Br Ir Orthopt J 2008; **5:** 47–53

Diagnostic accuracy of school vision screening

CATHERINE E. STEWART[1] PhD BMedSci (Orthoptics), JAMES M. GILCHRIST[2] PhD BSc, PAUL V. MCGRAW[3] PhD BSc AND BARRY WINN[4] PhD BSc

[1]*Department of Optometry & Visual Science, City University, London*
[2]*Department of Optometry, University of Bradford, Bradford*
[3]*School of Psychology, University of Nottingham, Nottingham*
[4]*Hull York Medical School, University of Hull, Hull*

Abstract

Aim: **Currently, the Department of Health recommends vision screening of all children between the ages of 4 and 5 years. The aim of this study was to evaluate the diagnostic accuracy of entry-level school-aged vision screening in the UK. The performance of vision screening protocols and referral criterion (Snellen chart) has been compared with an alternative screening tool (Crowded logMAR test). Diagnostic accuracy has been modelled in terms of the screening test and referral criterion.**

Methods: **A total of 494 children (mean age 5.25 ± 0.28 years) were screened, for the presence of visual anomalies, by school nurses using a Snellen visual acuity chart and Crowded logMAR visual acuity test. Referral was made on the basis of Snellen visual acuity measures using standard NHS protocols (visual acuity 6/9 or worse).**

Results: **Within the population screened, approximately 14% ($n = 68$) were referred for further investigation. Of the 68 children referred, over two-thirds were found to be visually normal (false-positive rate of 68%). No false-negatives were identified, indicating that the current screening criterion achieves maximum sensitivity. The Crowded logMAR test at the equivalent referral criteria (0.175 logMAR) also achieved maximum sensitivity but with significantly ($p = 0.0008$) fewer false-positive referrals (60%).**

Conclusion: **Current school nurse screening criteria produce a high percentage of false-positives leading to low referral accuracy. However, current protocols provide maximum sensitivity (no false-negatives) for detecting visual anomalies. In order to retain maximum screening sensitivity whilst reducing the rate of false-positive referrals, screening programmes should adopt more precise measures of visual acuity in the form of logMAR-based tests.**

Key words: Amblyopia, logMAR visual acuity, Positive predictive value (PPV), Referral criterion, School-aged vision screening

Correspondence and offprint requests to: Catherine E. Stewart, Department of Optometry & Visual Science, City University, CM466, Tait Building, Northampton Square, London, EC1V 0HB. e-mail C.E.Stewart@city.ac.uk

Introduction

Current screening guidelines in the UK recommend that every child between the ages of 4 and 5 years should have access to vision screening.[1] They suggest that this should incorporate a linear logMAR visual acuity test and preferably be administered by an orthoptist, although it is noted that orthoptists may not be in the position to deliver such a programme.[1] At this time, when many screening programmes are being established or modified, consideration must be given to the choice of screening test and the criterion level used for referral. Vision screening of school-aged children by school nurses is primarily carried out using the Snellen chart with a referral criterion of 6/9.

Conventional visual acuity charts based on the Snellen design suffer from many methodological limitations which prevent precise assessment of visual acuity, and this limits their effectiveness in screening.[2,3] Such limitations include unequal legibility of letters, irregular progression of letter sizes on the chart and variation in the number of letters per line. These factors produce variable crowding ratios (different spacing between letters on the chart) and variation in the precision with which different acuity levels may be estimated.[2]

The methodological limitations associated with such tests have been overcome with the introduction of charts that use letters of more equal legibility, employ equal number of letters per line, and scale letter size in equal steps according to the logarithm of the minimum angle of resolution (logMAR).[2–6] These charts allow visual acuity to be recorded with improved precision and provide an increased sensitivity for detecting changes in acuity levels between visits.[2–6] Such advances in acuity chart design are starting to filter into vision screening programmes in the UK.

In this study we present data on the diagnostic accuracy of visual acuity screening in a large sample of school-aged children using the conventional Snellen-based referral criterion, and demonstrate that screening

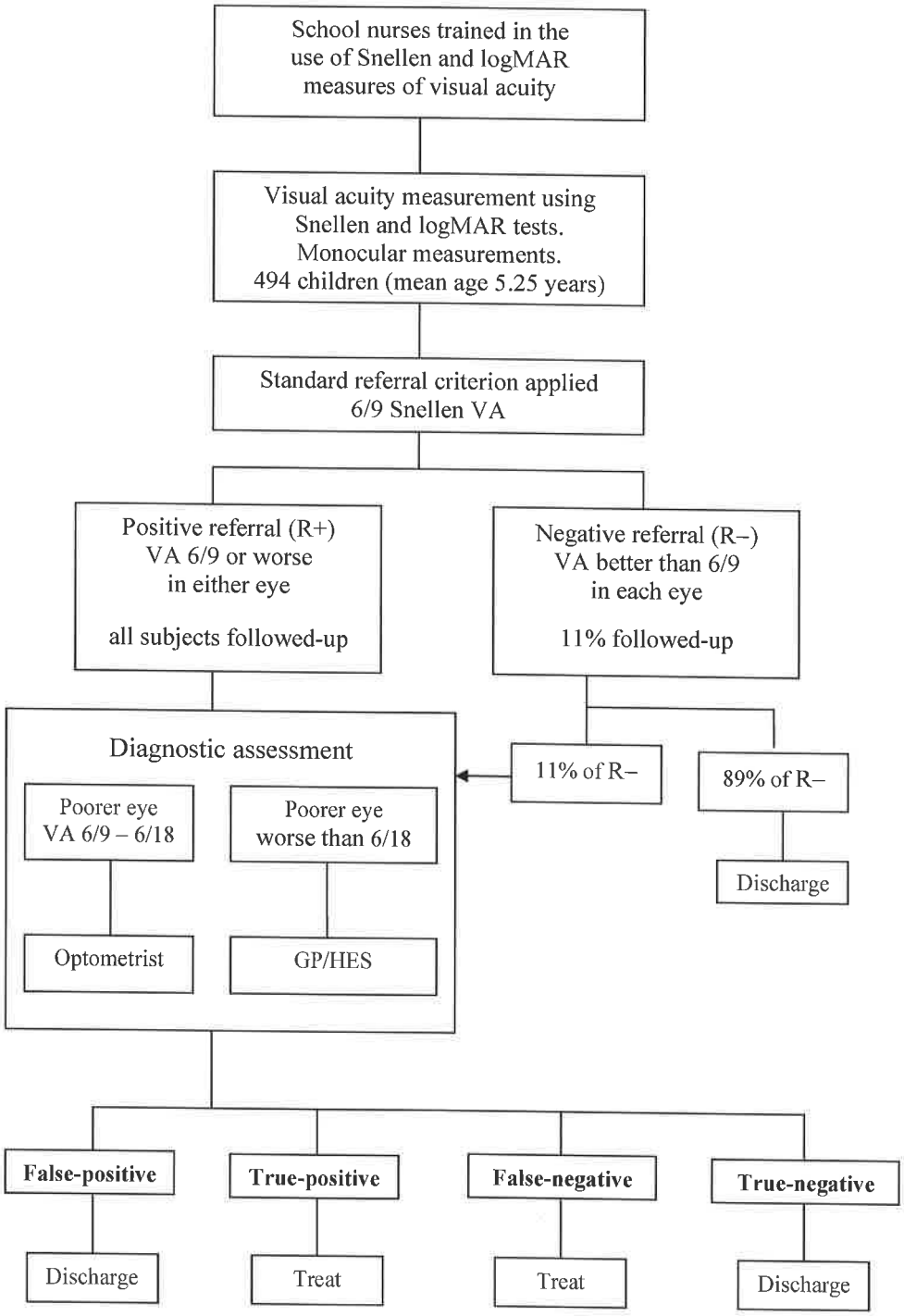

Fig. 1. Decision tree of the screening protocol.

performance can be improved significantly by simply changing the visual acuity test and referral criterion. This could have significant financial implications for the NHS.

Materials and methods

Screening protocols and procedures

Six school nurses working in the Bradford and Airedale region attended a training programme at the University of Bradford, introducing them to the use of logMAR visual acuity tests and establishing a common protocol for the study (Fig. 1). The testers (in this case 6 school nurses) were masked to the hypothesis of the study to prevent any possible influence on testing methods.

Following training, the six school nurses carried out visual acuity screening on 494 children in their normal school environment between April 1997 and June 1998. Screening was included as a component of general health interviews conducted on children of around 5 years of

age. Thus, the mean age of the children was 5.25 years with a standard deviation of 0.28 years. Informed consent was obtained from all parents of children participating in the screening programme and any subsequent diagnostic assessment.

Visual acuity in each eye was measured using a conventional Snellen chart and the Crowded logMAR test.[6] Tests were carried out in a random order. Threshold Snellen acuity was designated as the smallest letter size at which two-thirds of the letters on the line were correctly identified. The Crowded logMAR test always presents four letters per line, and a uniform step of 0.1 logMAR unit from one line to the next. Each letter is assigned a score of 0.025 logMAR, allowing scores to be determined by letter. All children tested were able to perform both tests.

Referral decisions were based on the criterion of 6/9 or worse in either eye as measured on the Snellen chart. All children whose Snellen acuity measurement produced a positive referral decision (R+) according to this criterion were referred for full investigation. Children with visual acuity between 6/9 and 6/18 in the poorer eye were referred to an optometrist, and those with vision worse than 6/18 in the poorer eye were referred to their general practitioner for subsequent assessment at their local hospital eye department. If diagnostic assessment revealed any significant refractive error warranting spectacle wear, manifest strabismus or decompensating heterophoria, amblyopia, disease of the eye or visual pathway, then the case was designated a true-positive (TP) referral, and the child was offered appropriate treatment. Unilateral amblyopia was defined as one Snellen line difference between the eyes, whereas bilateral amblyopia was defined as visual acuity of 6/9 or worse in both eyes. Significant refractive error warranting spectacle wear was left to the discretion of the clinician; however this tended to be hypermetropia $\geq$1.50 DS, astigmatism $\geq$0.75 DC, myopia $\geq$0.75 DS and anisometropia $\geq$0.75 D. This does mean that it is possible for children with the same refractive error to be classified differently, i.e. true-positive or false-positive. If none of these conditions was present then the case was designated a false-positive (FP) referral and the child was discharged without further assessment.

Evaluation of diagnostic accuracy

The diagnostic accuracy of dichotomous (pass/fail) screening is typically evaluated using statistics derived from a 2 × 2 table, which represents the comparison of screening referrals (R) with classification of diagnostic status (D).

Four popular indices of diagnostic accuracy may be calculated from values in the 2 × 2 table: sensitivity (Se), specificity (Sp), positive predictive value (PPV) and negative predictive value (NPV). In operational screening programmes, of which school vision screening is a typical example, the most useful measures of diagnostic accuracy are sensitivity (Se), which may be interpreted as a measure of the detection rate for disease or anomaly, and positive predictive value (PPV), which is a measure of referral accuracy.[6] In operational screening programmes, as discussed above, the majority of those screened will give a negative result and will not be followed-up for diagnostic assessment, and therefore the number of false-negatives and true-negatives (FN and TN) will not be known. This situation poses problems for screening evaluation, since it leaves the 2 × 2 table incomplete. A pragmatic solution is to carry out follow-up assessment on a subsample of those with negative screening outcomes, in order to provide estimates of true-negative and false-negative proportions. This approach has been used in the present study, in which a random subsample comprising 11% of those not referred (R−) underwent assessment by an orthoptist and optometrist. Subjects designated true-negative (TN) underwent no further investigation, while any designated false-negative (FN) were offered appropriate referral and treatment.

Results

Visual acuity data: whole sample

The distributions of visual acuity scores measured using Snellen and Crowded logMAR tests for right and left eyes are shown in Fig. 2. Distributions of right and left eye data are similar in both plots. Note the truncated Snellen scores due to the limited range of letter sizes at the lower (high acuity) end of the measuring scale for this particular screening model of the Snellen chart (illuminated Snellen charts often have 6/4 and 6/5 lines which would prevent truncation). Results for the logMAR test, however, approximate a normal distribution with no truncation of high acuity scores. A mean against difference plot depicting difference between Snellen scores and logMAR scores for the right eye of all subjects is shown in Fig. 3. Mean difference in visual acuity was 0.0 and standard deviation 0.08 (95% confidence interval 0.15).

Diagnostic accuracy of Snellen (6/9) screening

The outcomes of screening and diagnostic assessment, using the Snellen chart with 6/9 referral criterion, are shown in Table 1. Cell values indicate the numbers of subjects in each category, with corresponding proportions given in parentheses.

A total of 68 subjects were referred using the established Snellen acuity criterion, of which 46 proved to be false-positive and 22 true-positive. Children identified as true-positives presented with a range of

Table 1. Screening outcomes following a full eye examination based on current referral criteria of 6/9 or worse in either eye on the Snellen chart

		Referral decision: + (R)	Referral decision: −	
Diagnosis (D)	+	22 (0.045)	0 (0.0000)	22 (0.045)
	−	46 (0.093)	426 (0.862)	472 (0.955)
		68 (0.138)	426 (0.862)	494 (1.000)

Assumed sensitivity (Se) = 1.000.
Positive predictive value (PPV) = 0.324.

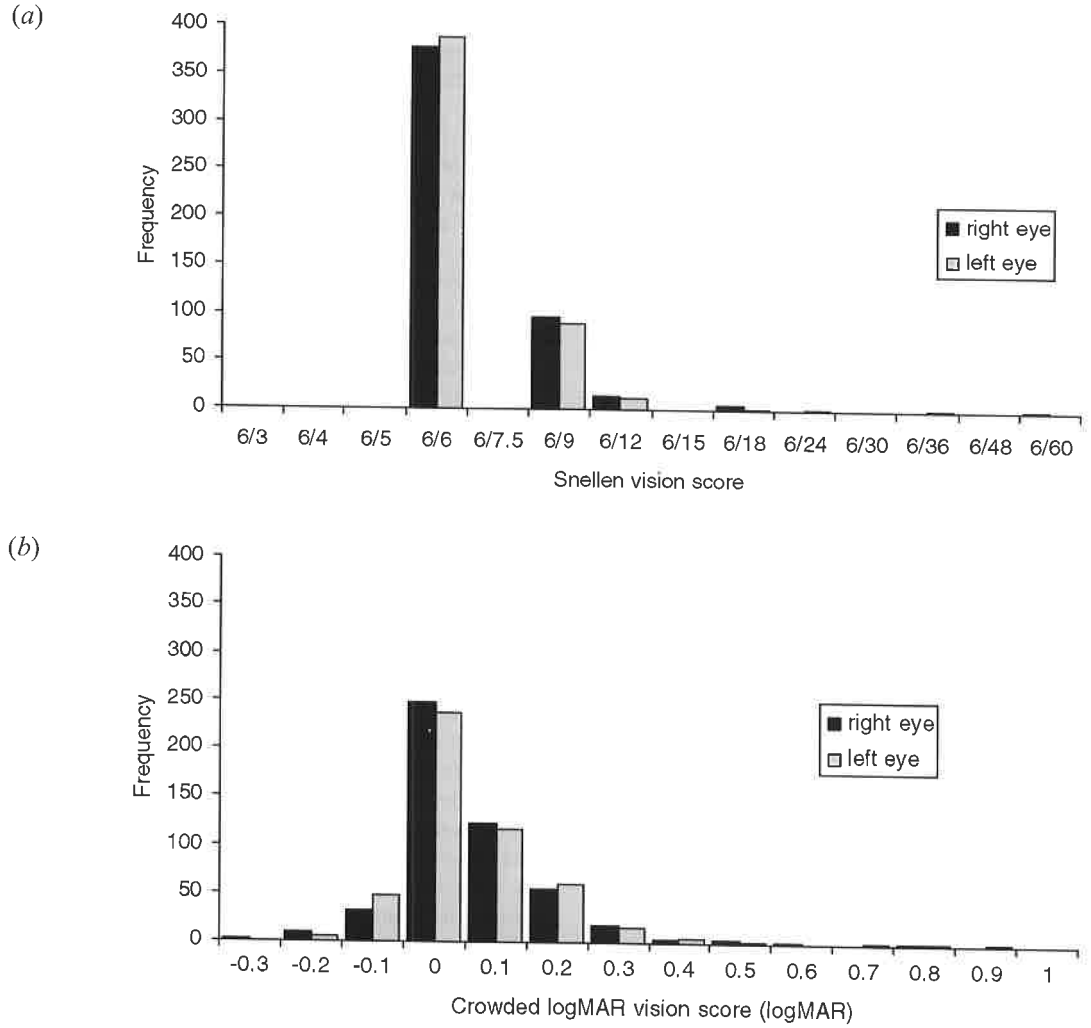

Fig. 2. Vision scores of each eye of the entire population screened ($n = 988$) measured using (*a*) Snellen chart and (*b*) Crowded logMAR test. Note: The Snellen chart has no 6/3, 6/4, 6/5, 6/7.5, 6/15, 6/30 or 6/48 line. These letter sizes are denoted to allow for direct comparison across charts.

ocular anomalies (Fig. 4). Spherical components of the more refractive eye ranged from +1.00 to +7.00, cylindrical components ranged from 0 to +1.25 DC and anisometropia ranged from 0 to 2.25 D.

Of the 426 children not referred from screening, a subsample of 47 (≈11%) were followed up for diagnostic assessment in order to obtain estimates of the proportions of true- and false-negatives. Four schools were randomly chosen from the 20 schools covered by the 6 school nurses, with the aim of obtaining a 10% sample of the 426 children not referred from screening. This follow-up revealed no cases of visual anomaly. Mean visual acuity was 0.04 in either eye with a range from −0.15 to 0.15. All children had previously passed screening and scored 6/6 in either eye. Therefore, a value of 0 is given for the number of false-negatives in Table 1, leaving 426 true-negative outcomes. Some caution is required in relation to the use of this estimate. A false-negative proportion of 0 in a sample of 47 subjects will relate to a population proportion of between 0 and 0.076 (95% CI). It is, therefore, conceivable that Snellen-based screening in the population may produce a proportion of false-negative referrals. However, we feel that there is justification in adopting the estimate of zero in the current analysis, and we provide reasons for this in the following paragraph.

Table 1 provides an estimate of 4.5% (0.045) for the prevalence of visual anomalies in 5-year-old children. This is similar to a prevalence estimate of 3.8% reported in a study of 1701 entry-level school-aged children by Bray *et al.*[7] Note that the estimated prevalence is the sum of true-positive and false-negative proportions. Therefore, if we assume a false-negative proportion greater than zero then the prevalence estimate will increase accordingly. Since the prevalence estimated by Bray *et al.*[7] is lower, and their sample considerably larger than ours, it does not appear that a prevalence of visual anomaly higher than 4.5% is likely.

Statistics of diagnostic accuracy show that the screening procedure achieves perfect sensitivity (Se = 1.00), but at the expense of a low positive predictive value (PPV = 0.324). This means that screening succeeds in referring all those who have a visual anomaly requiring treatment but also refers many who

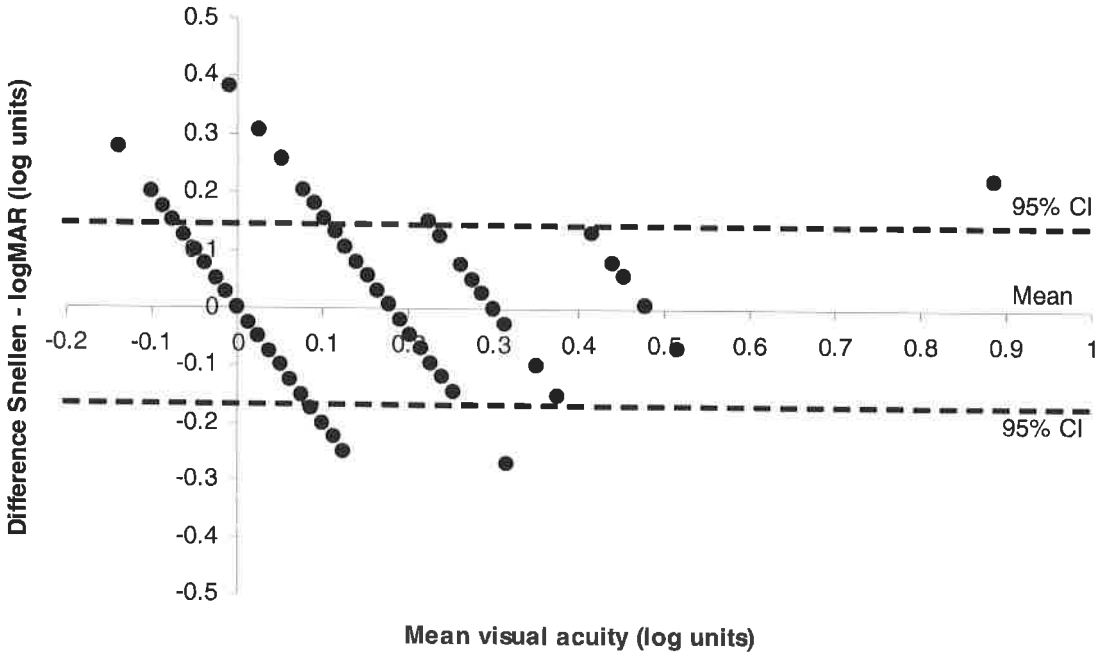

Fig. 3. Mean against difference plot of Snellen versus Crowded logMAR vision scores for the right eye only of all children ($n = 494$) Confidence intervals represent 1.96 standard deviation of differences.

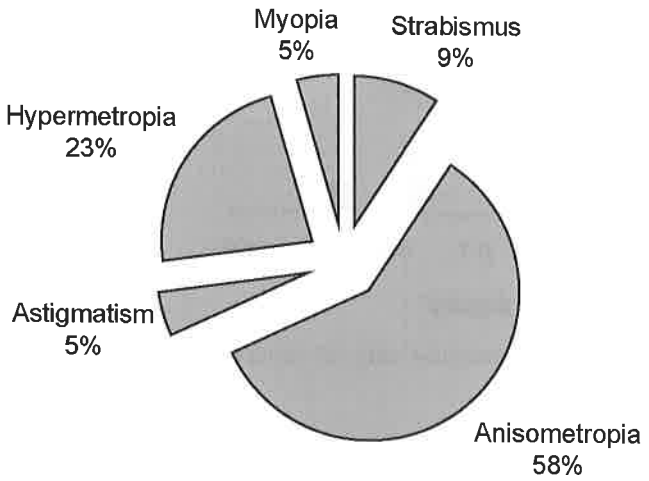

Fig. 4. Types of visual anomaly found in true-positive referrals.

do not have any anomaly. Indeed, the PPV value of 0.324 indicates that more than two-thirds (67.6%) of Snellen-based referrals are unnecessary false-positives.

Reappraising the screening criterion and test

Improvement in the diagnostic accuracy of visual acuity screening might be achieved by changing the referral criterion and/or the screening test. These two possibilities may be considered by examining the distributions of visual acuity data from the 68 children referred from screening (Fig. 5).

Fig. 5*a* shows Snellen acuity values of subjects who failed the screening test segregated as to whether they were determined to be anomalous (true-positives) or normal (false-positives). It is apparent that the majority of false-positive referrals may be avoided, at the expense of a slight increase in false-negatives, by shifting the referral criterion from 6/9 to 6/12. Fig. 5*b* shows the Crowded logMAR acuity distributions for the same subjects. Here we see that simply changing the screening test, but keeping the same referral criterion (Snellen 6/9 equivalent), has the effect of reducing the number of false-positives (change from 68% to 60%) and thus improving the PPV significantly ($\chi^2 = 11.24$, d.f. = 1, $p = 0.0008$) (change from 0.32 to 0.40; significance was tested using a score statistic[8] bearing some relation to the McNemar test for paired proportions). Many of those referred on the basis of their Snellen acuity achieved normal Crowded logMAR acuities of 0.100 log units (6/7.5), which is not a step size that appears on standard Snellen tests used in screening. In addition, each step on the logMAR abscissa represents a single letter change in acuity (0.025 log units) and it becomes apparent that a more precise indication of the balance between false-positive and false-negative referrals may be achieved by setting the criterion in terms of logMAR letters rather than Snellen lines. Results are presented in Table 2.

The optimum criterion to achieve maximum detection (or sensitivity) is the same for both Snellen (6/9) and Crowded logMAR (0.175). A referral criterion one step larger on either chart would fail to achieve 100% detection, while one step smaller would result in more false-positive referrals. Therefore the established Snellen-based criterion is accurately placed to avoid missing any cases of anomaly. We also see, however, that the Crowded logMAR test achieves this with fewer false-positive referrals.

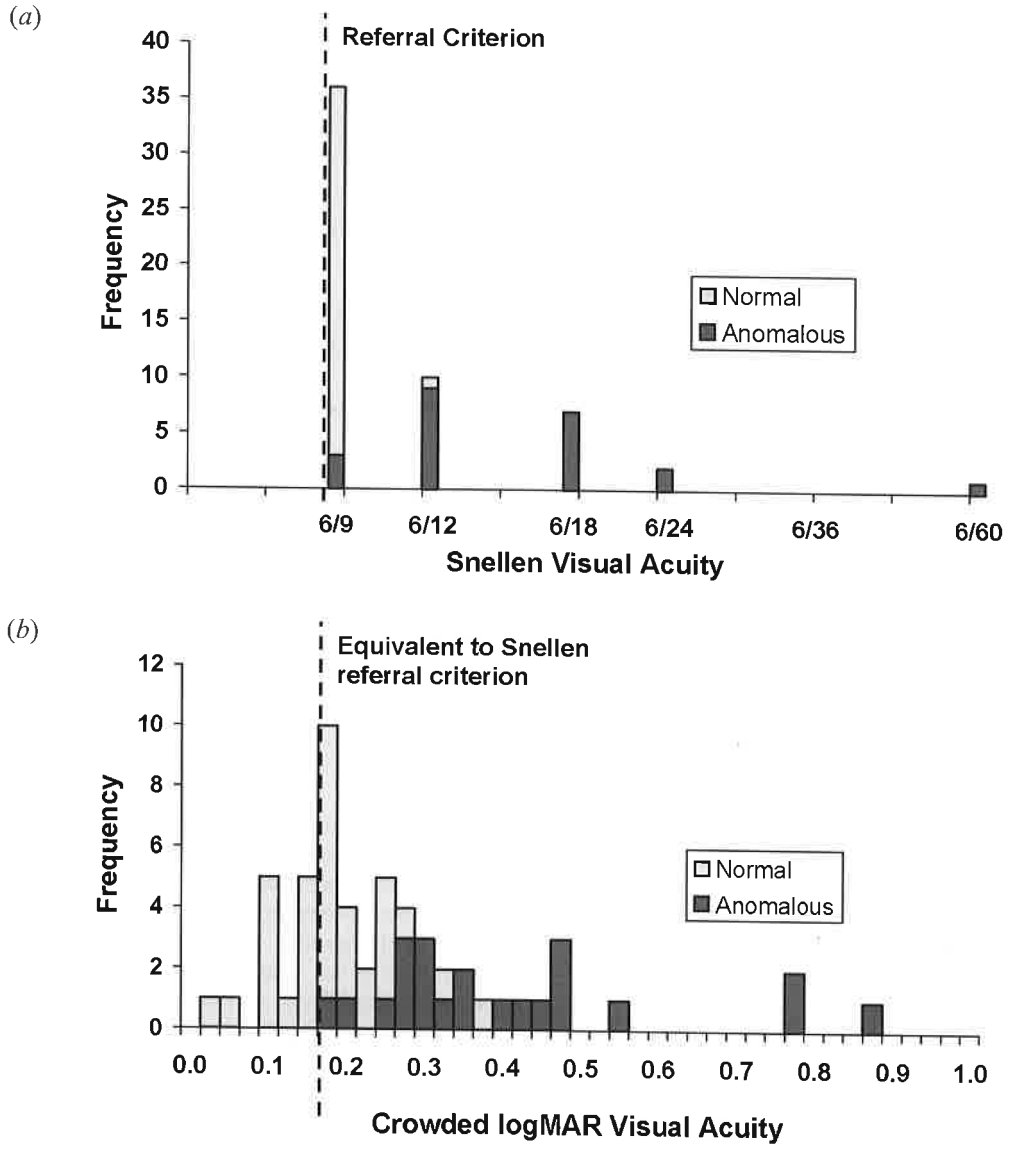

Fig. 5. Vision scores for the poorer eye of children referred from screening ($n = 68$) measured using (*a*) Snellen chart and (*b*) crowded logMAR test.

Table 2. Modelling of diagnostic accuracy of (*a*) Snellen chart and (*b*) Crowded logMAR test with different referral criteria

(*a*) Snellen chart

Criterion	Sensitivity	PPV
6/9	1.00	0.32
6/12	0.87	0.66
6/18	0.46	1.00
6/24	0.14	1.00

(*b*) Crowded logMAR test

Criterion	Sensitivity	PPV
0.175	1.00	0.40
0.200	0.96	0.48
0.225	0.91	0.51
0.250	0.91	0.54
0.275	0.86	0.61
0.300	0.73	0.67
0.325	0.59	0.68
0.350	0.55	0.75
0.375	0.46	0.83
0.400	0.46	0.91
0.500	0.18	1.00

Discussion

This study has shown that current school-entry school nurse vision screening using the Snellen test and criterion of 6/9 or worse for referral does achieve maximum sensitivity but produces a high number of false-positive referrals (68%). An improvement in the diagnostic accuracy of the procedure is desirable, both to save resources and to avoid raising unnecessary concerns in parents of children who are visually normal. We have identified two alternative strategies for reducing the number of false-positive referrals: firstly changing the visual acuity test and secondly changing the referral criterion.

With current referral criteria a child who is unable to read the 6/6 line but able to read 6/9 will be recorded as having 6/9 acuity and referred, even though testing has only established that the child's vision is less than 6/6 and greater than or equal to 6/9 (0.175 log units equivalent difference). The absence of any intermediate

grading scales (in particular a 6/7.5 line) is reflected in the high number of inappropriate referrals. However, a proportion (15%) of the visually normal referred children achieved an acuity score between 0 and 0.15 (between 6/6 and <6/9) on the Crowded logMAR test. This indicates that a change to the Crowded logMAR test would reduce the number of false-positives from 68% to 60%, improving the PPV from 32.4% to 40%. This measure would reduce false-positives without increasing the false-negative referrals.

LogMAR tests allow for a more robust, repeatable and accurate assessment of visual acuity over Snellen-based tests, largely because they allow for interpolated scoring (scoring by each letter seen), have finer and equal increment sizes of letters and present an equal demand at each acuity level (number of letters per line). The introduction of a new test for wide use in vision screening programmes will cause additional costs to the purchasers; however, existing tests will need updated at some point and now might be an opportune time to change over to a logMAR-based test. This simple measure would reduce the cost of screening to the NHS per annum.

Alternatively a change in the referral criterion of the existing Snellen chart test to 6/12 would considerably reduce the number of false-positives, improving PPV to 65.5%. However, this would result in a drop in sensitivity from 100% to 86.4% which would mean that 13.6% of anomalous children would be missed by screening.

A recent randomized controlled trial of treatment of unilateral visual impairment detected at preschool screening,[9] which adopted a referral criterion of $\leqslant$6/9 in the worse eye, concluded that children with mild visual impairment (6/9 or 6/12 measured at screening) benefit little from either conventional treatments (spectacles only, or spectacles and occlusion ('patching') therapy), and suggested that children with 6/9 in one eye and 'normal' vision in the other eye should not constitute screening failures. At this level of referral criterion the Crowded logMAR test offers little benefit over the Snellen chart in terms of detection and referral accuracy. We must bear in mind, though, that visual acuity measurements on pre-school children are less reliable, and 6/9 is within the normal range of visual acuity results on a crowded test (optotypes on a line) for the pre-school age group. However, for the age group we tested here (approximately the currently recommended age for vision screening), 6/9 would be considered on the edge of normality, with a percentage of those scoring 6/9 having a true visual deficit warranting treatment.[10]

Conclusion

Current school nurse screening criteria produce a high percentage of false-positives but no false-negatives, providing screening with low referral accuracy but maximal detection of visual anomalies. By changing visual acuity assessment to the Crowded logMAR test and adopting the equivalent logMAR referral criteria of 0.175 (6/9), referral accuracy can be improved and maximum sensitivity retained. At a time when screening programmes are being developed in line with current guidelines, logMAR-based visual acuity tests should be incorporated so that the referral criterion could be sensitively manipulated to accommodate the population to be screened.

We are grateful to the children and parents that took part in the study. We thank all the school nurses (Carol Dent, Sharon Wears, Jane Sheldon, Pat Scrivener, Miriam Heselwood and Jenny Coles) who recruited the subjects. C.E.S. is supported by a Department of Health post-doctoral award. P.V.M. is supported by a University Fellowship from the Wellcome Trust. The study was supported by a North Yorkshire Health grant to B.W. Results from this study were presented in poster format at the Association of Research into Vision and Ophthalmology (ARVO).

References

1. Hall DMB. Screening for vision defects. In: Hall DM, Elliman D (eds) *Health for All Children*, 4th edition. Oxford: Oxford Medical Publications, 2003: 230–244.
2. Bailey IL, Lovie JE. New design principles for visual acuity letter charts. *Am J Optom Physiol Opt* 1976; **53:** 740–745.
3. McGraw PV, Whitaker D, Winn B. Editorial: Reliability of the Snellen chart. *Br Med J* 1995; **310:** 1481–1482.
4. Lovie-Kitchin JE. Validity and reliability of visual acuity. *Ophthalmic Physiol Opt* 1988; **8:** 363–370.
5. Bailey IL, Bullimore MA, Raassch T, Taylor HR. Clinical grading and effects of scaling. *Invest Ophthalmol Vis Sci* 1991; **32:** 422–432.
6. McGraw PV, Winn B. Glasgow Acuity Cards: a new test for the measurement of letter acuity in children. *Ophthalmic Physiol Opt* 1993; **13:** 400–403.
7. Bray LC, Clarke MP, Jarvis SN, Francis PM, Colver A. Pre-school vision screening: a prospective comparative evaluation. *Eye* 1996; **10:** 714–718.
8. Leisenring W, Alonzo T, Pepe MS. Comparisons of predictive values of binary medical diagnostic tests for paired designs. *Biometrics* 2000; **56:** 345–351.
9. Clarke MP, Wright CM, Hrisos S, Anderson JD, Henderson J, Richardson SR. Randomised controlled trial of treatment of unilateral visual impairment detected at preschool vision screening. *Br Med J* 2003; **327:** 1251–1256.
10. Stewart CE. Comparison of Snellen and log-based acuity scores for school-aged children. *Br Orthopt J* 2000; **57:** 32–38.