Computerized Binocular Pupillography of the Swinging Flashlight Test Detects Afferent Pupillary Defects

Nicholas J. Volpe
Departments of Ophthalmology and Neurology, University of Pennsylvania School of Medicine, The Scheie Eye Institute, and the Department of Neurology, Albert Einstein Medical Center, Philadelphia, Pennsylvania, USA

Laila Dadvand, Shane K. Kim, Maureen G. Maguire, and Gui-Shuang Ying
Departments of Ophthalmology, University of Pennsylvania School of Medicine, The Scheie Eye Institute, Philadelphia, Pennsylvania, USA

Mark L. Moster
Department of Neurology, Albert Einstein Medical Center, Philadelphia, Pennsylvania, USA

Steven L. Galetta
Departments of Ophthalmology and Neurology, University of Pennsylvania School of Medicine, The Scheie Eye Institute, Philadelphia, Pennsylvania, USA

ABSTRACT

Purpose: To investigate the ability of a portable pupillometer, capable of 20-second binocular recordings of the swinging flashlight test (SFT), to detect relative afferent pupillary defects (rAPDs). Methods: Pupillary response curves were recorded from both eyes in healthy volunteers (n = 22) with and without simulated rAPDs (using neutral density filters (NDFs)) and in abnormal patients (n = 24) with clinically graded rAPDs. The light stimulus (0.2 sec on and 1 sec off, or 2 sec on and 0.4 sec off) alternated between both eyes, simulating the SFT. Constriction amplitude (CA), constriction velocity (CV), and pupillary release were calculated by computer algorithm. In abnormal patients, NDFs were used to neutralize inter-eye differences. Results: Significant correlation (Spearman’s ρ 0.71, 0.73) between NDF strength and absolute inter-eye differences was seen for CA and CV in simulated rAPDs. All abnormal patients (15/15) having rAPDs greater than 0.5 log units were distinguished from normals using either the upper bound of the one-sided 95% confidence interval (95% CI) value of CA or CV as determined from 22 healthy volunteers. Inter-eye variability in some normals prevented confident distinction of six abnormal patients with 0.3 log unit rAPDs. Using NDFs, subtle rAPDs were predicted in three patients having questionable rAPDs on clinical examination. CA and CV were more sensitive than pupillary release for all comparisons. Conclusions: This binocular pupillography identified all of our patients with >0.5 log unit rAPDs. Using NDFs, all of our abnormal patients were accurately identified and their rAPDs quantified. Variability in some normals makes them indistinguishable from patients with subtle rAPDs.

Keywords: Diagnostic tests; pupil; pupillography; relative afferent pupillary defect; swinging flashlight test

Received 24 September 2008; accepted 21 April 2009. Correspondence to: Nicholas J. Volpe, MD, Scheie Eye Institute, 51 North 39th Street, Philadelphia, PA 19104. E-mail: nickvolp@mail.med.upenn.edu

INTRODUCTION

Although clinical grading of relative afferent papillary defects (RAPDs) with neutral density filters (NDFs) has been shown to be an effective means of evaluating and following patients with optic neuropathies, this grading system is highly subjective in its endpoints. Modern pupillography has been proposed as an alternative to the clinical grading of RAPDs and has led to a more objective quantification of the swinging flashlight test (SFT). Different investigators have identified the pupillographic characteristics that are best for diagnosing RAPDs. Thompson confirmed that looking for pupillary escape with the SFT was superior to Marcus Gunn’s test for redilation under constant light exposure. Fison et al. and Cox placed NDFs of increasing density over one eye of normal subjects in order to create artificial RAPDs and used pupillography to detect differences in constriction amplitude (CA). Cox et al. and Bergamin et al. used pupillography to show that differences in CA upon light stimulation (direct versus consensual) is the most sensitive method of detecting RAPDs and that CA is the pupillary parameter that best differentiates abnormal patients from normal subjects.

Pupillography has been difficult to incorporate into clinical practice given its lack of availability and portability. Previously, we showed that a portable monocular pupillographer is capable of recording the SFT and that RAPDs of 0.9 log units could only be identified with 80% sensitivity and 92% specificity, making it of insufficient sensitivity and specificity to be used in clinical practice. In this study, we aimed to improve sensitivity and specificity, by using a portable pupillographer capable of binocular pupillography recordings for up to 20 sec.

METHODS

Study Design

The SFT was recorded with a portable computer-driven pupillographer in 22 healthy volunteers ages 15 to 53 with no known ocular pathology, normal corrected vision, and no RAPD on clinical examination, and in 21 non-consecutive patients ages 10 to 77 with RAPDs that were clinically quantified using NDFs and without different pupillary abnormalities. In addition, the SFT was recorded in 3 patients having questionable RAPDs. A total of 22 men and 24 women participated in the study, and the mean age (standard deviation) of the study participants in the normal subject and abnormal patient groups were 31 (10) years and 50 (20) years, respectively.

Pupillary diameter versus time curves were recorded for both eyes in all of the patients. After recordings were made of the 22 normal volunteers, in 10/22 normal volunteers, simulated RAPDs were created by placing an NDF in front of one of the eyes (5 OD, 5 OS). NDFs were placed in 0.3 log unit increments from 0.3 log unit to 1.2 log unit. In the other 12/22 normal volunteers, a single 0.3 log unit filter was placed over the right eye and subsequently over the left eye. The SFT was recorded for each of the simulated RAPDs.

Patients in the abnormal study group (n = 21) were examined by one of the authors (NJV), and the severity of the RAPD was determined using NDFs, which were placed over the healthy eye in increasing 0.3 log unit increments until the pupillary CA was equalized or the “pupillary escape/release” abolished. In the 3 patients having questionable RAPDs on examination, a 0.3 log unit filter was placed in front of each of the eyes to either accentuate or reduce subtle inter-eye differences.

The study was performed in accordance with the tenets of the Declaration of Helsinki and was reviewed and approved by the University of Pennsylvania’s Institutional Review Board. Verbal informed consent was obtained from each subject.

Pupillographer

Pupil responses to the SFT were recorded with the Procyon P2000D pupillographer (Procyon Instruments Ltd, London, England) (Fig. 1A). Each eyepiece was equipped with 4 infrared diodes (GaA1A’s type SFH485 with peak emission at 880 nm) mounted at a 10-degree angle. The light source was provided by green visible stimulus LEDs centered at 555 nm with an illuminance of approximately 5 lux. Images were captured at 25 frames/sec, 520 x 390 pixels/frame, and had approximately 30 pixels/mm resolution.

Pupillary responses to light were recorded at a short duration stimulus (0.2 sec light on, 1 sec light off) used in accordance with the stimulus cycles used by Kawasaki et al. and a long duration stimulus (2 sec light on, 0.4 sec light off) resembling the clinical SFT stimulus frequency. The SFT is displayed as a real-time video recording, as well as a pupillary diameter versus time curve of both eyes undergoing the SFT (Fig. 1B).

Recording the Swinging Flashlight Test

Each subject was given 2 min to adapt to the dark prior to recordings and was instructed to maintain distance fixation on binocular target points placed in the back of the pupillographer. If excessive blinking occurred during a recording, the process was repeated after a 1-min...
Figure 1. Pupillometer, pupil response curves, and measurement of CA and CV: (A) Procyon P2000D pupillometer. The pupillometer interfaces with a computer to display a real-time video image of the swinging flashlight test. (B) Pupillary response curves of a patient with a 0.6 log unit right rAPD. The shaded vertical bars represent the duration of light stimulus which alternates between OD and OS, thereby simulating the SFT. The curves represent changes in pupillary diameter in response to each light stimulus. (Top) Pupillary diameter versus time curves for the short duration stimulus (light on 0.2 sec, 1-sec dark interval, 7 of 8 stimulus cycles displayed). (Bottom) Pupillary diameter versus time curves for the long duration stimulus (light on 2 sec, 0.4-sec dark interval). (C) Calculation of CA (top) and CV (bottom) in an abnormal patient. (Top) Amplitude difference between maximum and minimum pupillary diameter is calculated. (Bottom) A straight line is fit to the constriction part of the line, with slope representing CV.

break. Single blinks were removed by manually altering the pupil diameter in the spreadsheet in order to smooth the pupil response trace. When creating simulated rAPDs or attempting to abolish rAPDs in abnormal patients, NDFs were placed between the stimulator diodes and the patient’s eye to dim the stimulus light. Each recording lasted 20 sec and was comprised of 8 cycles per intensity for the short-duration stimulus (Fig. 1B, top panel) and 4 cycles per intensity for the long-duration stimulus (Fig. 1B, bottom panel). The subject was given a 2-min rest interval in the dark between successive recordings.

Pupillary Response Curves

Raw data (pupillary diameter versus time) was analyzed using a computer algorithm, which calculates the change in pupillary diameter and constriction velocity (CV) for both right and left eye stimuli. The short
stimulus cycle was used for calculating the change in pupillary diameter and CV (Fig. 1C) in order to decrease variability in rAPD measurement by increasing the number of stimulus pairs per unit time. The long stimulus cycle was used instead of the short stimulus cycle for calculating pupillary release times in order to provide sufficient time following stimulation for pupillary release to be observed and measured. Pupillary release times were defined as the duration of time between the start of the light stimulus and the onset of pupillary redilation. The average value of each measurement parameter of both eyes for all of the right eye stimuli was compared with that for all of the left eye stimuli.6,8

Data Analysis

The data was analyzed using SAS statistical software (SAS Institute, Cary, NC, USA). For normal subjects, the absolute value of the differences between pupillary response during right eye versus left eye stimuli and the mean and upper bound of the one-sided 95% confidence interval (95% CI) values of these absolute differences were calculated for the three measurement parameters. The 95% CI values determined from normal subjects provided the upper limit of absolute differences that one would expect to observe in patients without eye disease. We calculated the sensitivity and specificity for detecting rAPDs using the 95% CI value as a threshold. These cut points, obtained using one-sided 95% confidence intervals based on a t-distribution analysis, are from a small sample of normals and are subject to change when more data are collected from normals.

We calculated the absolute values of the inter-eye differences in CA, CV, and pupillary release time in order to measure the magnitude of inter-eye differences for normal subjects and abnormal patients. In order to accommodate for the decreased distribution width, and hence the associated decreased variance resulting from using absolute values of the inter-eye differences of the pupillary parameters, we used the upper bound of the one-sided 95% CI instead of the upper bound of the traditional two-sided 95% CI as a cutoff point to define statistical significance.

For the simulated rAPDs, we performed a non-parametric analysis using Spearman’s correlation coefficient (ρ) to assess the association between absolute inter-eye differences and NDF strength. Sixty-six pupillary recordings were made in normals and 48 recordings were made in abnormals using different NDFs to either create artificial rAPDs or to abolish inter-eye differences.

RESULTS

Normal Volunteers

For CA, the mean absolute difference in pupillary response was 0.07 mm, and the upper bound of the 95% CI was 0.16 mm (Fig. 2A, “0 filter”). The mean absolute difference and the 95% CI value for CV are 0.11 mm/sec and 0.26 mm/sec, respectively (Fig. 2B, “0 filter”). For pupillary release, the mean absolute difference was 0.11 sec and the 95% CI value was 0.33 sec (Fig. 2C, “0 filter”).

Simulated Afferent Pupil Defects

A significant correlation between NDF strength and absolute inter-eye differences was observed for CA (ρ = 0.71, p < 0.0001; Fig. 2A). Furthermore, the absolute difference in CA for each grade of NDF was examined to determine what percentage had differences beyond the 95% CI for normal volunteers (0.16 mm). For artificial rAPDs created with a 0.3 log unit filter, 21/34 (62%) were above the 95% CI value for normals. The data for 0.6, 0.9, and 1.2 log unit filters are 6/10 (60%), 9/10 (90%), and 10/10 (100%), respectively, demonstrating that of our sample of artificial rAPDs, 62% of 0.3 log unit, 60% of 0.6 log unit, 90% of 0.9 log unit, and 100% of 1.2 log unit rAPDs could be reliably distinguished from normal.

The results for CV were similar to that obtained for CA (Fig. 2B). A significant correlation between absolute inter-eye differences and NDF strength was also observed (ρ = 0.73, p < 0.0001). Using the 95% CI value for our normal volunteers (0.26 mm/sec), the percentage of artificial rAPDs found to be greater than this value was similar to the data for CA: 59% of 0.3 log unit, 60% of 0.6 log unit, 90% of 0.9 log unit, and 100% of 1.2 log unit rAPDs.

The correlation for pupillary release times was not as strong as the results for CA and CV. There was a weak correlation (ρ = 0.44, p < 0.0001) between filter strength and inter-eye differences observed for pupillary release (Fig. 2C). The percentage of volunteers with artificial rAPDs that were found to have absolute differences greater than the 95% CI value for normals (0.33 sec) was significantly less than the percentages observed for CA and CV: 15% of 0.3 log unit, 40% of 0.6 log unit, 40% of 0.9 log unit, and 30% of 1.2 log unit.

Afferent Pupil Defects in Abnormal Patients

The data from our group of 21 abnormal patients with clinically graded rAPDs are presented in Figure 3. As
in the simulated rAPDs, the absolute difference (OD versus OS stimulation) in CA, CV, and pupillary release were compared to the grade of the rAPDs (NDF required to abolish difference) found in our abnormal patients prior to pupillography. For comparison, the absolute difference data obtained from the normal volunteers (“no APD”) is presented in Figure 3.

The data for CA and CV are very similar, and both parameters were equally useful in detecting rAPDs. In Figures 3A and 3B, the minimum absolute differences
in CA and CV for all 15 abnormal patients with \( \geq 0.6 \) log unit clinically graded rAPDs are all greater than the 95% CI value for normal volunteers ("no APD"). For both CA and CV, 1/6 patients with 0.3 log unit rAPDs were found to have an absolute difference greater than the 95% CI value for normal subjects.

The data using pupillary release was again inferior to that using CA and CV in terms of detecting rAPDs. Using this parameter, only 13/15 patients with \( \geq 0.6 \) log unit rAPDs were detected, and none of the patients with 0.3 log unit rAPDs were detected (Fig. 3C).

The pupillometer not only detected rAPDs based on absolute differences in our measurement parameters during right versus left eye stimulation, but also enabled us to quantify rAPDs in all 21 of our clinically graded rAPDs. We were able to adequately neutralize an absolute difference in CA and CV with an NDF having the same value as that used over the "good eye" to neutralize the rAPD during the clinical SFT, in 16/21 patients, such that the patient’s absolute inter-eye difference fell to a value below the 95% CI value of absolute inter-eye differences determined for normal volunteers. In the other 5/21 patients, an artificial rAPD was created in the healthy eye because the log unit of the NFD was clinically overestimated. Thus, in these 5 patients, we could only determine a quantifiable range that the rAPD could be classified based on the NDF used.

Three patients had questionable rAPDs in the left eye on clinical exam. Their inter-eye differences were not great enough to distinguish them from our normal volunteers. However, by placing a 0.3 log unit filter over the right eye and subsequently over the left eye, we were able to accentuate subtle rAPDs in these patients (Table 1).

### Table 1. Absolute inter-eye differences in CA in 3 patients with questionable rAPDs

<table>
<thead>
<tr>
<th>Patient</th>
<th>No Filter OD (mm)</th>
<th>0.3 filter OD (mm)</th>
<th>0.3 filter OS (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.05</td>
<td>0.12</td>
<td>0.24</td>
</tr>
<tr>
<td>2</td>
<td>0.04</td>
<td>0</td>
<td>0.29</td>
</tr>
<tr>
<td>3</td>
<td>0.01</td>
<td>0.07</td>
<td>0.39</td>
</tr>
</tbody>
</table>

*For example, patient 2 had an absolute inter-eye difference of 0.04 mm in CA. When a 0.3 log unit filter was placed over the right eye and subsequently over the left eye of patient 2, the inter-eye difference was 0 mm and 0.29 mm, respectively, suggesting the presence of a subtle, <0.3 log unit rAPD in the left eye.

**Contraction Anisocoria**

Since the pupillometer was capable of making binocular recordings, the presence of a difference between direct and consensual response (i.e., contraction anisocoria)\(^{13–17}\) was canceled out when both responses were averaged to calculate constriction amplitude.

**DISCUSSION**

By using the Procyon P2000D binocular pupillometer and recording for twice as long, we were able to improve the sensitivity and specificity of rAPD detection compared with that in our previous study using a monocular portable pupillometer.\(^9\) We were able to distinguish abnormal patients having clinically graded rAPDs greater than 0.5 log units from healthy volunteers, with 100% sensitivity and with 91% and 95% specificity, respectively, for inter-eye differences in CA and CV. As in our previous study,\(^9\) we observed that a significant correlation exists between NDF strength and absolute inter-eye differences for CA and CV in simulated rAPDs and that pupillary release times are not sufficiently sensitive to reliably and accurately detect rAPDs.

Attempts to develop a pupillometer accurate enough to screen for rAPDs has been limited by the variability of the SFT.\(^8,18\) Kawasaki et al.\(^8\) utilized pupillography combined with a computer analysis program to define the 95% CI of each determination of rAPD. They discovered that with a few light alternations, there was a large variability in the quantification of rAPDs (95% CI > 0.5 log unit). It required roughly 200 light stimulus pairs to reduce the 95% CI to 0.1 log unit. They used similar techniques to demonstrate that subjects with normal visual function have subtle and fluctuating rAPDs (up to 0.3 log units) when tested over three years.\(^18\)

This variability in the pupillary light reflex helps explain why a monocular pupillometer capable of recording only 4 stimulus pairs cannot distinguish rAPDs of 0.9 log unit or less from healthy volunteers.\(^9\) By utilizing a pupillometer capable of binocular recordings and by doubling the recording time to 20 sec, we were able to increase the number of stimulus pairs to 8 and the number of pupil recordings to 16 for the short stimulus sequence. By increasing the recording time to 20 sec, we were able to improve rAPD detection without making the test too long.

In order to test the ability of the pupillometer to detect rAPDs, we established a normal range of inter-eye differences for three measurement parameters. In normal volunteers, we established the 95% normal
threshold value (based on the 95% CI value) of inter-eye difference for CA, CV, and pupillary release as 0.16 mm, 0.26 mm/sec, and 0.33 sec, respectively. For instance, if there is a greater than 0.16-mm inter-eye difference for CA, then a rAPD is likely to be present. Because we wanted to determine whether this pupillometer could measure simulated rAPDs induced by NDFs in a population of normals, we simulated rAPDs in the same group of normals, as performed previously.3,7,9

Using the present pupillometer, we were able to distinguish abnormal patients with 0.6 log unit rAPDs or greater from normal volunteers, but were unable to reliably detect 0.3 log unit rAPDs. This limitation may never be adequately resolved, given the presence of subtle rAPDs of up to 0.39 log unit in subjects with normal visual function.18,19 In our group of 22 normals, the average difference in CA (OD versus OS stimulation) was 0.07 mm, and the 95% CI was 0.16 mm. In our 6 abnormal patients with 0.3 log unit rAPDs on clinical exam, the mean difference in diameter was 0.11 mm with a 95% CI value of 0.17 mm. By virtue of this statistical overlap, it is likely that people with normal visual function can have subtle rAPDs without detectable visual pathway pathology. Age differences between our normal subject group (n = 22) (mean age of 31 with a standard deviation of 10 years) and our abnormal patient group having 0.3 log unit rAPDs (n = 6) (mean age of 50, with a standard deviation of 19 years) may have contributed to this overlap in inter-eye differences for CA and CV for these two study groups (Fig. 3A and 3B). Thompson et al.3 state that younger pupils are more freely moving and older pupils are stiffer, and thus rAPDs tend to be overestimated in younger patients and underestimated in older patients. Our older abnormal patient group having 0.3 log unit rAPDs may have had stiffer pupils compared with our normal control population, possibly contributing to the overlap we observe in the inter-eye differences for CA and CV between these two subject populations. Benson et al.20 found there to be no positive correlation between subject age and pupillary response for a normal control group of subjects.

The presence of up to a 0.39 log unit rAPD in people with normal visual function19 suggests that the data obtained from our patients with artificial rAPDs cannot be compared to the data from our pool of abnormal patients without some caveats. In 10 of our normal subjects, NDFs were placed in front of one eye in a random fashion to create artificial rAPDs. Since many of our normal subjects had inter-eye differences at baseline, the placement of a NDF did not accurately simulate a rAPD of the log unit filter used. This observation combined with our small sample size, contributes to the fact that only 60% of 0.6 log unit and 90% of 0.9 log unit artificially created rAPDs could be distinguished from normal.

For our pool of abnormal patients, the ability to detect rAPDs of less than 0.6 log unit may not be possible by comparing differences in CA or CV. However, another way to quantify and measure rAPDs is to successfully eliminate the inter-eye difference with an appropriate NDF.3 We placed a NDF, whose value was equal to that used to quantify the rAPD during the clinical SFT, over the “good eye” to neutralize the recorded inter-eye differences in the 21 patients with abnormal rAPDs. The inter-eye differences were successfully neutralized in 16/21 patients, enabling us to confirm the clinical diagnosis and rAPD grading in these patients. In the remaining 5 patients, the rAPD was detected but clinically overestimated, and the inter-eye difference was reversed, signifying that the rAPD was less severe than was initially determined by clinical grading. Thus, observing differences in CA and CV during the SFT combined with the use of NDFs placed over the healthy eye enables this pupillometer to detect and grade all rAPDs. By placing a 0.3 log unit NDF over the right eye and subsequently over the left eye to accentuate or reduce inter-eye differences, even subtle rAPDs were identified.

Although commonly depended upon for the performance of the clinical SFT, pupillary release times recorded with this pupillometer were not sufficiently sensitive to reliably and accurately detect rAPDs. Cox found a similar limitation of “pupillary escape” in his study.7 The longer stimulus setting only allowed for 4 light swings to be recorded, which limited the ability to detect rAPDs of less than 0.6 log unit.

In this investigation, we demonstrate enhanced sensitivity and specificity in the detection of rAPDs, as compared to our previous study,9 by using a portable binocular pupillometer capable of 20-sec recordings. Future studies would include measuring the sensitivity and specificity of this pupillometer compared to an examiner who is blinded to the patient’s clinical status.

This binocular portable pupillometer, in conjunction with NDFs, may be used in clinical practice to detect rAPDs. From our study of abnormal patients with ≥ 0.6 log unit rAPDs, this binocular portable pupillometer was capable of detecting these rAPDs with 100% sensitivity and 91% and 95% specificity, respectively, for CA and CV. Furthermore, we found that rAPDs can be graded with this pupillometer by using NDFs placed in front of the healthy eye to abolish the baseline measurement difference observed. In most cases, subtle rAPDs may be confirmed by using 0.3 log unit filters placed over each eye to accentuate or abolish any
observed differences. As visual field perimetry testing is commonplace in the clinical setting to quantify and track visual field defects, as well as to provide a permanent record of these measurements, this pupillometer may similarly be used for the detection, quantification, recording, and follow-up of disease progression and recovery of rAPDs.4

**Declaration of interest:** The authors report no conflict of interest. The authors alone are responsible for the content and writing of the paper.

**REFERENCES**