1040-5488/05/8205-0432/0 VOL. 82, NO. 5, PP. 432–438
OPTOMETRY AND VISION SCIENCE

ORIGINAL ARTICLE

Sensitivity of Screening Tests for Detecting Vision In Preschoolers-targeted Vision Disorders When Specificity Is 94%

THE VISION IN PRESCHOOLERS STUDY GROUP

ABSTRACT: *Purpose.* To compare the sensitivity of 11 preschool vision screening tests administered by licensed eye care professionals for the detection of the 4 Vision in Preschoolers (VIP)-targeted vision disorders when specificity is 94%. *Methods.* This study consisted of a sample (n = 2588) of 3- to 5-year-old children enrolled in Head Start programs, 57% of whom had failed an initial Head Start vision screening. Screening results from 11 tests were compared with results from a standardized comprehensive eye examination that was used to classify children with respect to the four VIP-targeted vision disorders: amblyopia, strabismus, significant refractive error, and unexplained reduced visual acuity (VA). With overall specificity set to 94%, we calculated the sensitivity for the detection of each targeted vision disorder. *Results.* With the overall specificity set to 94%, the most accurate tests for detection of amblyopia were noncycloplegic retinoscopy (NCR) (88% sensitivity), the SureSight Vision Screener (80%), and the Retinomax Autorefractor (78%). For detection of strabismus, the most accurate tests were the MTI Photoscreener (65%), the cover–uncover test (60%), the Stereo Smile II stereoacuity test (58%), the SureSight Vision Screener (54%), and the Retinomax Autorefractor (54% in year 1, 53% in year 2). The most accurate tests for detection of significant refractive error were NCR (74%), the Retinomax Autorefractor (66%), the SureSight Vision Screener (63%), and the Lea Symbols VA test (58%). For detection of reduced VA, the most accurate tests were the Lea Symbols Distance VA test (48%), the Retinomax Autorefractor (39%), and NCR (38%). *Conclusions.* Similar to the previously reported results at 90% specificity, the screening tests vary widely in sensitivity with specificity set at 94%. The rankings of the sensitivities for detection of the 4 VIP-targeted vision disorders are similar to those with specificity set to 90%. (Optom Vis Sci 2005;82:432–438)

Key Words: sensitivity, specificity, VIP-targeted disorders, screening tests, VIP study

The Vision in Preschoolers (VIP) study is a multicenter clinical study designed to evaluate commonly used and/or commercially available preschool vision screening tests. The study is designed in three phases. The goal of phase I was to evaluate the performance of 11 preschool vision screening tests administered by licensed eye care professionals (LEPs; optometrists and ophthalmologists) experienced in working with children. Tests were evaluated for their ability to detect accurately the 4 VIP-targeted conditions: amblyopia, strabismus, significant refractive error, and unexplained reduced visual acuity (VA). In addition, the screening tests were evaluated for their ability to detect conditions categorized into three hierarchical groups by severity of condition.[1] Group 1 conditions are considered very important to detect and treat early. Group 2 conditions are considered very important to detect early (but with less urgency than group 1). Group 3 conditions are considered less urgent, but nonetheless are clinically useful to detect. All screening tests in phase I were administered to a selected population of preschool-aged children in a controlled environment. Phase II activities will evaluate test performance when tests are administered by pediatric nurses and lay screeners to a selected population of preschool-aged children in a more real-world screening environment. Finally, the goal of phase III will be to evaluate screening tests administered by real-world screening personnel to a nonselected population of preschool-aged children in realistic screening environments.

The initial results from the comparison of the performance of screening tests from phase I of the VIP study have been recently published.[1] With specificity set at 90%, the sensitivity for detecting children with ≥1 targeted conditions, with conditions categorized into three levels of severity (groups 1, 2, and 3), and with each of the four targeted conditions was reported. The sensitivity with specificity set to 94% for detecting children with ≥1 targeted conditions and hierarchical group 1 conditions was also reported, but not for each of the four targeted conditions as a result of space

DISCUSSION

This study points out high interocular corneal symmetry, which may also be helpful in routine clinical assessments. Corneal thickness, in addition to being obtained for corneal refractive surgery, has now become a routine test in glaucoma evaluation. Clinically, it may be significant to use high symmetry as a validation of accurate binocular data so that absence of high symmetry may warrant repeat clinical measurement of the eyes. An additional clinical use of this reported high symmetry in keratometry is in intraocular power calculations for a post laser-assisted *in situ* keratomileusis (LASIK) eye if only one eye was treated and preoperative data was not available.

For determination of the health and adequacy of corneal tissue for corneal procedures, topography and corneal thickness have been used to screen patients for conditions including excessively thin corneas, mild keratoconus or form fruste keratoconus, pellucid marginal degeneration, and other ectasias.[8] Mapping with the Orbscan topography system, the posterior elevation and power is readily achievable.[5] These parameters may allow earlier detection of keratoconus or identify those patients at risk for developing mild and marked keratectasia post LASIK or photorefractive keratectomy (PRK). Compilation of normative data on posterior corneal elevation may assist in detecting early keratoconus. Further study may show how the biomechanical nature of the cornea,[9, 10] including the posterior corneal elevation, may be an important variable along with the suggested 250-μm residual stromal bed[1] as risk factors in keratectasia.

In conclusion, although a large range of values exists in average simulated keratometry, central corneal corneal thickness, and posterior elevation, there was a very high interocular symmetry, an average interocular difference of 0.43 D in simulated keratometry, 8 μm in corneal thickness and 6 μm in posterior corneal elevation in this group of 121 consecutive patients. This normative data may be helpful in detecting inaccuracies in collected clinical corneal data. Further study may disclose how well variation in this high symmetry could predict abnormalities of the eye, including postoperative keratectasia and quality of visual outcomes after corneal refractive surgery.

Received October 22, 2004; accepted February 23, 2005.

REFERENCES

1. Seiler T, Koufala K, Richter G. Iatrogenic keratectasia after laser in situ keratomileusis. J Refract Surg 1998;14:312–7.
2. Rabinowitz YS. Videokeratographic indices to aid in screening for keratoconus. J Refract Surg 1995;11:371–9.
3. Joo CK, Kim TG. Corneal ectasia detected after laser in situ keratomileusis for correction of less than -12 diopters of myopia. J Cataract Refract Surg 2000;26:292–5.
4. Auffarth GU, Tetz MR, Biazid Y, Volcker HE. Measuring anterior chamber depth with Orbscan topography system. J Cataract Refract Surg 1997;23:1351–5.
5. Yaylali V, Kaufman SC, Thompson HW. Corneal thickness measurements with the Orbscan topography system and ultrasonic pachymetry. J Cataract Refract Surg 1997;23:1345–50.
6. Doughty MJ, Zaman ML. Human corneal thickness and its impact on intraocular pressure measures: a review and meta-analysis approach. Surv Ophthalmol 2000;44:367–408.
7. Bland JM, Altman DG. Measuring agreement in method comparison studies. Stat Methods Med Res 1999;8:135–60.
8. Auffarth GU, Wang L, Volcker HE. Keratoconus evaluation using the Orbscan topography system. J Cataract Refract Surg 2000;26: 222–8.
9. Andreassen TT, Simonsen AH, Oxlund H. Biomechanical properties of keratoconus and normal corneas. Exp Eye Res 1980;31:435–41.
10. Roberts C. The cornea is not a piece of plastic. J Refract Surg 2000; 16:407–13.

Elliott Myrowitz
The Wilmer Eye Institute
Refractive Eye Surgery Service
The Johns Hopkins School of Medicine
10753 Falls Rd.
Lutherville, MD 21093
e-mail: emyrowitz@jhmi.edu

constraints. Specificity levels higher than 90% are advocated by proponents of preschool vision screening programs in which the expense of confirmatory examinations is very high like in rural or remote areas.[2] Furthermore, given the current widespread use of photoscreeners, the sensitivity of the other tests for detecting each VIP-targeted disorder when specificity is the same (94%) as that of the photoscreeners is important in fully evaluating the performance of the photoscreeners. Therefore, it is of interest to know the sensitivity of these 11 vision tests for detecting each of the four VIP-targeted conditions when the specificity is set to 94%, and particularly whether the ranking of sensitivity stays the same as when the specificity was set to 90%. To provide an additional basis for selection of appropriate screening tests, this article investigates the sensitivity of the 11 vision screening tests for the detection of each targeted condition with specificity set to 94%.

METHODS

Details of the VIP study design and 11 screening tests administered in years 1 and 2 of VIP phase I have been published elsewhere[1] and are briefly described here.

Participants were children enrolled in Head Start programs collaborating with one of the five VIP Clinical Centers (Berkeley, California; Boston, Massachusetts; Columbus, Ohio; Philadelphia, Pennsylvania; and Tahlequah, Oklahoma). The number of children in the study population with vision problems was augmented by recruiting all the children who had failed and a random sample of children who had passed a regular Head Start vision screening.

All children were 3, 4, or 5 years old when screened. VIP-certified licensed eye care professionals experienced in working with children administered 11 commonly used or commercially available screening tests, including two VA tests, two stereoacuity tests, the cover–uncover test, three tests of refractive error, and three photoscreeners. They are briefly described as follows:

Visual Acuity Tests: Crowded, linear Lea Symbols and crowded, linear HOTV VA tests were each conducted at a 3-meter test distance.

Stereoacuity Tests: The Random Dot E (RDE) stereoacuity test was conducted at up to three distances (0.5, 1.0, and 1.5 m). The Stereo Smile II stereoacuity test was conducted at 40 cm.

Cover–Uncover Test: The cover–uncover test was conducted at both distance (3 m) and near (40 cm).

Refractive Error Tests: Retinoscopy (NCR), the Retinomax Autorefractor, and SureSight Vision Screener were conducted on each eye without cycloplegia.

Photoscreeners: The iScreen Photoscreener was conducted at ~68 cm with a binocular color image captured digitally for electronic transmission to the iScreen scoring center in Memphis for independent analysis. The MTI Photoscreener was conducted at 1 m with binocular black and white instant photos captured and sent to the Vanderbilt Ocular Imaging Center (VOIC) for independent analysis. The Power Refractor II video/photorefraction was conducted at ~1 m. The binocular color image was captured and analyzed by the computer system, which generated a report of refractive error magnitude (sphere, cylinder, axis) and eye alignment status (degree of deviation) used to categorize each child's screening status. Results from monocular threshold VA testing, cover testing, and cycloplegic refraction performed during a comprehensive gold standard eye examination were used to classify children with respect to the 4 VIP-targeted conditions: amblyopia, strabismus, significant refractive error, and unexplained reduced VA. The definitions of the hierarchy of four conditions are provided in Table 1. Children with more than one of the targeted conditions were included in only the one group that corresponded to the most severe condition.

For each screening test, the failure criteria were selected to maximize the overall sensitivity for detecting ≥1 of the four targeted conditions with specificity set to 94%. Children were considered as failing a screening test if they met the failure criteria for one or both eyes. Using these failure criteria, the sensitivities and their 95% confidence intervals (95% CI) for detecting each of the four targeted conditions, ≥1 targeted conditions, and the most severe (group 1) conditions were calculated. Pairwise comparisons of sensitivity at 94% specificity between screening tests from the same year were made using the McNemar chi-squared test for correlated data and using a modification of the Mantel-Haenszel procedure when children completed only one of the two screening tests.[3] Comparison of sensitivity among screening tests from different years were made using the chi-squared test of independence. Because a large number of pairwise comparisons were made between screening tests, comparisons with p value of 0.0008 (0.05/66) are considered to be statistically significant by using the conservative Bonferroni approach.

RESULTS

In years 1 and 2 of the VIP phase I, 2211 children who failed the regular Head Start vision screening and 1772 children who passed the screening were selected for enrollment. Of these, 2780 completed the VIP vision screening, 2666 children underwent gold standard examinations (GSE), and 2588 of these 2666 children completed the GSE to allow the child to be classified as having normal vision or one of the four targeted visual conditions. The frequency distribution of VIP-targeted vision disorders for year 1 and 2 is shown in Table 2. The failure criteria when specificity was set at 94% are summarized in Table 3 for the tests of VA and stereoacuity and in Table 4 for the tests involving refractive error. However, 94% specificity could not be obtained with the cover–uncover test, HOTV VA, and RDE Stereoacuity tests. The chosen failure criteria for these tests provided the specificities that were the closest to 94%. The failure criterion (refixation with removal of occluder paddle) for the cover–uncover (98% specificity) provides only one set of values for sensitivity and specificity. There are a limited number of failure criteria for the HOTV VA test (93% specificity); therefore, only a limited number of specificity values can be attained. Because of the relatively high percentage of children who could not perform the nonstereo-matching task in the RDE test, no specificity value above 92% could be attained.

The rankings and corresponding sensitivities (and their 95% CIs) of the tests for the detection of each targeted condition are shown in Table 5. For comparison, the rankings for detection of any condition and group 1 conditions are shown in Table 6. With the specificity set to 94%, the most accurate tests for detection of amblyopia were NCR (88%), the SureSight Vision Screener

TABLE 1.
Frequency of the hierarchy of VIP targeted disorders. Details of classification of targeted disorders have been published[1]

Condition	n*	%
Group 1: Very important to detect and treat early	311	12.0
Amblyopia	92	3.6
Presumed unilateral: ≥3 line interocular difference, a unilateral amblyogenic factor†, and worse eye VA ≤20/64	36	1.4
Suspected bilateral: a bilateral amblyogenic factor‡, worse eye VA <20/50 for 3 year olds or <20/40 for 4 year olds, contralateral eye VA worse than 20/40 for 3-year-olds or 20/30 for 4-year-olds	56	2.2
Strabismus: Constant in primary gaze	67	2.6
Refractive error	269	10.4
Severe anisometropia (Interocular difference >2 D hyperopia, >3 D astigmatism, or >6 D myopia)	43	1.7
Hyperopia ≥5.0 D	123	4.8
Astigmatism ≥2.5 D	146	5.6
Myopia ≥6.0 D	15	0.6
Group 2: Important to detect early	229	8.8
Amblyopia	27	1.0
Suspected unilateral: 2-line interocular difference and a unilateral amblyogenic factor†	20	0.8
Presumed unilateral: ≥3 line interocular difference, a unilateral amblyogenic factor†, and worse eye VA >20/64	7	0.3
Strabismus: Intermittent in primary gaze	29	1.1
Refractive error	202	7.8
Anisometropia, (Interocular difference >1 D hyperopia, >1.5 D astigmatism, or >3 D myopia)	64	2.5
Hyperopia >3.25 D and <5.0 D and interocular difference in SE ≥0.5 D	64	2.5
Astigmatism >1.5 D and <2.5 D	113	4.4
Myopia ≥4.0 D and <6.0 D	3	0.1
Group 3: Detection clinically useful	215	8.3
Reduced VA	173	6.7
Bilateral: no bilateral amblyogenic factor‡, worse eye VA <20/50 for 3 year olds or <20/40 for 4 year olds, contralateral eye VA worse than 20/40 for 3 year olds or 20/30 for 4 year olds	54	2.1
Unilateral: no unilateral amblyogenic factor†, worse eye VA <20/50 for 3-year-olds or <20/40 for 4-year-olds or ≥2 line difference between eyes (except 20/16 and 20/25)	119	4.6
Refractive error	51	2.0
Hyperopia >3.25 D and <5.0 D and interocular difference in SE <0.5 D	45	1.7
Myopia >2.0 D and <4.0 D	6	0.2
Normal	1,833	70.9
Total number of children	2,588	

* Children with multiple conditions appear only in the Group of the condition with the highest severity.

† Strabismus, anisometropia, and a difference in spherical equivalent of ≥0.50 D when ≥1 eye had >3.50 D of hyperopia were considered unilateral amblyogenic factors.

‡ Astigmatism of >2.50 D, hyperopia of >5.00 D, or myopia of >8.00 D in each eye were considered bilateral amblyogenic factors.

D, diopter; SE, spherical equivalent; VA, visual acuity.

(80%), and the Retinomax Autorefractor (78% in year 1, 77% in year 2). The remaining tests had sensitivities ranging from 65% (Lea Symbols VA) down to 27 to 28% (RDE and cover–uncover test). For detection of strabismus, the most accurate tests were the MTI Photoscreener (65%), the cover–uncover test (60%), the Stereo Smile II test (58%), the SureSight Vision Screener (54%), and the Retinomax Autorefractor (54% in year 1, 53% in year 2). The Power Refractor II (34%) and RDE (29%) had the lowest sensitivity for the detection of strabismus, whereas other screening tests ranged in sensitivity between 44% and 50%. The most accurate tests for detection of significant refractive error were NCR (74%), Retinomax Autorefractor (66% in year 1, 63% in year 2), SureSight Vision Screener (63%), and Lea Symbols VA (58%). For detection of reduced VA, all the screening tests had sensitivities below 50%, with the most accurate tests being Lea Symbols VA (48%), Retinomax Autorefractor (39% in year 1, 36% in year 2), NCR (38%), HOTV VA (36%), and the SureSight Vision Screener (35%). NCR, the Retinomax Autorefractor, the SureSight Vision Screener, and Lea Symbols VA were best for detection of both any condition and group 1 conditions.

The p values from pairwise comparison of sensitivities (at 94% specificity) of 11 tests for the detection of each targeted condition

TABLE 2.
The frequency distribution of the 4 VIP-targeted vision disorders in years 1 and 2

VIP-targeted vision Disorders	Year 1 (N = 1,142)	Year 2 (N = 1,446)
Amblyopia	75 (6.57)	88 (6.09)
Reduced VA	132 (11.6)	114 (7.88)
Strabismus	48 (4.20)	62 (4.29)
Significant refractive error	240 (21.0)	299 (20.7)
Any Condition	346 (30.3)	409 (28.3)
Group 1 conditions	139 (12.2)	172 (11.9)
No vision disorders	796 (69.7)	1037 (71.7)

All values are n (%). The percentages do not sum to 100% because each child can have more than one vision disorder.

TABLE 3.
Failure criteria for visual acuity and stereoacuity tests to maximize sensitivity when specificity was set at 0.94§

Test	Age (years)	Failure criterion* (Inability to pass)
Lea Symbols™ visual acuity	3	10/32 line
	4	10/25 line
	5	10/20 line
HOTV visual acuity	3	10/32 line
	4	10/32 line
	5	10/25 line
Random Dot E stereoacuity	3	Non-stereo card
	4	Non-stereo card
	5	Stereo card at 50 cm (550 arc sec)
Stereo Smile II stereoacuity	3	480 arc sec card
	4	480 arc sec card
	5	240 arc sec card

* Failure criteria were chosen to maximize overall sensitivity for detecting any targeted condition when specificity was set to 0.94.

§ 0.94 specificity cannot be reached for HOTV VA and Random Dot E, because their measurement is discrete. The failure criteria providing specificity closest to 0.94 were chosen for HOTV VA (0.93 specificity) and for Random Dot E (0.92 specificity).

are not shown. Although p values are dependent on the number of children with each targeted condition and the correlation between screening tests, generally speaking, a 0.25 difference in sensitivity in detecting amblyopia, 0.30 difference in detecting strabismus, 0.15 difference in detecting significant refractive error, and 0.20 difference in detecting reduced VA can be considered to be statistically significant after adjustment for multiple comparisons.

DISCUSSION

Because the refractive error screening tests have multidimensional failure criteria (amount of hyperopia, myopia, astigmatism, and anisometropia), the screening tests cannot be assessed using standard receiver operator characteristic (ROC) curves. The comparison of sensitivities between tests with specificity set at 90% has been previously reported in the initial report of the VIP study.[1] Because an equitable comparison of sensitivities between tests requires tests to be set at a common specificity, the specificity of 94% was chosen in this report to allow comparison of test performance with two of the screening tests (iScreen and MTI photoscreeners) that have established failure criteria, which achieved a specificity of 94% (a specificity remarkably similar to that independently established and published by Donahue et al. for use in vision screening).[4] This article investigates the tests' sensitivities for the detection of four VIP-targeted conditions at 94% specificity. For three tests that could not achieve 94% specificity, specificity was set as close as possible to 94% (cover–uncover test [98%], HOTV VA [93%], and RDE Stereoacuity [92%]).

Similar to 90% specificity, the failure criteria at 94% specificity for VA and stereoacuity tests are age-dependent.[1] Although 19.7% of the total population was 3 years old at the time of the VIP screening, only 1.4% of the population was between 36 and 41 mo; therefore, it is possible that the skewed distribution (toward older 3 years olds) of 3 year olds in this study population may have influenced the failure criteria and perhaps sensitivity for 3 year olds in this study. It has been reported that a difference in testability and perhaps sensitivity exists between younger and older 3 year olds.[5]

When compared with the sensitivities with specificity set to 90%, the sensitivities for detecting each of the four targeted conditions generally drop for each screening test when the specificity for all tests is set to 94%. However, the tests with the best performance in detecting the four targeted conditions are largely the same as those found with a specificity of 90%. At 94% specificity, the best tests for the detection of amblyopia, significant refractive error, any condition, and group 1 conditions were NCR, the two autorefractors (Retinomax and SureSight Vision Screener), and Lea Symbols VA. The same tests plus HOTV VA performed best for the detection of unexplained reduced VA, although all the screening tests had sensitivities below 50% for the detection of reduced VA. This supports the results presented in the previous report,[1] which showed that the tests with the best sensitivities for detecting group 1 conditions were similar regardless of whether specificity was set at 90% or 94%.

Identification of the best tests for the detection of strabismus is limited by the relatively low number of children with strabismus (48 in year 1 and 62 in year 2). With specificity set at 94%, there are no statistically significant differences in sensitivity among the tests except that the Power Refractor II and RDE performed significantly worse than the MTI. Of note, increasing the specificity from 90% to 94% decreased the sensitivity of the RDE test from 60% to 29%. Because a high percentage of children without any targeted conditions were unable to successfully complete the pretest (identifying the nonstereo E),[1] the failure criterion had to be set to failing the nonstereo pretest for 3 and 4 year olds to attain close to 94% specificity for the entire group. Thus, children could "pass" the RDE test by merely being able to do the nonstereo-matching task, thereby precluding the detection of eye misalignment through detection of low stereoacuity. Not surprisingly, the sensitivity of the RDE test was very low for detecting strabismus at 94% specificity.

Because the difficulty demonstrated by 3 and 4 year olds in responding to the RDE may have contributed to the poor performance of the RDE test, we also analyzed the data for 5 year olds

TABLE 4.
Failure criteria for retinoscopy, photorefraction, and autorefractor screening tests to maximize sensitivity when specificity was set at 0.94

Instrument	Hyperopia	Myopia	Astigmatism	Anisometropia*
Non-cycloplegic retinoscopy	≥ 2.50 D	≥ 2.75 D	≥ 2.00 D	≥ 1.50 D
Power Refractor II	≥ 5.00 D	≥ 3.75 D	≥ 2.25 D	≥ 2.75 D
SureSight Vision Screener**	≥ 4.25 D	≥ 1.00 D	≥ 1.75 D	≥ 3.50 D
Retinomax autorefractor				
Year 1	≥ 1.75 D	≥ 2.75 D	≥ 2.00 D	≥ 2.75 D
Year 2	≥ 2.50 D	≥ 2.75 D	≥ 1.75 D	≥ 2.50 D

* The maximum of inter-eye differences in the power of the most positive meridian, the most negative meridian, and the magnitude of cylinder was used to determine presence of anisometropia for all tests.
** Used in child mode, which adds a correction for accommodation.

TABLE 5.
Ranking of tests from highest to lowest sensitivity in detecting the 4 VIP-targeted disorders with specificity set at 0.94§

Ranking	Amblyopia		Strabismus		Refractive error		Reduced VA	
	Screening tests	Sensitivity (95% CI)*	Screening tests	Sensitivity (95% CI)*	Screening tests	Sensitivity (95% CI)*	Screening tests	Sensitivity (95% CI)*
1	NCR	.88 (.81, .95)	MTI	.65 (.53, .77)	NCR	.74 (.68, .80)	Lea VA	.48 (.39, .57)
2	SureSight	.80 (.72, .88)	Cover-uncover	.60 (.46, .74)	Retinomax Y1	.66 (.60, .72)	Retinomax Y1	.39 (.31, .47)
3	Retinomax Y2	.78 (.69, .87)	Stereo smile II	.58 (.46, .70)	Retinomax Y2	.63 (.58, .68)	NCR	.38 (.30, .46)
4	Retinomax Y1	.77 (.67, .87)	Retinomax Y1	.54 (.40, .68)	SureSight	.63 (.58, .68)	HOTV VA	.36 (.28, .44)
5	Lea VA	.65 (.54, .76)	SureSight	.54 (.42, .66)	Lea VA	.58 (.52, .64)	Retinomax Y2	.36 (.27, .45)
6	MTI	.63 (.53, .73)	Retinomax Y2	.53 (.41, .65)	iScreen	.43 (.37, .49)	SureSight	.35 (.26, .44)
7	iScreen	.62 (.52, .72)	iScreen	.50 (.38, .62)	MTI	.42 (.36, .42)	iScreen	.27 (.19, .35)
8	Stereo Smile II	.61 (.51, .71)	NCR	.50 (.36, .64)	Power refractor	.42 (.36, .48)	Power refractor	.27 (.19, .35)
9	Power Refractor	.57 (.47, .67)	Lea VA	.48 (.34, .62)	HOTV VA	.40 (.34, .46)	MTI	.24 (.16, .32)
10	HOTV VA	.52 (.41, .63)	HOTV VA	.44 (.30, .58)	Stereo smile II	.37 (.32, .42)	Random dot E	.24 (.17, .31)
11	Random Dot E	.28 (.18, .38)	Power refractor	.34 (.22, .46)	Random dot E	.23 (.18, .23)	Stereo smile II	.20 (.13, .27)
12	Cover-uncover	.27 (.17, .37)	Random dot E	.29 (.16, .42)	Cover-uncover	.16 (.11, .21)	Cover-uncover	.06 (.02, .10)

* The screening tests with tied sensitivities are listed alphabetically. A 0.25 difference in sensitivity in detecting amblyopia, 0.30 difference in detecting strabismus, 0.15 difference in detecting significant refractive error, and 0.20 difference in detecting reduced VA can be considered to be statistically significant.
§ 94% specificity cannot be achieved for HOTV VA, Random Dot E, and cover-uncover test. The current sensitivity comparisons were based on 93% specificity for HOTV VA, 92% specificity for Random Dot E, and 98% specificity for cover-uncover test. Their sensitivities for cover-uncover test, iScreen and MTI photoscreeners have been previously reported.[1]

only. When the failure criterion was inability to pass the stereo card at 50 cm, the specificity was 95%, the sensitivity was 32% for detecting ≥1 targeted conditions, 49% for group 1 conditions, 52% for amblyopia, 65% for strabismus, 34% for refractive error, and 30% for reduced VA. Thus, even for the older children, the RDE test still had relatively low sensitivity for detecting children with any targeted condition or with a group 1 condition. However, the sensitivity for detecting children with strabismus was comparable to the best of the other screening tests.

Although it is not surprising that the test that performs best for the detection of a specific targeted disorder may not perform best for other targeted conditions, it is interesting to note that the MTI Photoscreener, the cover–uncover test, and the Stereo Smile II test are among the best tests for the detection of strabismus but are not among the best tests to detect ≥1 targeted conditions or group 1 conditions (NCR, Retinomax Autorefractor, SureSight Vision Screener, and Lea Symbols VA). Furthermore, the best tests to detect ≥1 targeted conditions or group 1 conditions are also the best tests for detecting amblyopia, significant refractive error, and reduced VA. This can be attributed to the relatively low number of children with strabismus and the fact that both strabismus and strabismic amblyopia are frequently associated with significant refractive error.[1,6,7]

As reported in the initial results of VIP study–phase I,[1] the performance of screening tests varies widely in detecting the four targeted disorders. Overall, the best tests for the detection of amblyopia, significant refractive error, any condition, or group 1 conditions were NCR, the two autorefractors (Retinomax and SureSight Vision Screener), and Lea Symbols VA, the same four tests that performed best with specificity set at 90%.

ACKNOWLEDGMENTS

This study is supported by grants from the National Eye Institute, National Institutes of Health, Department of Health and Human Services, Bethesda,

TABLE 6.
Ranking of tests from highest to lowest sensitivity in detecting any condition and Group 1 conditions with specificity set at 0.94§

Ranking	Any condition		Group 1 conditions*	
	Screening tests	Sensitivity (95% CI)	Screening tests	Sensitivity (95% CI)
1	NCR	.57 (.52, .62)	NCR	.90 (.85, .95)
2	Retinomax Y1	.52 (.47, .57)	Retinomax Y1	.87 (.81, .93)
3	Retinomax Y2	.52 (.47, .57)	Retinomax Y2	.81 (.75, .87)
4	SureSight	.51 (.46, .56)	SureSight	.75 (69, .81)
5	Lea VA	.49 (.44, .54)	Lea VA	.65 (.57, .73)
6	iScreen	.37 (.32, .42)	iScreen	.57 (.50, .64)
7	MTI	.37 (.32, .42)	Stereo smile II	.57 (.50, .64)
8	HOTV VA	.36 (.31, .41)	Power refractor	.56 (.49, .63)
9	Power refractor	.36 (.31, .41)	MTI	.55 (48, .62)
10	Stereo smile II	.33 (.28, .38)	HOTV VA	.48 (.40, .56)
11	Random dot E	.22 (.18, .26)	Random dot E	.30 (.22, .38)
12	Cover-uncover	.16 (.12, .20)	Cover-uncover	.24 (.17, .31)

The screening tests with tied sensitivities are listed alphabetically.

§ 94% specificity cannot be reached for HOTV VA, Random Dot E, and cover-uncover test. The current sensitivity comparisons were based on 93% specificity for HOTV VA, 92% specificity for Random Dot E, and 98% specificity for cover-uncover test.

* Conditions very important to detect and treat early.

Maryland: U10EY12534, U10EY12545, U10EY12547, U10EY12550, U10EY12644, U10EY12647, and U10EY12648.

The Vision In Preschoolers Study Group:

Executive Committee: *Paulette Schmidt OD, MS (Chair); Agnieshka Baumritter, MA; Elise Ciner, OD; Lynn Cyert, PhD, OD; Velma Dobson, PhD; Beth Haas; Marjean Taylor Kulp, OD, MS; Maureen Maguire, PhD; Bruce Moore, OD; Deborah Orel-Bixler, PhD, OD; Ellen Peskin, MA; Graham Quinn, MD; Maryann Redford, DDS, MPH; and Janet Schultz, RN, MA, CPNP; Gui-Shuang Ying, PhD.*

Advisory Committee (voting members): *Roy Beck, MD, PhD (Chair); Susan Cotter, OD; Jonathon Holmes, MD; Maureen Maguire, PhD; Maryann Redford, DDS, MPH; Melvin Shipp, OD, MPH, DrPH; and Sheila West, PhD.*

Writing Committee: *Gui-Shuang Ying, PhD (Chair); Marjean Taylor Kulp, OD, MS; Maureen Maguire, PhD; Elise Ciner, OD; Lynn Cyert, PhD, OD; and Paulette Schmidt, OD, MS.*

Participating Centers: *AA, Administrative Assistant; BPC, Backup Project Coordinator; GSE, Gold Standard Examiner; LS, Lay Screener; LPS, LEP Screener; PI, Principal Investigator; PC, Project Coordinator; PL, Parent Liaison; PR, Programmer; VD, Van Driver; NHE, Nurse/Health Coordinator.*

Berkeley, California, University of California at Berkeley School of Optometry: *Deborah Orel-Bixler, PhD, OD (PI/GSE/LPS); Pamela Qualley, MA (PC); Dru Howard (BPC/PL); Sarah Fisher, OD (GSE/LPS); Darlene Fong, OD (GSE); Cindy Hsiao, OD (GSE); Selim Koseoglu, MD (GSE); A. Mika Moy, OD (GSE); Sharyn Shapiro, OD (GSE); Lisa Verdon, OD (GSE); Tonya Watson, OD (GSE); Sara Frane, OD (LPS); Nina Friedman, OD (LPS); Jennifer Seino, OD (LPS); Sean McDonnell, (LS); Erika Paez (LS); Coriemae Perea (LS); Darlene Sloan (LS); Evelyn Smith (LS); Leticia Soto (LS); and Angela Stelly-Leonard (LS/NHE).*

Boston, Massachusetts, New England College of Optometry: *Bruce Moore, OD (PI GSE/LPS); Joanne Bolden (PC); Sandra Umaña (PC); Justin Smith (BPC); Nicole Quinn, OD (GSE/LPS); Nancy Carlson, OD (GSE/LPS); Melissa Suckow, OD (GSE/LPS); Amy Croteau, OD (GSE); Barry Kran, OD (GSE); Jean Ramsey, MD (GSE); Erik Weissberg, OD (GSE); Daniel Kurtz, PhD, OD (LPS); Daniel Laby MD (LPS); Stacy Lyons, OD (LPS); Marthedala Chery (LS/PL); Leticia Gonzalez (LS/PL); Edward Braverman (LS/VD); Susan Crowley (NHE); Paul Dennehy (VD); and Benny Jaramillo (VD).*

Columbus, Ohio, The Ohio State University College of Optometry: *Paulette Schmidt OD, MS (PI); Marjean Taylor Kulp, OD, MS (Co-Investigator, GSE/LPS); Molly Biddle (PC); Jo Haynes (PC); Jason Hudson (BPC); Kristyne Edwards, OD, MS (GSE/LPS); Heather Gebhart, OD (GSE/LPS); Ann Hickson, OD, MS (GSE/LPS); LeVelle Jenkins, OD (GSE/LPS); Sandra Anderson, OD (GSE); Nancy Evans, OD (GSE); Jay Henry, OD, MS (GSE); Richard Hertle, MD (GSE); Jeffrey Hutchinson, DO (GSE); Andrew Toole, OD, MS (GSE); Michael Earley, OD, PhD (LPS); Sherry Crawford, OD, MS (LPS); Kathy Reuter, OD (LPS); Keith Johnson (LS/VD); Beth Haas (LS), Tonya James (LS); Denise Martin (LS); Sandra Dorton (NHE); Youlanda Grace (NHE); Trina Hisle (NHE); Cheryl Jones (NHE); Betty Smith (NHE); and Robert Bower (VD).*

Philadelphia, Pennsylvania, Pennsylvania College of Optometry: *Elise Ciner, OD (PI, GSE/LPS); Angela Duson (PC/LS); Lydia Parke (BPC); Mark Boas, OD (GSE/LPS); Shannon Burgess, OD (GSE/LPS); Penelope Copenhaven, OD (GSE/LPS); Ellie Francis, PhD; OD (GSE/LPS); Michael Gallaway, OD (GSE/LPS); Graham Quinn, MD (GSE/LPS); Janet Schwartz, OD (GSE/LPS); Brandy Scombordi-Raghu, OD (GSE/LPS); Janet Swiatocha, OD (GSE/LPS); Edward Zikoski, OD (GSE/LPS); Jennifer Lin, MD (GSE); Sheryl Menacker, MD (GSE); Rose Little (LS/PL); Geneva Moss (LS/PL); Jose Figueroa (LS/VD); Barbara Hall (LS); Eric Nesmith (LS); Gwen Gold (BPC; NHE;PL); Deborah Ciner (PL); Elizabeth Jordan (PL); and David Harvey (VD).*

Tahlequah, Oklahoma, Oklahoma Northeastern State University College of Optometry: *Lynn Cyert, PhD, OD (PI, GSE/LPS); Linda Cheatham (PC/VD); Anna Chambless (BPC; PL); Colby Beats, OD (GSE); Debbie Coy, OD (GSE); Jeffrey Long, OD (GSE); Jerry Carter, OD (GSE); Shelly Rice, OD (GSE); James Dunn, OD (LPS); Elisabeth Harrington, OD (LPS); Leslie Trimble, OD (LPS); Shelly Dreadfulwater, (LS/PL); Cindy McCully (LS/PL); Rod Wyers (LS; VD); Edith Bingham (LS); Vicky Taylor (LS); Glenda Byfield (PL); Pat Gower (NHE); Kathryn Roastingear (NHE); and Elizabeth Ross (NHE).*

Study Center: Columbus, Ohio, The Ohio State University College of Optometry: *Paulette Schmidt, OD, MS (PI); and Beth Haas (Study Coordinator).*

Coordinating Center: Philadelphia, Pennsylvania, University of Pennsylvania, Department of Ophthalmology: *Maureen Maguire, PhD (PI); Ellen Peskin, MA (Project Director); Agnieshka Baumritter, MA (Projector Director); Mary Brightwell-Arnold (Systems Analyst); Christine Holmes (AA); Andrew James (PR); Aleksandr Khvatov (PR); Lori O'Brien (AA); Claressa Whearry (AA); and Gui-Shuang Ying, PhD (Biostatistician).*

National Eye Institute, Bethesda, Maryland: *Maryann Redford, DDS, MPH.*

Received September 24, 2004; accepted February 22, 2005.

REFERENCES

1. The Vision In Preschoolers Study Group. Comparison of preschool vision screening tests as administered by licensed eye care professionals in the Vision In Preschoolers Study. Ophthalmology 2004;111: 637–50.
2. Arnold RW. Highly specific photoscreening at the Alaska State Fair. Valid Alaska Blind Child Discovery photography and interpretation. Alaska Med 2003;45:34–40.
3. Duffy SW, Rohan TE, Altman DG. A method for combining matched and unmatched binary data. Application to randomized, controlled trials of photocoagulation in the treatment of diabetic retinopathy. Am J Epidemiol 1989;130:371–8.
4. Donahue SP, Arnold RW, Ruben JB and the AAPOS Vision Screening Committee. Preschool vision screening: what should we be detecting and how should we report it? Uniform guidelines for reporting results of preschool vision screening studies. J AAPOS 2003;7:314–6.
5. Woodruff ME. Observations on the visual acuity of children during the first five years of life. Am J Optom Arch Am Acad Optom 1972; 49:205–15.
6. Atkinson J, Braddick O, Robier B, et al. Two infant vision screening programmes: prediction and prevention of strabismus and amblyopia from photo- and videorefractive screening. Eye 1996;10:189–98.
7. Repka MX, Beck RW, Holmes JM, et al. A randomized trial of patching regimens for treatment of moderate amblyopia in children. Arch Ophthalmol 2003;121:603–11.

The Vision In Preschoolers Study Group
The Ohio State University
College of Optometry
320 West Tenth Avenue
PO Box 182342
Columbus, OH 43218-2342