Chapter 1

Evidence-based eye examinations

David B. Elliott

CHAPTER OUTLINE

1.1 Several tests are widely used because of tradition and not evidence

Evidence-based optometry means integrating individual clinical expertise with the best currently available evidence from the research literature.[1] For the majority of primary eye care procedures described, the evidence base for when and how they should be measured is provided. This may be from clinical experience (i.e., clinical pearls), which could be the authors' own experience or supported by citations to clinical textbooks and articles or from research evidence. What should always be avoided is the use of examination procedures based on tradition, anecdotal evidence, or habit. Three procedures that are widely used in several countries but seem principally used because of tradition are discussed below.

1.1.1 Visual acuity should not be measured with Snellen charts

The principal reason for using a Snellen chart nowadays would appear to be tradition or habit. Visual acuity (VA) charts using the logMAR system are now widely available and are much superior to Snellen charts, providing VA measurements that are twice as repeatable and over three times more sensitive to interocular differences in VA (see section 3.1.1).[2,3] VA can still be recorded and reported to others in the more familiar Snellen format of 6/6, 20/20, or 1.0.

1.1.2 The von Graefe prism dissociation test does not deserve its widespread use

The faith in this test has been such that a publication in a leading international optometry journal pronounced it as the gold standard test of heterophoria assessment[4] and used it to assess the usefulness of the cover test. The gold standard in this area should be the cover test and not the von Graefe[5–8] as it is objective and not reliant on subject responses or subject to prism adaptation and subsequent studies have shown it to be far more repeatable than the von Graefe prism dissociation test, which has been shown to be unreliable.[5–8] The study should have used the cover test as the gold standard and they would then have reported the limitations of the von Graefe prism dissociation test. Several studies have now reported that the

modified Thorington test is much simpler, faster, and more reliable than the von Graefe prism dissociation test,[5–8] and if an additional test to the cover test is required, it should be the modified Thorington.

1.1.3 Lengthy tentative reading addition tests are not necessary

A final example of tests used because of tradition in phoropter-based refractions is the use of tentative reading addition tests of binocular cross-cylinder followed by negative and positive relative accommodation (NRA-PRA). These are lengthy tests given that the only study to have properly investigated the various tentative addition tests indicated that asking patients their age provided more accurate information than either of those two tests.[9] Using an ingenious study design, Hanlon et al.[9] examined patients who returned to an optometry practice because they were dissatisfied with the near vision in their new glasses. Each patient's reading addition was determined using four methods (age, binocular cross-cylinder, NRA-PRA balance, $^1/_2$ amplitude of accommodation). The review (recheck) examination also determined whether the near addition in the glasses the patient disliked was too low or too high. The percentage of adds for each tentative add test that gave the same result as the incorrect add or worse (higher than an improper add determined too high or lower than an improper add determined as too low) was calculated. They reported that the simplest and quickest test, asking the patient their age, accounted for the fewest errors (14%). The other techniques gave errors in 61% (binocular cross-cylinder), 46% (NRA-PRA), and 30% ($^1/_2$ amplitude) of cases.[9] This suggests that the tentative addition should simply be based on patient age. Subsequent research suggests that the tentative addition estimate can be further improved by considering both the patient's age and working distance and/or symptoms with their current near correction (see 4.13.1). There is little need for the lengthy binocular cross-cylinder followed by NRA-PRA.

1.2 Tips on reviewing the evidence base for clinical tests

Currently professional bodies provide clinical guidelines that are based on research evidence, and expert clinicians and researchers write review articles and books and give lectures, and this has been reported to be the preferred source of information for many optometrists.[10] Reviewing the research literature yourself should become more common in future years,[10,11] particularly for the literature pertaining to clinical procedures, although there are clear difficulties to overcome. Medicine has been using evidence-based practice for many more years than optometry, yet numerous primary care doctors still seldom practice it because of lack of time, difficulties in searching for, appraising, and applying evidence and preference for using guidelines provided by professional bodies.[12]

1.2.1 PubMed, Google Scholar, and international optometry research journals

If you wish to review the literature,[11] two very useful free access databases are PubMed (www.pubmed.com; provided by the US National Library of Medicine) and Google Scholar. They both include the abstracts or summaries of papers from all the main optometry and ophthalmology research journals. Questions from clinicians on optometric internet/e-mail discussion groups can often be fully answered by a quick PubMed or Google Scholar search that can provide a much stronger level of evidence than anecdotal suggestions from colleagues based on one or two patient encounters.

Full access to one or more of the international optometry research journals is provided by membership of various professional bodies: *Ophthalmic & Physiological Optics* (College of Optometrists, UK), *Optometry & Vision Science* (American Academy of Optometry), *Clinical & Experimental Optometry* (Optometry Australia, New Zealand Association of Optometrists, Hong Kong Society of Professional Optometrists and the Singapore Optometric Association), *Journal of Optometry* (Spanish General Council of Optometrists), *Contact Lens & Anterior Eye* (British Contact Lens Association), *African Vision & Eye Health* (South African Optometric Association), the *Chinese Journal of Optometry & Ophthalmology*, and the *Canadian Journal of Optometry*. Some professional bodies also have their own library and provide full access to a wide range of online optometry and ophthalmology journals.

1.2.2 How do you assess the usefulness of optometric tests?

The usefulness of optometric tests is typically assessed by either comparing the test against an appropriate gold standard[13] and/or assessing its repeatability[14] and/or its discriminative ability. For example, a test that is being used as an objective measure of subjective refraction should be assessed by how closely the results match subjective refraction results[13] and new tonometers are assessed by their agreement with the results of Goldmann Applanation Tonometry (GAT, although this is not ideal when GAT has flaws, see section 7.7.1).

1.2.3 How to assess tests that are part of the subjective refraction

The use of subjective refraction as a gold standard assessment of refractive error has meant that there has been little or no comparison of the various methods used in subjective refraction. Previous studies have tended to compare the various tests against each other. For example, West and Somers[15] compared the various binocular balancing tests and found that they all gave similar results and concluded that they were therefore all equally useful. Johnson et al.[16] reported a similar finding when comparing subjective tests for astigmatism. However, the size of the differences found with the different balancing and astigmatic tests in these studies would likely have led to symptoms and patient dissatisfaction with some of the refractive prescriptions if they had been worn, so that the conclusions seem incorrect and the study design poor.[13] An inventive but underutilised approach is to use some measure of patient satisfaction[17] or dissatisfaction[9] as the gold standard. For example, Strang et al.[17] compared the refractive corrections provided by subjective refraction and autorefraction by randomly allocating glasses in a double-blind protocol. Subjects wore each prescription for 2 weeks and completed a questionnaire that assessed visual performance and ocular comfort following each period of wear.

1.2.4 Bland-Altman plots vs. correlation coefficients

In the past, test comparison studies tended to quantify the relationship between the test and gold standard using correlation coefficients. This is not appropriate for a variety of reasons,[14,18] including that correlation coefficients are very much affected by the range of values used in the analysis.[19] If a small range of values is used in calculations, the correlation coefficient is likely to be much smaller than if a larger range is used (Fig. 1.1).[19] A much better analysis, commonly known as a Bland-Altman plot, shows the 95% confidence limits of the difference between the test and gold standard (Fig. 1.2).[14,18] Bland-Altman assessments are limited in that values are in the units of measurement rather than a unitless 0 to 1 scale like correlation coefficients, so it may not be obvious what is a good result. For this reason, the extent to which the 95% Bland-Altman agreement figures are clinically acceptable should be discussed by the authors of a paper and acceptable limits should be determined prior to any assessment.[14]

1.2.5 How best to assess test-retest repeatability

Repeatability assesses the ability of a measurement to be produced consistently. It is best to assess repeatability in

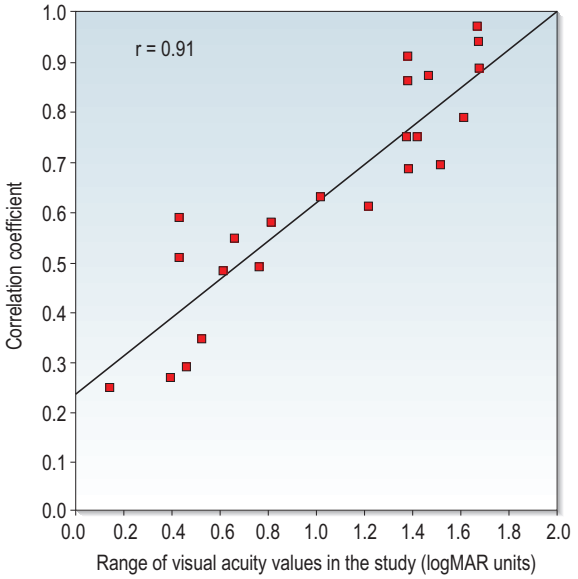

Fig. 1.1 Correlation coefficients from the literature between high-contrast visual acuity and other spatial vision measures are plotted as a function of the range of high-contrast acuities in those studies. (Redrawn with permission from Haegerstrom-Portnoy G, Schneck ME, Lott LA, Brabyn JA. The relation between visual acuity and other spatial vision measures. *Optometry and Vision Science.* 2000;77:653–62. ©The American Academy of Optometry, 2000.)

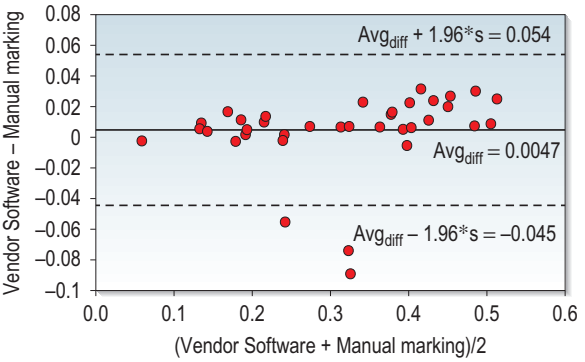

Fig. 1.2 A Bland- Altman plot of the difference (diff) between the Foveal Avascular Zone (FAZ) area (mm^2) measured by the custom vendor software and manual marking as a function of the mean of the FAZ area from the two techniques. (Reprinted with permission from Arthur E, Papay JA, Haggerty BP, Clark CA, Elsner AE. Subtle changes in diabetic retinas localised in 3D using OCT. *Ophthalmic & Physiological Optics* 2018;38:477–91. ©The College of Optometrists, 2018).

terms of the coefficient of repeatability, which represents the 95% confidence limits of the difference between the test and retest scores and can be displayed using Bland-Altman plots (see Fig. 1.2).[14,18] Intraclass correlation coefficients can be used when comparing tests that do not use the same units, but their limitations need to be realised.[14] In particular, a large range of values should be used, so that correlation coefficients are not artificially low (Fig. 1.1).

Concordance values (the percentage of patients getting exactly the same score on test and retest) have also been used to indicate that a test is repeatable. However, a high proportion of patients often obtaining exactly the same score on follow-up visits indicates that the step sizes on the test are too big rather than that the test is repeatable.[20] For example, a VA chart containing only 20/20 (6/6), 20/60 (6/18), and 20/200 (6/60) lines would provide very high concordance, but would be of very little value. Many 4- to 5-point grading scales have similar problems and scores are best interpolated between grades, if possible, with a 0.1 scale being used.[21]

Repeatability appears to be a very important quality of a test, because an unreliable test is likely to correlate poorly with a gold standard and has poor discriminative ability.[22] Because these studies are also relatively quick and simple, the results of repeatability studies should be available for all clinical tests.

1.3 "Screen everybody, don't miss disease" vs. "what about false positives?"

Optometrists detect glaucoma and other eye diseases by 'opportunistic case finding' in that patients are self-selecting and they are detected as part of an eye examination that includes some assessment of ocular health and visual function.[23] Professional bodies within different countries generally provide evidence-based guidelines that tend to suggest which tests are appropriate for different patient demographics and perhaps for certain signs and symptoms. There has been a tendency, however, particularly with the increased use of clinical assistants within optometric practice, to increasingly screen patients with tests such as visual fields and non-contact tonometry to attempt to 'not to miss anything.'[24] This approach is examined below and highlights the importance of understanding diagnostic indices of optometric tests.

1.3.1 Do you really understand a test's diagnostic ability?

New diagnostic tests must have their diagnostic ability compared with a gold standard reference. The research study will therefore determine how well a test can correctly identify 'abnormal' or 'normal' eyes as classified independently by a gold standard test or battery of tests. For example, new instruments or techniques that attempt to identify patients with primary open-angle glaucoma (POAG) are typically assessed against classifications of patients into glaucomatous and control groups by clinical evaluation of optic nerve head assessment, anterior chamber angle, and visual fields.[25]

Table 1.1 Possible outcomes of a screening test

	Diseased eye	Normal eye
Test says diseased	True positive, TP (hit)	False positive, FP (false alarm)
Test says normal	False negative, FN (miss)	True negative, TN

Please note that the following figures of sensitivity, specificity, and prevalence have been simplified to help the explanation. Imagine a POAG test that correctly detects patients with POAG 95% of the time (the sensitivity of the test is 95%); if the test indicates that a patient has POAG, what are the chances that they actually have the disease? Is it 95%? If lower, how much lower? When considering this question, you must not only consider how good the test is at identifying POAG, but you must also consider how good the test is at correctly identifying someone as normal. Unfortunately all tests provide false–positive findings: patients who have normal, healthy eyes for whom the test results suggest are abnormal. There are four possible outcomes from the results of a diagnostic test (Table 1.1), and this information is used to quantify how well the test discriminates between 'normal' and 'abnormal' eyes, by providing sensitivity and specificity values.

- Sensitivity is the ability of the test to identify the disease in those who have it.
- Sensitivity = TP / (TP + FN).
- Specificity is the ability of the test to correctly identify those who do not have the disease.
- Specificity = TN / (TN + FP).
- The false–positive rate is simply 1 minus the specificity.
- Another important term to understand is the predictive value (PV), which has positive (PPV) and negative (NPV) forms.
- PPV or +PV is the proportion of people with a positive test result who have the disease. PPV = TP / (TP + FP).
- NPV or −PV is the proportion of people with a negative test result who do not have the disease. NPV = TN / (TN + TP).

The reported sensitivity and specificity of a test will differ depending on the pool of patients examined, the gold standard used to determine the presence or absence of disease, and the cut-off criteria used. Sensitivity and specificity values and plots of one against the other for a range of cut-off values in receiver operating characteristic (ROC) curves (Fig. 1.3) are usually presented and are often quantified using the area under the ROC curve.

The ability of a diagnostic test to identify correctly those patients with disease is highly dependent on how prevalent the condition is (known as Bayes Theorem). For example, let us consider POAG and assume a prevalence in the over age 40 population of 1%, and a diagnostic test for glaucoma with 95% sensitivity and 95% specificity. The first column in Table 1.2 shows the likely outcomes from 1000 patients. Nine or all 10 patients with POAG have a positive test result, but so have 50 patients with normal, healthy eyes. Returning to the question at the beginning of this section, if a POAG test that correctly detects patients with POAG 95% of the time (95% sensitivity) indicates that a patient has POAG, the chances that they actually have the condition (given a test specificity of 95%) is 16%! Detecting disease that has a low prevalence is very difficult no matter how good your diagnostic tests are because there are so few patients with the disease and so many people who do not have that disease. This also highlights that with diseases with low prevalence, you are better off using tests (or cut-off scores for a test) that have the highest specificity (limiting false–positive results) even if this lowers sensitivity and a small number with POAG (in its early stages) are missed.

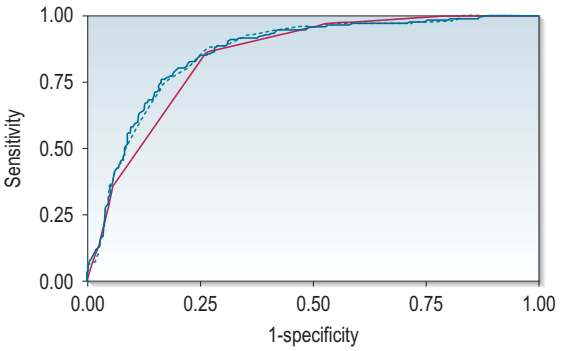

Fig. 1.3 Receiver operating characteristic (ROC) curves for prediction of high myopia in myopic children at age 11 years. (n = 910) Age myopia onset (*red line*); Age of myopia onset, gender, race, and school (*dashed green line*); Age of myopia onset, gender, race, school, books per week, and parental myopia (*blue line*). (Redrawn with permission from Chua et al. Age of onset of myopia predicts risk of high myopia in later childhood in myopic Singapore children. *Ophthalmic & Physiological Optics.* 2016;36:388–94. ©The College of Optometrists, 2016.)

1.3.2 The number of false–positive referrals

The number of false–positive referrals varies depending on the disease and its prevalence, the structure and funding eye care system; the level of clinical experience,

Table 1.2 Results for a 'glaucoma test' with different values of prevalence of primary open-angle glaucoma (POAG)

	SENSITIVITY 95% AND SPECIFICITY 95%		
	POAG prevalence, 1%	**POAG prevalence, 10%**	**Repeated testing (POAG, 1%)**
A. Patients with POAG	10	100	10
B. Patients without POAG	990	900	990
C. True positive (95% of A)	9.5 (i.e., 9 or 10)	95	9
D. False positive (5% of B)	50	45	2.5
E. True negative (95% of B)	940	855	47.5
F. False negative (5% of A)	0.5 (i.e., 0 or 1)	5	0.5
G. Number with positive result (C+D)	60	140	11.5
H. Number with negative result (E+F)	941	860	48
Positive predictive value: C/G	9.5/60 (16%)	95/140 (68%)	9/11.5 (78%)
Negative predictive value: E/H	940/941 (~100%)	855/860 (99.4%)	47.5/48.5 (99%)

expertise, and equipment; the introduction of locally agreed guidelines; and so forth.[26–30] False–positive rates for diseases such as POAG with a low prevalence will tend to be high, as discussed earlier. For example, Bowling et al.[31] reported a 46% false–positive rate for suspected glaucoma from 2505 optometric referrals to the Oxford Eye Hospital in the United Kingdom over a 10-year period (1994–2004).

1.3.3 False–positive eye referrals cause worry and stress

Elmore et al.[32] reported the false–positive rate of the two main breast cancer screening tests to be 6.5% and 3.7%. These translate to very good specificity values of 93.5% and 96.3%. Despite this good specificity, over a 10-year period, nearly one-third of the women screened had at least one false–positive mammogram or clinical breast examination. This highlights that if you test healthy people often enough, they will sooner or later obtain a false–positive test result. It has been shown that these false–positive results have negative psychological effects on these women and their families.[33] Similarly, considerable worry and stress are caused by a false–positive result leading to referral to a secondary eye care system.[34] Patients should not be referred to secondary eye care on the basis of a slightly high intraocular pressure using a non-contact tonometer or a single positive visual field screening result. In addition to the psychological effects on patients and their families, the costs in terms of secondary eye care staff and patient time (including the delay that other patients will suffer because of busy clinics) prompted by a positive screening result must be considered.

1.3.4 Reducing false positives 1: only screen 'at risk' patients

Owing to the high number of false–positive results when screening patients for a disease with low prevalence (Bayes Theorem), it may be better to screen only those patients who are 'at risk.' In these patients, the prevalence of the disease is higher than in the general population. The middle column in Table 1.2 considers the likely outcomes using the same test discussed earlier, on patients with a family history of POAG where the prevalence of the disease is higher and for simplicity we will assume a figure of 10%. In total, 140 patients have positive results, of which 95 have the disease (PPV = 68%). Note how much better the test performs when it is used in patients with a higher prevalence of the disease. The positive predictive value is also significantly improved if you just perform screening on all patients over 75 years of age or patients over 40 years of age who are African American or African Caribbean or those with suspicious optic discs or high intraocular pressure. Burr et al.,[23] in their systematic review, suggested that screening of patients with 'minor' risk factors, including myopia and diabetes, did not improve the PPV sufficiently and was not cost-effective.[23]

1.3.5 Reducing false positives 2: repeat testing

Another way of keeping false–positive referrals to a minimum, and imperative if you are intending to screen more than 'at risk' patients, is to repeat positive results. For example, as part of the ocular hypertension treatment study, Keltner et al.[35] found 703 Humphrey visual field test results that showed abnormal (positive glaucoma hemifield test and/or corrected pattern standard deviation, $P < 0.05$) and reliable visual fields.[35] On retesting, abnormalities *were not* confirmed for 604 (86%)! The vast majority of visual field abnormalities were not verified on retest, and confirmation of visual field abnormalities is essential for distinguishing reproducible visual field loss from long-term variability.

If the same glaucoma diagnostic test from Table 1.2, which suggested that 60 patients had POAG (only 10 did), was repeated on these 60 patients, 9 of the 10 patients with glaucoma would be identified, but 95% of those 50 with false–positive results (47.5) would now give a normal result, with only 2.5 (i.e., 2 or 3) false positives on retest. On retesting, positive results are found for 11.5 patients, of whom 9 have the disease (PPV = 78%). Of course, you could also combine both approaches by only screening patients at risk and repeating positive tests.

1.3.6 Reducing false positives 3: mentoring inexperienced clinicians

Studies have shown that a majority of false–positive referrals are from inexperienced clinicians, typically in the first few years as a qualified optometrist.[28,30] An obvious strategy would be for newly qualified clinicians to continue to be mentored, particularly for referrals. Targeted continuing education and training for these clinicians would also likely be of benefit.[30]

1.3.7 Reducing false positives 4: intermediate/enhanced/collaborative care schemes

Given the high cost of referrals to secondary eye care, plus the burgeoning elderly population leading to ever-increasing referrals, a variety of schemes have been developed to decrease false–positive referrals. These are typically clinics intermediate between primary and secondary eye care that provide

enhanced optometric services often in collaboration with ophthalmology.[27,29]

1.4 Should I perform a database, system, and/or problem-oriented eye examination?

The primary eye care examination must first and foremost adhere to the legal requirements where you are working, although these tend to be provided in very broad terms. Some professional organisations to which you may belong may also provide clinical guidelines of what your eye examination should include. These may be prescriptive or for guidance only. Three main styles for a primary eye care examination could be used singularly or in combination: (1) the database format, which uses a predetermined series of tests; (2) the systems approach, which ensures an assessment of several systems; and/or (3) the problem-oriented approach, which focuses mainly on the patient's problems.[36] In addition, some parts of the eye examination could be performed by clinical assistants.

1.4.1 The database examination

The database examination style means using essentially the same set of clinical procedures in every examination. A large 'complete' database of information is collected to ensure that most patients' problems can be addressed using the information provided. This is the style of examination that will be used by students, because they need to practice the various clinical techniques to gain technical competence. Technical competence should be the aim for students in the early years of clinical learning. A much greater task is gaining clinical competence and understanding the tests and their results, how they interact and how they can be used in differential diagnosis, and to solve the patient's problems. Only once a student/practitioner has gained a high level of clinical competence should the database style of examination be abandoned and another approach used.

Although the database examination style is ideal for students, it is not for experienced practitioners. Often, if a large database is used, some data collected provide no useful information regarding the clinical diagnosis or treatment options. If patients require additional testing, because of the inflexibility of the database examination approach, practitioners either perform the tests at the end of the examination, which can lead to them being late for subsequent examinations, or another appointment is made at a later date. At its worst, this style of examination could be said to provide some test data which are not used and of little value and provide a bias against performing additional procedures which may be of real benefit.

1.4.2 Systems examination

A systems examination style includes an assessment of visual function, and of the refractive and binocular systems and an ocular health assessment. The optometric examination is defined not by tests used, but by the systems that are assessed (Table 1.3). This approach is much more flexible as it does not demand that a certain collection of tests be used. In such an examination style, a minimal database has been gathered when each system has been tested. In summary, think in terms of assessing systems and not of using individual tests.

Table 1.3 Classification of tests/procedures into one of four clinical oculovisual systems

Visual[a]	Binocular[a]	Refractive	Ocular Health
Case history	Case history	Case history	Case history
Visual acuity	Visual acuity	Visual acuity	Visual acuity
Disability glare	Cover test	Retinoscopy	Biomicroscopy
Photostress recovery	Convergence tests	Autorefraction	Ophthalmoscopy
Contrast sensitivity	Accommodation tests	Subjective refraction	Tonometry
Colour vision	Suppression tests	Reading add	Gonioscopy
Visual fields	Stereopsis	Keratometry	Pupil responses
	Motility		Imaging

[a]Other classifications discuss the sensory and motor systems rather than the visual and binocular systems and place suppression and stereopsis within the sensory system.

1.4.3 Problem-oriented examinations

The problem-oriented examination aligns the examination around the problems reported by the patient. However, it does not only use tests that help solve the patient's problems because it is built on a systems examination approach.[36] In addition, an array of tests are typically used as screening or 'entrance' tests,[37] which provide an initial assessment of each system. Some of the entrance tests may be part of pretesting conducted by clinical assistants, so that a period of time is required prior to meeting the patient for this information and any previous record cards to be reviewed. Clinical assistants can provide data from automated procedures including focimetry, autorefraction, fundus photography, ocular coherence tomography (OCT), ultra-widefield imaging (e.g., Optos), automated visual fields, non-contact tonometry, and pachymetry plus simple tests such as colour vision, stereopsis, and interpupillary distance (PD) measurement; the tests used may differ depending on the age of the patient.

From information from previous records (if available), the preliminary entrance tests, and the initial case history information, you will develop a mental list(s) of tentative diagnoses. These will likely lead to further questioning as part of differential diagnosis (e.g., particularly for symptoms such as red eye) and the subsequent tests used in the remainder of the problem-oriented examination will be determined by which tests best help the differential diagnosis process.

1.5 The evidence to support routine dilated fundus examinations is limited

Two main arguments are proposed in favour of the routine dilated fundus examination (DFE). The first is that a DFE increases the number of anomalies detected within the central retina.[38,39] In two studies, a non-dilated fundus examination with direct ophthalmoscopy was compared with a DFE using headband binocular indirect ophthalmoscopy and direct ophthalmoscopy.[38,39] Siegel et al.[38] also used a monocular indirect ophthalmoscope examination as part of a non-dilated examination. The poor field of view of the direct ophthalmoscope was particularly blamed for missing anomalies in the posterior pole because it is too small to examine the area quickly and easily. However, fundus biomicroscopy is now the standard of care for central fundus examination and provides a much better field of view than direct ophthalmoscopy as well as a stereoscopic view and can be conducted non-dilated (section 7.10.1).

The second argument for routine DFEs is that significant anomalies would otherwise be missed in the peripheral retina. Two studies have retrospectively reviewed record charts to determine the extent that significant fundus lesions were detected in DFEs.[40,41] Pollack and Brodie reviewed 1094 ophthalmologic records of DFEs of asymptomatic patients without risk factors and found three (0.3%) with clinically significant fundus lesions outside the vascular arcades and unlikely to have been detected using direct ophthalmoscopy.[40] Varner identified one case (0.2%) of an asymptomatic patient with a clinically significant peripheral lesion after 10 years of follow-up of 592 older (mean age approximately 70 years) patients and also concluded that the value of routine DFEs is very low.[41]

Disadvantages of routine DFEs include their inconvenience to the patient (e.g., glare and blur problems, they may not be able to drive home or return to work)[42] and the increased possibility of false–positive referrals,[43] particularly given the high prevalence of benign peripheral retina conditions in asymptomatic patients.[41] The cost of false–positive referrals must be recognised as it is not just the cost of the unnecessary secondary care appointment but the anxiety and inconvenience caused to the patient (section 1.3.3).

In summary, minimal evidence exists to support routine DFEs when compared against a routine non-dilated fundus examination with fundus biomicroscopy and DFEs in patients with pertinent signs and/or symptoms including symptoms of flashes and floaters, high myopia, recent cataract surgery, a family history of retinal detachment, and a small non-dilated pupil that would restrict the stereoscopic view of the central retina.[42]

1.6 Lessons to learn from the Honey Rose case

The potential for mistakes with overreliance on ocular imaging, particularly when taken by clinical assistants, was tragically illustrated in the United Kingdom in 2012. Details were provided in the 2016 trial of Honey Rose, a locum optometrist who was charged with gross negligence manslaughter (subsequently quashed on appeal)[44,45] owing to missing bilateral papilloedema in 8-year-old Vincent Barker who subsequently died of hydrocephalus 5 months after the eye examination. Experts agreed that his life could have been saved if the papilloedema had been spotted at the eye examination and the patient referred.[44] Although the patient had no symptoms to suggest a problem at the time of the eye examination, symptoms of a bout of unexplained headaches from a few weeks previously were reported to the optometrist.[44] Honey Rose claimed that she did not perform ophthalmoscopy because the patient closed his eyes and looked away from

the bright light.[45] However, she did not make any record of this difficulty. It would also appear that the optometrist included a note on the patient's records to indicate she had viewed his inner eye.[45] Fundus images were taken at the time of the eye examination by a clinical assistant, and an expert witness indicated that they clearly showed bilateral papilloedema.[44,45] Honey Rose claimed that she had received insufficient training in the information technology (IT) systems and that she could not work the screen on the camera and asked a colleague to display it. She also claimed at different points of the trial that she viewed (1) the photograph of Vincent Barker from a previous appointment or (2) the photograph of a different patient.[44]

There are many lessons that optometry must learn from this tragic case, and these include:

- Patients can have life-threatening conditions and be asymptomatic.
- Optometrists must not rely solely on images to view the fundus.
- It is the responsibility of the optometrist to learn how to work imaging equipment and other tests usually performed by clinical assistants.
- It is the responsibility of eyecare practices to ensure that optometrists, including locum optometrists, fully understand the imaging (and other) processes used in their practice.
- It must be understood that mistakes linking images to patients (and specific examinations) can occur and great care must be taken to avoid them.

Although the evidence from this case suggests that the errors were those of the optometrist, they also highlighted other potential errors in the system, including:

- Technical problems, including a malfunctioning shutter, faulty display, and power cut on the day in question that were not dealt with in an appropriate manner.[45] Appropriate processes need to be put in place and regularly audited to ensure equipment works properly.
- Anecdotal evidence indicates that clinical assistants can fail to add patient details to images during busy clinics, thus connecting the previous patient's details to the following patient's image. This highlights the requirement for high-quality training of clinical assistants and the need for the clinician to be aware that such errors can occur and to look out for them.

References

1. Greenhalgh T. *How to Read a Paper: The Basics of Evidence-Based Medicine*, ed 3. Oxford: Blackwell Publishing; 2006.
2. Bailey IL, Lovie-Kitchin JE. Visual acuity testing. From the laboratory to the clinic. *Vision Res*. 2013;90:2–9.
3. Lovie- Kitchin JE. Is it time to confine Snellen charts to the annals of history? *Ophthalmic Physiol Opt*. 2015;35: 631–6.
4. Calvin H, Rupnow P, Grosvenor T. How good is the estimated cover test at predicting the von Graefe phoria measurement? *Optom Vis Sci*. 1996;73:701–6.
5. Rainey BB, Schroeder TL, Goss DA, Grosvenor TP. Inter-examiner repeatability of heterophoria tests. *Optom Vis Sci*. 1998;75:719–26.
6. Cebrian JL, Antona B, Barrio A, Gonzalez E, Gutierrez A, Sanchez I. Repeatability of the modified Thorington card used to measure far heterophoria. *Optom Vis Sci*. 2014;91:786–92.
7. Wong EP, Fricke TR, Dinardo C. Interexaminer repeatability of a new, modified prentice card compared with established phoria tests. *Optom Vis Sci*. 2002;79:370–5.
8. Casillas EC, Rosenfield M. Comparison of subjective heterophoria testing with a phoropter and trial frame. *Optom Vis Sci*. 2006;83:237–41.
9. Hanlon SD, Nakabayashi J, Shigezawa G. A critical view of presbyopic add determination. *J Am Optom Assoc*. 1987;58:468–72.
10. Suttle CM, Jalbert I, Alnahedh T. Examining the evidence base used by optometrists in Australia and New Zealand. *Clin Exp Optom*. 2012;95: 28–36.
11. Graham AM. Finding, retrieving and evaluating journal and web-based information for evidence-based optometry. *Clin Exp Optom*. 2007;90:244–9.
12. Hisham R, Ng CJ, Liew SM, Hamzah N, Ho GJ. Why is there variation in the practice of evidence-based medicine in primary care? A qualitative study. *BMJ Open*. 2016;6:e010565.
13. Elliott DB. What is the appropriate gold standard test for refractive error? *Ophthalmic Physiol Opt*. 2017;37: 115–7.
14. McAlinden C, Khadka J, Pesudovs K. Statistical methods for conducting agreement (comparison of clinical tests) and precision (repeatability or reproducibility) studies in optometry and ophthalmology. *Ophthalmic Physiol Opt*. 2011;31:330–8.
15. West D, Somers WW. Binocular balance validity: a comparison of five different subjective techniques. *Ophthalmic Physiol Opt*. 1984;4:155–9.
16. Johnson BL, Edwards JS, Goss DA, et al. A comparison of three subjective tests for astigmatism and their interexaminer reliabilities. *J Am Optom Assoc*. 1996;67: 590–8.
17. Strang NC, Gray LS, Winn B, Pugh JR. Clinical evaluation of patient tolerance to autorefractor prescriptions. *Clin Exp Optom*. 1998;81:112–8.
18. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;1:307–10.
19. Haegerstrom-Portnoy G, Schneck ME, Lott LA, Brabyn JA. The relation between visual acuity and other spatial vision measures. *Optom Vis Sci*. 2000;77: 653–62.
20. Bailey IL, Bullimore MA, Raasch TW, Taylor HR. Clinical grading and the

effects of scaling. *Invest Ophthalmol Vis Sci.* 1991;32:422–32.
21. Chylack Jr LT, Wolfe JK, Singer DM, et al. The Lens Opacities Classification System III. The longitudinal study of cataract study group. *Arch Ophthalmol.* 1993;111:831–6.
22. Elliott DB, Bullimore MA. Assessing the reliability, discriminative ability, and validity of disability glare tests. *Invest Ophthalmol Vis Sci.* 1993;34:108–19.
23. Burr JM, Mowatt G, Hernández R, et al. The clinical effectiveness and cost-effectiveness of screening for open angle glaucoma: a systematic review and economic evaluation. *Health Technol Assess.* 2007;11:1–190.
24. Vernon SA. The changing pattern of glaucoma referrals by optometrists. *Eye (Lond).* 1998;12:854–7.
25. Dabasia PL, Fidalgo BR, Edgar DF, Garway-Heath DF, Lawrenson JG. Diagnostic accuracy of technologies for glaucoma case-finding in a community setting. *Ophthalmology.* 2015;122:2407–15.
26. Ang GS, Ng WS, Azuara-Blanco A. The influence of the new general ophthalmic services (GOS) contract in optometrist referrals for glaucoma in Scotland. *Eye (Lond).* 2009;23:351–5.
27. Baker H, Ratnarajan G, Harper RA, Edgar DF, Lawrenson JG. Effectiveness of UK optometric enhanced eye care services: a realist review of the literature. *Ophthalmic Physiol Opt.* 2016;36:545–57.
28. Davey CJ, Scally AJ, Green C, Mitchell ES, Elliott DB. Factors influencing accuracy of referral and the likelihood of false positive referrals by optometrists in Bradford, United Kingdom. *J Optom.* 2016;9:158–65.
29. Ly A, Nivison-Smith L, Hennessy M, Kalloniatis M. The advantages of intermediate-tier, inter-optometric referral of low risk pigmented lesions. *Ophthalmic Physiol Opt.* 2017;37:661–8.
30. Parkins DJ, Benwell MJ, Edgar DF, Evans BJW. The relationship between unwarranted variation in optometric referrals and time since qualification. *Ophthalmic Physiol Opt.* 2018;38:550–61.
31. Bowling B, Chen SD, Salmon JF. Outcomes of referrals by community optometrists to a hospital glaucoma service. *Br J Ophthalmol.* 2005;89:1102–4.
32. Elmore JG, Barton MB, Moceri VM, Polk S, Arena PJ, Fletcher SW. Ten-year risk of false positive screening mammograms and clinical breast examinations. *N Engl J Med.* 1998;338:1089–96.
33. Brett J, Austoker J. Women who are recalled for further investigation for breast screening: psychological consequences 3 years after recall and factors affecting re-attendance. *J Public Health Med.* 2001;23:292–300.
34. Davey CJ, Harley C, Elliott DB. Levels of state and trait anxiety in patients referred to ophthalmology by primary care clinicians: a cross sectional study. *PLoS One.* 2013;8:e65708.
35. Keltner JL, Johnson CA, Quigg JM, et al. Confirmation of visual field abnormalities in the Ocular Hypertension Treatment Study. *Arch Ophthalmol.* 2000;118:1187–94.
36. Amos JF. The problem-solving approach to patient care. In: Amos JF, ed. *Diagnosis and Management in Vision Care.* Boston: Butterworths; pp 1–8, 1987.
37. Carlson NB, Kurtz D, McGraw-Hill Education. *Clinical Procedures for Ocular Examination.* New York: McGraw-Hill Education; 2016.
38. Siegel BS, Thompson AK, Yolton DP, Reinke AR, Yolton RL. A comparison of diagnostic outcomes with and without pupillary dilatation. *J Am Optom Assoc.* 1990;61:25–34.
39. Parisi ML, Scheiman M, Coulter RS. Comparison of the effectiveness of a non-dilated versus dilated fundus examination in the pediatric population. *J Am Optom Assoc.* 1996;67:266–72.
40. Pollack AL, Brodie SE. Diagnostic yield of the routine dilated fundus examination. *Ophthalmology.* 1998;105:382–6.
41. Varner P. How frequently should asymptomatic patients be dilated? *J Optom.* 2014;7:57–61.
42. Batchelder TJ, Fireman B, Friedman GD, et al. The value of routine dilated pupil screening examination. *Arch Ophthalmol.* 1997;115:1179–84.
43. Ly A, Nivison- Smith L, Hennessy M, Kalloniatis M. The advantages of intermediate-tier, inter-optometric referral of low risk pigmented lesions. *Ophthalmic Physiol Opt.* 2017;37:661–8.
44. Mullock A. Gross negligence (Medical) manslaughter and the puzzling implications of negligent ignorance: Rose v R [2017] EWCA crim 1168. *Med Law Rev.* 2018;26:346–56.
45. Storey T. Whether 'obvious and serious' risk of death in cases of gross negligence manslaughter to be determined both objectively and prospectively: R v Rose [2017] EWCA Crim 1168. *J Crim Law.* 2017;81:343–6.