Prediction of Early-Onset Esotropia From Components of the Infantile Squint Syndrome

Clifton M. Schor,* Robert E. Fusaro,*† Nance Wilson,†† and Suzanne P. McKee‡

Purpose. To examine the association between components of the infantile squint syndrome (ISS) and age of onset of esotropia among subjects in the Cooperative Amblyopia Classification Study (CACS).

Methods. Fifty subjects were classified retrospectively as having early-onset esotropia (EOE) and 150 subjects were classified as having late-onset esotropia (LOE), depending on whether symptoms of (or treatment for) strabismus occurred before the first birthday or between the first and ninth birthdays, respectively. The authors compared the degree to which latent nystagmus (LN), dissociated vertical deviation (DVD), monocular asymmetry of optokinetic nystagmus (MOKN), monocular asymmetry of smooth pursuit (MSP), and perceived monocular speed bias (MSB) predicted EOE.

Results. Slow-phase velocity of MOKN and MSP were faster in response to nasal than to temporal target motion. In contrast, MSB revealed that targets of equal velocity were perceived as moving faster temporally than nasally. The authors evaluated MOKN, MSP, and MSB as dichotomous and as continuous predictors. Dichotomous analysis showed significant associations between DVD and asymmetries of MOKN in the preferred eye of subjects with EOE. Univariate logistic regression models, based on DVD and LN as well as on continuous measures of MOKN, MSP, and MSB, revealed predictive power for all ISS components except LN. In the preferred eye, MSP asymmetry was the strongest single predictor of EOE; multivariate analysis revealed that prediction of EOE improved with the inclusion of DVD.

Conclusions. Multivariate analysis indicated that dichotomous measures of DVD and continuous measures of MSP were independent predictors of EOE in a population of 8- to 40-year-old subjects with strabismus. In the preferred eye, MOKN asymmetry was predictive of EOE in the absence of information about MSP. Predictive values of all ISS components depended heavily on the baseline prevalence of EOE in the target population. Invest Ophthalmol Vis Sci. 1997;38:719-740.
velocity discrimination provides feedback for MSP and MOKN, it has been suggested that motion asymmetry could be the basis of the horizontal motor asymmetries that characterize ISS.11

No comprehensive, population-based studies of the ISS components have been reported. Studies of large clinical populations of EOE have investigated the prevalence of certain ISS components, such as MOKN6,7,18,19 or DVD with LN.5,30-32 Several small studies have documented MSP asymmetries in subjects with amblyopia, esotropia, or both.9-11,23-25 In addition, several small studies of motion sensitivity in subjects with strabismus, amblyopia, or both, yielded mixed results concerning the presence of motion asymmetries and their direction26-29; however, large population-based studies of asymmetric MSP and MSB, and the prevalence of these conditions in association with strabismus, have yet to be conducted.

Information about age of onset is useful in forming a prognosis for functional correction of strabismus when it is associated with amblyopia in adults.30 Uncorrected amblyopia interferes with the binocular sensory feedback necessary for maintaining binocular eye alignment. Accordingly, the prognosis for functional correction of strabismus in adults hinges on the prognosis for functional correction of amblyopia, which is greater for binocular disorders of late onset than for early onset. The association between EOE and various components of the ISS can be used to estimate retrospectively the age of onset of esotropia when it is unknown.31

In this study, we investigated the degree to which each of the five ISS components predicted EOE in 200 subjects with strabismus from the Cooperative Amblyopia Classification Study (CACS)31 using two data analysis approaches. In the first analysis, we categorized continuous ISS measurements as normal or abnormal based on the 95th percentile of nonpreferred eye measurements of a normal control group from CACS (n = 69). We estimated the positive and negative predictive values for EOE of each ISS component. Although such dichotomous prediction rules are convenient, they may be inefficient and even misleading given the crucial role of the diagnostic cutoff. Therefore, in the second analysis, we used logistic regression to calculate MOKN, MSP, and MSB as continuous measurements, without dichotomization, in univariate and multivariate prediction models of EOE. Analysis revealed that DVD, MSP, and MOKN were strong predictors of EOE. Predictive values of any ISS component ultimately depended on the prevalence of EOE in the target population. Our results address the problem of forming a prognosis for functional correction of strabismus in subjects with a history of amblyopia who ranged in age from 8 to 40 years.

SUBJECTS AND METHODS

We evaluated the data base records of 200 subjects with strabismus and 69 normal control subjects recruited for the CACS.31 We retrospectively classified 50 of the subjects with esotropia as having early-onset esotropia and 150 subjects as having late-onset esotropia, depending on whether a history was reported of symptoms of (or treatment for) strabismus during the first year or between the first and ninth birthdays, respectively. During childhood, the strabismus could have been associated with amblyopia, as indicated by a history of monocular patching therapy. In the current clinical evaluation, amblyopia was defined as a best-corrected acuity that was worse than or equal to 20/40 in one eye. For 19 additional subjects with strabismus in the CACS data base, age of onset was indeterminate; hence, these subjects were not included in our study. Ocular histories were obtained from medical records and from parents and subjects. Because subjects were recruited primarily from public address advertisements, age of onset was estimated from patient- and parent-reported histories. Ninety-six percent of the early-onset group, and 89% of late-onset group, either had or had recovered from amblyopia at the time of their participation in this study. Those who recovered from amblyopia had a history of patching and a best-corrected visual acuity of better than 20/40 in both eyes. Anisometropia of at least 2 D was observed in 26% of the early-onset group and in 42% of the late-onset group. All 200 subjects with strabismus had undergone some form of treatment for strabismus and amblyopia (if present). Known treatments included surgery in 101 subjects and patching in 173 subjects. Strabismus was classified as accommodative if the horizontal angle of strabismus, as measured by alternate cover test, was classified in a higher esotropia or a lower exotropia category at the near rather than at the far test distance. Generally, if the angle of strabismus increased at the near distance by one diagnostic category (the near angle of deviation was at least 3 base out prism dipters greater than the far angle of deviation), the AC/A ratio was at least 7Δ/1 D. Age of onset was determined during the clinical examination. Laboratory tests were conducted by a different examiner who had no knowledge of the age of onset. Subjects were informed fully about the nature of the procedures, which were performed according to the guidelines of the Declaration of Helsinki, and gave their written consent before the beginning of the experiment.

Table 1 lists the prevalence of the various categories of strabismus present at the time of examination in the CACS study for early- and late-onset groups. Constant esotropia was more prevalent in EOE than it was in LOE; corrected strabismus was more preva-
All subjects enrolled in the study underwent clinical examination of refractive error measured under dry and cycloplegic conditions; visual acuity evaluated with Bailey–Lovie LogMar test with the best optical correction; horizontal and vertical angles of deviation quantified with prism-cover test; DVD manifest by cover test; latent nystagmus observed with cover test and visuoscopy; and ocular health (assessment of the optic disks and maculae). The presence of DVD and strabismus were evaluated using standard criteria. During unilateral and alternating cover tests, each eye was occluded for at least 5 seconds. Latent nystagmus was diagnosed if a conjugate jerk nystagmus was present on covering either eye with the slow phase directed toward the covered eye. Latent nystagmus was not diagnosed if a temporal drift of the unoccluded-fixating eye was present on covering either eye. The preferred eye of subjects with and without strabismus was defined using a monocular sighting task of a distant target viewed through a small aperture held near the subject’s face.

Clinical Testing Procedure

All subjects enrolled in the study underwent clinical examination of refractive error measured under dry and cycloplegic conditions; visual acuity evaluated with Bailey–Lovie LogMar test with the best optical correction; horizontal and vertical angles of deviation quantified with prism-cover test; DVD manifest by cover test; latent nystagmus observed with cover test and visuoscopy; and ocular health (assessment of the optic disks and maculae). The presence of DVD and strabismus were evaluated using standard criteria. During unilateral and alternating cover tests, each eye was occluded for at least 5 seconds. Latent nystagmus was diagnosed if a conjugate jerk nystagmus was present on covering either eye with the slow phase directed toward the covered eye. Latent nystagmus was not diagnosed if a temporal drift of the unoccluded-fixating eye was present on covering either eye. The preferred eye of subjects with and without strabismus was defined using a monocular sighting task of a distant target viewed through a small aperture held near the subject’s face.

TABLE 1. Qualitative Characteristics of Subjects With Strabismus in CACS Sample of EOE and LOE

<table>
<thead>
<tr>
<th>Strabismus</th>
<th>Early Onset (%)(n = 50)</th>
<th>Late Onset (%)(n = 150)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unilateral</td>
<td>82</td>
<td>72</td>
</tr>
<tr>
<td>Alternating</td>
<td>12</td>
<td>10</td>
</tr>
<tr>
<td>Intermittent</td>
<td>6</td>
<td>10</td>
</tr>
<tr>
<td>Constant</td>
<td>74</td>
<td>54.7</td>
</tr>
<tr>
<td>Accommodative</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(high AC/A)</td>
<td>14</td>
<td>16.6</td>
</tr>
<tr>
<td>Esotropic</td>
<td>64</td>
<td>46.7</td>
</tr>
<tr>
<td>Exotropic</td>
<td>30</td>
<td>34.7</td>
</tr>
<tr>
<td>Orthotropic-recovered</td>
<td>6</td>
<td>18.7</td>
</tr>
</tbody>
</table>

CACS = Cooperative Amblyopia Classification Study; EOE = early-onset esotropia; LOE = late-onset esotropia.

All subjects underwent a series of oculomotor and psychophysical tests to determine the functioning of each eye. Only three of these tests are relevant for the this article: the speed of the OKN response to drifting sinusoidal gratings, the speed of smooth pursuit to a moving target, and the speed discrimination for drifting sinusoidal gratings. Stimuli for all three tests were displayed on a Princeton Max-15 monitor (Princeton Graphic Systems, Princeton, NJ) (P-4 phosphor) with screen dimensions of 19.7 × 25.5 cm. Mean luminance for the displays was 90 cd/m². Tests were made monocularly while the other eye was covered with an opaque eye patch. Subjects wore their best optical corrections, as determined by clinical examination preceding laboratory testing.

Eye movements for MOKN and MSP were measured with an Eye-Trac Model 210 Limbus Tracker (Applied Science Laboratories, Waltham, MA). This device consists of two infrared detectors for each eye that respond to the amount of light reflected from the limbal-scleral boundary; the difference in reflected light signaled by the detectors is converted into degrees of eye rotation per volt. Before oculomotor measurements were taken, the tracker was calibrated for each eye by asking the patient to look at a large flashing X presented first in the center of the monitor and then 4.5° to the left or the right of center. Differential changes in voltage produced by the changes in gaze were used to scale subsequent eye movements.

Monocular Asymmetry of Optokinetic Nystagmus

The target for the MOKN measurements was a low-frequency sinusoidal grating moving in one of two directions (left or right) at 9.4°/second. The spatial frequency of the grating was scaled with visual acuity; for LogMar acuities better than 20/100 (Snellen equivalent), the frequency was 1 cyc/deg; for acuities between 20/100 and 20/600, the frequency was 0.5 cyc/deg, and for acuities poorer than 20/600, the frequency was 0.25 cyc/deg. Corresponding changes in target temporal frequency were made to keep the speed approximately constant at 9.4°/second for all subjects. Stimulus duration was 15 seconds. Subjects were instructed to look at the moving grating and to allow their eyes to follow it. Viewing distance was 1 meter, at which the stimulus field subtended a visual angle of 13°.

Eye position was sampled every 5 msec and stored for further analysis. Raw data were smoothed by averaging with a moving window 11 samples wide (average the 11 samples; assign the mean to the center of the window; move the window by one sample; repeat the procedure). Abrupt changes in position between
nearly samples that exceeded a threshold value (14°/second for adjacent samples or 12°/second for samples three steps apart) were deemed saccades and were removed from the record; these saccadic segments were replaced with segments indicating no change of eye position. Optokinetic nystagmus speeds were calculated from the desaccaded record by averaging with a moving window 50 samples wide (250 msec). The speeds for each 50-sample segment were sorted by direction; speeds < 0.26°/second were discarded. The mean speed for the accepted segments was calculated for each direction.

**Monocular Smooth Pursuit**

The test stimulus for pursuit was a large moving X, 1° in height for subjects with acuities better than 20/40 and 2° in height for subjects with acuities of 20/40 or worse. The target moved left and right at 3.5°/second for 40 seconds; traverse length was 7°. Subjects were instructed to keep looking at the moving X. Viewing distance was 1 m. Eye position was sampled every 5 msec and was stored on disk for subsequent analysis.

Pursuit data were analyzed using a procedure similar to the one used for analyzing OKN. The initial 250 msec of each record, corresponding to the latency and acceleration phases of pursuit, was discarded. After initial smoothing and desaccading, pursuit samples were sorted by stimulus direction. As for OKN records, pursuit speed was calculated by averaging with a moving window 50 samples wide; samples containing speeds < 0.26°/second or >15°/second were discarded. Mean speed then was calculated from the accepted smooth samples for each direction.

**Speed Discrimination and Monocular Speed Bias**

The target for speed discrimination was a vertical sinusoidal grating of 1 cyc/deg drifting at a mean speed of 3°/second. A small 1° fixation spot was superimposed on the screen to control fixation. The grating was 11.2 cm high × 23.2 cm wide; viewing distance was 0.67 m for all subjects. Target duration was varied at random by 25% from trial-to-trial to obscure information about the distance moved by the bars because the distance moved would necessarily covary with incremental changes in target speed. Duration was 800 msec ± 200 msec.

The speed increment threshold was estimated from 100 two-alternative, forced-choice trials. Subjects judged which of two stimulus presentations contained the faster speed. The first presentation contained a target moving at the reference speed (3°/second) always drifting toward the left. The second presentation contained a target moving at the test speed, drifting either left or right, with the direction chosen at random from trial to trial. The interstimulus interval was 500 msec. The test speed could be one of four values chosen at random (one or two steps faster or one or two steps slower than the reference speed). The step size was varied adaptively, depending on the performance, within a 20-trial block. The minimum step was 5% of the reference, whereas the maximum step size was 80%. If the patient's performance was so poor that the step was increased to the maximum, only two steps were presented, corresponding to the fastest and the slowest test speeds (0.6°/second and 5.4°/second).

Responses were binned by the direction of the second (test) interval. Separate biases and thresholds were estimated from the psychometric functions for each direction by probit analysis and from the whole response set (both directions pooled). The bias, also called the point of subjective equality, was estimated as the stimulus value corresponding to the 50% point on the cumulative normal function fitted by probit analysis; threshold was defined as half the difference between the stimulus values corresponding to the 25% and 75% points on the fitted function (d' = 0.675).

For the MOKN and MSP analysis, the ratio of the nasal speed to the sum of the nasal and temporal speed (N/[N + T]) was considered. If OKN nasal speed were faster than temporal speed, this ratio was greater than 0.5. For the analysis of MSB in the speed discrimination task, the ratio of the temporal speed bias to the sum of the temporal and nasal biases (T/[T + N]) was considered. If the patient perceived the nasal speed as faster than the temporal speed, the ratio was greater than 0.5. From a comparison of these two ratios, we determined whether perceived speed was related to either MOKN or MSP in individual subjects.

**Data Analysis Methods.** Two broad categories of analyses were conducted on these data to evaluate the usefulness of ISS components for predicting whether the age of strabismus onset was early or late. The first, more traditional approach evaluated ISS components when considered as dichotomous predictors—i.e., diagnostic rules that predict EOE (or LOE) depending on the presence (or absence) of an ISS abnormality. Because the anomalies are either absent or present, DVD and LN are inherently binary. Several of the ISS (MOKN, MSP, MSB) are continuous and present a spectrum of abnormalities. To evaluate them along with DVD and LN in this dichotomous framework, the continuous ISS measurements were dichotomized as either normal or abnormal compared to diagnostic cutoff points calculated from the 95th percentile of measured responses of the nonpreferred eyes in a group of 69 control subjects without strabismus. For each ISS component, subjects with abnormal and normal ISS measurements were then cross-classified with respect to EOE or LOE. This permitted calculation of several measures of diagnostic usefulness described below directly from the 2 × 2 tables within Table 2.
TABLE 2. Numbers of Subjects Classified With Normal and Abnormal Characteristics of Infantile Squint Syndrome

<table>
<thead>
<tr>
<th></th>
<th>Preferred</th>
<th></th>
<th>Nonpreferred</th>
<th></th>
<th>Preferred and Nonpreferred</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>EOE</td>
<td>LOE</td>
<td>EOE</td>
<td>LOE</td>
<td>EOE</td>
<td>LOE</td>
</tr>
<tr>
<td>LN</td>
<td>9</td>
<td>30</td>
<td>9</td>
<td>30</td>
<td>9</td>
<td>30</td>
</tr>
<tr>
<td>Normal</td>
<td>41</td>
<td>120</td>
<td>41</td>
<td>120</td>
<td>41</td>
<td>120</td>
</tr>
<tr>
<td>DVD</td>
<td>12</td>
<td>5</td>
<td>12</td>
<td>5</td>
<td>12</td>
<td>5</td>
</tr>
<tr>
<td>Normal</td>
<td>38</td>
<td>145</td>
<td>38</td>
<td>145</td>
<td>38</td>
<td>145</td>
</tr>
<tr>
<td>MOKN bias</td>
<td>24</td>
<td>26</td>
<td>21</td>
<td>36</td>
<td>15</td>
<td>14</td>
</tr>
<tr>
<td>No bias</td>
<td>26</td>
<td>124</td>
<td>29</td>
<td>114</td>
<td>35</td>
<td>136</td>
</tr>
<tr>
<td>MSP bias</td>
<td>32</td>
<td>49</td>
<td>28</td>
<td>54</td>
<td>25</td>
<td>26</td>
</tr>
<tr>
<td>No bias</td>
<td>18</td>
<td>101</td>
<td>22</td>
<td>96</td>
<td>25</td>
<td>124</td>
</tr>
<tr>
<td>MSB bias</td>
<td>13</td>
<td>23</td>
<td>19</td>
<td>35</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>No bias</td>
<td>37</td>
<td>127</td>
<td>31</td>
<td>115</td>
<td>44</td>
<td>147</td>
</tr>
</tbody>
</table>

EOE = early-onset esotropia; LOE = late-onset esotropia; LN = late nystagmus; DVD = dissociated vertical deviation; MOKN = monocular optokinetic nystagmus; MSP = monocular smooth pursuit; MSB = monocular speed bias.

Dichotomous predictors often are evaluated in terms of their sensitivity and specificity. For the purposes of our study, the sensitivity of an ISS component is the probability that a person with EOE will exhibit an abnormal value of the component in question. Specificity is the probability that a person with LOE will exhibit a normal value for the component. For example, the sensitivity of DVD is defined as the probability that a subject with EOE displays DVD. The sensitivity of DVD can be estimated as the proportion of subjects with strabismus in our sample with EOE who have DVD (12 of 50). Similarly, the specificity of DVD is defined as the probability that a subject with LOE does not have DVD, which can be estimated by the proportion of CACS subjects with LOE who do not have DVD (145 of 150). High sensitivity implies low false-negative rates; high specificity implies low false-positive rates.

Although they are useful in many respects, sensitivity and specificity do not address directly the following clinical concern: What is the probability that a subject with (or without) an ISS anomaly in fact has EOE (or LOE)? This aspect of a diagnostic test's performance is embodied in its positive and negative predictive values (± PV). Positive predictive value is the probability of EOE based on the presence of an abnormal ISS measurement; negative predictive value represents the probability of LOE (not EOE) based on the presence of a normal ISS measurement. The +PV of DVD is defined as the probability of EOE given DVD; the −PV is defined as the probability of LOE given no DVD. The prevalence of EOE in CACS, without regard to ISS measurements (the marginal, or what we call the “baseline,” prevalence), is simply the proportion of subjects with strabismus with EOE (0.25). Although ±PVs are of substantial clinical interest, they depend explicitly on the baseline prevalence as well as on the sensitivity and specificity of the test in question (see Appendix A).

The dichotomous analyses are limited in several
important respects. First, although dichotomous tests are clinically convenient, they may discard useful information because all “abnormal” values are considered to pose the same risk. In addition, it is difficult to evaluate the usefulness of multiple ISS components, e.g., DVD and MSP, using the dichotomous approach. Therefore, we also fit logistic regression models to the CACS data using standard maximum likelihood techniques. We represented the continuous ISS components using piecewise linear splines. The logistic regression parameter estimates enable us to estimate the probability of EOE for a specific value of the continuous ISS component (for a given baseline prevalence of EOE). We selected among models by comparing their prediction errors estimated using a computer-intensive resampling technique called cross-

---

**FIGURE 1.** Box plots illustrate the distributions of the bias ratios in the preferred and nonpreferred eyes for monocular optokinetic nystagmus, monocular smooth pursuit, and monocular speed bias for the early-onset, late-onset, and normal control groups. Lower, middle, and upper lines of each shaded box represent the 25th, 50th, and 75th percentiles, respectively. Vertical lines extend to the most distant observation within 1.5 times the interquartile range; outliers are represented by isolated dashes beyond these “whiskers.”
validation. Although we speak of “prediction” throughout this article, strictly speaking, the goal is to ascertain correctly the subject’s age of onset of esotropia as early or late using subsequent measurements of the ISS components.

RESULTS

Dichotomous Analysis: Positive and Negative Predictive Value

The box-and-whisker plots in Figure 1 show the distributions of the MOKN, MSP, and MSB ratios of the preferred and nonpreferred eyes for the early onset, late onset, and normal control groups. The lower, middle, and upper lines of each shaded box represent the 25th, 50th, and 75th percentiles, respectively; the vertical lines extend out from the box to the most distant observation within 1.5 times the interquartile range; outliers are represented by isolated dashes beyond these “whiskers.”

Diagnostic cutoffs for abnormal nasal-temporal ratios of MOKN, MSP, and MSB, set at the 95th percentiles of measurements of the nonpreferred eyes of the normal control group (n = 69), were 0.636, 0.559, and 0.458, respectively. The number of subjects with EOE and LOE with normal and abnormal values of the five ISS components are shown in the 2 × 2 contingency tables within Table 2. These data were used to compute prevalences, sensitivity, specificity, positive predictive values, and negative predictive values of the ISS components. Inspection of Table 2 shows that all the ISS components, with the exception of LN, were more prevalent in EOE than in LOE. The prevalence of DVD and the asymmetries of MOKN, MSP, and MSB in preferred and nonpreferred eyes were higher in the early group than in the late-onset group. Latent nystagmus was equally prevalent in the two groups, (0.18 versus 0.20). The greatest difference in prevalence between early- and late-onset groups was for DVD (0.24 versus 0.03), but its overall prevalence in both groups was low.

TABLE 3. Estimates of Sensitivity and Specificity for Components of ISS in EOE and LOE

| ISS Components | Preferred Eye | | Nonpreferred Eye | | Both Eyes |
|----------------|---------------|-------------------|-------------------|-------------------|
|                | Sensitivity | Specificity | Sensitivity | Specificity | Sensitivity | Specificity |
| DVD            | 0.24        | 0.97       | 0.24        | 0.97       | 0.24        | 0.97       |
| LN             | 0.18        | 0.80       | 0.18        | 0.80       | 0.18        | 0.80       |
| MOKN           | 0.48        | 0.83       | 0.42        | 0.75       | 0.50        | 0.91       |
| MSP            | 0.64        | 0.67       | 0.56        | 0.64       | 0.50        | 0.83       |
| MSB            | 0.26        | 0.85       | 0.38        | 0.77       | 0.12        | 0.98       |

ISS = infantile squint syndrome; EOE = early-onset esotropia; LOE = late-onset esotropia; LN = latent nystagmus; DVD = dissociated vertical deviation; MOKN = monocular optokinetic nystagmus; MSP = monocular smooth pursuit; MSB = monocular speed bias.

For all five components, sensitivity and specificity were estimated (Table 3). Dissociated vertical deviation displayed high specificity (0.97) but relatively low sensitivity (0.24). Monocular optokinetic nystagmus in the preferred eye displayed relatively high specificity (0.83) and moderate sensitivity (0.48).

Dichotomous analysis focused on the positive and the negative predictive values for all five ISS components, which are displayed, along with 95% confidence intervals, in Figures 2A (preferred eyes) and 2B (nonpreferred eyes). The confidence intervals for the ±PV estimates in Figures 2 and 3 are computed as ±2 times the standard errors of these estimates. In Figure 2, the ±PV for populations with the EOE prevalence of 0.25 observed in CACS and, consequently, the standard errors, are estimated directly from Table 2 using standard methods for proportions. In Figure 3, however, the ±PVs (for populations with other EOE prevalences) are estimated using the nonlinear functions of sensitivity and specificity in Appendix A; hence, we derived the standard errors for ±PVs using the delta method, which relies on a Taylor series approximation. As shown in Appendix A, several factors determine the ±PV of a given ISS component, including baseline prevalence of EOE in a given target population and sensitivity and specificity of the test, which are determined in part by the diagnostic cutoff criteria adopted for dichotomous classification of an abnormal response. To be a valuable predictor of EOE, a given ISS component must have a predictive value greater than the baseline prevalence of EOE in a given clinical population. The baseline prevalence of EOE in our population was 50 of 200 subjects, or 0.25. Conversely, the baseline prevalence of LOE was 150 of 200 subjects, or 0.75. Horizontal lines in Figure 2 indicate the baseline prevalence of EOE (0.25) and LOE (0.75) in our sample for the + and −PV plots, respectively. Positive predictive values for DVD (0.71) and for MOKN (0.48) in the preferred eye were substantially greater than the baseline prevalence of EOE (0.25). Although the estimated +PVs for several of
the other components were statistically significantly different from the baseline prevalence of 0.25, none appeared substantial enough to suggest clinical usefulness. The —PVs for all ISS components were close to 0.75, suggesting that, in a population with an EOE prevalence of 0.25, a normal value for an ISS component provides relatively little added information.

Generally, positive PVs of MOKN and MSP were higher for preferred than for nonpreferred eyes. Abnormalities of both ISS components were more prevalent in preferred than in nonpreferred eyes of the EOE group (see Table 2), leading to elevated +PV and —PV. The greater prevalence of abnormal ratios of MOKN and MSP in EOE for preferred than for nonpreferred eyes probably resulted from more degraded MOKN and MSP in both the nasal and the temporal directions in nonpreferred eyes. Abnormal values for nasal and temporal OKN and smooth pursuit responses were set by the 95th percentiles of the measurements of the nonpreferred eyes of the normal control group. Among the EOE group, MOKN slow-phase velocities were abnormally low in nasal and temporal directions in the nonpreferred eyes of eight subjects and in the preferred eyes of only two subjects. Similarly, among the EOE group, MSP velocities were abnormally low in both directions in the nonpreferred eyes of three subjects and none of the preferred eyes. The same trend occurred in the LOE group, in which MOKN was reduced in both directions in the nonpreferred eyes of 16 subjects and in the preferred eyes of eight subjects; MSP was reduced in both directions in the nonpreferred eyes of 11 subjects and the preferred eyes of seven subjects. Large and nearly equal abnormal responses in the two directions resulted in ratios
FIGURE 3. Plot-estimated positive (top) and negative (bottom) predictive values of the preferred eyes for biases of dissociated vertical deviation, monocular optokinetic nystagmus, monocular smooth pursuit, and monocular speed bias as a function of baseline prevalence of early-onset esotropia. B and D in the monocular optokinetic nystagmus graphs represent measures of +PV and −PV from Bourron-Madignier and Demer and von Noorden.

that fell below the diagnostic cutoff, whereas greater loss in the temporal direction resulted in larger ratios that failed the diagnostic cutoff. If the bidirectional anomalies are included in the group of subjects with abnormal ratios, the motor anomalies either are equally prevalent in the two eyes or are more prevalent in the nonpreferred eye.

Positive and negative PVs depend explicitly on the baseline prevalence of the disease in question (see Appendix A). Consequently, the +PV and −PV for any of the ISS components estimated from the CACS data base are not universally applicable. For example, the values of +PV shown in Figure 2 are specific to the baseline prevalence of our clinical population. If the baseline prevalence of EOE were greater than in the CACS study, the +PV for the ISS component would be higher and the −PV would be lower, as illustrated by Figure 3.

Figure 3 plots the estimated +PV and −PV in the preferred eyes for measures of DVD, MOKN, MSP, and MSB as a function of baseline prevalence of EOE. The diagonal dashed lines in the upper set of figures represent the EOE baseline prevalence for various hypothetical populations with strabismus; the dashed lines in the lower figures represent the complementary prevalences of LOE in the same populations. The curved functions describe the corresponding ±PV of individual ISS components surrounded by the 95% pointwise confidence intervals. Figure 3 shows that DVD is a useful predictor of EOE over a large range of baseline prevalence values. An abnormal MOKN measurement is moderately informative. The −PV plots suggest that normal ISS measurements appear to provide little predictive information beyond that provided by the baseline prevalence of EOE.

The estimates of +PV, based on asymmetric MOKN in the current study sample, were compared with the empirical results of two other studies of asymmetric MOKN prevalence in early- and late-onset esotropia. Figure 3 plots the +PV and −PV for MOKN reported by Bourron-Madignier et al and Demer and von Noorden against the

Figure 4. Sensitivity and specificity of monocular optokinetic nystagmus in the preferred eye are plotted for diagnostic cutoffs ranging from 0.5 to 1.0 in the Cooperative Amblyopia Classification Study database. Values are compared with the performance reported by (B) Bourron-Madignier18 and (D) Demer and von Noorden.19 The cutoff used in Tables 2 and 3 and in Figures 2 and 3 was 0.64. As expected, as the cutoff criterion was altered from 0.5 to 1.0, the sensitivity of the test decreased (fewer subjects with early-onset esotropia were diagnosed with abnormal monocular optokinetic nystagmus, and false-negative results increase), whereas the specificity increased (more subjects with late-onset esotropia were diagnosed with normal monocular optokinetic nystagmus, and false-positive results decreased).

Baseline prevalence for EOE and LOE reported in their samples. Although our +PV estimate agrees with Bourron-Madignier's for a population at their baseline prevalence for EOE (0.54), our -PV is substantially smaller. Our -PV estimate agrees with Demer and von Noorden at their baseline prevalence for EOE (0.58); however, our +PV is substantially less. Therefore, discrepancies between our predictive value estimates cannot be explained entirely by the different baseline prevalences of EOE in the three study samples.

A possible factor that could contribute to residual differences in predictive values shown in Figure 3 is that the three studies may have used different diagnostic cutoff criteria for categorizing MOKN biases as abnormal. For example, a more stringent or a higher diagnostic cutoff criterion for asymmetric MOKN might increase the +PV of MOKN. The effect of the cutoff can be evaluated independently of baseline prevalence by comparing sensitivity and specificity of MOKN in the three studies. Different cutoffs will define diagnostic tests of different sensitivities and specificities that are independent of baseline prevalence. Because the cutoffs in the other two studies are unknown, we tested this hypothesis by reanalyzing our data using cutoffs ranging from 0.5 to 1.0 for categorizing MOKN biases as abnormal (as shown in Fig. 4). We found that the sensitivities and specificities reported by Bourron-Madignier et al18 and Demer and von Noorden19 lay substantially off our performance curve. This discrepancy suggests that differences in sensitivity and specificity (and, consequently, the differences in +PV and -PV after accounting for baseline prevalence) are not simply the result of differences in diagnostic criteria. As will be presented in the Discussion, factors such as MOKN stimulus parameters may account for the differences in results.

Although several factors accounted for the discrepancies between +PV and -PV reported in these three studies, a major factor was probably the baseline prevalence of EOE. Demer and von Noorden19 reported a +PV for MOKN (0.91) that was nearly twice our estimate of +PV for the preferred eye (0.48). Differences in baseline prevalence for EOE (0.40) accounted for most of the difference in +PV observed between these three studies. Clearly, it is unwise to apply the ±PV values from any of these studies to other clinical populations without first considering differences in baseline prevalence of EOE.

Logistic Regression Analysis of the Association Between Infantile Squint Syndrome Components and Early-Onset Esotropia

In this section, the predictive usefulness of the ISS components was assessed using the more powerful tool of logistic regression. Biomedical researchers have used logistic regression models extensively to represent the association between binary outcomes (e.g., early versus late onset of esotropia) and various "predictor" or "explanatory" variables (e.g., MOKN, MSP, MSB).35 The 2 × 2 contingency table analyses of the previous section, though simple to perform and interpret, required dichotomization of the continuous ISS measurements. As cutoffs, we chose the upper 95th percentiles of normal CACS subjects' measurements. Although this choice of cutoff is reasonable, it is not necessarily optimal. Moreover, reducing the continuous measurement to a dichotomous variable, even using an optimal cutoff, is likely to discard prognostic information. The foregoing dichotomous analysis treats all ISS measures above the diagnostic cutoff as equally predictive of EOE. However, it is likely that MOKN asymmetries, which are extremely biased, are associated more strongly with EOE than with smaller asymmetries, even though both may fail the diagnostic cutoff. Thus, in this section, logistic regression uses the continuous ISS mea-
TABLE 4. Univariate Logistic Regression Estimates

<table>
<thead>
<tr>
<th>ISS Component</th>
<th>a Estimate</th>
<th>Standard Error</th>
<th>b Estimate</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dichotomous Predictors</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DVD Latent nystagmus</td>
<td>-1.3391</td>
<td>0.1818</td>
<td>2.2146</td>
<td>0.5624</td>
</tr>
<tr>
<td>Continuous Predictors</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MOKN Preferred Nonpreferred</td>
<td>-1.9241</td>
<td>0.2659</td>
<td>0.8202</td>
<td>0.1800</td>
</tr>
<tr>
<td>MSB Preferred Nonpreferred</td>
<td>-1.4794</td>
<td>0.2155</td>
<td>1.5955</td>
<td>0.5202</td>
</tr>
<tr>
<td>MSP Preferred Nonpreferred</td>
<td>-2.1068</td>
<td>0.2728</td>
<td>1.2944</td>
<td>0.2506</td>
</tr>
</tbody>
</table>

ISS = infantile squint syndrome; DVD = dissociated vertical deviation; MOKN = monocular optokinetic nystagmus; MSB = monocular speed bias; MSP = monocular smooth pursuit.

The patient and the clinician. Finally, logistic regression facilitates a multivariate analysis, enabling us to consider simultaneously the association of several ISS dichotomous and continuous components with EOE.

Univariate Logistic Regression Models. If \( D \) is an indicator function of a binary outcome, the general form of a logistic regression model for the probability of the outcome given a single predictor, \( X \), is as follows:

\[
Pr[D = 1|X = x] = \frac{\exp(a + bx)}{1 + \exp(a + bx)}
\]  (1)

In our case, \( D = 1 \) if the subject had EOE, and \( D = 0 \) if the subject had LOE. The ISS predictor under consideration is represented by \( X \) (and its observed value by \( x \)). The parameters, \( a \) and \( b \), are estimated by fitting such models to data from the CACS subjects, \((D,x)\), using standard maximum likelihood techniques.35 This model is easily generalized to accommodate multivariate predictors by viewing \( X \) as a matrix and \( b \) as a vector of parameters. Table 4 presents the \( a \) and \( b \) parameter estimates for eight univariate logistic regression models, each using one of the ISS components to predict EOE.

Conditional Probability Interpretation of the Logistic Model Parameters. The \( b \) parameter in equation 1 readily translates into a measure of association known as the odds ratio. The odds ratio is used commonly in epidemiology, in part because its interpretation does not depend on the prevalence of the disease in question; in fact, the odds ratio can be estimated from case–control studies in which the proportion of cases is fixed by design. Although odds ratio estimates have provided epidemiologists with considerable insight...
TABLE 6. Multivariate Logistic Regression Estimates

<table>
<thead>
<tr>
<th>ISS Component</th>
<th>Model 1</th>
<th></th>
<th>Model 2</th>
<th></th>
<th>Model 3</th>
<th></th>
<th>Model 4</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Est</td>
<td>SE</td>
<td></td>
<td>Est</td>
<td>SE</td>
<td>Est</td>
<td>SE</td>
<td>Est</td>
</tr>
<tr>
<td>MSP</td>
<td>1.2944</td>
<td>0.2506</td>
<td></td>
<td>1.1365</td>
<td>0.2604</td>
<td>0.8905</td>
<td>0.3079</td>
<td>0.8984</td>
</tr>
<tr>
<td>DVD</td>
<td>1.4313</td>
<td>0.6273</td>
<td></td>
<td>1.3865</td>
<td>0.6297</td>
<td>1.3940</td>
<td>0.6314</td>
<td></td>
</tr>
<tr>
<td>MOKN</td>
<td>0.3520</td>
<td>0.2345</td>
<td></td>
<td>0.3607</td>
<td>0.2434</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSB</td>
<td>-0.0933</td>
<td>0.7041</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>-2.1068</td>
<td>0.7278</td>
<td></td>
<td>-2.1269</td>
<td>0.2769</td>
<td>-2.2774</td>
<td>0.3026</td>
<td>-2.2720</td>
</tr>
<tr>
<td>-2 log likelihood</td>
<td>191.113</td>
<td>185.753</td>
<td></td>
<td>183.534</td>
<td>185.594</td>
<td>183.517</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

ISS = infantile squint syndrome; MSP = monocular smooth pursuit; DVD = dissociated vertical deviation; MOKN = monocular optokinetic nystagmus; MSB = monocular speed bias; Est = estimates; SE = standard errors.

TABLE 7. Prediction Error Estimates for Multivariate Logistic Regression Models

<table>
<thead>
<tr>
<th>ISS Components (preferred eye)</th>
<th>Prediction Error</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>MSP</td>
<td>Total</td>
<td>0.1552</td>
<td>0.4449</td>
<td>0.0587</td>
<td></td>
</tr>
<tr>
<td>MSP, DVD</td>
<td></td>
<td>0.1524</td>
<td>0.4372</td>
<td>0.0574</td>
<td></td>
</tr>
<tr>
<td>MSP, MOKN</td>
<td></td>
<td>0.1606</td>
<td>0.4539</td>
<td>0.0628</td>
<td></td>
</tr>
<tr>
<td>MSP, MSB</td>
<td></td>
<td>0.1707</td>
<td>0.4987</td>
<td>0.0614</td>
<td></td>
</tr>
<tr>
<td>MSP, DVD, MOKN</td>
<td></td>
<td>0.1518</td>
<td>0.4331</td>
<td>0.0580</td>
<td></td>
</tr>
<tr>
<td>MSP, DVD, MOKN, MSB</td>
<td></td>
<td>0.1544</td>
<td>0.4395</td>
<td>0.0594</td>
<td></td>
</tr>
</tbody>
</table>

ISS = infantile squint syndrome; MSP = monocular smooth pursuit; DVD = dissociated vertical deviation; MOKN = monocular optokinetic nystagmus; MSB = monocular speed bias.
TABLE 8. Monocular Optokinetic Nystagmus Asymmetry in Preferred Eyes of Accommodative and Nonaccommodative EOE and LOE

<table>
<thead>
<tr>
<th></th>
<th>Accommodative ET</th>
<th>Nonaccommodative ET</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Median</td>
</tr>
<tr>
<td>EOE</td>
<td>7</td>
<td>0.54</td>
</tr>
<tr>
<td>LOE</td>
<td>25</td>
<td>0.54</td>
</tr>
</tbody>
</table>

EOE = early-onset esotropia; LOE = late-onset esotropia; ET = esotropia; IQR = interquartile range.

than are the LOE data; this corresponds to the strong association of EOE with higher nasal MSP asymmetries. In contrast, the distributions of both MOKN and MSB measurements were more similar among EOE and LOE subjects. The paucity of MSB observations < 0.4 suggests that the support for the apparent association of MSB with EOE is weak (also reflected in the very wide confidence intervals for this component).

It is essential to note that the estimates displayed in Figure 5 assume that the baseline prevalence of

![Figure 5](https://iovs.arvojournals.org/pdfaccess.ashx?url=/data/journals/iovs/933197/)
Figure 6. Estimates of $P[D|x]$ for four values of monocular smooth pursuit are plotted as a function of the baseline prevalence of early-onset esotropia. The diagonal line represents $P[D|x]$ equal to baseline prevalence.

EOE among the target population of adult subjects with strabismus is 0.25. The estimated conditional probabilities of EOE in an adult population with strabismus with an EOE baseline prevalence other than 0.25 would differ. Figure 6 displays the estimated $P[D|MSP]$ for various prevalences of EOE. For reference, the light continuous line at 45° represents the predicted $P[EOE]$ based solely on baseline prevalence. For example, an MSP bias of 0.60 elevates the probability of EOE very little above any baseline prevalence. In contrast, if there was a dramatic asymmetry of MSP = 0.80 (the bold series of long and short dashes), the probability of EOE would be very high at any baseline prevalence. The bold continuous line in Figure 6 for an MSP ratio of 0.5 indicates that if MSP is symmetric (i.e., there is no bias), the probability that a patient has EOE is less than the baseline predictive value.

Univariate Model Selection Based on Prediction Error. Several of the ISS components appear associated with EOE when considered as univariate predictors in logistic regression models. How can we choose among them? One of our primary objectives is to assess how the ISS components might facilitate clinical assessment of the age of strabismic onset. Thus, a natural way to appraise the clinical usefulness of the ISS components is to estimate the prediction error that would result if we actually used the ISS measurements, through logistic regression, to estimate the probability that subjects had EOE. As described in Appendix C, we quantify the prediction error for a subject as the squared difference between that estimate and 1 (if the subject had EOE) or 0 (if the subject had LOE).

We estimated such prediction errors from the CACS data by using a resampling technique called tenfold cross-validation (see Appendix C). Table 5 presents the mean prediction error results for each of seven univariate logistic regression models. Total prediction error (TPE) represents the estimated prediction error for an adult with strabismus drawn at random from the CACS sampling population. Because TPE essentially weights the prediction errors for subjects with EOE and LOE according to the relative prevalence of EOE and LOE, TPE reflects, in part, the CACS baseline EOE prevalence. Therefore, Table 5 also reports estimates of prediction error averaged separately among the 50 subjects with EOE and then among the 150 subjects with LOE, which we term "negative" and "positive" prediction errors (−PE and +PE), respectively. Partitioning the prediction error into ±PE provides some insight into how the ISS components might perform in other populations of adults with strabismus with different baseline prevalences of EOE. Additionally, this allows model evaluation to reflect the differential consequences of misclassifying subjects with EOE or LOE. For example, a patient with LOE might not be treated if he or she were misclassified as having EOE, which typically has a lower prognosis for functional correction than LOE.

As with our earlier analyses, it is instructive to compare the performance of prediction rules based on ISS measurements to a simple rule based solely on the baseline prevalence. Because the EOE baseline prevalence in CACS was 0.25, if we had not considered the ISS measurements, we would have concluded that there was a probability of 0.25 that any individual sampled from the CACS data base would have an EOE. Baseline TPE and ±PE for such a prediction rule, without consideration of ISS components, are calculated as follows:

Baseline $- PE = (1 - 0.25)^2 = 0.5625$ (2)
Baseline $+ PE = (0 - 0.25)^2 = 0.0625$ (3)
Baseline TPE = $[50 \times (-PE) + 150 \times (+PE)]/200 = 0.1875$ (4)

These baseline prediction errors help us assess the performance of the prediction models based on the ISS components.

The TPE column in Table 5 indicates that prediction errors were smaller for models using ISS components in preferred eyes than for models using nonpreferred eyes. The prediction error was not significantly reduced by using MSB in either eye (0.1814 and 0.1871, preferred eyes and nonpreferred eyes, respec-
Retrospective Prediction of Early-Onset Esotropia

In a population with a relatively low prevalence of EOE, the +PE displayed in Table 5 would be weighted more heavily; thus, the TPE for DVD should be smaller. Finally, our estimated TPE assumes that negative and positive prediction errors bear equal costs; for example, if the consequences of a false negative are substantially less than of a false positive, Table 5 suggests that DVD may be more useful than MOKN.

Multivariate Logistic Regression Analyses

Multivariate Logistic Regression Estimates. Our univariate analyses suggested that MSP in the preferred eye is the single best ISS predictor of EOE among adult subjects with strabismus. We fit a series of multivariate models with additional ISS variables to determine whether the prediction of EOE could be improved by using other ISS components in conjunction with MSP asymmetry in the preferred eye. Table 6 presents four models, starting with the univariate model based on MSP nasal asymmetry alone and successively adding DVD, MOKN, and MSB (the latter two measures were for preferred eyes). Adding DVD to the univariate MSP model resulted in a statistically significant improvement in model fit ($P = 0.02$); the $b$ parameter estimates for both these ISS components remained large, which indicates independent and substantial associations with EOE. Adding MOKN to this bivariate model (yielding model 3) produced no significant improvement in fit ($P = 0.14$) but did reduce the strength of association between MSP and EOE resulting from partial colinearity of MOKN and MSP. Finally, adding MSB to the model produced no significant improvement in fit ($P = 0.90$) and had little effect on the other parameter estimates. Thus, although MOKN and MSB showed substantial and statistically significant associations with EOE when considered univariately, these associations virtually disappeared when we controlled for MSP and DVD by their inclusion in the model.

Prediction Error Estimates for Multivariate Logistic Regression Models. Table 7 presents the estimated prediction errors for six multivariate models. Using a model with DVD and MSP (preferred eye) reduces the TPE by 0.0028, or only an additional 2%. Adding MOKN or MSB (preferred eyes) to the MSP model actually worsened predictions slightly; adding MOKN to a model including both MSP and DVD reduced the TPE by 0.0006, or only an additional 0.4%. These results suggest that it is useful to consider MSP in the preferred eye and in DVD but that adding the other ISS variables provides little additional information in predicting EOE among adult subjects with strabismus. Although the predictive power added by these laboratory measurements appears small, knowing their magnitude in a given adult patient may substantially reduce uncertainty about the age of onset in certain circumstances. For example, knowing a patient has an extreme pursuit asymmetry raises the probability that EOE substantially above the baseline prevalence, whereas moderate values of MSP asymmetry provide marginal support for EOE above the baseline prevalence (see Fig. 6).

Estimates of the Probability of Early-Onset Esotropia as a Function of Monocular Smooth Pursuit Bias, With and Without Dissociated Vertical Deviation. Figure 7 displays the estimated probability of EOE based on monocular smooth pursuit (MSP) for subjects with (solid line) and without (dotted line) dissociated vertical deviation (DVD). Clearly, including the DVD test results improves the predictive usefulness of MSP measurements for a wide range of continuous measures of MSP bias.
illustrates the improved predictive power of MSP when it is considered along with DVD. The probability of EOE, based on measures of DVD and MSP asymmetry, in an adult with strabismus selected at random from a population with a 0.25 baseline prevalence of EOE, could be estimated from this plot. The probability of EOE for a patient with DVD and MSP asymmetry, sampled from another population or clinic with a different baseline prevalence for EOE, would differ from that seen in Figure 7.

DISCUSSION

Dichotomous Analyses

Dichotomous analysis illustrates that early- and late-onset groups exhibited clinical signs of ISS, indicating that the syndrome components are not exclusively associated with EOE. Consequently, a prediction of EOE can only be made with a limited degree of certainty when a subject has one or more ISS components. To ascertain whether the ISS components could be used as predictors of EOE, we analyzed these components in terms of their sensitivity, specificity, and ±PV.

Although MOKN, MSP, and MSB had higher sensitivities to EOE, they had substantially lower specificities than did DVD (Table 3). Of all ISS components, DVD had the highest +PV. However, the prevalence of DVD is generally low, and, thus, it has low sensitivity, which diminishes its clinical usefulness in predicting EOE. We estimated the +PV of MOKN to be somewhat higher than the baseline prevalence of EOE, indicating that MOKN increased the likelihood of a correct prediction of EOE only slightly above chance. The dichotomized versions of MSP and MSB had even lower +PV than did MOKN. None of the −PV for the ISS components were significantly different from chance (baseline prevalence), indicating that they were of little additional value for diagnosing LOE. The only ISS component that was relatively uncommon in EOE and LOE (0.18 and 0.20, respectively) was LN; it has a higher prevalence in preadolescence and may be an excellent predictor of EOE in younger populations.36

Our predictive value analyses focused on unilaterally defined failure criteria—i.e., where the preferred and nonpreferred eyes are evaluated separately with respect to the cutoff. An alternative diagnostic strategy would classify subjects as having EOE only if they fail the criterion for a given ISS component in preferred and nonpreferred eyes. Table 3 indicates that, as expected, for such a bilateral diagnostic rule, sensitivity declines and specificity increases relative to the performance of unilaterally defined criteria because abnormal values were less prevalent in the “both” than in the preferred or nonpreferred eye categories. In settings in which false-positive results are much more costly, this suggests that bilateral failure criteria for MOKN, MSP, and MSB may be appropriate.

Logistic Regression Analysis

Logistic regression analysis provided a more powerful technique for exploring the association between MOKN, MSP, and MSB and EOE than did the dichotomous analysis of these ISS components. We found that MOKN and MSP nasal asymmetries, as well as MSB temporal asymmetries, all revealed an association with EOE. Logistic regression enabled us to estimate the probability of EOE given the magnitude of individual ISS components. Admittedly this is clinically attractive, but these estimates are sensitive to the baseline prevalence of EOE in any given population.

Our prediction error analysis of continuous measurements of ISS components revealed that MSP in the preferred eye was the single strongest predictor of EOE, in contrast to the dichotomous analysis that relied on a diagnostic cutoff based upon the 95th percentile of the control population. Because logistic regression analysis can use continuous measures and does not depend on a cutoff, it revealed the importance of MSP in predicting EOE. This suggests that if MSP is measured clinically by direct observation, considerable care is needed to establish a cutoff.

The ±PE columns in Table 5 indicated that a logistic prediction model based solely on MSP in the preferred eye had the smallest −PE and +PE among the seven univariate models. Accordingly, MSP is likely to have the smallest TPE, even in populations with different EOE baseline prevalence. Again, TPE measures the average (squared) difference between the probability of EOE estimated by the logistic regression model and the actual diagnosis from the case history (EOE or LOE) (see Appendix C). Our estimated TPE assumes that +PE and −PE bear equal costs, which may be incorrect. Treatment might be more aggressive for a patient with a high rather than a low prognosis. Consequently, a patient with LOE who is classified incorrectly as having EOE might undergo a less effective treatment regimen, which could reduce the prognosis for functional correction of strabismus. However, the reverse error would enhance the initial treatment effort by overestimating the prognosis.

The dichotomous analysis suggested that among MOKN, MSP and MSB, MOKN was the best indicator of EOE. Dichotomizing the ISS data at the 95th percentile of normal subjects’ measurements led to a loss of information; however, these cutoffs may not have been optimal, and this approach treated all measures of ISS above the cutoff as equal. Subsequent logistic regression analyses using continuous measures of MOKN, MSP, and MSB revealed that MSP (preferred eye) was the strongest single predictor of EOE.
Logistic regression also facilitated a multivariate analysis to determine whether any of the ISS components considered jointly provided independent predictive information. Both dichotomous and logistic analyses agreed that DVD is a useful indicator of EOE. Because of the low overall prevalence of DVD, testing for this alone would have been of limited value. That is, although most persons who had DVD had EOE (thus yielding a high +PV), relatively few subjects had DVD. We found that DVD and MSP (preferred eyes) were the best set of two predictors. Adding MSB, MOKN, or both to this model produced neither a significantly better fit to the data nor a reduction in prediction error.

**Generalizability of Results.** Several factors could limit the generalizability of our results to other populations. Most of our subjects were adults with relatively long histories of strabismus, most had a history of amblyopia, and all had undergone treatment for strabismus that may have modified their ISS components. Clearly, our results are not generalizable to the problem of determining whether the age of onset was early or late among children younger than 8 with strabismus. However, our results should apply to subjects with strabismus who range from 8 to 40 years of age, have a history of amblyopia, and seek functional correction for strabismus. Other factors that might limit the generalizability of our results include potential misclassification, recall and selection biases, and measurement error.

Possible misclassification of our subjects as having EOE or LOE could bias estimates of the association between ISS components and age of onset. Our subjects with strabismus (age range, 8 to 40 years) were classified largely on the basis of their recollection of the age of onset, earliest treatment, or both. Consequently, our determination of age of onset is subject to error. If subjects were misclassified randomly, this would generally attenuate the relationship between the age of onset, earliest treatment, or both. Consequently, our determination of age of onset is subject to error. If subjects were misclassified randomly, this would generally attenuate the relationship between the age of onset and the ISS components. It is more likely, however, that subjects who did have EOE would be misclassified as having LOE than the converse (for example, if there was no recollection of the age of onset in a subject with EOE for whom treatment occurred after the first birthday, our study would classify this subject mistakenly as having LOE). If this misclassification was independent of ISS status at the time of our study, this again would attenuate our estimate of the b parameter in equation 1 and would bias our +PV and −PV estimates down and up, respectively. If, however, the probability of misclassification differed according to ISS status, the potential biases could be in either direction.

In addition, there is the issue of potential selection bias. We have discussed how the +PV, −PV, and conditional probability estimates are all dependent on the baseline prevalence of EOE in the target population. If this baseline prevalence is known, and it is assumed that our sensitivity and specificity estimates are unbiased, we could extend our results to other populations as shown in Figures 3 and 6.

Finally, the ISS measurements themselves are subject to error. However, because our objective is to evaluate the clinical usefulness of the ISS components as measured in practice, the same error is inherent in any clinical evaluation of ISS. A separate issue, of course, is whether a more precise measure of MSB, for example, might show a stronger association with EOE; alternatively, less precise measurements of MSP would be less predictive of EOE.

**Earlier Studies and Relevance to the Current Investigation.** *Monocular Optokinetic Nystagmus.* Figure 1 confirms earlier studies by Bourron-Madignier et al,18 Demer and von Noorden,19 and Mein,7 which reported that the slow phase of MOKN and MSP were faster, on the average, for the EOE and the LOE in response to nasal motion. These studies also indicated that the MOKN and MSP ratios were larger for the early-onset than for the late-onset group. Differences between +PV of MOKN observed in our study sample and in earlier studies by Bourron-Madignier et al18 and Demer and von Noorden19 result in large part from differences in the clinical baseline prevalence of EOE (Fig. 3). Other potential factors include differences in cutoff selection as well as possible differences in population age, treatment history, and association with amblyopia. In addition, part of the discrepancy between the +PV measures in our study and the studies of Bourron-Madignier et al18 and Demer and von Noorden19 may have been caused by differences in stimulus parameters.

Bourron-Madignier et al18 and Demer and von Noorden19 stimulated MOKN with a standard handheld OKN drum. The velocity (9°/second) and the field size (13°) used in our study are consistent with the use of the standard handheld OKN drum (20 cm diameter) held a distance of 87 cm from the patient and rotated approximately 13 times/minute. Bourron-Madignier et al18 used a slower stimulus velocity (5.5°/second) and a larger field (22.6°) with a 20 cm drum held 50 cm from the patient and rotated 13 times/minute. Demer and von Noorden19 used a larger field size (37°) and a higher stimulus velocity (38°/second) with a 20 cm drum held 30 cm from the patient and rotated 20 rpm. Variations in these stimulus parameters could either increase or decrease the +PV, depending on the severity of MOKN asymmetries in the patient population under evaluation.

The +PV of asymmetric MOKN also is influenced by an interaction between characteristics of the test stimulus and the slow phase velocity of OKN. Increasing stimulus velocity reduces slow phase gain for
MOKN both nasally and temporally.\textsuperscript{57,58} In addition, asymmetric MOKN varies with the spatio-temporal parameters of the stimulus during infancy.\textsuperscript{39,40} In strabismus and amblyopia, the nasal asymmetry of MOKN increases with stimulus velocity\textsuperscript{14} up to a limit when MOKN fails to respond to higher stimulus velocities in either the nasal or the temporal direction (at which point the N:T ratio equals 1.0). The upper velocity limit that exhibits asymmetric MOKN and the optimal stimulus velocity for revealing asymmetric MOKN will vary between subjects, depending on the severity of their MOKN disorder. Increasing stimulus velocity will enhance the asymmetry in some subjects and reduce it in others. In the extreme, at very high stimulus velocities, asymmetric MOKN might become more apparent in the late-onset group than in the early onset group if the stimulus exceeds the response range of the early onset group. The result of very high stimulus velocity would be to decrease the +PV of the test. Indeed, the stimulus velocity used in the current study produced more bidirectional reduction of MOKN in nonpreferred than preferred eyes and resulted in greater +PV for MOKN in predicting EOE for the preferred eye category. Finally, the upper velocity limit for stimulating MOKN is influenced by field size\textsuperscript{57}; hence, lower stimulus velocities should be used to reveal MOKN asymmetries for small test fields (i.e., with handheld OKN drums than for large, full-field stimulators).

A Comparison of Monocular Optokinetic Nystagmus in Accommodative and Nonaccommodative Strabismus. We considered the possibility that two classes of strabismus (accommodative and nonaccommodative) might contribute to differences in the prevalence of MOKN asymmetries found in the EOE and LOE groups. Accommodative esotropia was classified on the basis of an abnormally high AC/A ratio. Traditionally, these patients are classified as having nonrefractive accommodative esotropia\textsuperscript{88}; however, we have no information concerning our adult subjects’ refractive status during infancy, so we hesitate to categorize it as either refractive or nonrefractive. Current measures of refractive errors reveal that 6 of 7 subjects with EOE and 18 of 25 subjects with LOE had greater than 1 D of anisometropia. If hyperopia is defined as greater than 2 D of positive error in both eyes, one subject in the EOE group and six subjects in the LOE group could be categorized as having refractive accommodative esotropia. If hyperopia is defined as greater than 2 D of positive error in either the preferred or the nonpreferred eye at the time of participation in the CACS study, three subjects in the EOE group and 12 subjects in the LOE group could be categorized as having refractive accommodative esotropia. Differences in etiology of accommodative and nonaccommodative esotropia could result in different characteristics of MOKN in these two classes of esotropia. Table 8 compares the median values of MOKN asymmetry in the preferred eyes of accommodative and nonaccommodative EOE and LOE. The 32 subjects with accommodative esotropia occurred with approximately equal prevalence in EOE and LOE (0.14 and 0.17, respectively); seven had EOE and 25 had LOE. The median MOKN bias for subjects with accommodative esotropia was the same (0.54) for EOE and LOE. In contrast, the median MOKN bias for subjects with nonaccommodative esotropia was substantially higher for EOE (0.66) than for LOE (0.54). These results are in agreement with those of Aiello et al,\textsuperscript{41} who observed symmetrical MOKN in three children with accommodative esotropia. Caution is recommended in using MOKN to predict EOE in populations with a high prevalence of accommodative esotropia.

Monocular Smooth Pursuit. As mentioned, several small population studies have documented MSP asymmetries in subjects with amblyopia, esotropia, or both,\textsuperscript{5–11,23–25} but large population-based studies of asymmetric pursuit and its prevalence in association with EOE have yet to be conducted. Dichotomous measures did not reveal a significant relationship between MSP asymmetry and EOE because of the difficulty in selecting an appropriate diagnostic cutoff. Our logistic regression analysis, however, revealed a strong relationship between MSP and EOE. The best predictor of EOE was DVD in association with MSP asymmetry in the preferred eye. These results suggest that continuous objective measures, as opposed to dichotomous measures obtained from subjective observation of the patient’s MSP, are recommended in a clinical setting as the strongest indicator of EOE. The continuous measures of MSP must be interpreted in terms of the baseline prevalence of EOE in the clinical population from which they are sampled. Figure 6 may be useful in different clinical populations by providing an estimate of EOE (contingent on the magnitude of the MSP asymmetry for different baseline prevalence).

Monocular Speed Bias. As was true for MSP, no large population-based studies have been conducted to determine the association of MSB with EOE. In our study, the psychophysical measures of MSB illustrate that subjects with strabismus perceived temporal target motion as faster than nasal target motion (see Fig. 1). These temporal velocity biases were larger for the EOE group than for the LOE group, suggesting that MSB might be included as a sensory component of the ISS. The temporal MSB observed in the CACS subjects with esotropia followed the same direction as reported in two recent studies. Hague et al\textsuperscript{80} observed higher thresholds for motion detection in the nasal than in the temporal direction in 5 of 7 subjects with
Retrospective Prediction of Early-Onset Esotropia

EOE. None of these subjects had LN or a history of amblyopia. Brosnahan et al.\textsuperscript{17} investigated MSB in 18 subjects with EOE and in 11 subjects with LOE with a higher standard velocity (10°/second) than was used in the current investigation (3°/second). Using diagnostic cutoff criteria similar to ours, they also observed the same temporal MSB. Roberts and Westall\textsuperscript{28} found no velocity bias in six subjects with strabismus and amblyopia with unknown age of onset. Similarly, Hartmann et al.\textsuperscript{26} used a motion-nulling paradigm and found that velocity biases in four infants with esotropia were small and mixed in direction or were nonexistent in the presence of larger MOKN biases.

The temporal velocity biases reported here, as well as those reported by Hague et al.\textsuperscript{25} and Brosnahan et al.\textsuperscript{17} are in the opposite direction to those reported for two subjects by Tychsen and Lisberger.\textsuperscript{11} The conflicting reports regarding the direction of the velocity bias could result from differences in the subjects examined, as well as from differences in the test stimulus parameters. A minority of our subjects (n = 15) did have the same directional velocity bias (nasal) as reported by Tychsen and Lisberger,\textsuperscript{11} but these were uncommon compared with the number of subjects with a temporal bias (n = 45) (7.6% nasal bias and 22.3% temporal bias). The stimulus velocity used by Tychsen and Lisberger\textsuperscript{11} (13.5°/second to 16.5°/second) is equivalent to velocities used by Brosnahan et al.\textsuperscript{17} and Hartmann et al.\textsuperscript{26} Exposure durations and psychometric procedures used by Tychsen and Lisberger\textsuperscript{11} were identical to those used in the current study and by Brosnahan et al.\textsuperscript{17} An important difference between the study by Tychsen and Lisberger\textsuperscript{11} and those listed above is that Tychsen and Lisberger used a single moving bar to test MSB, whereas all the other studies used drifting gratings. Spatial parameters such as field size, number of bars, and spatial frequency content might help to explain the differences.

The temporal velocity bias observed by us and by Hague et al.\textsuperscript{25} and Brosnahan et al.\textsuperscript{17} could result from an interaction between the asymmetric MOKN and the drifting grating stimulus for testing velocity discrimination. Normally, perceived velocity is lower when the eyes pursue the moving stimulus than when fixation is held stationary, as observed in the Aubert–Fleischel phenomenon.\textsuperscript{42–44} Perhaps the moving grating stimulus for MSB engaged OKN during the motion test, and, when MOKN was biased, the viewing eye followed the drifting grating at a higher velocity when it moved in the dominant direction. The single bar used by Tychsen and Lisberger\textsuperscript{11} would be less likely to engage OKN. Our study and that of Brosnahan et al.\textsuperscript{17} used a fixation spot to suppress OKN. However, small, following eye movements might still have occurred.

Clinical Relevance

It is usually difficult, if not impossible, to obtain any documented information about age of onset of strabismus or amblyopia from an adult patient. The ISS components provide a way of estimating the probability of EOE when the age of onset is unknown. Logistic regression indicated that DVD, MSP, MOKN, and MSB were useful in discriminating between early- and late-onset esotropia in adult subjects. This information could be helpful in forming a prognosis for the functional correction of strabismus in adults when it is associated with amblyopia.\textsuperscript{40} After correction of strabismus, well-maintained functional binocular alignment requires that disparity be sensed by the two eyes. This is possible only if both eyes have adequate visual acuity. Thus, the prognosis for functional correction of strabismus hinges on the prognosis for correction of amblyopia that may be associated with strabismus. Worth\textsuperscript{45} demonstrated that when treatment is delayed to adulthood, the prognosis for functional correction of amblyopia is reduced drastically for patients with EOE, whereas many patients with LOE achieved good visual acuity. There are numerous current reports of corrected amblyopia in adult patients.\textsuperscript{46–49} The age of onset influences the prognosis for functional correction that, according to comparative studies,\textsuperscript{50,51} is attributable to plasticity of the adult visual cortex. Therefore, the association of the ISS with EOE provides a means to assess age of onset from measures of sensory and motor visual functions.

These measures can be used to form a prognosis for functional correction of strabismus associated with amblyopia in adults for whom there is no documentation of the age at which binocular vision was disrupted. The prognosis is based on the assumption that EOE is more difficult to treat than LOE. Probability estimates of EOE and LOE can be derived from current measures of ISS components; however, the predictive values of these measures must be tempered by the baseline prevalence of EOE and LOE, perhaps obtained from the patient records of a given clinic. Baseline prevalence can vary as a result of many factors, including the specialties of the practicing clinicians, the age of patients, the socioeconomic structure of the community served by the practice, and access to and thoroughness of the health care available to the population.

Key Words

motion perception, optokinetic nystagmus, pursuit eye movements, statistics, strabismus

References


37. Schor CM, Narayan V. The influence of field size upon
APPENDIX A

Simple application of Bayes’ Theorem demonstrates that both $+PV$ and $-PV$ depend explicitly on the prevalence of the outcome in question.

Let $EOE$ represent an early-onset of esotropia (i.e., younger than 12 months), and let $ISS$ represent an abnormal value of the infantile squint syndrome component under consideration. Then, the sensitivity and specificity of infantile squint syndrome for early-onset esotropia are defined as follows (where $Pr[A|B]$ is the probability of $A$ given $B$):

\[
\text{Sens} = Pr[ISS | EOE] \\
\text{Spec} = Pr[\text{not ISSN} | \text{not EOE}]
\]

If $\phi$ represents the marginal, or baseline, prevalence of early-onset esotropia (i.e., $\phi = Pr[\text{early-onset esotropia}]$), then by Bayes’ Theorem

\[
+PV = Pr[EOE | ISS] = \frac{Pr[ISS | EOE] \times Pr[EOE]}{Pr[ISS]} = \frac{\text{Sens} \times \phi}{[\text{Sens} \times \phi] + \left((1 - \text{Spec}) \times (1 - \phi)\right)}
\]

Similarly

\[
-PV = \frac{\text{Spec} \times (1 - \phi)}{[\text{Spec} \times (1 - \phi)] + \left((1 - \text{Sens}) \times \phi\right)}
\]

Assuming sensitivity and specificity estimates generalize beyond the study population, one can use these formulas to estimate $\pm PV$s for populations with other prevalences of early-onset esotropia.

APPENDIX B

Below we describe how to choose a functional form for the continuous infantile squint syndrome components.

Fitting logistic regression models using the continuous infantile squint syndrome components requires choosing an appropriate scale. A simple linear model of the association of these predictors with the log odds (logit) of early-onset esotropia across the full range of asymmetries (from temporal to symmetric to nasal) could lead to biased estimates. For example, suppose that subjects with monocular optokinetic nystagmus ratios of 0.3 and 0.4 are likely to be at equally low risk for early-onset esotropia, but that subjects with monocular optokinetic nystagmus ratios of 0.8 may be at substantially higher risk than those with ratios of 0.7. The linear logistic model applied in such a case falsely presumes that a unit difference in monocular optokinetic nystagmus in the preferred eye assume that per-
sons with temporal monocular optokinetic nystagmus biases (ratios < 0.5) are at the same risk for early-onset esotropia. However, for those with nasal asymmetries (ratios > 0.5), the logit of early-onset esotropia increases linearly with the magnitude of the asymmetry, and the slope is estimated by fitting the model using maximum likelihood techniques.

This parsimonious representation was suggested through the use of flexible nonparametric logistic regression techniques indicating that the piecewise linear function captured the essential relationship between monocular optokinetic nystagmus nasal biases and risk for early-onset esotropia. Based on similar reasoning and analogous nonparametric regression results, the same piecewise linear transformations were used for monocular optokinetic nystagmus in the nonpreferred eyes and for monocular smooth pursuit and monocular speed bias in preferred and nonpreferred eyes. However, temporal monocular speed biases appeared to be more strongly associated with early-onset esotropia than for nasal biases; therefore, temporal rather than nasal biases in monocular speed bias represent the region of linearly increasing logits.

**APPENDIX C**

Estimating the prediction errors for various models requires choosing a metric for prediction error and choosing an estimation method.

**Choosing A Prediction Error Metric**

Consider the prediction error for a candidate model with respect to a single subject with strabismus, \(i\), drawn from the same population as the Cooperative Amblyopia Classification Study sample, whose age of onset is known as early (\(D_i = 1\)) or late (\(D_i = 0\)). A natural approach to assessing the model’s predictive usefulness is the following:

- Measure subject \(i\)’s infantile squint syndrome component(s), \(X_i\), denoting the actual value by \(x_i\).

Estimate the probability that subject \(i\) had an early onset given the measurement, \(Pr(D_i | X = x_i)\), by substituting \(x_i\) and the parameter estimates for \(a\) and \(b\) from the candidate model into equation 1. This is the model’s “prediction” for \(i\).

The model’s prediction error, with respect to subject \(i\), can then be quantified as \((D_i - Pr(D_i | X = x_i))^2\).

For example, if subject \(i\) had an early onset (\(D_i = 1\)), the larger the estimated \(Pr(D_i | X = x_i)\), the smaller the prediction error would be. Similarly, if \(D_i = 0\), larger estimates of \(Pr(D_i | X = x_i)\) result in larger prediction errors. Using data from multiple subjects, we could estimate the average prediction error for the candidate model by averaging \((D_i - Pr(D_i | X = x_i))^2\) for all subjects.

**Estimating Prediction Error.**

If sufficient data are available once a metric is chosen, ideally one would fit the various candidate models to a “learning” data set and then estimate the prediction error for each model by applying it to a new “test” data set. Instead, we use the following resampling technique known as tenfold cross-validation:

- Divide the subjects randomly into 10 groups.
- Exclude one group containing 10% of the data, and fit the candidate models to the remaining 90% of the subjects.
- Using parameter estimates obtained in the previous step, calculate the prediction error for the excluded subjects.
- Repeat the second and third steps 10 times, and exclude another group of 10% of the subjects in each iteration until each subject has served in a test set.
- Summarize the prediction error for each candidate model by the mean prediction error for all 200 subjects.