Screening Considerations

Screening for open-angle glaucoma (OAG) has two potential objectives:

- to identify people with manifest OAG so that affected individuals can be treated before becoming visually impaired, and
- to identify people with ocular hypertension (OH), the major risk factor for OAG, so that these individuals can be treated and thus reduce their risk of developing OAG.

Although current screening programs frequently combine these two objectives, they are important to distinguish for three reasons. First, the accuracy of a screening test at correctly identifying people depends on which of these characteristics is the purpose of the screen. (See box A for a description of the basic components of screening test accuracy). A test that is very good at identifying people with OH may be very poor at identifying people with manifest OAG, and vice versa. Second, OH is much more prevalent in the population than OAG, and the yield of an OH test (i.e., the proportion of all positives who are true positives) is thus likely to be much higher. And third, the potential costs and medical benefits of screening depend on which groups of people are identified. This last consideration is discussed further in chapter 5.

Description of Screening Technologies

OTA reviewed three different screening technologies that have been used, alone or in combination, in large-scale screening for OH and OAG. These are:

- *tonometry*, which measures intraocular pressure (IOP);
- *ophthalmoscopy*, which enables the examiner to see abnormalities in the optic disc; and
- *perimetry*, which measures the extent of visual field loss.

Tonometry is the most familiar of the three screening methods. It may be performed by physicians, optometrists, or (less commonly) by opticians.[1] Tonometers work by measuring the resistance of the eye to a force, which may be applied by direct contact with the eye or by shooting a puff of air (non-contact tonometry). Tonometry is the method most often used in large screening programs (78,93) and is often part of routine visits to eye care professionals.

Ophthalmoscopy requires the analysis of the appearance of the surface of the optic nerve by a physician or optometrist using an ophthalmoscope (a tool that enables the examiner to look through the pupil at the retina). This procedure can identify characteristics of the optic disc indicative of OAG. It can be performed by trained non-ophthalmologic physicians (e.g., family practitioners).[2] It has been used in community screening efforts in the United States, but it is usually used in conjunction with tonometry (78). Its use in this role has been promoted in at least one textbook on glaucoma (67).

In contrast to tonometry and ophthalmoscopy, *perimetry* identifies actual visual loss. In perimetry, dots of light of varying brightness are introduced at a pattern of points in the visual field. The patients' responses to these stimuli are recorded by the perimetrist (in manual perimetry) or by a computer (in automated perimetry). A lack of response (i.e., the patients' inability to see the light) indicates a blind spot in the visual field. (When the visual field defect is less severe, the patient may see the light but only if it is very bright.)

Perimetry can be time-consuming and expensive and has only rarely been used as a screening tool (63). It is most commonly

1 Opticians can perform only non-contact tonometry.

2 However, not all family practice residencies require some ophthalmologic training, and only two-thirds include routine glaucoma screening as a part of the care residents must provide (113).

Box A--- Components of Screening Test Accuracy

Two basic attributes are used to compare the accuracy of a screening test in identifying patients for followup: the sensitivity of the test (expressed as the proportion of people with the condition--i.e., OH or OAG--who actually test positive) and the specificity of the test (expressed as the proportion without the condition who actually test negative). Sensitivity and specificity do not depend on the prevalence of OH or OAG in the population.

Sensitivity and specificity often vary inversely. A test that is very sensitive (i.e., identifies most of the people with the disease) often also falsely identifies many people who actually do not have the disease, giving it a low specificity. However, this is not always the case. Some tests are both more sensitive and more specific than others. And, an inexperienced examiner may cause a test performed by that examiner to be both less sensitive and less specific than the same test performed by a more experienced examiner. Sensitivity is generally considered more important in OAG screening than specificity, since false positives (people falsely identified as having glaucoma) can be eliminated in the diagnostic workup, but false negatives (people falsely identified as not having glaucoma) are not referred and thus are not diagnosed. However, if specificity is very low, a program can incur substantial followup costs, lead to unnecessary treatment, and cause much distress for people who are disease-free.

Two other attributes used to compare screening tests are the positive and negative predictive values of a test, the ability of the test to correctly predict disease or health. The predictive value of a positive test is the proportion of all test-positives who actually have the disease, while the negative predictive value is the proportion of all test-negatives who do not. Thus, a positive predictive value of 5 percent means that of 100 persons who test positive, 5 have the disease.

Unlike sensitivity and specificity, predictive values do depend on the prevalence of the condition. "Given a fixed level of sensitivity and specificity, predictive values increase as the prevalence increases...[S]creening will lead to a large number of overreferrals if carried out in a population where the disease is rare; conversely, false-positives are greatly reduced when screening is done in population where the disease is common" (71). When a condition is very rare, even a very sensitive and specific test may have a modest positive predictive value.

Since true positives are a higher proportion of total positives when a condition is common in the population, the cost per case detected is lower than would be the case if a large number of false positives incurred followup costs. Thus, high prevalence is one factor leading to low cost per case detected through a screening program. (Other factors can also lead to low costs per case detected by compensating somewhat for large numbers of false positives referred for followup--for example, low followup visit costs.)

used, in conjunction with ophthalmoscopy, as a diagnostic procedure (e. g., to confirm that a person testing positive on tonometry, due to high IOP, actually has OAG). Visual field defects detected by perimetry are used as the "gold standard" definition of OAG when evaluating the accuracy of tonometry and ophthalmoscopy in identifying people with OAG.

Screening Test Accuracy[3,4]

Although glaucoma screening has been common for many years, there have been few rigorous studies of the accuracy of screening tests in correctly identifying either OH or manifest OAG. Table 8 summarizes relevant estimates of the accuracy of tonometry, ophthalmoscopy, and perimetry as screening tools for detecting OH (by tonometry) and manifest OAG (for all three technologies). Most notable in this table is the scarcity of estimates available and the enormous variation among the estimates that do exist. Particularly lacking are methodical studies of variation in accuracy among different types of examiners. The estimates that do exist, and the characteristics of the respective technologies that may affect accuracy, are described in greater detail below.

3 "Accuracy" is used here in its more general sense, the ability to identify something correctly, rather than in its strict statistical meaning.

4 These accuracies generally apply to the over-40 population. The diagnostic accuracy of the various screening technologies has not been reported in the literature for the elderly alone (over age 65).

Table 8--- A Comparison of Estimates of Accuracy for Three Glaucoma Screening Technologies

Technology	Sensitivity	Specificity	Setting/context	Source
Tonometry >21 mm Hg				
accuracy compared to elevated pressure on Goldmann tonometry	71%	97%	study of Schiotz tonometry	Bengtsson, 1972
accuracy compared to "glaucoma", defined in various ways	75%	81%	study in hospital clinic	Packer et al., 1965
accuracy compared to confirmed visual field defects	50%	..	population survey	Leske et al., 1982
	72%	30%	mass screening study	Ford et al., 1982
Ophthalmoscopy				
accuracy in live eyes compared to confirmed visual field defects	72%	64%	mass screening study	Ford et at., 1982
	76%	..	population survey	Leske et al., 1982
	44- 53%[a]	69- 77%[b]	study of examiner accuracy	Wood and Bosquanet, 1987
	84%	97%	study of ophthalmoscopy accuracy using expert examiners	Hoskins and Gelber, 1975
	100%	..	population survey with expert examiner	Graham, 1969
Perimetry				
accuracy compared to confirmed visual field defects	96%	89%	manual perimetry in population survey	Rock et al., 1972
	93%	88%	study of automated perimetry	Sommer et al., 1987
	92%	46%	automated perimetry in mass screening study	Ford et al., 1982

[a] 44% for consultants, 53% for junior doctors.
[b] 69% for junior doctors, 77% for consultants.

SOURCES: See references

Tonometry

Tonometry may be used for either of the two different screening purposes: to identify people with OH who, due to their high IOPs, are at high risk of developing **OAG;** and to identify people with manifest OAG. It is fairly successful at identifying people with OH, since IOP is the characteristic it is designed to measure. It is much less accurate in identifying people with manifest OAG, since only a small proportion of people with elevated IOP also have manifest OAG (and, conversely, a substantial minority of people with OAG do not have elevated TOPS at the screening visit).

Screening for OH. Since high IOP--the definitive characteristic of OH--is the quality measured by a tonometer, tonometry might be expected to be quite accurate in identifying people with OH. Although the potential accuracy is high, in practice a number of factors can have a substantial impact on the number of individuals in whom OH is correctly identified.

These factors fall into two categories: those associated with the technology, and those associated with the nature of IOP itself. Technology-associated factors that affect the accuracy of tonometry include the type of tonometer and the person performing the test. Of the three main types of tonometers (Schiotz, applanation, and non-contact), applanation tonometry is generally considered the most accurate (128), although all have been used for screening (56,78,93). As might be expected, tonometry is more accurate when performed by someone who does it frequently, yielding fewer false positive readings (108).

Even when applanation tonometry is performed by skilled examiners, however, factors associated with the intrinsic nature of IOP affect tonometric accuracy in identifying OH. For example, there is substantial daily variation in IOP, and two tonometry measurements of the same person during different times of the day can lead to quite different conclusions (4).

The net result of these factors, combined with random errors, is an accuracy far below the ideal. A 1968 review of studies of the prevalence of OH found five studies that reported both the number of people with OH found at the screening visit and the number in whom OH was confirmed at a followup v is it. In these studies, the proportion of people with unconfirmed OH--i. e., false positives --ranged from 27 to 86 percent of the people referred. Sensitivity and specificity could not be calculated, since people testing negative were not retested (96).

A Swedish study attempted to estimate false negatives as well as false positives. In this study, readings on a Schiotz tonometer were compared with readings with an applanation tonometer (considered the "gold standard" for the purposes of this study). The investigators found that the Schiotz readings had a 71 percent sensitivity and 98 percent specificity for identifying high IOP (11).

The accuracy of modern applanation tonometry in screening for OH remains unclear. Although it is considered more accurate than Schiotz tonometry, it still produces false positives and negatives due to factors such as interexaminer variation and the variation of IOP with the time of day. Thorburn found that when two examiners performed separate applanation tonometry tests, the readings differed by 2 mm Hg or more in 40 percent of the paired measurements(lll). The variation did not depend on the level of IOP. Armaly found that approximately 30 percent fewer eyes had IOPs of 20 mm Hg or greater in the afternoon than in the morning (4).

Screening for manifest glaucoma. Although most people with OAG have elevated IOP, tonometry alone cannot distinguish between people suffering optic nerve damage and those with equivalent IOPs who are not. Thus, it has a naturally high false positive rate for manifest OAG (and a low specificity). This rate can only be reduced by raising the IOP designated as the cutoff criterion for referral (which, in turn, raises the false negative rate and decreases sensitivity).

In an early study of the accuracy of tonometry, Packer et al. found that tonometry was 75 percent sensitive and 81 percent specific at identifying persons with glaucoma at a cutoff level of 22 mm Hg (90). These results are not directly applicable to identifying manifest OAG, since the definition of "glaucoma" used by the investigators in this study was quite broad and included many characteristics not necessarily indicative of OAG (e.g., people with narrow angles, abnormal disks, or certain results on other tests were considered to have glaucoma for the purposes of the study (89)). However, the study does provide a good example of the effect on test accuracy of changing the cutoff level for referral. The investigators found that raising the screening cutoff level from 22 mm Hg to 26 mm Hg decreased sensitivity from 75 to 59 percent, while specificity increased from 81 to 95 percent (90).[5]

Two more recent studies of the accuracy of tonometry used glaucomatous visual field defects to define OAG. In a community screening study in New Orleans, tonometry demonstrated a 72 percent sensitivity for OAG by this definition, but only a 30 percent specificity (cutoff level 22 mm Hg) (35).[6] Leske et al., in a study of the Framingham population found that tonometry (cutoff level 22 mm Hg) had a 50 percent sensitivity for detecting OAG in glaucoma suspects (i.e., the group of people testing positive on one or more OAG screening tests), implying a somewhat lower sensitivity for the entire screened population (75). Specificity could not be estimated, since people who tested negative were not rescreened.[7]

Ophthalmoscopy

Ophthalmoscopy can identify people in whom OAG has already caused visible nerve damage. It can identify people who have developed OAG earlier than can perimetry, since considerable nerve damage occurs before defects in the visual field become apparent (98). Its primary disadvantage is that analysis of the optic disc is highly subjective, and abnormalities can be very difficult to interpret. Consequently, it is difficult to be confident that a person reported to have an abnormal optic disc actually has OAG until visual field defects have also become apparent.

The sensitivity and specificity of ophthalmoscopy in identifying people with manifest OAG (i.e., visual field defects) depend on such factors as:

- whether the pupils are dilated for the procedure (less common in community screening settings, more common in office settings);
- the criteria used to define a positive test (e.g., a cup:disc ratio of .6, or an examiner's overall impression that the disc is "abnormal"); and
- the skill and experience of the examiner.

Ford et al. found the sensitivity of ophthalmoscopy for identifying manifest OAG to be 72 percent and specificity to be 64 percent in community mass screening (35). Leske et al. examined several specific test criteria and found that vertical optic cup:disc ratios gave the best results for the Framingham population; this criterion had a sensitivity of 76 percent (75).

---

5 The community screening programs organized by the National Society Prevent to Blindness usually use a cutoff of 24 mm Hg to keep the number of referrals manageable (95).

6 The results from the study by Ford and colleagues have been published only in abstract form. According to one of the investigators, all subjects screened subsequently underwent a comprehensive ophthalmic examination in order to determine the true disease status of the sujects and calculate the sensitivity and specificity of the screening procedures (130). A similar, larger study is now ongoing. It will be of interest to see if the results of the larger study confirm those of the original one, since the study cited above found surprisingly high sensitivity but surprisingly low specificity.

7 The positive predictive value of IOP in this study was 5 percent (i.e., of the people with IOP of 22 mm Hg or greater, 5 percent had visual field defects) (75). A large proportion of the people in this study were elderly.

Examiner skill is crucial to the accuracy of ophthalmoscopy in identifying manifest OAG. In studies in which glaucoma specialists performed ophthalmoscopy, or studied photographs of the optic nerve, accuracy is generally much higher than that reported in the two screening studies above. American investigators have reported a sensitivity of 84 percent and a specificity of 97 percent for ophthalmoscopy performed by glaucoma specialists (54). Studies in which photographs of optic discs were examined by specialists have found sensitivities of 71 to 88 percent and specificities of 75 to 97 percent (29,50,53,54). An English population survey to determine glaucoma prevalence found that all cases of manifest OAG identified in the survey were detectable by ophthalmoscopy (performed by a glaucoma specialist) (42).

One exception to the reports of high accuracy when ophthalmoscopy is performed by specialists is a recent English study of examiner skill. In this study, ophthalmoscopy performed by consultants was more specific for OAG--but less sensitive--than the same procedure performed by junior physicians (126).

Perimetry

Perimetry can identify people in whom actual visual field defects have occurred, but who are not yet visually impaired. It is for people with visual field defects that a diagnosis of OAG can be made with the most confidence. In part for this reason, this paper has considered OAG to be manifestly present only if visual field defects are among the signs of disease. However, it should be noted that the disease is established, and optic nerve damage is underway, before visual field defects occur. Thus perimetry will naturally be more accurate than ophthalmoscopy at detecting manifest OAG, but it detects the disease at a later stage.

The accuracy of perimetry for screening is determined by comparing screening results with multiple comprehensive perimetric examinations. Accuracy depends heavily on how much of the visual field is screened and the manner in which it is done. The accuracy of manual perimetry is more variable than that of automated perimetry, since it depends more on the skill and consistency of the person performing the test.

A method of large-scale screening by manual perimetry was developed by Armaly (4) and subsequently modified by Drance and colleagues (99). The modified test is reported to have a sensitivity for manifest OAG of 96 percent, a specificity of 89 percent, and a positive predictive value of 83 percent (99). These numbers are probably maximums; in general practice, the accuracy of manual perimetry could be much lower depending on how the examiner carried out the procedure.

Automated perimetry has some advantages over manual perimetry for mass screening (63), although it is not necessarily more accurate. It has been reported to have a sensitivity varying from 80 to 96 percent, depending on how extensive a test is performed; the main problem is a potentially high false positive rate (as high as 33 percent) if test conditions are not properly established (63).[8] The type of perimeter also affects accuracy.

Three studies demonstrate the variation in accuracy that may be found when automated perimetry is used to detect OAG. Under careful research conditions, Sommer et al. found automated perimetry to have a sensitivity and specificity of 93 and 88 percent, respectively (103). Ford et al., on the other hand, found that automated perimetry in a community screening program had a 92 percent sensitivity but only a 46 percent specificity (35). According to one of the authors in this latter study, the difficulty of many subjects in understanding how to respond correctly when undergoing perimetry for the first time led to a high false-positive rate in the community setting (130). Bengtsson and Krakau found that 3 percent of eyes in a careful mass screening program in Sweden tested positive on automated perimetry, but only half the positive eyes ac-

8 For example, false positives can be reduced by automatically retesting any point in the visual field not reported as seen by the patient on the first try.

tually had visual field defects that were glaucomatous or meriting medical attention (14). The authors did not retest people who initially tested normal and therefore could not report completely on the sensitivity and specificity of the test.