Diagnostic Tools for Glaucoma Detection and Management

Pooja Sharma, BS,1 Pamela A. Sample, PhD,2 Linda M. Zangwill, PhD,2 and Joel S. Schuman, MD3

1Drexel University College of Medicine, Philadelphia, Pennsylvania, USA; 2Hamilton Glaucoma Center, Department of Ophthalmology, University of California, San Diego, La Jolla, California, USA; and 3Department of Ophthalmology, UPMC Eye Center, Eye and Ear Institute, Ophthalmology and Visual Science Research Center, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania, USA

Abstract. Early diagnosis of glaucoma is critical to prevent permanent structural damage and irreversible vision loss. Detection of glaucoma typically relies on examination of structural damage to the optic nerve combined with measurements of visual function. To aid the clinician in evaluation of visual function and structure, computer-based devices such as confocal scanning laser ophthalmoscopy, scanning laser polarimetry, and optical coherence tomography provide quantitative assessments of structural damage, and visual function testing includes standard automated perimetry as well as selective techniques, including short-wavelength automated perimetry and frequency-doubling technology perimetry are available. This article will review current literature on diagnostic modalities available for glaucoma with emphasis on the best evidence available in the literature to support their use in clinical practice. (Surv Ophthalmol 53:S17–S32, 2008. © 2008 Elsevier Inc. All rights reserved.)

Key words. confocal scanning laser ophthalmoscopy • diagnosis • frequency-doubling technology perimetry • glaucoma • optical coherence tomography • scanning laser polarimetry • short-wavelength automated perimetry • standard automated perimetry • stereophotography

Glaucoma is a chronic neurodegenerative disease characterized by loss of retinal ganglion cells, resulting in distinctive changes in the optic nerve head (ONH) and retinal nerve fiber layer (RNFL). Early diagnosis is critical to prevent permanent structural damage and irreversible vision loss. Detection of glaucoma typically relies on examination of structural damage to the optic nerve combined with measurements of visual function. Because clinical examination of the ONH and RNFL is subjective and therefore prone to variability, recent research has focused on objective methods to aid in the diagnosis of glaucoma. Techniques such as confocal scanning laser ophthalmoscopy, scanning laser polarimetry, and optical coherence tomography have been extensively studied as adjuncts to subjective ONH evaluation. Similarly, selective perimeter techniques, including short-wavelength automated perimetry (SWAP) and frequency-doubling technology (FDT) perimetry, are being explored as replacements to standard automated perimetry (SAP) to provide earlier detection of visual field deficits. This article will review current literature on diagnostic modalities available for glaucoma with
emphasis on the best level I and level II evidence available in the literature to support their use in clinical practice.

Following are the evidence levels recommended by the US Preventive Services Task Force (www.ahrq.gov/clinic/uspstfix.htm), given to provide a framework for evaluating the current peer-reviewed literature (through October 2007) on both measures of structure and function in the diagnosis and follow-up of glaucoma:

- Level I: (Interventional) Evidence obtained from at least one properly done, well-designed randomized controlled trial or meta-analysis of high-quality randomized controlled trials.
- Level I: (Observational) Evidence obtained from well-done, population-based prevalence or incidence studies.
- Level II: (Interventional) Evidence obtained from well-done, non-randomized comparative trials or well-done, systematic literature reviews summarizing primarily level II publications.
- Level II: (Observational) Evidence obtained from high-quality, case-control and cohort studies.
- Level III: (Interventional or Observational) Evidence obtained from non-comparative case series, case reports, and expert or consensus opinion.

**Evaluation of Glaucomatous Structural Damage**

**INTRODUCTION**

Irreversible loss of retinal ganglion cells characteristic of glaucoma manifests as ONH cupping as well as focal and diffuse RNFL loss. Current evidence suggests that in many eyes significant optic nerve damage may precede visual field loss; in the Ocular Hypertension Treatment Study (OHTS), for example, disk change was detected earlier than visual field abnormalities in over half of patients progressing to an initial diagnosis of glaucoma.\(^4^4\) As a result, over the past decade, ONH and RNFL imaging has gained widespread use in the diagnosis and management of glaucoma patients. Currently no level I data exists for the use of ONH and RNFL measurement devices.\(^8^2\) Most level II studies use a definition of glaucoma that incorporates glaucomatous visual field loss defects, although the use of visual field testing as a reference standard for glaucoma has recently been challenged in favor of progressive optic disk change.\(^9^4\)

**OPTIC DISK PHOTOGRAPHY**

Stereoscopic ONH photography is a simple and low-cost method providing a three-dimensional full-color view of the ONH; in practice, it is the most commonly utilized technique to objectively document structural damage in glaucoma suspects.\(^3^7\) Stereoscopic views of the optic nerve via ophthalmoscopy or slit-lamp biomicroscopy, documented by drawings in the patient’s chart, are also an important method to detect glaucomatous neuropathy. However, due to the inherent subjectivity of a qualitative assessment, there is considerable variability in classifying the ONH as normal or glaucomatous both within and between graders.\(^1^3^2,1^3^5\) Even among glaucoma specialists, there can be high intra- and interobserver variability in clinically assessing the optic disk.\(^3^9\) Optic disk damage based on photograph assessment has been used as an endpoint in three randomized clinical trials (level I evidence): the OHTS, the Early Manifest Glaucoma Trial (EMGT), and the European Glaucoma Prevention Study (EGPS). These studies have shown that by standardizing optic disk evaluation, photographs can be reproducibly evaluated.\(^1^0^7,1^5^2\) Other level II studies have also overcome some of this variability by using a variety of methods to standardize optic disk evaluation.\(^3^2,3^7,4^1,5^2,9^4\)

Recently, there have been tremendous advances in the development of computer-based technologies with the ability to provide reproducible, quantitative assessments of the ONH. An advantage of subjective assessment over quantitative analysis is a comprehensive evaluation of the ONH, including parameters that cannot be quantified, such as disk hemorrhages and pallor. In fact, given the wide range of normal variations of the ONH, qualitative variables have been shown to have higher specificity than quantitative parameters in separating normal from glaucomatous eyes. Furthermore, subjective ONH evaluation provides the clinician with the opportunity to assess the impact of other nonglaucomatous processes that may impact functional testing.

Manual, subjective examination of the ONH via ophthalmoscopy, slit-lamp biomicroscopy or stereoscopic optic nerve head photography remain mainstays in the evaluation of a glaucoma patient, with objective documentation of optic disk damage preferred whenever possible.\(^8^2\)

**CONFOCAL SCANNING LASER OPHTHALMOSCOPY (CSLO)**

CSLO is an imaging tool designed to provide the examiner with a quantitative three-dimensional composite image of the ONH and posterior segment. The commercially available instrument, the Heidelberg Retina Tomograph (HRT; Heidelberg Engineering, Heidelberg, Germany), works by emitting a 670-nm diode laser beam to sequentially...
scan the retinal surface in horizontal and vertical directions at multiple focal planes, generating a stack of 64 coronal planes, each with 384 × 384 pixels. These stacks are then re-assembled to enable height measurements of the retinal and ONH surface topography. A number of stereometric parameters are generated by CSLO, including rim area, rim volume, cup shape measure, linear cup/disk (C/D) ratio, retinal height variation along the contour line, and RNFL thickness. These parameters are automatically generated by the device software after user identification of the ONH margin (the creation of a contour line), and some require a reference plane automatically generated by the CSLO software 50 microns below the mean retinal surface inferior-temporal 6 degrees on the user defined contour line. All structures within the contour line and above the reference plane are considered neuro-retinal rim, and below the reference plane, cup.

HRT stereometric parameters, discriminant analysis and the Moorfields Regression Analysis (MRA) have demonstrated the ability to discriminate between healthy patients and early glaucoma patients diagnosed by stereophotography. The MRA divides the ONH into six sectors, and classifies each sector as well as the overall ONH as within normal limits (WNL), borderline (BL), or outside normal limits (ONL) based on comparison to an age and ethnicity specific normative database. Additionally, HRT assessments of the ONH are capable of distinguishing between healthy eyes and eyes with glaucomatous visual field defects, with a range of sensitivities from 51–97% and a range of specificities from 75–95%. Furthermore, the OHTS reported many HRT parameters to be associated with the development of glaucoma by univariate and multivariate analysis; the most predictive values were mean height contour, rim area, and mean cup depth.

Advantages of HRT include good image quality through undilated pupils (though dilation may be necessary at times), and the ability to upgrade existing machines with newer software, allowing the clinician to build upon long-term databases. Most importantly, the sophisticated registration capability of HRT to superimpose baseline and follow-up images allows for automated detection of change to the ONH. The use of HRT in the ancillary study to OHTS has resulted in a well-characterized data set, beneficial for future investigations of this technique.

A limited number of level II studies compare the abilities of stereophotographic grading and HRT to detect glaucomatous damage. Some level II studies have shown that stereophotographic grading provides greater diagnostic efficacy when compared to CSLO; however, in these studies, glaucoma specialists graded the stereophotographic ONH images, which may not reflect ONH assessment in general clinical practice. Other level II studies have shown that the diagnostic accuracy of photographs is comparable to HRT MRA classification and that HRT linear C/D ratio measurements can be used interchangeably with stereophotograph C/D ratio measurements in the OHTS predictive models and risk calculator.

Limitations of the HRT include the requirement for the operator to manually outline the disk margin and the use of a reference plane in the calculation of many stereometric parameters. This has been addressed in the newer generation HRT version 3.0 software Glaucoma Probability Score (GPS), based on the work of Swindale and colleagues, which provides automated interpretation of ocular topography, eliminating the need for an operator-drawn contour line and reference plane. This advancement reduces a source of variability in CSLO measurements. The GPS has been shown in level II studies to have comparable overall diagnostic accuracy to the MRA, with the GPS tending toward higher sensitivity and lower specificity.

Furthermore, the diagnostic accuracy of both MRA and GPS improves with increasing disk size and severity of disease; however, very small disks and very large disks tend to reduce the sensitivity and specificity of the device, respectively. An additional improvement to the HRT version 3.0 software is a larger and ethnicity-specific normative database.

Detection of glaucomatous change is important for early detection of the disease as well as for management of glaucoma patients, yet is often difficult to assess. Change may be the first sign of glaucoma, particularly in patients with suspicious often large optic disks that are difficult to determine whether they are physiologic or glaucoma. The HRT Topographic Change Analysis (TCA) provides localized, objective, and quantitative information on the volume, area, and depth of retinal height changes using sophisticated statistical analysis that automatically identifies repeatable change greater than the variability of the superpixels comprising an individual’s images. Although further studies are needed to document the ability of the HRT to detect change over time, the HRT is a promising tool for early glaucoma diagnosis.

**SCANNING LASER POLARIMETRY (SLP)**

SLP is a non-invasive method to objectively measure the RNFL; RNFL thickness corresponds to a decrease in the ganglion cell layer from the...
fovea to the optic disk. SLP has gained popularity as a potential diagnostic tool for glaucoma particularly in response to studies indicating that RNFL damage may precede optic nerve damage in early glaucoma.\textsuperscript{112} The instrument consists of confocal scanning laser ophthalmoscope with a polarized laser beam; when the polarized light passes through the birefringent RNFL, a measurable phase shift is created, which can be correlated to the RNFL thickness.\textsuperscript{127}

SLP was first commercially available as the GDx Nerve Fiber Analyzer (Laser Diagnostic Technologies, Inc., San Diego, CA). This instrument contained a fixed anterior segment compensating device to compensate for the polarization effects of other ocular birefringent structures, such as the cornea and lens. In light of evidence that the parameters for corneal compensation are different for different subjects,\textsuperscript{17} a new device with variable corneal compensation (GDx-VCC) (Zeiss Meditec, Dublin, CA) has been developed to allow for individualized eye-specific compensation of anterior segment birefringence. Several studies have shown that the addition of VCC to GDx substantially enhances its discriminating power for glaucoma detection\textsuperscript{15,123} and correlation with visual field loss.\textsuperscript{15,123}

GDx-VCC has been shown to have good diagnostics accuracy, with reported area-under-the-curve values (AUCs) for glaucoma detection ranging from 0.90–0.978.\textsuperscript{12,36,97} A study comparing GDx-VCC with RNFL photography found that although both techniques correlate with damage in corresponding hemiretinas, the best GDx-VCC parameter had a higher degree of discriminant ability than the best RNFL photographic parameter.\textsuperscript{96}

Recent data suggests that GDx-VCC may be useful for earlier glaucoma diagnosis. In a cross-sectional analysis by Medeiros et al, the GDx-VCC was able to detect structural abnormalities in preperimetric eyes with progressive optic disk changes as compared to controls.\textsuperscript{94} A study of glaucoma suspects by Mhammadi et al demonstrated that the original GDx Nerve Fiber Analyzer RNFL thickness measurements at baseline were predictive of future glaucomatous visual field loss.\textsuperscript{104}

Limitations of older SLP devices include fixed corneal compensation; newer models such as the GDx-VCC and GDx-ECC individualize anterior segment birefringence compensation. Anterior and posterior segment pathology does affect the accuracy of SLP measurements. Unreliable values for RNFL thickness have been reported in patients with media opacities, ocular surface diseases, peripapillary atrophy, and in those who have had keratorefractive surgery.\textsuperscript{55} The presence of vitreous opacities, optic nerve crescents, and other non-glaucomatous retinal distortions may induce erroneous RNFL measurements.\textsuperscript{55} Additionally, some GDx-VCC scans are characterized by problematic atypical birefringence patterns (ABPs); ABPs result from artifact introduced by the device’s attempt to compensate for poor noise-to-signal ratio.\textsuperscript{5} An updated technique, enhanced corneal compensation (ECC), was developed to reduce this artifact. Recent studies found that GDx-ECC significantly reduces the frequency and severity of ABPs and improves the correlation between RNFL measures and visual function as compared to GDx-VCC.\textsuperscript{13,90,127} As with many rapidly advancing technologies, the need to upgrade hardware with each advancement to the GDx has made it challenging for clinicians to follow patients over time with SLP. Although progression software is now available, images must be acquired with the current generation of instruments to be included in the analysis. Additional research is still needed on the new GDx-ECC to evaluate its utility as a diagnostic test for glaucoma, and to evaluate progression analysis in large longitudinal cohorts.

OPTICAL COHERENCE TOMOGRAPHY (OCT)

OCT uses low-coherence interferometry to perform high-resolution cross-sectional imaging of tissue morphology, providing an optical biopsy. It is analogous to ultrasound except it uses light instead of sound. Low-coherence near-infrared light is transmitted from a diode light source to the retina via a fiber optic delivery system.\textsuperscript{60} Backscatter from the retina is captured and used to construct a cross-sectional tomographic image of the retina. OCT permits direct, real-time visualization of retinal pathology and also provides quantitative measurements of retinal architecture at higher resolutions than CSLO and SLP.\textsuperscript{60} The current commercially available ophthalmic time domain OCT (Stratus OCT; Carl Zeiss Meditec, Inc.) has ~10-micron axial image resolution.

Limitations of OCT include the need to dilate some patients who undergo OCT, as well as the lack of an ethnicity-specific normative database and progression analysis software, although the latter are currently in production. In addition, the Stratus OCT does not have the ability to automatically register follow-up to baseline scans to ensure that the measurements are obtained at the same location for analysis of change. As for all optical imaging technologies, image quality may be compromised in patients with ocular opacities.

RNFL measurements by OCT prototype, I/II, and Stratus have demonstrated good reproducibility.\textsuperscript{109,125}
and OCT prototype and I and II measurements have shown correlation with glaucoma status\textsuperscript{124} and RNFL appearance.\textsuperscript{110} Additionally, OCT is able to identify RNFL defects in areas corresponding to visual field deficits.\textsuperscript{139} Some studies have shown that RNFLs of the superior and inferior quadrants have been shown to have the best discriminating power between eyes with glaucomatous visual field loss from controls (AUC 0.79–0.952, superior; AUC 0.863–0.971, inferior).\textsuperscript{11,19,70,79,106} Specifically, the inferior/inferotemporal (6 and 7 clock hours) and superior/superotemporal (11 and 12 clock hours) had the highest AUCs. Other studies have found that mean RNFL thickness has the highest AUC.\textsuperscript{141,144}

Although OCT was originally designed to evaluate retinal thickness, software development has also permitted ONH analysis; the highest performing ONH parameters for Stratus OCT such as C/D ratio and integrated rim volume have AUCs equivalent to the best RNFL parameters.\textsuperscript{79,80,84,126,144} OCT is also useful for macular volume assessment,\textsuperscript{48} which may be technically easier to measure than RNFL thickness.\textsuperscript{48} However, studies comparing the ability of Stratus OCT to distinguish between normal and glaucomatous eyes using macular, RNFL, and ONH assessments found that the ONH and RNFL thickness parameters provided the best discrimination between normal and glaucomatous eyes, whereas measurements of total retinal thickness in the macula lacked discriminating power.\textsuperscript{21,35,78} In order for macular thickness to have clinical value in glaucoma discrimination segmentation of the intraretinal layers is necessary.\textsuperscript{62}

Only one cohort study by Wollstein et al has been published to date evaluating the ability of a prototype OCT device to detect change in RNFL thickness longitudinally. In comparison to SAP, OCT detected more progression events (defined by a RNFL thickness decrease of 20 microns) during the 5-year course of the investigation; 22% progressed by OCT only, 9% by SAP alone, and 3% by both SAP and OCT. Whereas the sensitivity of OCT was higher than that of SAP, the relative specificity of the structural and functional measures requires further study.\textsuperscript{145}

Newer developments in OCT technology have enabled increases in scanning speed to 20,000–50,000 A-scans/second in commercial devices, allowing the creation of three-dimensional datasets. This technology, known variously as Fourier-domain OCT, spectral OCT, frequency-domain OCT, and high-speed, high-resolution OCT, has recently become commercially available. Three-dimensional imaging promises to permit registration from scan session to session, and to allow arbitrary analysis of three-dimensional datasets. Further, recent development of an ultrahigh-resolution (UHR) OCT in commercially available high-speed OCT devices enables intraretinal imaging comparable to conventional histopathology, including visualization of the ganglion cell layer, photoreceptor layers, and retinal pigment epithelium. UHR OCT uses a femtosecond laser light source to obtain axial resolutions of ~3 mm in the human eye.\textsuperscript{33,40} Studies comparing UHR OCT with standard OCT imaging found that ultrahigh-resolution enables improved visualization of intraretinal morphology and may enhance the clinical utility of standard OCT imaging for glaucoma.\textsuperscript{142}

Higher resolution OCT and larger datasets allow segmentation of retinal layers with high precision. Newer means of glaucoma assessment based on these innovations are likely to be forthcoming. The promise of enhanced sensitivity, specificity, and reproducibility will require validation in future studies.

STRUCTURAL ASSESSMENT COMPARISONS (ALL LEVEL II EVIDENCE)

Studies directly comparing imaging techniques have shown no significant differences in their abilities to distinguish glaucomatous from control eyes. Two articles comparing OCT with GDx revealed similarly high AUCs for glaucoma detection,\textsuperscript{35,78} although in another analysis, Stratus OCT parameters correlated with functional loss better than GDx VCC parameters by regression analyses.\textsuperscript{81} Although high correlations among ONH measurements have also been demonstrated in studies comparing HRT I and OCT II and Stratus OCT,\textsuperscript{126} and HRT II and Stratus OCT,\textsuperscript{54} significant differences between the actual measurements were reported, so that the optic disk measurements should not be used interchangeably for patient care. Medeiros et al compared the most widely utilized versions of each of the three imaging technologies (HRT II ONH scan, GDx VCC RNFL scan, and Stratus OCT fast RNFL scan), and reported nearly equivalent AUCs for the best parameters from each device.\textsuperscript{98}

STRUCTURE SUMMARY

Further research is required to evaluate the abilities of imaging technologies for the diagnosis of glaucomatous structural damage and progression of the disease. Nonetheless, several important conclusions can be drawn from the studies done to date.

1. Glaucoma causes structural damage to the ONH, RNFL, ganglion cell layer, and inner plexiform layer.
2. Photographic assessment of the ONH remains a mainstay in the diagnosis and management of glaucoma suspects and glaucoma patients; however, the technologies discussed herein are powerful tools that may assist the clinician in the early diagnosis of glaucoma. They provide objective and quantitative analysis and standardization of the interpretation of ocular structure at an expert level. Furthermore, these technologies facilitate earlier detection of functional loss and enhance assessment of structure-function correspondence.

3. An evidence-based medicine review of ophthalmic glaucoma diagnostic imaging technologies from the American Academy of Ophthalmology concluded as follows:

The ONH and RNFL imaging devices provide quantitative information for the clinician. Based on studies that have compared the various available technologies directly, there is no single imaging device that outperforms the others in distinguishing patients with glaucoma from controls. Ongoing advances in imaging and related software, as well as the impracticalities associated with obtaining and assessing optic nerve stereophotographs, have made imaging increasingly important in many practice settings. The information obtained from imaging devices is useful in clinical practice when analyzed in conjunction with other relevant parameters that define glaucoma diagnosis and progression.

---

Evaluation of Glaucomatous Functional Damage

INTRODUCTION

Over the past several years, a number of clinical trials have incorporated the development or progression of vision loss measured with standard visual fields as a primary endpoint. The OHTS evaluated the effects of lowering intraocular pressure in eyes with ocular hypertension that had no loss of vision on SAP or evidence of abnormality on simultaneous stereophotographs at baseline. Endpoints for this study were repeatable abnormality on three consecutive SAP fields or evidence of glaucomatous optic neuropathy (GON) on two consecutive stereophotos. This level I study of IOP-lowering also provided some post-hoc analyses (level II evidence) of the predictive value of various clinical measures. This analysis showed that a larger SAP pattern standard deviation (PSD), even though within the normal range, was predictive of the development of glaucoma in the risk model. This finding was validated in an independent dataset of ocular hypertensives from the Diagnostic Innovations in Glaucoma Study (DIGS) and again with another independent cohort from the EGPS. The OHTS and the EGPS also found (level II) that visual field loss was the first endpoint reached by converting eyes in 35% and 60% of the eyes, respectively. The EMGT, a level I study comparing treatment to observation in cases of early glaucoma, also found in a level II post-hoc analysis that change in visual fields was the first endpoint reached in 86% of eyes. There were differences between the two OHT studies and the EMGT in the demographics and disease state of the cohorts, diagnostic criteria, severity of disease, and photographic methodology that may have lead to the higher percentage in EMGT, but the finding that visual field loss is an important component in the diagnosis and management of glaucoma was evident in all.

STANDARD AUTOMATED PERIMETRY (SAP)

Although SAP is the standard used in the previous clinical trials to assess visual function, there are some limitations to SAP. SAP evaluates differential light sensitivity using a small (0.47 degree) white flash (200 msec) on a dim (31.5 asb) white background. Because all the primary retinal ganglion cell types responsible for vision respond to this stimulus, SAP is a non-selective test. Due to the inherent redundancy of the visual system, SAP may not provide adequate sensitivity to detect early glaucomatous changes. In some patients, a significant amount of ganglion cell loss has occurred (25–50%) before SAP can detect functional deficits. Another concern with SAP, and all visual function tests described here, is its high test-retest variability, particularly in regions of visual field deficits, making it difficult to assess whether the visual field is worsening on serial examination. For example in the OHTS, the majority of initial visual field abnormalities detected by SAP were not confirmed on repeat visual field testing and the study endpoint was reset in year 2 to require three consecutive abnormal results rather than the initially proposed two. Similarly, the EMGT study required progression on three consecutive fields to reach endpoint. Many factors influence the variability of visual field results, such as patient performance and reliability, fixation losses, fatigue, learning effects, changes in pupil size, improper refractive correction, and true physiological variability. In a multifactor model of variability in SAP or SWAP fields,
Blumenthal and colleagues found the three most important contributors were defect severity, location of defect, and patient's diagnosis. However, all factors together accounted for only one-third of the variability found. Some improvements in both test time and the variability of SAP have been made by applying the Swedish interactive thresholding algorithm (SITA), a strategy that significantly decreases testing time (to 4–5 min) without compromising accuracy for detecting visual field defects as compared to much longer full-threshold testing. Although data from level I clinical trials with SAP-SITA are not yet available, SAP-SITA has become the standard for clinical use with the Humphrey Visual Field Analyzer (HFA; Carl Zeiss Meditec, Inc.). Additionally, it is the only visual field test that offers an analysis for identifying progression of existing defect, the guided progression analysis (GPA) on the HFA (Carl Zeiss Meditec, Inc.), which is based on the progression analysis developed for the EMGT. SAP is also available on the Octopus perimeter (Haag-Streit, Koeniz, Switzerland).

Psychophysical tests of specific visual functions have been developed to measure visual performance and to understand the underlying glaucomatous changes in retinal ganglion cell function. Whereas SAP is non-selective, each visual function specific perimetric test attempts to isolate a sub-population of ganglion cells by evaluating a specific visual function characteristically processed by that cell subtype.

**SHORT-WAVELENGTH AUTOMATED PERIMETRY (SWAP)**

As an example, SWAP requires detection by the short-wavelength cones. The stimulus information is then processed through the blue-yellow (aka small bistratified) ganglion cells that project their axons to the koniocellular (interlaminar) layers of the lateral geniculate nucleus of the thalamus. SWAP is commercially available in both Humphrey and Octopus perimeters, and the test is used in several clinical studies cited throughout this review. SWAP utilizes a 440-nm narrow band 1.8-degree target at 200-msec duration on a bright 100 cd/m² yellow background. The target locations and thresholding procedures are identical to SAP now that the SITA-version of the test is very recently available. SWAP-SITA takes about 4 minutes on average.

The amount of isolation is unknown for the other function specific tests, but we know that SWAP provides isolation of about 15 dB. This means the blue-yellow ganglion cell system would have to lose 15 dB of sensitivity before another cell type could assist in responding to the SWAP stimulus.

It should also be noted that SWAP has limitations. Studies have shown higher test–retest variability with SWAP as compared to SAP, however, the use of SWAP-SITA decreases testing time and inter-test variability, although further studies are required to elucidate the diagnostic performance of SWAP-SITA. It is important to explain the appearance of the SWAP target and to give practice before performing the patient’s first SWAP examination. An additional disadvantage of SWAP is that testing results appear to be more heavily influenced by ocular media opacities than SAP, complicating the diagnosis particularly in elderly patients.

Despite the disadvantages, early studies found SWAP-full threshold (FT) to have a higher sensitivity for early glaucomatous damage than SAP, with the ability to detect visual field loss 3–5 years earlier than SAP. For example, one study compared glaucomatous structural damage to function measured with SAP and SWAP on 479 eyes followed for several years. All had normal SAP fields at baseline with 17.5% developing confirmed visual field loss on follow-up. It was noted that 75–80% of these conversions had baseline glaucomatous optic disk damage. Twelve percent of the convert eyes also had confirmed SWAP deficits at baseline, and an additional 8% developed confirmed SWAP defects on follow-up. Although participants in the OHTS had normal SAP visual fields and normal-appearing optic disks at baseline, the SWAP ancillary arm found 21% had SWAP abnormalities (Johnson CA, et al: Short-wavelength automated perimetry [SWAP] in the Ocular Hypertension Treatment Study [OHTS]. Investigative Ophthalmology and Visual Science 43 [ARVO abstract]: #2138, S86, 2002). A study comparing SAP-FT, SWAP-FT, FDT-N30, and motion automated perimetry (MAP) in the same eyes with early signs of GON found that SAP identified only 46% of the eyes as abnormal whereas FDT perimeter identified 70%, SWAP showed 61%, and MAP 52%. This study also showed that when more than one visual function is affected the loss of function occurs in the same retinal area for all and that the first measurable loss of function in an individual can occur for any of the retinal ganglion cell (RGC) subtypes. The finding, which suggests that glaucoma damage is non-selective for RGC type has been supported by recent work in the lateral geniculate nucleus with other psychophysics and with more recent inter-function comparison studies.

These findings suggest the premise that both structure and function are important for following patients, and that a percentage of these eyes may have been found earlier with SWAP. However, there is a problem with the early selective function studies using both SWAP and FDT (see discussion at end of the function section).
FREQUENCY-DOUBLING TECHNOLOGY PERIMETRY (FDT)

FDT is based on the frequency-doubling illusion, which occurs when viewing a grating with a low spatial frequency and a high temporal rate.67,83 This test is thought to measure the magnocellular retinal ganglion cells (about 10%) of the population. FDT is measured on a commercially available device, the Matrix perimeter, developed by Welch-Allyn (Skaneateles, NY) and marketed by Carl Zeiss Meditec, Inc. Contrast thresholds are obtained. The C-20 or N-30 versions have 17 or 19 targets, respectively, of 10-degree visual angle undergoing 25-Hz counterphase flicker. Thresholding is done with a modified binary search technique. Test time is approximately 5 minutes. Most studies have used these versions of the test.

More recently a 24-2 version using the zippy estimation of sequential testing (ZEST) thresholding technique at 54 locations within the central 24 degrees of the visual field,66 plus one foveal location, has been developed.133,134 The stimuli subtend 5 degrees of visual angle and comprise gratings whose contrasts are modulated sinusoidally with a spatial frequency of 0.5 cycles per degree undergoing 18-Hz counterphase flicker. Test time also averages about 5 minutes.

Advantages of FDT are a lower test–retest variability compared to SAP22 and SWAP. Disadvantages include unreliable FDT testing results seen in patients with age-related and posterior subcapsular cataracts.24

Early evidence has shown the test may be more sensitive to early glaucomatous defect than is standard perimetry.17,23,25,61,66,91,116,117,145 FDT has been used for screening for glaucoma more often than other visual field test types and results support its usefulness.53,85 Results obtained with Matrix perimetry have been shown to correlate highly with FDT perimetry,113 suggesting that Matrix perimetry can be used similarly to FDT-N30 perimetry for detection of early glaucomatous visual field loss with the advantage of better spatial comparison with SAP.

LIMITATIONS TO EARLY STUDIES

When comparing two or more visual field tests, visual fields should not be used to classify the study participants. The main limitation to inter-function comparison studies is that many use SAP either to classify the subjects included in the study or as the gold standard against which the other tests are compared. This assumes that SAP is best and no other test will ever perform as well. Conversely, when patients are selected as normal on SAP and then the percentage of those found abnormal on the visual function specific tests is computed, SAP suffers by comparison.

Because there is no true gold standard for glaucoma diagnosis, the presence of glaucomatous optic neuropathy or even more stringently, progressive optic neuropathy, have been suggested as good surrogate standards for use when comparing functional tests.

An additional limitation is caused by improvements in technology. To our knowledge, other than the two articles by Sakata115 and Racette113 discussed in the following section, there are no published level II studies comparing the most recent versions of these tests, SAP-SITA, SWAP-SITA, and FDT Matrix 24-2, although there are several clinical studies underway to do so.

INTERFUNCTION COMPARISONS (ALL LEVEL II EVIDENCE)

Studies that have used a non-function gold standard are few, but they are increasing. Sample and colleagues used both the presence of glaucomatous optic neuropathy and progressive optic neuropathy as two gold standards for glaucoma when comparing SAP-SITA, SWAP-FT, FDT-N30, and high-pass resolution perimetry.120 This study verified earlier findings that glaucoma does not selectively affect one ganglion cell subtype first, and when abnormality is found on more than one test type, the area of the retina affected is the same. No significant differences were found in receiver operating characteristic (ROC) areas for SAP, SWAP, and FDT. Other level II studies using GON to classify eyes agree with these findings,22,115 whereas a study by Racette et al suggests that the Matrix FDT 24-2 test was somewhat better than SAP-SITA at discriminating between healthy and glaucomatous eyes.113 Others found the prototype FDT Matrix 24-2 and SWAP-FT were both useful for early detection; SAP was not compared in this study.129

Another study addressed the gold-standard issue in a different way.120 Separate evaluation of SWAP-FT and FDT-N30 parameters and of structural OCT and SLP parameters was done using two different gold standards, one based on optic disk appearance (consensus masked review of stereophotographs) and one based on SAP fields. Results found that the most sensitive FDT parameters tended to be more sensitive than SWAP parameters at set specificities. When optic disk appearance was used as the gold standard, ROC areas for FDT and SWAP were 0.88 and 0.78, respectively. When SAP was used as the gold standard, the ROC areas were 0.87 and 0.76, respectively. Structural measures based on OCT were more sensitive than SWAP measures.120
Studies have also found that a combination of test types may be beneficial for increasing sensitivity to early damage without significant drops in specificity. For example, SAP-SITA with either SWAP-FT or FDT-N30,120 or SAP-SITA with Matrix 24-2,115 FDT, and SWAP were better together than either alone.59

FUNCTION SUMMARY
In summary, more study is needed to determine the clinical utility of the newest versions of each test for diagnosis and management of glaucoma. However, some very important findings have come from the studies done to date:

1. Glaucoma causes loss of all three primary retinal ganglion cell subtypes, the parvocellular, magnocellular, and small-bistratified.
2. There are individual differences in which test will first identify loss of vision, both among the different functional tests and across structure and function.
3. Glaucoma affects the same area of the retina first when two or more visual field tests show abnormality.
4. The newer, faster thresholding algorithms correlate well with older versions of each test and have the advantages of decreased testing time and somewhat improved variability.
5. Repeatable results are necessary to confirm the diagnosis of glaucoma based on visual fields. Results can be repeated within a test or across tests looking for evidence of damage in the same area of the retina.

The Relationship Between Structural and Functional Damage for Glaucoma Detection

INTRODUCTION
Understanding the relationship between structural and functional changes can lead to improvements in glaucoma detection and the management of individual patients, and advance our understanding of the nature of glaucomatous progression. This section will provide an update on what is known about the relationship between structural and functional changes in glaucoma, and how this information can be used to diagnose and monitor glaucoma patients. A summary of the recent literature on the strength or shape (linear versus non-linear) of the correlation between quantitative measures of structure and function is beyond the scope of this review.

There have been several level II evidence cross-sectional studies and a limited number of level I and level II longitudinal studies examining the relationship between structural and functional damage for glaucoma diagnosis. Cross-sectional studies have compared the diagnostic accuracy of structural and functional tests alone and in combination. Longitudinal studies have estimated the predictive accuracy of baseline structural and functional measures for future glaucomatous change and assessed the detection of structural and functional change over time. Comparison of results across studies is difficult because the relationship between structural and functional damage is influenced by many factors. The definition of glaucoma, the severity of disease and study inclusion criteria, the type of structural and functional assessment utilized, and the measurement scales and parameters used in the analysis each can influence the association between structural and functional measurements. Inconsistencies in study results can be explained, at least in part by differences in these study design factors. Of particular importance for comparing structural and functional tests is the unavoidable problem of how glaucoma is defined. If glaucoma is defined based on glaucomatous visual field damage alone, then functional measures may show better diagnostic accuracy than structural measures, because they are measuring aspects of glaucoma (visual field damage) close to the criteria used to define the disease. Similarly, when the definition of glaucoma is based on structural measures alone, then structural tests may show biased and higher diagnostic accuracy. As discussed previously, many studies now try to use definitions of glaucoma that are as unrelated as possible to the test being evaluated. The question of how to define glaucoma is particularly problematic when comparing the diagnostic accuracy of structural and functional tests, as the definition used will either be biased toward one type of test, or if both glaucomatous optic disk and visual field damage are required, then it is less likely that the glaucoma patients included are at an early stage of the disease. An imperfect solution to reduce this possibility of biased estimates is to use two definitions of glaucoma, one based on structure, the other based on function, and compare their results to the two definitions14,128 or require both.87 An additional, albeit imperfect solution using the statistical technique of latent class or latent variable analysis may offer an approach to evaluation of glaucomatous damage or progression using multiple different types of observations, without the requirement for an a priori gold standard.43

CROSS-SECTIONAL STUDIES
Few level II cross-sectional studies have directly compared the diagnostic accuracy of structural and
functional tests or examined whether combining results from structural and functional tests improve early glaucoma detection. Using earlier generation instruments, Bowd et al reported that areas under the ROC curve tended to be highest for OCT1 (AUROC = 0.89 and 0.91) and FDT N-30 (0.88 and 0.87) parameters, followed by GDx Nerve Fiber Analyzer (0.79 and 0.81) and SWAP-FT (0.78 and 0.76), regardless of the whether glaucoma was defined based on SAP visual field or stereophotographic based optic disk damage, respectively.14 Moreover, there was only modest agreement among structural and functional tests on which glaucoma patients were identified. Using recent generation instruments, Hong et al reported that the diagnostic accuracy of the best parameters for FDT Matrix (AUROC = 0.99), and SLP-VCC (AUROC = 0.91), were better than Stratus OCT (AUROC = 0.79), and semi-quantitative assessment RNFL photography (AUROC = 0.75) for differentiating between eyes with early glaucoma (SAP MD = 2.2 ± 1.1 dB).56

Shah et al, Mardin et al, and Hong et al reported that combining functional tests with structural tests can improve the diagnostic accuracy over using one test alone.56,87,128 Specifically, Mardin et al used several machine-learning classifiers, including a linear separation (stabilized linear discriminant analysis), a tree-based classifier called bagging,86 and a combination of these two methods called double-bagging, to combine HRTII rim area, volume, and cup-shape measurements with Octopus visual field indices to differentiate between normal eyes and eyes with moderate glaucoma. Moderate glaucoma was defined as having glucomatous appearance of the optic disk, followed by the presence of glucomatous visual function defects (Octopus MD = 7.1 ± 4.8 dB). The diagnostic accuracy of combining functional and HRT indices was maximized at a sensitivity of 95% and a specificity of 91% using the double-bagging approach.87 Shah et al reported sensitivities and specificities for two definitions of glaucoma, one based on stereophotograph assessment of glucomatous optic disk and the other based on SAP damage. For HRT, optic disk, GDx, and OCT RNFL measurements, comparison to each instruments’ normative database was used to identify values outside normal limits. Pattern standard deviation (PSD) outside 99% limits was used to identify FDT damage, and PSD or glaucoma hemifield test (GHT) outside normal limits was used to identify SWAP damage. Sensitivity and specificities for glucomatous based on visual field damage for GDx-VCC, Stratus OCT, HRTII, FDT N-30, and SWAP full threshold were 41.9% and 98.3%, 58.1% and 98.3%, 58.1% and 84.5%, 44.2% and 98.3%, and 65.1% and 86.2%, respectively. Adding FDT N-30 to each of the best structural parameters led to a significant (p < 0.05) increase in sensitivity without a significant change in specificity compared with structural parameters alone. In contrast, adding SWAP-FT to each of the best structural parameters led to a significant increase in sensitivity and also a significant decrease in specificity compared with each structural parameter alone.128 Similar results were reported when glucomatous optic disk damage was used as the definition of glaucoma. In addition, the change in likelihood ratios after combining FDT with OCT, GDx, and HRT suggest that adding FDT may be more helpful for excluding a diagnosis of glaucoma, than confirming it. Hong et al compared the sensitivity and specificity of FDT Matrix, SLP-VCC, OCT, and RNFL photography and found that the combination of FDT Matrix (more than five points depressed below 5% level of pattern deviation plot) and SLP-VCC nerve fiber indicator may offer the best criteria for early glaucoma detection (SAP MD = 2.2 ± 1.1 dB), as the specificity increased from 92.5% to 100%, whereas the sensitivity remained the same (90%) compared to FDT Matrix alone.56

LONGITUDINAL STUDIES: DOES STRUCTURAL DAMAGE PRECEDE FUNCTIONAL DAMAGE IN GLAUCOMA?

There have been a limited number of level I and level II longitudinal studies that include both structural and functional measures of glucomatous change.3,50,71,102,108,130,143 The OHTS and EGPS (both level I evidence) were designed to evaluate the effect of ocular hypotensive treatment on the development of glaucoma in participants with ocular hypertension. The OHTS and EGPS showed that, in many eyes, repeatable structural defects (assessed using standardized, masked qualitative assessment of stereophotographs) are detectable before repeatable SAP functional defects (55% and 40% respectively), whereas in some eyes, functional defects are detected first (35% and 60%, respectively). Only 10% of eyes in the OHTS and none in the EGPS showed structural and functional change at the same time.71,102 Using very different techniques for detection of optic disk changes than the OHTS and EGPS, the EMGT (level I evidence) utilized flicker chronoscopy and side-by-side comparisons of non-stereoscopic fundus photographs for detecting change on photographs and glaucoma change probability (based on PD changes) for identifying visual field changes, and reported progression by visual function only in 86%, by optic disk change in 1% and by both structural and functional change in 13% of progressing glaucoma patients.50
Level II observational cohort studies have also shown that in most eyes (17–60%), glaucomatous structural change using a variety of photography and imaging instruments is detected before functional change based on a variety of functional tests. However, these same studies reported that in a substantial proportion (18–51%) of eyes, functional change was detected before structural change. A major point of agreement in these studies, regardless of whether structural assessment is completed using photographs, HRT, OCT, or GDx, and functional assessment based on SAP or SWAP, is that very few eyes (3–24%) have structural and functional changes that are detected at the same time. In a very thorough analysis of the issue using SAP, high-pass resolution perimetry, and HRT topographic change analysis (level II evidence), Artes and Chauhan demonstrated that these indicators of structural and functional change provide largely independent measures of progression, and they suggest a new technique evidence of change analysis to facilitate comparison across tests. Specifically, utilizing three criteria of change (liberal, intermediate and conservative) for each test, they showed rather limited (<24%) agreement between SAP and HRT for detecting change after an average of 7 years of follow-up. There are several possible explanations for why in some eyes, structural change is detected first, whereas in others, functional change is detected first. Study design factors mentioned previously include the definition of glaucomatous change, the instrument and measurement scale (linear or non-linear), the variability of the measurements, and the frequency of testing. Moreover, it is possible that diseased retinal ganglion cells begin to malfunction before dying, resulting in reduced visual sensitivity without a detectable structural loss. It is also likely that ganglion cell loss precedes detectable visual sensitivity loss in areas of the visual field where redundancy of ganglion cells is high. Moreover, reproducibility of the measurements varies by individual, type of test, and stage of disease so that the test with the lowest variability (structural or functional) may show the first sign of glaucomatous change. There also may be clinical, demographic, or ocular characteristics that can help predict whether structural or functional changes will be the first sign of glaucoma.

Hood and Kardon have elegantly outlined a formal framework to better understand what is meant by “Does structural damage precede functional damage in glaucoma?” In short they argue that it is important to differentiate between 1) the statistical correlation and the theoretical mathematical relationship and 2) whether retinal ganglion cells or structural tests (as surrogates) are being measured. Moreover, they provide strong evidence that which test detects glaucoma first depends in large part on the standard deviation of the measurement in relation to the normative data to which they are being compared. The model suggests that structural tests can show statistical significance before SAP loss because the normal confidence limits are greater for the functional test compared to the structural test. When SAP damage is detected first, it may be that the patient started out with more structure to begin with, and are still within the normal limits of the structural test utilized.

LONGITUDINAL STUDIES—PREDICTION

Longitudinal studies have consistently identified both baseline structural and functional factors that predict the development of glaucomatous change in ocular hypertensive and glaucoma suspect eyes. This is important because economic analyses from OHTS suggested that ocular hypotensive treatment is cost-effective in the subgroup of ocular hypertensive individuals with an IOP ≥ 24 mm Hg and an annual risk of POAG ≥ 2%. Unfortunately, it is difficult to assess the annual risk of developing POAG in an individual patient. Prediction models, and in particular, risk calculators can provide this information using baseline demographic and clinical and ocular measurements. The OHTS and EGPS studies report that baseline older age, higher IOP, thinner central corneas, worse SAP pattern deviation, and larger stereophotograph-based C/D ratios were predictive of the development of glaucomatous optic disk and visual field changes in ocular hypertensive participants with normal-appearing optic disks and visual fields at study entry. A joint analysis of the OHTS and EGPS untreated arm showed that the predictive factors for developing POAG, including both SAP PD and stereophotograph C/D ratio as well as other factors, were remarkably similar in the two studies. The data from the untreated arm from the two studies were then merged to produce a new risk model that was then developed into an online risk calculator that clinicians can use to assess the risk of the development of POAG for an individual ocular hypertensive patient. Medeiros et al has demonstrated that using HRT linear C/D ratio measurements can be used interchangeably with the stereophotographic-based C/D ratio measurements in the original OHTS risk model.

Several studies (level II evidence) have found also that baseline optical imaging structural (OCT) and HRT, GDx-VCC, and GDx Nerve Fiber Analyzer measurements are predictive of the development of glaucoma.
glaucomatous visual field or optic disk changes in ocular hypertensive or glaucoma suspect eyes. The largest study of ocular hypertensive eyes, the Confocal Scanning Laser Ophthalmoscopy Ancillary Study to the OHTS, reported that several HRT baseline stereometric measurements were predictive of the development of a repeatable optic disk or visual field endpoint. Specifically, baseline MRA results outside of normal limits in any sector had a 2.4-times greater chance of developing an endpoint than eyes within normal MRA limits at baseline. It is important to note that most subjects with baseline MRA outside of normal limits did not develop glaucoma during the follow-up period studied. Specifically, in the OHTS the positive predictive value (the probability of developing glaucoma during follow-up with a baseline MRA outside of normal limits) ranged from 14% to 40%, depending on the MRA sector selected. The negative predictive value (probability of not developing glaucoma with a baseline MRA within normal limits or borderline), on the other hand, was as high as 93%.

Conclusions

Glucoma is a disease that requires a clinical diagnosis. There is no gold standard for the presence or progression of the disease. Because of the lack of a definitive measure, it is difficult to have certainty with regard to the relative sensitivities and specificities of either our current structural or functional techniques. Clearly, advances in functional and structural evaluation techniques provide more objective documentation and precision for diagnosis and progression detection than the more subjective and coarse methods of the past.

Structural imaging technologies have proven in level II studies to be at least as good as stereoscopic disc photography read by expert observers in the discrimination between health and glaucoma. This indicates that structural imaging technologies provide clinicians at all levels with the ability to assess the optic nerve and RNFL in a standardized, objective and quantitative fashion at the level of an expert observer. Correspondence between structural and functional assessment allows higher certainty for the health or glaucomatous status or for stability or glaucoma progression. Unfortunately, both cross-sectional and longitudinal studies have shown that the correlation between detectable structural and functional damage and change in early glaucoma is at best modest. It is important to include both structural and functional examinations for assessment of glaucoma at each stage of the disease.

The evaluation of progression is plainly most mature for function as measured by SAP and for structure as assessed by CSLO. This is a rapidly evolving field for the development of both structural and functional technologies. We are likely to see further technological advances to permit earlier detection of disease and its progression with higher levels of certainty than currently available.

Method of Literature Search

The majority of the articles for this review were identified by searching the Medline database, years 1950–2007, using the following key words: glaucoma, diagnosis, diagnostic techniques, ophthalmological, ophthalmoscopy, optical coherence tomography, visual fields, perimetry. Other articles were identified from the references of the articles in the Medline search. Relevant articles written in languages other than English were included only when an English abstract was available. A few select articles published before 1992 are included for historical purposes, but the review is based mainly on clinical trials published in the past 15 years. Case reports were not included.

References


37. Fingeret M, Medeiros FA, Susanna R, et al: Five rules to evaluate the optic disc and retinal nerve fiber layer for glaucoma. Optometry 76:661–8, 2005


GLAUCOMA DETECTION AND MANAGEMENT


