Diagnosis and Grading of Papilledema in Patients With Raised Intracranial Pressure Using Optical Coherence Tomography vs Clinical Expert Assessment Using a Clinical Staging Scale

Colin J. Scott, MD; Randy H. Kardon, MD, PhD; Andrew G. Lee, MD; Lars Frisén, MD, PhD; Michael Wall, MD

Objectives: To compare and contrast 2 methods of quantitating papilledema, namely, optical coherence tomography (OCT) and Modified Frisén Scale (MFS).

Methods: Digital optic disc photographs and OCT fast retinal nerve fiber layer (RNFL) thickness, fast RNFL map, total retinal thickness, and fast disc images were obtained in 36 patients with papilledema. Digital optic disc photographs were randomized and graded by 4 masked expert reviewers using the MFS. We performed Spearman rank correlations of OCT RNFL thickness, OCT total retinal thickness, and MFS grade from photographs.

Results: OCT RNFL thickness and MFS grade from photographs correlated well (R=0.85). OCT total retinal thickness and MFS grade from photographs had a similar correlation of 0.87. Comparing OCT RNFL thickness with OCT total retinal thickness, a slope of 1.64 suggests a greater degree of papilledema thickness change when using the latter.

Conclusions: For lower-grade abnormalities, OCT compares favorably with clinical staging of optic nerve photographs. With higher grades, OCT RNFL thickness processing algorithms often fail, with OCT total retinal thickness performing more favorably.


Papilledema, or optic disc swelling due to raised intracranial pressure, has been graded using the Frisén Scale. This scale uses visual features of the optic disc and peripapillary retina to stage optic disc edema. Its reproducibility has been validated, but it is limited by use of an ordinal scale.

Optical coherence tomography (OCT) is a potential tool to quantify changes in the degree of papilledema and to monitor the efficacy of treatment interventions. First described in 1991 by Huang et al, OCT is a cross-sectional imaging technique that quantitatively assesses multiple layers of the retina, allowing measurement of the retinal nerve fiber layer (RNFL) with a resolution of approximately 10 µm using time-domain variables. Direct measurements are calculated by a computer algorithm to quantify the nerve fiber layer and total retinal thickness. OCT offers several advantages over conventional photographic imaging such as use with small pupil sizes and nuclear cataracts. OCT findings are reasonably reproducible on successive measurements in normal eyes and in eyes with glaucoma. Despite these advantages, limitations of OCT include the requirement for patient fixation stability during imaging without eye tracking and the failure of segmentation algorithms for defining RNFL thickness and total retinal thickness in eyes with severe papilledema.

Although fundus photographs require subjective interpretation, they show detailed topographic relationships of optic disc edema that may be difficult to quantify and document on clinical examination. Unlike OCT, photographic artifacts (e.g., defocus or incorrect light exposure) that could interfere with interpretation are usually obvious.

OCT is useful in evaluating various ophthalmic disorders such as band atrophy, glaucoma, diabetic retinopathy, cystoid macular edema, central serous retinopathy, macular hole formation, optic nerve head pits, and ischemic optic neuropathy. Most of these conditions involve RNFL thinning. Few studies evaluated the use of OCT in measuring RNFL or total retinal thickening due to papilledema. A 2007 study compared the findings of OCT vs disc photographs in children with idiopathic intracranial hypertension. Another study assessed the importance of RNFL thickness analysis in intracranial hypertension using imaging laser polarimetry. In pa-

Author Affiliations:
Department of Ophthalmology and Visual Sciences, Carver School of Medicine, University of Iowa, and Veterans Affairs Hospital, Iowa City (Drs Scott, Kardon, Lee, and Wall); and Institute of Neuroscience and Physiology, Department of Clinical Neurosciences, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden (Dr Frisén). Dr Lee is now with the Department of Ophthalmology, The Methodist Hospital and the Departments of Ophthalmology, Neurology, and Neurosurgery, Weill Cornell Medical College New York, New York.
dopapilledema but did not directly compare OCT vs fundus photograph grades. To our knowledge, this is the first study to compare OCT RNFL thickness and OCT total retinal thickness with fundus disc photographs in adults having raised intracranial pressure. Our objectives were to compare and contrast OCT and Modified Frisén Scale (MFS) grading with the peripapillary RNFL in quantifying optic disc edema.

Methods

Digital optic disc photographs of the right or left eye were selected for review if corresponding OCT images of the RNFL were available and if photographs and OCT images were of sufficient quality. For OCT analysis, the circular image had to be centered on the nerve, the signal strength had to be at least 5 (range, 1[lowest]-10 [highest]), and the proprietary algorithm used by the manufacturer (OCT3; Zeiss Meditech, Dublin, California) had to successfully delineate the borders of the RNFL on manual inspection of the images. The proprietary algorithm creates a best-fit estimate of the inner and outer borders of the RNFL and the total retinal thickness to calculate the mean thickness, as demonstrated by the overlaid white lines of the B-scan in Figure 1. Algorithm failure occurred when the borders of the retinal layers delineated by the software were qualitatively incompatible with the borders inferred by experts (C.J.S. and R.H.K.). For fundus photographs, the optic nerve had to be in focus to allow evaluation of the RNFL at the disc border. Analysis was monocular rather than binocular because of some poor photographic quality or failure to capture the entire extent of the z-axis by OCT due to peripapillary elevation and cropping of the B-scan in 1 of 2 eyes. Furthermore, use of both eyes would violate a principle of parametric statistics, as the results of one eye are not independent of the other eye. Although the relationship between eyes could be accounted for with nested analysis of variance, using both eyes in some patients would have complicated our statistical analysis. Patients retrospectively identified as having raised intracranial pressure at some point in their clinical course from January 2004 to March 2008 were included in the study, and the imaging studies chosen were from the patient’s first presentation to our clinic for evaluation. While most cases were idiopathic, the specific etiology was not addressed, as the pathogenesis of disc edema pathogenesis of papilledema was secondary to raised intracranial pressure in all cases is static and should not affect comparison of disc appearance with OCT values. No other cause of visual loss or optic disc edema was present such as optic nerve compression, inflammatory optic neuropathy, or ischemic optic neuropathy. Monocular digital photographs were randomized and graded by 4 masked expert reviewer neuroophthalmologists (3 from the University of Iowa and 1 from the University of Gothenburg, Sweden). Photographs were viewed at 1600 × 1200 resolution (SyncMaster 213T; Samsung, Ridgefield Park, New Jersey).

We used an adaptation of a clinical staging scale initially described by one of us (L.F.) (Table 1 and Figure 1).1 The first 2 grades are primarily dependent on changes surrounding the optic disc such as decreased translucency and opacification of the peripapillary nerve fiber layer (the term blurring is avoided because it has many meanings, including lack of image focus). Grade 0 represents no optic disc edema, and grades 1 through 5 represent increasing severity of optic disc edema. We modified the Frisén Scale by identifying a key feature for each grade, with the other criteria considered minor. For grade 1, the key feature is the presence of a C-shaped halo of opacification of the RNFL at the region of the optic disc border; a gap remained at the location of the maculopapular bundle. Rarely, a thin smooth-edged,
ring-shaped peripapillary light reflex may be present in normal anatomy; when present, it is a continuous reflex. Therefore, pathologic designation requires irregularity and breakup of any ring-shaped peripapillary reflex (grade 1 in Figure 1). Grade 2 is characterized by a circumferential halo with decreased translucency (opacification of the RNFL), with complete visualization of each major retinal artery and vein along its length as it courses over the optic disc and its border. Grade 3 has loss of visualization (discontinuity) along a segment of at least 1 major vessel as it courses over the optic disc margin but continues in the center of the optic nerve. Grade 4 is characterized primarily by discontinuity due to opacification of at least 1 but not all major vessels on the optic disc.

In the original description by Frisén, grade 5 is classified as follows: “Anterior expansion of the nerve head now dominates over sideways expansion. The nerve head assumes a relatively smooth, dome-shaped protrusion, with a narrow and smoothly demarcated halo. The major retinal vessels climb steeply over the dome surface. Segments of these vessels may or may not be totally obscured by overlying swollen tissue.” Because this grade did not have a clear-cut primary feature, we added one in which grade 5 is reached when there is obscuration of all major vessels on the optic disc. We excluded grade 5 cases from our analysis because the few examples did not have OCT images in which the algorithm did not fail.

Corresponding OCT images of each grade are shown in Figure 1. Compared with OCT RNFL thickness, OCT total retinal thickness borders may demonstrate a greater degree of neurosensory detachment with increasing edema, but the software algorithm can fail at severe papilledema levels, with borders appearing beyond the range of visualization by reviewers. For MFS grading of the optic disc photographs, if there was uncertainty in assigning a grade between 2 increments, the lesser grade was given an additional “+” symbol. For example, if it was difficult to determine whether there was a C-shaped halo or a 360° circumferential halo, the disc would be graded as 1+. Because the first sign of papilledema in an experimental monocular pure optic nerve edema study was a Key model was optic disc elevation, grade 0+ was used to characterize elevation that exceeded grade 0 but lacked the characteristics of grade 1.

OCT fast RNFL, fast RNFL map, and fast optic disc images were obtained at the time of digital color photography (Stratus OCT3, Zeiss Meditec). The mean RNFL thickness and total retinal thickness were obtained if the proprietary algorithm seemed to follow the expected borders in their entirety on the displayed B-scan and if the circular image seemed to be centered on the video display. Most RNFL circular image data represented 3 trials of each eye, for which the data were averaged. The trial with the most accurately appearing fit for RNFL borders along with the highest signal strength was chosen for further analysis. OCT data with low signal strength (<5) or for which no trials had adequate RNFL border tracing over 360° of the circular image were excluded and tabulated.

Disc grade for each patient was based on majority rule among 3 of us (R.H.K., A.G.L., and M.W.) using the same computer monitor for evaluation. If all 3 reviewers were not in agreement, grading by 2 reviewers was used, with the third reviewer’s grade excluded. To evaluate possible higher correlation, 2 of us (R.H.K. and M.W.) arranged the photographs in ranked order from lowest to highest amount of disc edema. As a cross-check, grading was performed by one of us (L.F.) and correlated with OCT findings independent of the 3 reviewers. We performed Spearman rank correlations of OCT RNFL thickness, OCT total retinal thickness, and MFS grade from photographs.

Because the Stratus OCT3 was not specifically designed to measure papilledema, the software failure rate of the algorithm for determining RNFL thickness and total retinal thickness was assessed as a function of papilledema grade. A separate cohort of 101 patients was randomly selected from a database of 224 patients with known papilledema at initial presentation from the University of Iowa using the same Stratus OCT3 to determine if borders were captured accurately. The image failed if the OCT algorithm followed the expected borders of RNFL thickness and total retinal thickness on the B-scan as evaluated by one of us (C.J.S.). The image failed if the algorithm did not follow the expected borders. The respective disc photograph for each OCT image was evaluated and excluded if of insufficient quality. Digital photographs were reviewed and graded by one of us (C.J.S.) using the MFS as previously described. The failure rate was calculated for each photographic MFS grade similarly as for each OCT image.

Among 36 patients with papilledema, 7 (19%) had grade 0, 7 (19%) had grade 1, 10 (28%) had grade 2, 4 had grade 3 (11%), and 8 (22%) had grade 4 disc edema using majority rule. The 3 reviewers agreed on grade most of the time, and only once was there more than a 1-grade difference by 1 reviewer compared with the other 2 reviewers (Figure 2A). One-grade differences occurred between all grades, and differences exceeding 1 grade occurred between grades 2 and 4 (Figure 2B). If the “+” notation was added to a grade, it was omitted to establish a numeric scheme for further quantitative analysis. OCT RNFL thickness and total retinal thickness showed significant correlation with the MFS grade from photographs.

In grading papilledema, we found strong correlation between the MFS and OCT findings. When OCT RNFL...
thickness was compared with MFS grade from photographs (using majority rule), Spearman rank correlation was 0.85 ($P < .001$). When OCT total retinal thickness was compared with MFS grade from photographs, Spearman rank correlation was 0.87 ($P < .001$) (Figure 3). In cross-check grading by one of us (L.F.), Spearman rank correlations were 0.79 ($P < .001$) and 0.83 ($P < .001$) for the latter.

Two of us (R.H.K. and M.W.) also ranked photographs based on lowest to highest amount of disc edema. This was done to increase stratification of the data to demonstrate potential higher correlation when using a finer scale. Spearman rank correlation increased to 0.88 ($P < .001$) when OCT RNFL thickness was compared with ranked order of MFS grade (range, 1-36) (Figure 4). Spearman rank correlation increased to 0.90 ($P < .001$) when OCT total retinal thickness was compared with ranked order of MFS grade. Interobserver Spearman rank correlation between the 2 reviewers was highly significant at 0.87. In majority rule when comparing MFS grading from photographs, OCT total retinal thickness had higher correlation than OCT RNFL thickness. When comparing OCT RNFL thickness with OCT total retinal thickness, a strong linear relationship was observed at $R^2 = 0.90$ ($y = 1.64x + 105.5$) ($P < .001$) (Figure 5). The slope of 1.64 suggests a greater degree of papilledema thickness change when using OCT total retinal thickness. Among 199 eyes in 101 patients evaluated to analyze whether the algorithm was able to capture expected borders of the RNFL and total retinal thickness, the highest failure rates were seen with higher grades of disc edema (Figure 6). Higher failure rates were found with OCT RNFL thickness than with OCT total retinal thickness. All photographs were of sufficient quality to assign an MSF grade without failure, as most patients had several photographs taken at the same time and the highest-quality photograph was used for assessment.

**COMMENT**

We found strong Spearman rank correlation between OCT RNFL thickness, OCT total retinal thickness, and MFS
Papilledema grading differences in 36 patients. A, Using optical coherence tomography (OCT) retinal nerve fiber layer (RNFL) thickness vs using ranked order of Modified Frisén Scale grade. B, Using OCT total retinal thickness vs using ranked order of Modified Frisén Scale grade.

The correlation between OCT and fundus appearance of the disc was greatest when photographs were ranked in continuous order of edema severity. Variation in OCT measurements at each grade could be partly explained by the ordinal character of the MFS, as ranking the photographs in order within each grade resulted in higher correlation between OCT findings and MFS grade. We modified the Frisén Scale in an attempt to improve its reproducibility. We avoid the term blur, as the cause could be multifactorial, including media opacification or other optical anomalies. Instead, we prefer the term obscuration to more clearly delineate the pathologic effect at the peripapillary nerve fiber layer. We no longer include retinochoroidal folds as criteria for grade 2 because they can occur at any disc stage. The scale has been modified to no longer include domed disc in grade 5, as this has been difficult to define. Instead, grade 5 criteria include complete obscuration of all vessels on the disc and leaving the disc. While the MFS is useful for daily practice and for clinical trials, limitations remain (Table 2). Ophthalmoscopic evaluation and grading of papilledema can show significant variability among observers and can be influenced by media opacities, cataract formation, and pigment epithelium, which can also affect fundus photographs. With digital photography, variability may occur for several reasons, including characteristics and calibration of the viewing monitor. Variability among examiners could also occur because of direct examination vs photographic evaluation and due to magnification and computer monitor variables. OCT also has limitations (Table 2). For example, the quality of each image is eye (ie, media opacity) dependent and operator dependent and can be affected by improper disc centering and poor fixation. Low signal strength can also significantly reduce data quality, which was often seen in higher grades of disc edema; however, data with signal strength of less than 5 were excluded in our study.

Our study also demonstrated increased variability in the mean OCT RNFL thickness and OCT total retinal thickness with higher grades of disc edema. One of the greatest limitations of the OCT3 is the proprietary algorithm to trace the RNFL thickness and the total retinal thickness, which often cannot accurately outline thicknesses in advanced stages of papilledema. Although our study found this to be less problematic when measuring total retinal thickness, the failure rate remained high for grade 5 papilledema at 33%. Disc hemorrhages can also affect accuracy, as they may influence reflectivity in an inconsistent manner. Even with evolving technology to
improve OCT, fundus photographs will likely remain an important and reliable method for the assessment of disc edema because of their availability, use among multiple medical disciplines, and few but obvious artifacts, as well as the ability of the examiner to individually assess edema using multiple cues.

Our study has several limitations. Our sample size is small, especially at the higher papilledema grades. Measurement error can occur with OCT and with assessment of fundus photographs, and OCT algorithm failure rates can become significant with high grades of papilledema. This study demonstrates that OCT and MFS are complementary methods that can be used to follow up patients with papilledema. However, OCT failure rates increase as disc edema increases, suggesting that with current algorithms OCT may be more useful at lower grades of papilledema. The algorithm defining total retinal thickness had a lower failure rate than that defining RNFL thickness, indicating that OCT total retinal thickness measurement may be more robust, especially at higher grades of papilledema. The stronger correlation of MFS grade with total retinal thickness compared with RNFL thickness suggests that more emphasis should be placed on total retinal measurements.

Optical coherence tomography can be a useful tool in the evaluation of papilledema by confirming or complementing an examiner’s clinical examination. The advantages are listed in Table 2 compared with digital optic disc photographs; particularly, OCT quantification uses a continuous scale, whereas MFS grading from photographs uses an ordinal scale (grade range, 0-5). We do not recommend OCT in all patients with papilledema but believe it may be useful to confirm results when diagnosis is challenging. OCT may be especially useful when there is uncertainty about the presence of optic disc edema. We are investigating its use in evaluating disc edema change over time, and good correlation between photographic grading and OCT analysis is encouraging. Our ability to measure high-grade papilledema should be improved with new segmentation algorithms providing more robust software analysis of disc edema in conjunction with higher-quality images, which are easier to obtain with spectral-domain OCT. Current OCT measurement can be useful as an adjunct to clinical staging.

Submitted for Publication: January 8, 2009; final revision received September 2, 2009; accepted October 5, 2009.

Correspondence: Michael Wall, MD, Department of Ophthalmology and Visual Sciences, Carver School of Medicine, University of Iowa, 200 Hawkins Dr, Iowa City, IA 52242 (michael-wall@uiowa.edu).

Financial Disclosure: None reported.

Funding/Support: This study was supported by a post-doctoral fellowship grant from Fight for Sight (Dr Scott), by merit and rehabilitation grants from the Department of Veterans Affairs (Drs Kardon and Wall), and by an unrestricted grant from Research to Prevent Blindness.

Table 2. Advantages and Disadvantages of Optical Coherence Tomography and Modified Frisén Scale Grade From Photographs for Quantifying Papilledema

<table>
<thead>
<tr>
<th>Optical Coherence Tomography</th>
<th>Modified Frisén Scale Grade From Photographs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Advantage</td>
<td>Advantage</td>
</tr>
<tr>
<td>Reliable at lower grades</td>
<td>More reliable at higher grades</td>
</tr>
<tr>
<td>Continuous as opposed to ordinal measurement scale</td>
<td>Obvious artifacts can be ignored when observing key features</td>
</tr>
<tr>
<td>Not as limited by pupil size or cataracts</td>
<td>Use among many specialties without specialized equipment</td>
</tr>
<tr>
<td>Requires less observer experience</td>
<td></td>
</tr>
<tr>
<td>Disadvantage</td>
<td>Disadvantage</td>
</tr>
<tr>
<td>In higher grades, possible incorrect retinal nerve fiber layer thickness or total retinal thickness borders</td>
<td>Cataracts and media opacities limit quality</td>
</tr>
<tr>
<td>More patient cooperation required</td>
<td>Requires pupil dilation Subjective and dependent on experience</td>
</tr>
<tr>
<td>No algorithm for optic disc edema</td>
<td></td>
</tr>
<tr>
<td>Various artifacts such as those induced by poor centering, myopia, or movement may not be obvious and may interfere with interpretation</td>
<td>Increased variability and measurement failure at higher grades</td>
</tr>
<tr>
<td>Increased variability and measurement failure at higher grades</td>
<td></td>
</tr>
</tbody>
</table>

Figure 6. Pass rates and fail rates using optical coherence tomography (OCT) measurements vs Modified Frisén Scale grade. A total of 199 eyes were evaluated in 101 patients. A, Using retinal nerve fiber layer (RNFL) thickness. B, Using total retinal thickness.