is still not assessed by this instrument.

Although the Wisconsin cataract grading system is reproducible, a skilled grader requires at least a year’s training, and adjudication by senior ophthalmologist/researcher is commonly needed for junior graders. The total average grading time spent on one image is therefore approximately five minutes (training period and grading time of slit-lamp photographs based on the experience of the Blue Mountains Lens Eye Study; unpublished data, Wang JJ, Mitchell P, 2010). By contrast, the training to perform CAD is one day, and the grading time for each image is approximately 30 seconds, substantially less than the time needed for the manual grading method. The demand on human resources can also be further reduced by using overnight processing of batches of images. In addition to its efficiency, several other key advantages are offered by the CAD imaging method. First, the user interface of the CAD grading system is easy and simple. Second, we showed that 99.4% (5717 of 5750) of the slit-lamp images could be successfully graded by the current CAD program, which is higher than the proportion gradable by the manual Wisconsin system (97.0%, 5576 of 5750). Third, as a large majority of the lens images were graded automatically (94.9%, 5459 of 5750) without manual intervention, this CAD program avoids human bias and inter-

**Table 2. Agreement between Computer-Aided Diagnosis (CAD) and Wisconsin Nuclear Cataract Grading Systems in Cataract Severity Categories**

<table>
<thead>
<tr>
<th>Computer-Aided Diagnosis (CAD) Grading</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Wisconsin</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>80</td>
</tr>
<tr>
<td>2</td>
<td>51</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>131 (−2.40%)</td>
</tr>
</tbody>
</table>

Weighted kappa = 0.925 (0.009).

**Table 3. The Area under the Receiver Operating Characteristic Curves (AUCs) and the Highest Average of Sensitivity and Specificity of the Computer-Aided Diagnosis (CAD) and Wisconsin Nuclear Cataract Grading Systems versus Clinical Diagnosis of Cataract**

<table>
<thead>
<tr>
<th>Nuclear Cataract* Cases</th>
<th>Computer-Aided Diagnosis (CAD) Grading</th>
<th>Wisconsin Nuclear Cataract Grading</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AUC† (95% CI)</td>
<td>Sensitivity % (95% CI)</td>
</tr>
<tr>
<td>n = 1473 (26.6%)</td>
<td>0.892 (0.884–0.900)</td>
<td>81.8 (79.7–83.7)</td>
</tr>
</tbody>
</table>

*Clinical diagnosis of cataract was defined as Lens Opacities Classification System III (LOCS III) Nuclear Opalescence (NO) scores of 4 or higher, or LOCS III Nuclear Color (NC) scores of 4 or higher.
† AUC, area under the receiver operating characteristic curve.
‡ Pairwise comparison of AUCs between Computer-aided diagnosis grading and Wisconsin nuclear cataract grading.
TABLE 4. Correlations between Selected Imaging features of the Computer-aided Diagnosis (CAD) system and Clinical Grading of Cataract based on the Lens Opacities Classification System III (LOC III)

<table>
<thead>
<tr>
<th>Selected CAD Imaging Features</th>
<th>LOCS III Nuclear Opalescence</th>
<th>LOCS III Nuclear Color</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$r$</td>
<td>$P$</td>
</tr>
<tr>
<td>Mean intensity of nucleus</td>
<td>0.399</td>
<td>$&lt;0.001$</td>
</tr>
<tr>
<td>Mean hue in HSV color space of nucleus</td>
<td>-0.601</td>
<td>$&lt;0.001$</td>
</tr>
<tr>
<td>Mean saturation in HSV color space of nucleus</td>
<td>-0.375</td>
<td>$&lt;0.001$</td>
</tr>
<tr>
<td>Intensity of sulcus</td>
<td>0.443</td>
<td>$&lt;0.001$</td>
</tr>
</tbody>
</table>

HSV, hue, saturation, and value; $r$, Pearson correlation coefficient.

observer variability. Fourth, other lens and nuclear features (e.g., lens color) and quantitative data (e.g., intensity of the sulcus) extracted from the slit-lamp images\(^{23-25}\) can provide the opportunity to conduct comprehensive and quantitative analyses on cataract morphology. Finally, this new CAD imaging method may provide reliable assessments on longitudinal changes in the nuclear opacity for study samples with multiple lens images taken over time. Thus, this CAD method may be used not only for lens opacity assessment in large epidemiologic studies, but also potentially for mass screening for cataract, for clinical trials evaluating potential preventive approaches for age-related cataract, and for precise nuclear cataract phenotypic definition in genetic research.\(^{15-18}\)

This current version of the CAD nuclear cataract grading program, however, has limitations and there is room for further improvement. First, the program was developed in alignment with the Wisconsin Cataract Grading System, using a decimal grading scale from 0.1 to 5.0, which may not permit direct comparison with the results provided by other grading systems. Further development to incorporate other classification systems (e.g., LOCS III and the Age-Related Eye Disease Study [AREDS] system) with our current CAD method would provide potential benefit in clinical practice and enable comparison with other research studies. Second, although only 5% of images needed manual intervention in CAD grading, the variability of this manual intervention has not yet been assessed. Third, despite the fact that overall agreement between CAD and Wisconsin grading score was generally good, the agreement for severe cataract was only moderate. Further refinement of the CAD program for more severe cataract types is therefore needed. Finally, while we have standardized the slit-lamp camera setting during photography in the SIMES participants, the lens features and grading scores by the CAD program may vary in relation to different slit-lamp camera settings.

In summary, we demonstrate the repeatability and validity of a new CAD nuclear cataract imaging program against the current Wisconsin Cataract Grading System in a large population-based study. This CAD program appears to provide significant advances over current manual and subjective cataract grading protocols with an objective, efficient, and reliable assessment on the nuclear opacity. This new CAD imaging program will be of substantial benefit not only to population-based epidemiologic studies, but also to genetic studies and interventional trials that require information on precise changes in the lens opacity over time.