Telemedicine Diagnosis of Cytomegalovirus Retinitis by Nonophthalmologists

Michael Yen, BS; Somsanguan Ausayakhun, MD; Jenny Chen, MD; Sakarin Ausayakhun, MD; Choeng Jirawison, MD; David Heiden, MD; Gary N. Holland, MD; Todd P. Margolis, MD, PhD; Jeremy D. Keenan, MD, MPH

**IMPORTANCE** Cytomegalovirus (CMV) retinitis continues to be a leading cause of blindness in many developing countries. Telemedicine holds the potential to increase the number of people screened for CMV retinitis, but it is unclear whether nonophthalmologists could be responsible for interpreting fundus photographs captured in a telemedicine program.

**OBJECTIVE** To determine the accuracy of nonophthalmologist photographic graders in diagnosing CMV retinitis from digital fundus photographs.

**DESIGN, SETTING, AND PARTICIPANTS** Fifteen nonexpert graders each evaluated 182 mosaic retinal images taken from the eyes of patients with AIDS who were evaluated at the Ocular Infectious Diseases Clinic at Chiang Mai University in Chiang Mai, Thailand.

**MAIN OUTCOMES AND MEASURES** Graders diagnosed each image as CMV retinitis present, CMV retinitis absent, or unknown. The results from each grader were compared with those of an indirect ophthalmoscopic examination from an experienced on-site ophthalmologist as well as with the consensus grade given by a panel of CMV retinitis experts.

**RESULTS** Relative to the on-site ophthalmologist, the sensitivity of remote CMV retinitis diagnosis by nonexpert graders ranged from 64.0% to 95.5% (mean, 84.1%; 95% CI, 78.6%-89.6%), and the specificity ranged from 65.6% to 92.5% (mean, 82.3%; 95% CI, 76.6%-88.0%). Agreement between nonexpert and expert graders was high: the mean sensitivity and specificity values of nonexpert diagnosis using expert consensus as the reference standard were 93.2% (95% CI, 90.6%-95.8%) and 88.4% (95% CI, 85.4%-91.1%), respectively. Mean intrarater reliability also was high (mean Cohen κ, 0.83; 95% CI, 0.78-0.87).

**CONCLUSIONS AND RELEVANCE** The sensitivity and specificity of remote diagnosis of CMV retinitis by nonexpert graders was variable, although several nonexperts achieved a level of accuracy comparable to that of CMV retinitis experts. More intensive training and periodic evaluations would be required if nonexperts are to be used in clinical practice.
C

tomegalovirus (CMV) retinitis is a well-known complication of AIDS. Although the incidence of CMV retinitis has declined in the United States and Europe owing to widespread use of highly active antiretroviral therapy (HAART), it is still common in countries with a high burden of human immunodeficiency virus (HIV) and limited access to HAART. At a tertiary ophthalmologic center in Chiang Mai, Thailand, CMV retinitis was found in 33% of HIV-positive patients who received screening and was one of the leading causes of blindness among all patients seen at the clinic. In this resource-poor setting, patients often have severe disease when they first present to an ophthalmologist despite receiving HAART, which is suggestive of inadequate screening.9,10

There are significant challenges to screening the population at risk for CMV retinitis. Chief among these is a relative shortage of ophthalmologists. For example, in Thailand nearly half of all ophthalmologists practice in Bangkok, leaving the outlying provinces relatively understaffed. Uneven distribution of ophthalmologists is likely to be similar in other resource-poor settings. Improving coverage of CMV retinitis screening will require addressing the relative lack of ophthalmology care for patients with HIV.12

Telemedicine, which is used to screen for conditions such as diabetic retinopathy (DR) and retinopathy of prematurity (ROP), is a promising method for screening at-risk patients with AIDS for CMV retinitis. A prior study showed that telemedicine diagnosis by expert graders has high sensitivity and specificity for CMV retinitis. However, ophthalmologists may have difficulty giving prompt diagnosis because of time demands on their schedules. Trained nonexpert graders at dedicated reading centers may provide faster turnaround at a fraction of the cost of an ophthalmologist. Previous research found that nonexperts can be trained to remotely detect hypertensive retinopathy, DR, and ROP with high accuracy. The purpose of the present study was to evaluate how accurately individuals other than ophthalmologists were able to diagnose CMV retinitis using fundus photographs.

Methods

Ethics Statement
Approval for the study was obtained from the Committee on Human Research at the University of California, San Francisco, and the Faculty of Medicine Research Ethics Committee at Chiang Mai University. This study adhered to the tenets of the Declaration of Helsinki. All participants gave written informed consent. No financial compensation was provided.

Study Design
Fundus photography was performed in a series of patients with AIDS by a trained ophthalmic photographer, and indirect ophthalmoscopy was conducted by an experienced ophthalmologist. Digital fundus photographs were read by 3 CMV retinitis experts (D.H., G.N.H., and T.P.M.), as reported previously. In the present study, we performed a training session for 15 nonophthalmologists (including M.Y.) and subsequently asked them to evaluate the same set of fundus photographs reviewed by the CMV retinitis experts. We compared the diagnosis given by the nonophthalmologists with that of the in-clinic ophthalmologist and CMV retinitis experts to determine the accuracy of nonexpert diagnosis of CMV retinitis.

Patient Population
We used the same images from a prior study of patients being screened for CMV retinitis at the Ocular Infectious Diseases Clinic at Chiang Mai University. Between August 7, 2008, and April 9, 2009, we enrolled newly referred patients with AIDS who had evidence of CMV retinitis on indirect ophthalmoscopy, as well as patients who had received a CMV retinitis diagnosis in the month before the start of the study. From August 7, 2008, to January 7, 2009, we also enrolled patients with AIDS who had no evidence of CMV retinitis on indirect ophthalmoscopy. An experienced attending ophthalmologist performed indirect ophthalmoscopy on both eyes of each study participant. One-time fundus photography was performed on both eyes by a trained retinal photographer (TRC-NW6S digital fundus camera, Topcon). Using the camera’s preset internal fixation light to dictate the position of each picture, the photographer captured a standard set of 9 overlapping 45° fundus images. The photographer was masked to all clinical information, including whether the study participant had CMV retinitis and whether the patient was being evaluated for the first time or during a follow-up visit. Automated mosaics covering an 85° retinal field were created using software that was included with the camera. There were no adverse events associated with either the fundus photography or indirect ophthalmoscopy. Eyes were excluded from this study if the attending ophthalmologist was unable to determine with certainty the presence or absence of CMV retinitis.

Nonexpert Graders
Fifteen nonexpert graders from US and Thai medical centers were grouped into 3 categories based on prior clinical and research experience. The first group had no clinical or research experience and included 5 ophthalmology secretaries and 1 undergraduate university student. The second group consisted of individuals with research experience but no clinical experience and included 3 research assistants, 1 research coordinator, and 1 laboratory technician. Graders in the third group had clinical but no research experience and included 1 ophthalmic nurse and 3 medical students (including M.Y.). All graders attended one 2-hour training session that covered eye anatomy, retinal abnormalities, CMV retinitis, and reasons for ophthalmology referral. Trainings were conducted in English or Thai depending on the setting, and all participants viewed the same training materials. Graders were given a copy of the English-language training materials to use as reference.

Image Evaluation
Similar to the procedures in the prior study, graders examined the fundus photographs of each eye independently, in random order, without accompanying clinical information. Nonexpert graders assessed only the montage retinal images (ie, not the individual frames). Graders first evaluated the quality of each photograph, with defined as an image in fo-
Fig. 1. Sensitivity and Specificity of Remote Nonexpert Diagnosis of Cytomegalovirus (CMV) Retinitis on Mosaic Fundus Photographs

Estimates are shown for each of 15 nonexperts and 3 CMV retinitis experts as well as the mean of the nonexperts. Graders had clinical experience (clinical), research experience (research), or neither clinical nor research experience (neither).

Statistical Analysis
We calculated the sensitivity and specificity of CMV retinitis diagnosis for each nonexpert grader using the on-site ophthalmology examination as the reference standard. In practice, nonexpert graders would provide initial screening to determine which patients should be referred to an ophthalmologist for definitive diagnosis. Therefore, both CMV retinitis present and unknown counted as a positive diagnosis because both would warrant ophthalmology referral. To determine patterns of false-positive and false-negative photographic assessments, we identified eyes that were incorrectly categorized by most (≥8) of the nonexpert graders relative to the indirect ophthalmoscopic examination and compared the results with the consensus diagnosis of 3 expert graders from the prior study16 (ie, the majority diagnosis of the 3 experts). We calculated the positive predictive value for images graded as CMV retinitis present and unknown separately and the negative predictive value for images graded as CMV retinitis absent.

To compare the performance of nonexperts with that of experts, we calculated the sensitivity and specificity of each of the nonexperts using the expert consensus as the reference standard. Images were excluded from this analysis if there was no majority consensus or if the consensus was a diagnosis of unknown. We chose to express the agreement between expert and nonexpert graders as sensitivity and specificity because grades from these 2 groups of individuals could not reasonably be thought of as interchangeable data points—we would not expect nonexperts who received a brief training workshop to perform as well as expert ophthalmologists.

We used nonparametric tests (eg, Wilcoxon rank sum and Kruskal-Wallis) to perform comparisons between subgroups of graders. We used the Cohen κ coefficient to estimate (1) the intrarater reliability for each nonexpert grader from a set of 50 randomly selected and randomly presented duplicate images and (2) the interrater reliability among the 15 nonexpert graders. To assess the relationship between intrarater agreement and diagnostic accuracy, we calculated the Pearson correlation coefficient for intrarater agreement and a summary measure combining sensitivity and specificity (Youden index). To account for nonindependence of eyes from the same person, we calculated bootstrap 95% CIs for diagnostic test statistics with resampling at the patient level (9999 repetitions). All statistical analyses were performed with Stata, version 13 (StataCorp).

Results
Characteristics of the participants included in this study have been described previously.16 The on-site ophthalmologist was able to definitively determine the presence or absence of CMV retinitis in 182 distinct eyes of 94 patients; 167 of these eyes were evaluated as a first examination, and 15 eyes were seen during a follow-up visit. Of the 182 eyes included in the study, CMV retinitis was diagnosed in 89 (48.9%) by the attending
ophthalmologist. Each nonexpert grader reviewed all 182 retinal photographs.

The nonexpert graders recorded a diagnosis of CMV retinitis present in a mean of 71 eyes, CMV retinitis absent in a mean of 91 eyes, and unknown in a mean of 20 eyes. Remote diagnosis of CMV retinitis by nonexperts had sensitivities ranging from 64.0% to 95.5% (mean, 84.1%; 95% CI, 78.6%-89.6%) and specificities ranging from 65.6% to 92.5% (mean, 82.3%; 95% CI, 76.6%-88.0%). In comparison, the CMV retinitis experts achieved sensitivities ranging from 88.8% to 91.0% and specificities ranging from 84.9% to 88.2% (P = .14 and P = .17, respectively, Wilcoxon rank sum test) (Figure 1). The 3 groups of nonexperts did not have markedly different mean sensitivities (85.4% for graders with clinical experience, 82.9% for graders with research experience, and 84.3% for graders with neither clinical nor research experience; P = .59, Kruskal-Wallis test) or specificities (86.6%, 82.6%, and 79.2%, respectively; P = .28).

Because the predictive value of a test depends on the prevalence of disease, we depicted the positive predictive value and negative predictive value over a range of disease prevalence estimates (Figure 2). As described previously, the prevalence of CMV retinitis in eyes being screened for that disease during the study period was 35.6%. Using this 35.6% prevalence, a nonexpert diagnosis of CMV retinitis present indicates an 86.4% probability that the patient has the disease, whereas a diagnosis of unknown indicates a 37.4% chance of the patient having CMV retinitis. At this same 35.6% prevalence, a diagnosis of CMV retinitis absent has a 90.5% chance of being a true negative. As shown in Figure 2, nonexpert graders had lower mean predictive values than did expert graders.

Intrarater agreement for the 15 nonexpert graders ranged from κ values of 0.48 to 1.00 (mean, 0.83; 95% CI, 0.78-0.87) (Figure 3). Mean intrarater agreement was 0.94 for the group of nonexperts with clinical experience, 0.78 for those with research experience, and 0.76 for those with neither clinical nor research experience (P = .18, Kruskal-Wallis test). Graders with low intrarater agreement tended to also have lower sensitivities and specificities (correlation coefficient between intrarater κ and Youden index, 0.82; 95% CI, 0.53-0.94). Intrarrater agreement between the 15 graders was substantial, with an κ value of 0.63 (95% CI, 0.57-0.68). For comparison, the previous study16 found intrarater κ values of 0.93 for each expert grader and an intrarater κ of 0.86.

According to the expert consensus diagnosis, CMV retinitis was present in 73 eyes and absent in 91 eyes; no consensus could be established for the remaining 18 eyes. As shown in Figure 4, the diagnoses given by the 15 nonexperts were generally in agreement with the consensus diagnosis: mean sensitivities were 93.2% (95% CI, 90.6%-95.8%) and mean specificities were 88.4% (95% CI, 85.4%-91.1%). The 3 groups had similar mean sensitivities (94.5% for the group with clinical experience, 92.1% for the group with research experience, and 93.4% for the group with neither clinical nor research experience; P = .54, Kruskal-Wallis test) and specificities (90.4%, 88.6%, and 87.0%, respectively; P = .78).

The Cohen κ value is shown for each nonexpert grader based on grades from 50 randomly selected duplicate images; graders had clinical experience (clinical), research experience (research), or neither clinical nor research experience (neither). Limit lines indicate 95% CI.
We identified 9 eyes recorded as not having CMV retinitis by indirect ophthalmoscopy, which received a diagnosis of either CMV retinitis present or unknown by most nonexperts. An additional 9 eyes received a diagnosis of CMV retinitis by indirect ophthalmoscopy but not by most nonexperts (Table). For these discrepant cases, we compared the grades given by nonexperts with those given by the expert photographic graders and found that the diagnosis of most of the nonexperts agreed with that of the consensus of the experts (Table). Most of the false-positive nonexpert grades were the result of poor-quality photographs.

<table>
<thead>
<tr>
<th>Indirect Ophthalmoscopy</th>
<th>Expert Consensus</th>
<th>Nonexpert Diagnosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>CMV retinitis absent (false-positives)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Good</td>
<td>Absent</td>
</tr>
<tr>
<td>2</td>
<td>Good</td>
<td>Present</td>
</tr>
<tr>
<td>3</td>
<td>Good</td>
<td>Present</td>
</tr>
<tr>
<td>4</td>
<td>Poor</td>
<td>Unknown</td>
</tr>
<tr>
<td>5</td>
<td>Poor</td>
<td>Unknown</td>
</tr>
<tr>
<td>6</td>
<td>Poor</td>
<td>Unknown</td>
</tr>
<tr>
<td>7</td>
<td>Poor</td>
<td>Unknown</td>
</tr>
<tr>
<td>8</td>
<td>Poor</td>
<td>Unknown</td>
</tr>
<tr>
<td>9</td>
<td>Poor</td>
<td>Unknown</td>
</tr>
<tr>
<td>CMV retinitis present (false-negatives)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Good</td>
<td>Absent</td>
</tr>
<tr>
<td>11</td>
<td>Good</td>
<td>Absent</td>
</tr>
<tr>
<td>12</td>
<td>Good</td>
<td>Absent</td>
</tr>
<tr>
<td>13</td>
<td>Good</td>
<td>Absent</td>
</tr>
<tr>
<td>14</td>
<td>Good</td>
<td>Absent</td>
</tr>
<tr>
<td>15</td>
<td>Good</td>
<td>Absent</td>
</tr>
<tr>
<td>16</td>
<td>Good</td>
<td>Absent</td>
</tr>
<tr>
<td>17</td>
<td>Good</td>
<td>Absent</td>
</tr>
<tr>
<td>18</td>
<td>Good</td>
<td>Unknown</td>
</tr>
</tbody>
</table>

Abbreviation: CMV, cytomegalovirus.

Discussion

Relative to an in-clinic ophthalmology examination by an attending ophthalmologist, nonexpert graders diagnosed CMV retinitis from fundus photographs with a mean sensitivity of 84.1% and specificity of 82.3%. Nonexperts usually had lower accuracy compared with a trio of CMV retinitis experts, although several of the nonexperts had sensitivity and specificity estimates that clustered near those of the expert graders (Figure 1). Among the 15 nonexperts, 7 had a sensitivity level and 7 had a specificity level equal to or greater than that of at least 1 of the experts. However, only 2 nonexperts had both sensitivity and specificity matching or exceeding that of 1 of the expert graders.

The results from this study are consistent with those from a study of telemedicine diagnosis of ROP, in which nonexpert graders exhibited considerable variability (sensitivity ranged from 73% to 87% and specificity ranged from 73% to 91%). Moreover, personnel other than ophthalmologists are already used in clinical practice at dedicated reading centers to stage DR. In these cases, graders complete multiple-day training with strict oversight and quality assurance protocols. Given the variability exhibited in the present study, if nonexperts are to be used for remote diagnosis of CMV retinitis they likewise must be carefully selected and undergo more intensive training than was offered in this study. Moreover, given the inconsistent intra-rater agreement, periodic testing of nonexperts would be mandatory. Nonetheless, this study demonstrates that at least some nonexperts were capable of grading fundus photographs for CMV retinitis at a level similar to that of ophthalmologist experts, suggesting that nonexpert assessment could be useful for a CMV retinitis telemedicine program.

The prevalence of CMV retinitis during the study period (35.6%) was lower than the proportion of eyes in the present study that received a diagnosis of CMV retinitis (48.9%) because patients without CMV retinitis were not enrolled in the later months of the study. Although this discrepancy should not affect our estimates for sensitivity and specificity, the site-specific prevalence will affect the estimates of the positive and negative predictive values. In fact, because the study site is a tertiary referral center, it is likely that even the 35.6% prevalence estimate is higher than what would be observed at a typical HIV clinic in Thailand. At a 10% prevalence, which is a more realistic estimate for a nontertiary health care setting, the mean positive predictive value would be 61.9% for a diagnosis of CMV retinitis present and 11.5% for a diagnosis of unknown, and the mean negative predictive value would be 97.9%. These results demonstrate that a telemedicine program with nonexpert graders would be effective in ruling out CMV retinitis in patients who do not have the disease.
Telemedicine Diagnosis by Nonophthalmologists

The present study had several limitations. The study was designed to determine the sensitivity and specificity of an average nonexpert grader and therefore was not powered to be able to detect a difference between expert and nonexpert graders. Although training was performed in the grader’s native language, training materials were in English; Thai graders may have had a more difficult time referencing these materials after the training session. Thai-language materials could easily be developed if a telemedicine grading system based on nonexpert graders was to be implemented. The study included both patients presenting for an initial examination as well as those who received the CMV retinitis diagnosis in the previous month and had begun receiving treatment. We believe it unlikely that photographers were biased by this because they did not participate in this clinic before the study and were masked to all patient information. Moreover, because CMV retinitis resolves slowly and leaves a scar, we do not believe that receiving several weeks of antiviral treatment would have greatly altered the sensitivity and specificity of fundus photographs for detecting CMV retinitis.

Conclusions

Some, but not all, nonexpert graders in the present study were able to diagnose CMV retinitis using fundus photographs at a level that closely matched that of an expert grader. Strict training protocols and methods of evaluation would be required to select the best nonexpert graders and ensure accurate and consistent diagnosis. Nonexpert graders could potentially be used in primary care settings to screen at-risk HIV-positive patients for CMV retinitis using telemedicine, with referral to an ophthalmologist for definitive diagnosis. Studies of the cost-effectiveness of using nonexpert graders in a telemedicine program for CMV retinitis screening would be helpful before such a program were implemented.
Apivatthakul, Thanarat Chaichana, BS, Alissara Chaisong, CMA, Saringkarn Intachai, BA, Duanwarang Inthawichai, BA, Nutchanun Kitsanaruxamee, BA, Jaruwan Kongnern, BS, Phuririwat Maksongkhla, CMA, Chantawan Papsisanu, BSN, and Napatsawan Sangkun (Chiang Mai University); Sangita Annamalai, BA (University of California, Berkeley); and Kathleen Apakupakul, MS, Nicole Stoller, MPH, and Sophia Swanson, MD (University of California, San Francisco). Graders from Chiang Mai University were financially reimbursed for their contributions.

REFERENCES