Autorefraction vs. Retinoscopy

A comparison of non-cycloplegic measures in a pediatric sample

John P. Lowery, O.D., M.Ed.a
Amy Joachim, O.D.b
Ronda Olson, O.D., c
Jessica Peel, O.D., d
Nicole N. Pearce, O.D., e

a. Pacific University, Forest Grove, Oregon
b. Winsted, Minnesota
c. Kayenta, Arizona
d. Beaverton, Oregon
e. Decatur, Georgia

Abstract

The purpose of this study was to determine the accuracy of the autorefractor compared to retinoscopy under non-cycloplegic conditions, such as occur in vision screenings and triage-level eye care clinics. Seventy-five children evaluated during a humanitarian eye care clinic were included in the study. Autorefraction (Nikon Retinomax) and distance retinoscopy measures were taken pre and post-cycloplegia. Spherical equivalent refractive errors were derived for comparison.

The mean difference between non-cycloplegic autorefraction and cycloplegic retinoscopy was -1.35D OD and -1.15D OS. The mean difference between non-cycloplegic retinoscopy and cycloplegic retinoscopy was -0.47D OD and -0.25D OS. The case specific range of discrepancy from cycloplegic values was up to -6.63D for non-cycloplegic autorefraction and up to -2.50D for non-cycloplegic retinoscopy. Screening sensitivity for hyperopia (2D) was 43% using non-cycloplegic autorefraction and 67% using non-cycloplegic retinoscopy.

This study provides direct evidence that non-cycloplegic autorefraction underestimates hyperopia to a clinically unacceptable degree in children. Retinoscopy, with appropriate fogging technique is much more likely to yield a clinically acceptable measure of hyperopia in children. Guidelines for appropriate use of the autorefractor in pediatric care are also presented.

Key Words

autorefraction, cycloplegia, hyperopia, myopia, pediatric refraction, retinoscopy, vision screening

Introduction

Refractive conditions represent the most prevalent class of disabling visual disorders in the pediatric population.1-7 Given the potentially adverse developmental effects of uncorrected significant refractive conditions, it is imperative that pediatric eye care providers utilize accurate methods of objective refraction when examining young children. Retinoscopy and autorefraction are the most commonly employed methods of refraction for pediatric care. Autorefraction is favored by many eye care providers because it is relatively quick and easy to perform and can be delegated to trained technicians. In the last decade, portable, pediatric friendly autorefractors have entered the ophthalmic marketplace, encouraging the replacement of retinoscopy with autorefraction. Due to increased public health policy focused on early detection of visual disorders, autorefractors have become more widely used.

Refractive error may be more accurately measured in children if a cycloplegic agent is used to inhibit accommodation. This is especially true for hyperopes who often maintain a significant latent refractive component. Significant uncorrected hyperopia is often associated with esotropia and amblyopia8,9 and has also been implicated as a risk factor for poor academic performance.10-12 Therefore, accurate refractive measurement of hyperopia is clearly imperative in pediatric care.

There are many large-scale public health settings where cycloplegia is contraindicated due to time and medical-legal constraints. In children’s vision screenings, autorefraction is often chosen over retinoscopy to determine refractive status due to its relative speed and facility. Autorefraction is also used as a stand-alone refractive measure in humanitarian eye care clinics in which prescription lenses are dispensed. However, many pediatric eye care specialists prefer distance retinoscopy in these settings because accommodation may be sufficiently controlled with fogging lenses and an appropriate distance fixation target. Retinoscopy is often the only option for refracting very young children who are unable to maintain fixation for accurate autorefraction measures.13 Unfortunately, professionals with the skill to accurately perform retinoscopy are in short supply considering the multitude of public health initiatives aimed at detecting vision disorders in children.

The purpose of this study was to determine which method, distance retinoscopy with optical fogging or autorefraction, would yield more accurate results in pediatric subjects without the use of cycloplegic agents. Emphasis is placed on comparison of the average and range of error in underestimating hyperopia. Both retinoscopy and autorefraction (Nikon Retinomax®) were measured without cycloplegia followed by a cycloplegic control measure for comparison within the same subjects. A literature review revealed a number of pediatric studies from...
which some comparative data could be drawn. Research comparing the accuracy of non-cycloplegic to cycloplegic findings with the Nikon Retinomax has demonstrated a strong tendency for the Retinomax to underestimate hyperopia in children. Significant instrument myopia in pediatric patients has been demonstrated with many other autorefractor models. Only one study was found that provided a direct comparison of non-cycloplegic autorefraction vs retinoscopy measurements. In the discussion of this paper, our findings are compared to previous studies in order to produce a summary of research to date. The data is also analyzed to guide vision care providers and public health officials in both screening and prescribing for refractive conditions in children.

**METHODS**

Each child was examined with the Retinomax and retinoscopy prior to installation of 1 drop proparacaine (0.5%), and 2 drops of tropicamide (1%). After 25 minutes, both autorefraction and retinoscopy were performed for cycloplegic comparison. Distance retinoscopy was performed at 67 cm with +1.50 fogging lenses and age-appropriate fixation targets placed at approximately 20 feet. The two optometrists performing retinoscopy were blind to all autorefractor measures taken and children were randomly assigned to retinoscopists.

All Retinomax readings were taken with the automatic fogging mechanism on. The instrument automatically calculates the average of eight readings taken in quick succession. As recommended by

<table>
<thead>
<tr>
<th></th>
<th>Mean Difference and Std Deviation in Spherical Equivalent (dry – wet)</th>
<th>High</th>
<th>Low</th>
<th>% differing by more than 1 D</th>
</tr>
</thead>
<tbody>
<tr>
<td>AutoRx OD</td>
<td>-1.35 1.14 (p&lt;.0001)</td>
<td>1.125</td>
<td>-6.63</td>
<td>57%</td>
</tr>
<tr>
<td>AutoRx OS</td>
<td>-1.15 0.90 (p&lt;.0001)</td>
<td>1.625</td>
<td>-4.00</td>
<td>57%</td>
</tr>
<tr>
<td>Ret OD</td>
<td>-0.47 0.68 (p&lt;.0001)</td>
<td>1.25</td>
<td>-2.50</td>
<td>17%</td>
</tr>
<tr>
<td>Ret OS</td>
<td>-0.25 0.65 (p=.0013)</td>
<td>0.875</td>
<td>-2.50</td>
<td>6.60%</td>
</tr>
</tbody>
</table>

*AutoRx=Autorefraction
Ret=Retinoscopy*
the manufacturer, only final readings with a confidence level of at least eight were accepted.

Tropicamide was chosen over alternative cycloplegic agents (cyclopentolate or atropine) for this study because of its minimal potential for side effects and shorter duration of action. Furthermore, the difference in cycloplegic effects between 1% tropicamide and 1% cyclopentolate has been shown to be clinically insignificant for measuring distance refraction in non-strabismic, hyperopic children.

RESULTS

The sample consisted of 69 hyperopic and six myopic children with a mean spherical equivalent of +1.32D OD, and +1.29D OS as determined by cycloplegic retinoscopy. The spherical equivalent of all refractive measures ranged from +10.00D to -7.25D. Cylinder powers measured for all eyes ranged up to -4.25D with 13% greater than 1.00D.

The data gathered was evaluated using both screening and prescriptive criteria in order to compare the validity of the two objective refractive methods across the range of clinical settings for which they are used. To evaluate the validity of each method for prescribing lenses in children, spherical equivalents were calculated for each measured refraction on all eyes to yield four comparison values: non-cycloplegic (dry) and cycloplegic (wet) measures by each method. Cycloplegic retinoscopy was used as the gold standard measure of refractive error in determining the accuracy of non-cycloplegic measures.

Prescriptive Validity

Table 1 summarizes all non-cycloplegic measures compared to cycloplegic retinoscopy. The mean difference and standard deviation between dry auto-refraction and wet retinoscopy was -1.35 ± 1.14D OD (p<.0001) and -1.15 ± 0.90D OS (p<.0001). The mean difference between dry and wet retinoscopy was -0.23 ± 0.68D OD (p<.001) and -0.25 ± 0.65D OS (p=.0013). The difference between dry autorefraction and wet retinoscopy was -0.23 ± 1.00D OD and -0.25 ± 1.00D OS, with 17% in the right eyes and 6.6% in the left eyes. Figures 1 and 2 illustrate the differences in spherical equivalent between dry and wet findings (OD) for both retinoscopy and autorefraction across the range of refractive errors in our study group. Dry values were also compared to wet autorefraction with similar results being found. There was a strong correlation between cycloplegic autorefraction and cycloplegic retinoscopy with the mean difference in spherical equivalents being statistically significant but not clinically significant at 0.19D ± 0.61 OD, and

| Table 2. Non-cycloplegic (dry) Measures Compared to Cycloplegic (wet) Autorefraction |
|---------------------------------|---------------------------------|-----------------|
| Mean Difference in Spherical Equivalent (dry – wet) | High | Low | % differing by more than 1 D |
| AutORx OD | -1.15 | 0.75 | -5.88 | 41% |
| AutORx OS | -0.86 | 0.50 | -3.13 | 33% |
| Ret OD | -0.27 | 1.88 | -2.88 | 20% |
| Ret OS | 0.03 | 3.00 | -2.75 | 14% |

AR=Autorefraction
Ret=Retinoscopy

| Table 3. Comparison of Cycloplegic Measures [Sphere and Cylinder Powers] |
|---------------------------------|---------------------------------|-----------------|---------------|-----------------|
| Mean Difference and standard deviation in Spherical Equivalent | High | Low | R* | % differing by more than 1 D |
| Sphere Power OD | 0.20 | 0.61 (p=0.0065) | 2.375 | -1.38 | 0.96 | 8% |
| Sphere Power OS | -0.28 | 0.70 (p=0.0008) | 1.625 | -3.00 | 0.96 | 12% |
| Cylinder Power OD Mean diff. AR – Ret = -0.23 | 1.00 | -3.00 | 0.8 | 4% |
| Cylinder Power OS Mean diff. AR – Ret = -0.23 | 2.25 | -3.25 | 0.85 | 4% |

*correlation coefficients of cycloplegic retinoscopy vs cycloplegic autorefraction

| Table 4. Screening Results: Comparison of non-cycloplegic (dry) findings to cycloplegic (wet) retinoscopy |
|---------------------------------|---------------------------------|-----------------|---------------|
| Sensitivity | Specificity | PPV | NPV |
| Hyperopia 2 D in either eye |
| Dry Ret | 67% | 96% | 90% | 83% |
| Dry AR | 43% | 100% | 100% | 76% |
| Myopia 1 D in either eye |
| Dry Ret | 100% | 99% | 86% | 100% |
| Dry AR | 83% | 94% | 56% | 99% |
| Astigmatism 1 D in either eye |
| Dry Ret | 79% | 98% | 94% | 93% |
| Dry AR | 88% | 97% | 88% | 97% |
| Anisometropia 1.5 D |
| Dry Ret | 100% | 93% | 44% | 100% |
| Dry AR | 100% | 92% | 40% | 100% |

AR=Autorefraction
Ret=Retinoscopy

Screening Epidemiology Statistics:
Sensitivity measures the percentage of people who truly have the condition who test positive.
Specificity measures the percentage of people who do not have the condition who test negative.
Positive Predictive Value (PPV): Measures the likelihood that a person with a positive test result actually has the disease or condition.
Negative Predictive Value (NPV): Measures the likelihood that a person with a negative test result actually does not have the disease or condition.
-0.28D± 0.70 OS. The mean difference in cylinder component between wet autorefraction and wet retinoscopy was clinically insignificant with 4% differing by more than 1D.

**Screening Validity**

All non-cycloplegic refractive measures were compared within subjects to cycloplegic retinoscopy findings using the following screening criteria for failure:

Spherical values for hyperopia 2 diopters and myopia 1 diopter
Cylindrical values 1 diopter in either eye
Anisometropia 1.5 diopters in equivalent sphere

The same values were used to determine true positives from cycloplegic findings. Table 4 summarizes the comparison of methods for the screening criteria outlined above. According to cycloplegic retinoscopy, 27 children would have failed our screening for hyperopia and six would have failed for myopia. Based on dry autorefraction measures, 12 children would have failed for hyperopia (sensitivity = 43%, specificity = 100%), and six would have failed for myopia (sensitivity = 100%, specificity = 94%). Based on dry retinoscopy measures, 18 children would have failed the screening for hyperopia (sensitivity = 67%, specificity = 96%), and six would have failed for myopia (sensitivity = 100% and specificity = 99%). The sensitivity and specificity values for determination of astigmatism were accurately determined to a significant degree in some children.

**Discussion**

The present study gives evidence that the Nikon Retinomax produces a myopia effect which can result in significant underestimation of hyperopia in children when a cycloplegic agent is not used. Several other studies have found similar results when comparing non-cycloplegic to cycloplegic measures with the Retinomax. The largest study involved 4973 school-age children in the Shunyi District of China. Mean differences between dry and wet autorefraction measures were -1.23 ± 0.97 D. For children whose cycloplegic refraction was at least +2.00 D, the difference increased to -2.98 ± 1.65 D. For those whose refraction was -2.00 or more myopic, the mean difference between wet and dry measures was only -0.41 ± 0.46 D. Harvey, et al. demonstrated with children ages 3.6 to 5.6 that dry measures with the Retinomax were 1.15D more minus on average than wet measures. Wesemen and Dick performed a comprehensive evaluation of the accuracy of the Retinomax with both adult and pediatric subjects. In comparing dry autorefraction to wet retinoscopy measures in pediatric subjects, they found 12 of 79 eyes were over-minus in 2 to 4 diopters and seven eyes were found to be over-minus from 4 to 10 diopters. The mean difference in equivalent sphere between dry autorefraction and wet retinoscopy was -1.13D. They also found that measures taken with the Retinomax fogging system were more accurate with it off (Quick Mode). El Defrawy et al. using the Retinomax on 102 pediatric subjects age 5 to 72 months found dry autorefraction findings over-minus subjects by up to 8.00D.

There is evidence that the Retinomax creates a particularly significant instrument myopia effect due to its compact design. In a recent study evaluating various refractive instruments specifically designed for pediatrics, Suryakumar and Bobier measured 43 pre-school children before and after cyclopegia with retinoscopy, Retinomax, Welch Allen Suresight and PowerRefractor (off-axis photorefractor). In comparing dry measures to wet retinoscopy, they found the following mean differences in equivalent sphere:

- Retinomax -1.149 ± 1.47 D
- Welch Allen DAV Suresight +0.49 ± 1.06 D
- PowerRefractor (LED view) -0.85 ± 0.77 D, (diffuse target) -0.32 ± 0.56 D
- Retinoscopy -0.64 ± 0.48 D

They concluded that the instrument myopia effect was related to target design and distance. It is important to clarify that other instruments evaluated in their study use strategies that differ significantly from standard infrared autorefractor design. The PowerRefractor is an off-axis photorefractor which allows the child to view a distance target while taking measurements. The Welch Allen DAV Suresight is a wavefront sensing device with a 14 inch working distance. It uses an assumed accommodative posture factored into pediatric measurements. When Schimitzke and Wesemann evaluated the Welch Allen Suresight in adult mode (no calibration for accommodation) they found that 47% of their pediatric subjects were over-minus by more than 2 D with a range up to -6.13 D. While the results of the Suryakumar and Bobier study indicate that Suresight strategy for compensating accommodative posture is effective for most subjects, the range of error found in the previous study indicates that hyperopia may still be underestimated to a significant degree in some children.

Several other studies evaluating other standard infrared autorefractors with optical fogging strategy for accommodative control have demonstrated a clinically significant myopic bias of non-cycloplegic measures in pediatric subjects. Evans evaluated the accuracy of the Rx 1 autorefractor in a study that included 50 children from 5 to 12 years of age. Pre and post-cycloplegic measurements were taken for comparison. The mean difference in spherical equivalent values between dry and wet auto-refraction was -1.13D. Without cycloplegia, 26% of the autorefracton findings showed a difference of more than 1.00 D from the equivalent sphere obtained with wet retinoscopy. With cycloplegia only 11.5% of the findings showed the same discrepancy between autorefracton and retinoscopy.

Silverberg et al. evaluated the Nidek 1600 on 89 pediatric patients (178 eyes). Compared to wet retinoscopy, dry autorefractor findings were considered accurate if they were within 0.50D of sphere, 0.50D of cylinder and 15 degrees of axis. The dry autorefraction findings were found to be accurate for sphere in 25 eyes (14%), cylinder in 124 eyes (69.6%), and axis in 91 eyes (51.1%). Agreement between dry autorefracton and wet retinoscopy was particularly low for children younger than 5 years. Cycloplegic autorefracton met the accuracy criteria for sphere in 100 eyes (56.2%), cylinder in 143 eyes (80.3%), and axis in 91 eyes (51.1%). In a study of 222 children less than 8 years old, Williams, et al. found the Topcon PR2000 Pediatric Refractometer to yield a mean difference between dry autorefracton and wet retinoscopy.
findings of $-1.16 \pm 1.52 \text{D}$ with a range of error up to $-6.5 \text{D}$. Helveston, et al. evaluated the accuracy of the Nidek 3000 including 50 children under 8 years of age in the study. Although the data was not used to directly compare dry autorefraction to wet retinoscopy, they did report an error of up to $8.00 \text{D}$ in some children due to instrument myopia. Additional evidence can be found in the literature indicating significant instrument myopia effect using other autorefractor models without cycloplegia in children. 24-26

The findings of the present study demonstrate that non-cycloplegic retinoscopy provides a more accurate measure of refractive error compared to non-cycloplegic autorefraction. The mean difference (less than $-0.5 \text{D}$) between dry and wet retinoscopy findings is clinically insignificant in most cases. Yet caution must be used when relying on dry retinoscopy findings, especially for prescriptive measurements. In our study, up to 17% of dry measures differed by more than 1 D from the wet measures.

The standard deviation of mean differences in our study and those from Suryakumar and Bobier indicate that non-cycloplegic retinoscopy has significantly less variability than non-cycloplegic Retinomax measurements. Given the trend of greater disparity between dry and wet findings with increasing hyperopia, the ultimate cycloplegic refraction may be predicted more reliably from dry retinoscopy measurements. Previous studies evaluating the accuracy of dry retinoscopy have demonstrated this relationship more clearly. Young et al. performed a study comparing dry and wet retinoscopy measurements with 328 subjects ranging from 6 to 15 years old. Two hundred six children were hyperopic from plano to 3.00D, and 31 showed hyperopia greater than 3.00D. The mean difference between measures (wet – dry) retinoscopy was $+0.67 \text{D}$ for the low hyperopes, and $+2.06 \text{D}$ for the moderate to high hyperopes. In a parallel study, Schultz found a nearly identical range of differences ($+0.75$ to $+2.00$) on a smaller population of hyperopic children. Hiatt performed retinoscopy before and after cycloplegia on 149 hyperopic eyes of patients aged 6.0 to $10.0$ years. It was concluded that 25% to 33% more hyperopia is measured after cycloplegia, with a more pronounced difference in the younger patient. Our results comparing wet and dry retinoscopy showed a mean difference of $-0.125$ for myopes, $+0.51$ for hyperopes less than $2.00 \text{D}$, and $+0.64$ for hyperopes measuring $2.00 \text{D}$ or greater. This relationship of increasing discrepancy between wet and dry findings with increasing hyperopia was not as clearly demonstrated by our autorefraction findings. In myopes the mean difference was $+0.001$; in hyperopes less than $2.00 \text{D}$ the mean difference was $+1.53$; and in hyperopes greater than $2.00 \text{D}$, the mean difference between wet and dry autorefraction was also $+1.53$. The expected relationship between dry and wet retinoscopy findings with increasing hyperopia allows the clinician to more accurately predict the true degree of hyperopia. Dry autorefraction findings may not hold this level of predictive value due to the high variability of accommodation between subjects.

Analysis of screening findings from this study along with those of similar studies demonstrates that dry retinoscopy provides a greater degree of sensitivity than dry autorefraction for detecting significant hyperopia in children. Both retinoscopy and autorefraction appear to be useful for detecting all other refractive conditions. Other studies evaluating the Retinomax as a screening instrument have demonstrated similar results. Cordonnier and Dramaix took into account the instrument myopia effect of the autorefractor by modifying their screening failure and true positive criteria to maximize sensitivity and specificity values. They found the best predictive values for hyperopia could be obtained with a failure criteria for non-cycloplegic autorefraction findings at $+1.50 \text{D}$ of hyperopia in order to detect absolute hyperopes $>3.50 \text{D}$ (as determined by wet autorefraction measures). In the initial study with 220 pediatric subjects, a sensitivity of 70.2% and specificity of 94.6% was achieved using these criteria. In a follow-up study with 302 children, the same criteria yielded a sensitivity and specificity of 46% and 97% respectively. A relatively low sensitivity (66%) for detecting anisometropia of $1.50 \text{D}$ or greater was also noted. Data reported by the Vision in Preschoolers Study Group indicates that modification of screening failure criteria for hyperopia may be a valid means of utilizing the Retinomax and Welch Allen Suresight Vision Screener for vision screening in young children. The screening failure criteria for hyperopia was modified to $\geq 1.50 \text{D}$ for Retinomax, $\geq 4.00 \text{D}$ for the Suresight and $\geq 2.75 \text{D}$ for non-cycloplegic retinoscopy in order to detect hyperopia $>3.25 \text{D}$ in any meridian. These modifications allowed specificity to be set at a minimum of 90% for each method. For detecting any of the highest risk conditions (amblyopia, constant strabismus, high ametropia (including hyperopia $\geq 5 \text{D}$) or anisometropia), non-cycloplegic retinoscopy provided an overall sensitivity at 90%, compared to Retinomax at 87% and Suresight at 81%. No data analysis is reported that specifically evaluates the sensitivity of each method in determining hyperopia alone. Therefore, direct comparison to the findings of this study is limited. However, the researchers found it necessary to lower the failure criteria for hyperopia by $0.50 \text{D}$ for non-cycloplegic retinoscopy and $1.50 \text{D}$ for non-cycloplegic autorefraction in order to achieve similar sensitivity and specificity with each method. This correlates well with the mean differences between non-cycloplegic and cycloplegic measures determined by this study.

Summary

While both non-cycloplegic autorefraction and retinoscopy tended to underestimate hyperopia in our subjects, this effect occurred with greater amplitude and variance for autorefraction measurements. The underestimation of hyperopia with non-cycloplegic autorefraction represents a clinically significant error for both screening detection and prescribing lenses in children. The low screening sensitivity for hyperopia limits the efficacy of infrared autorefraction in children’s vision screening without modification of the screening failure criteria to adjust for the instrument myopia effect. Despite these limitations, the autorefractor can be quite useful in pediatric refractive care. With adequate cycloplegia, the Retinomax and other infrared autorefractors have been shown to yield similar measures to cycloplegic retinoscopy in children. 20,21,25,37-41

Autorefraction may be particularly useful under cycloplegia because of the difficulty in determining the precise axis of astigmatism with retinoscopy due to the aberrations manifested with a dilated pupil. Autorefraction has also been proven to measure astigmatic power and axis on
children with clinically acceptable accuracy without cycloplegia.

The results of this study confirm that an experienced retinoscopist can provide valid measurements of refractive error for children's vision screenings. Distance retinoscopy is also the most accurate refractive method when prescribing lenses for children in situations where cycloplegia is contraindicated. However, the clinician should use appropriate optical fogging techniques and be observant for indicators of latent hyperopia such as fluctuating accommodation, pupillary constriction, esophoria or esotropia.

None of the researchers involved in this study have any financial or proprietary interest in the Retinomax.

Source
Nikon Retinomax
Right Medical Products
200 Expressway Court
Virginia Beach, VA 23462.

References


Corresponding author: John P. Lowery, O.D., M.Ed. Pacific University College Of Optometry 2043 College Way Forest Grove, OR 97116 lowery@pacificu.edu Date accepted for publication: December 15, 2004