Reliability of subjective, linear, ratio and area cephalometric measurements in assessing adenoid hypertrophy among different age groups

Marcelo Quiroga Souki; Bernardo Quiroga Souki; Leticia Paiva Franco; Helena Maria Gonçalves Becker; Eustáquio Afonso Araújo

ABSTRACT

Objective: To test the validity of four different types of lateral cephalometric radiograph (LCR) measurements as a diagnostic test of adenoid hypertrophy in different age groups of mouth-breathing children.

Materials and Methods: Eighty-six mouth-breathing children (male 54.65%, mean age 7.0 ± 2.2 years) were randomly selected from a hospital population. Adenoid obstruction of the nasopharynx was evaluated by subjective, linear, ratio, and area LCR measurements. Each measurement was compared with flexible fiberoptic endoscopy diagnosis.

Results: Kendall correlation coefficients for agreement between tests were ≥0.67 and kappa scores were substantial (≥0.64). Higher correlation coefficients and agreement values were found in older age groups. When the sample was stratified by age, the 3- to 5-year-old age group showed lower correlation coefficients and agreement strength for subjective, linear, and ratio measurements. The sensitivity of LCR varied from 71% (ratio) to 84% (linear). The specificity varied from 83% (linear) to 97% (ratio). The positive predictive value varied from 88% (linear) to 97% (ratio). The negative predictive value varied from 70% (ratio) to 78% (linear). The validity of each measure was different among the age groups.

Conclusions: LCR is a valid method for measuring adenoid hypertrophy in children from 6 to 12 years old. The diagnosis of adenoid hypertrophy, based on LCR measurements, in children with primary dentition (3–5 years old) should be made with caution. The combination of linear and ratio LCR measurements is a reliable screening tool to determine the need for an ear, nose, and throat evaluation. (Angle Orthod. 2012;82:1001–1007.)

KEY WORDS: Airway; Cephalometry; Adenoids; Nasal endoscopy

INTRODUCTION

The association between upper airway impairment and abnormal dentofacial growth of children has been reported frequently in the literature, yet some controversies still exist regarding whether it is a cause-effect relationship. Among the etiologic factors involved with severe nasal impairment, adenoid hypertrophy plays an important role. In recent decades, there has been a growing awareness of severe upper airway obstruction due to its association with sleep-disordered breathing and especially with obstructive sleep apnea in children. There is evidence that severely obstructed children should be promptly and effectively treated. Though early identification of severely obstructed children is recommended, the diagnosis of such breathing dysfunction is sometimes postponed because parents are unaware of the problem. As children visit pediatric dentists and orthodontists early in life, where children are subjected...
to a thorough medical history and radiologic survey, such practitioners have the chance to identify children with respiratory problems, especially if they have access to reliable diagnostic tools.

Lateral cephalometric radiography (LCR), used to evaluate orthodontic craniofacial morphology, has been advocated for at least 55 years as a method to identify nasopharyngeal airway obstruction. Since then, subjective, linear, ratio, and area measurements have been proposed. Despite the inability of LCR to provide volumetric data, and several claims against its use in the diagnosis of airway problems, there is evidence that LCR is a useful screening tool for evaluating nasopharyngeal obstruction. As is the case with all diagnostic tests, LCR has imperfections, and thus, its clinical reliability should be critically evaluated by four indices: sensitivity, specificity, positive predictive value, and negative predictive value. According to a systematic review, previous papers have not presented an adequate sensitivity-specificity analysis of LCR. Additionally, no study has compared the sensitivity and specificity of LCR as a tool for diagnosing nasopharynx obstruction to nasal flexible fiberoptic endoscopy (FFE), the current gold standard for diagnosing upper airway obstruction. As previous studies have not tested the strength of agreement between LCR and FFE in different age groups, a question arises: Are cephalogram measurements a reliable tool in the early diagnosis of adenoid hypertrophy?

The aim of this study, therefore, was to evaluate the sensitivity, specificity, accuracy of positive predictive value, and accuracy of predictive negative value of four types of cephalometric measurements in the assessment of severe adenoid hypertrophy in mouth-breathing children. In addition, we tested the correlation coefficients and agreement between LCR and FFE in different age groups. The null hypothesis was that correlation coefficients and agreement between LCR measurements and FFE findings were similar in all age groups.

MATERIALS AND METHODS

Subjects

The study protocol was approved by the Ethics Committees of the Federal University of Minas Gerais, Brazil (ETIC 291/03) and Pontifical Catholic University of Minas Gerais, Brazil (CAAE 0077021300006). An informed written consent was obtained from the parents before the subjects entered the study.

The parent sample involved 444 children who were referred consecutively by physicians between 2002 and 2007 to the Outpatient Clinic for Mouth-Breathers at the Federal University of Minas Gerais University Hospital. Based on the clinical and endoscopic ear, nose, and throat (ENT) examination performed by two otorhinolaryngologists in the first consultation, mouth breathing was confirmed, and nasopharyngeal obstruction by adenoidal tissue was classified into one of three categories: mild (<50%), moderate (50%–75%), and severe (>75%). Ninety children were then randomly selected, 30 from each of the three nasopharyngeal obstruction categories. To test the correlation coefficient and agreement between LCR and FFE, children were grouped into three age categories (3–5 years, 6–8 years, and 9–12 years).

ENT Assessment

The nasopharynx was evaluated with a standard nasopharyngeal videofiberoptic examination. The nasopharynx was classified in a dichotomous fashion as either obstructed (≥75% nasopharyngeal obstruction) or not obstructed (<75% obstruction).

Cephalometry

LCR was performed within 1 week of endoscopy using the same device for each exam (Siemens Orthophos CD, Bensheim, Germany), following the strict definition of a LCR. The cephalometric measurements used to evaluate airways are presented in Figure 1. To test the validity, cephalometric nasopharynges were classified as “obstructed” or “not obstructed” according to the guidelines presented in Figure 1.

Subjective measurement was based on the Holmberg and Linder-Aronson five-index scale. Grades 1, 2, and 3 were considered “not obstructive,” whereas grades 4 and 5 were classified “obstructive.” To evaluate linear, ratio, and area measurements, cephalograms were traced manually by an experienced orthodontist in a single-blinded fashion, using a 0.5-mm lead pencil on a standard light box. The examiner did not have prior knowledge of the patients’ endoscopic diagnoses, histories, or characteristics. Linear and ratio measurements were obtained by a ruler. The cut-off point of linear measurement was 5 mm. The ratio measure had six cut-off points based on the age groups and genders of the children.

Tracings were scanned (300 dots per inch on gray scale), imported into a commercially available software system (AutoCAD 2006, San Rafael, Calif.) and processed for area measurement (mm²). Cut-off points defining nasopharyngeal obstruction were based on a previous description.
Data Analysis

To calculate the measurement error, cephalograms of 25 randomly selected children were traced and measured twice, with a 2-week interval between the measurements. A Wilcoxon paired-sample test showed no significant difference between measurements ($P > .05$).

The endoscopy interobserver Spearman agreement correlation coefficient was 0.94. A Wilcoxon paired-sample test showed no significant difference between measurements ($P = .705$).

To compare data obtained from the two exams, a Kendall correlation coefficient and a kappa ($\kappa$) statistic were applied to these variables. Values of $\kappa$ were interpreted according to the patterns of agreement established by Landis and Koch. $P$ values $\leq .05$ were considered statistically significant. The area under the receiver operating characteristic (ROC) curve was evaluated for all cephalometric measurements at a significance level of 1%.

The validity of the diagnosis of nasopharyngeal obstruction was analyzed by measuring sensitivity, specificity, positive predictive value, negative predictive value, false positive rate, and false negative rate.

RESULTS

After quality control, four radiographs were excluded from the sample. Therefore, data from 86 children were included in the analysis (boys 54.65%, mean age $7.0 \pm 2.2$ years, range 3.0–12.7 years). Table 1 shows the distribution of children according to age groups and gender. There was no statistically significant difference in the distribution of genders ($P > .05$), and there was no statistically significant difference observed between the type of measurement (subjective, linear, ratio, or area) and gender ($P > .05$, Table 1); that is, analysis was gender-independent.

Table 2 shows the correlation coefficient ($r$) and $\kappa$ statistic resulting from the comparison of the diagnosis made using cephalometric radiography and that made by endoscopic examination. All nasopharyngeal cephalometric measurements showed significant positive correlations with endoscopic measurements ($r > 0.67$) when the total sample was analyzed. However, when the sample was age stratified, correlation varied from 0.44 to 0.76. The agreement was "substantial" ($\kappa > 0.64$) for interpreting the radiographic nasopharyngeal obstruction by adenoid hypertrophy in the total sample. Among the three age groups, $\kappa$ scores varied from 0.44 to 0.73.

![Figure 1](http://meridian.allenpress.com/angle-orthodontist/article-pdf/82/6/1001/1393247/010612-13_1.pdf)

Figure 1. Schematic drawings of LCR measurements and the cut-off points of each one. (A) Holmberg and Linder-Aronson subjective method. (B) McNamara’s linear method. (C) Ratio method of Kemaloglu et al. (D) Handelman and Osborne area method.

Table 1. Distribution of Children According to Gender and Age Groups

<table>
<thead>
<tr>
<th>Age Groups, y</th>
<th>Boys</th>
<th>Girls</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>3–5</td>
<td>19</td>
<td>10</td>
<td>29</td>
</tr>
<tr>
<td>6–8</td>
<td>20</td>
<td>17</td>
<td>37</td>
</tr>
<tr>
<td>9–12</td>
<td>8</td>
<td>12</td>
<td>20</td>
</tr>
<tr>
<td>Total</td>
<td>47</td>
<td>39</td>
<td>86</td>
</tr>
</tbody>
</table>

$^a \chi^2 = 3.11 (2 df); P = .210.$

Angle Orthodontist, Vol 82, No 6, 2012
Results showed a strong correlation between LCR and FFE, with an area under the ROC curve of 0.84 (95% CI, 0.75–0.93) at a significance level of $P < .01$ (Figure 2).

Differences in the pattern of correlation coefficients and agreement were found in the three age levels investigated. The 3- to 5-year-old age group showed weak correlations ($r$ and $k$), while older children (6–8 and 9–12) showed a strong correlation (Table 2).

The sensitivity of the cephalometric measurements varied between 71% and 84%. The specificity was higher than 83%. The positive predictive value of the three cephalometric measurements was very high (88% to 97%). Thus, the chance of a false positive is small. The ratio measurement showed a 3% chance of a false positive. However, the negative predictive value for all measurement types was not elevated (70% to 78%) (Table 3). The subjective measurement showed moderate sensitivity (76%) and high specificity (91%).

Validity indices varied within each type of measurement among the three age groups.

**DISCUSSION**

Over the past few decades, obstructive sleep apnea (OSA), a life-threatening disorder, has become recognized as the most extreme variety of mouth breathing on the wide spectrum of upper airway obstruction. Thus, early diagnosis and treatment of children with OSA is indicated to normalize breathing function.

The major cause of OSA in children is adenotonsillar hypertrophy and the severity of OSA is associated with the size of the hypertrophic adenoids. Though the use of cavum and lateral cephalometric radiography has been described in the medical and dental literature for years, it has also been the target of criticism in some studies. These skeptical conclusions, though, were questioned in a systematic review on this topic. According to this literature review on cephalometric diagnosis of posterior nasopharynx obstruction, study issues were present, including serious deficiencies in sample size, absence of blinding, age factors, and the absence of a rigorous gold standard. Knowing that LCR is routinely used in the orthodontic clinic and that the early recognition of severe airway problems is of interest to this specialty, a more thorough assessment of the ability of LCR to diagnose nasopharyngeal obstruction by the adenoids is critical.

Accordingly, the present investigation was designed and used the comprehensive methodologic scoring guidelines proposed by Major et al. in their systematic review. The purpose of this study was to overcome the aforementioned limitations present in previous investigations. This study intended to recruit a well-described randomly selected sample. We opted to include children from 3 to 12 years of age, rather than 5 to 17 years, as suggested by the guideline, because the former age group is at a higher risk for developing nasopharyngeal obstruction due to growth characteristics of lymphoid tissues. Diagnosing adenoid hypertrophy in very young
children (3–5 years old) would prevent facial growth abnormalities and the associated risks of OSA, while adolescents rarely have nasopharyngeal obstruction.

The larger sample size, inclusion of a control group (mouth breathing without adenoid hypertrophy), prospective design (instead of retrospective), and use of strict LCR methods were in accordance with the recommended guidelines. Four well-established measurements from the orthodontic literature were chosen: the Holmberg and Linder-Aronson subjective method,\textsuperscript{15} McNamara's linear method,\textsuperscript{16} the Kemaloglu et al.\textsuperscript{17} ratio, and the Handelman and Osborne\textsuperscript{18} area method.

Recently, three-dimensional imaging has progressively become a more accessible tool, and some investigations have used it in the assessment of the pediatric upper airway.\textsuperscript{22,28} However, because the current gold standard in otorhinolaryngology used to evaluate adenoid hypertrophy and upper airway obstruction is flexible nasoendoscopy,\textsuperscript{19} this method was chosen for our comparisons.

Our results showed a $\kappa$ score $\geq 0.64$, which represents substantial agreement according to the Landis and Koch interpretation scale.\textsuperscript{26} The degree of correlation (Kendall test) was very similar among the different radiographic measurements and nasal endoscopy ($r = 0.67–0.69$). Other authors,\textsuperscript{23,30} despite using other methods for comparison, found similar correlation coefficients. The area under the ROC curve of 0.84 suggests a strong correlation between diagnosis made according to LCR and FFE. Because lymphoid tissues and the nasopharynx have a singular growth pattern\textsuperscript{14–18} over time, the evaluation of upper airway problems should take age into account, but this has not been the case in previous reports. Therefore, we tested the hypothesis that correlation coefficients and agreement between LCR measurements and FFE findings were similar in all age groups. When children were grouped by age, the correlation coefficients as well as the strength of agreement were different. Young children (3–5 years old) showed weak values (r and $\kappa$) when compared to older children (6–8 and 9–12 years old) and to the total scores. These findings suggest that LCR should be used with caution in children with primary dentition (under 6 years old), though this reduces the benefit of very early diagnosis of nasopharyngeal obstruction by orthodontists and pediatric dentists.

Nevertheless, our findings show that in children with early mixed dentition (6–8 years old), the agreement between linear and ratio LCR measurements is very consistent ($\kappa = 0.73$); therefore, LCR can be used with more confidence. Unfortunately, the most susceptible children (3–5 years old) for tonsils and adenoid hypertrophy, as well as those whose diagnosis is more important, since it may prevent dentofacial alterations, are the ones when LCR most failed in correct diagnosis. These are the ones for whom diagnosis is more important, since it may prevent dentofacial alterations. As this is the first investigation reporting the strength of agreement between LCR and FFE across different age

<table>
<thead>
<tr>
<th>Cephalometric Measurements</th>
<th>Sensitivity, %</th>
<th>Specificity, %</th>
<th>Positive Predictive Value, %</th>
<th>Negative Predictive Value, %</th>
<th>False Positive, %</th>
<th>False Negative, %</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Subjective</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>76</td>
<td>91</td>
<td>93</td>
<td>73</td>
<td>7</td>
<td>23</td>
</tr>
<tr>
<td>Age groups</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3–5</td>
<td>73</td>
<td>86</td>
<td>94</td>
<td>50</td>
<td>6</td>
<td>50</td>
</tr>
<tr>
<td>6–8</td>
<td>80</td>
<td>94</td>
<td>94</td>
<td>80</td>
<td>6</td>
<td>20</td>
</tr>
<tr>
<td>9–12</td>
<td>78</td>
<td>91</td>
<td>88</td>
<td>83</td>
<td>12</td>
<td>17</td>
</tr>
<tr>
<td><strong>Linear</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>84</td>
<td>83</td>
<td>88</td>
<td>78</td>
<td>12</td>
<td>22</td>
</tr>
<tr>
<td>Age groups</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3–5</td>
<td>73</td>
<td>86</td>
<td>94</td>
<td>50</td>
<td>6</td>
<td>50</td>
</tr>
<tr>
<td>6–8</td>
<td>80</td>
<td>94</td>
<td>94</td>
<td>80</td>
<td>6</td>
<td>20</td>
</tr>
<tr>
<td>9–12</td>
<td>78</td>
<td>91</td>
<td>88</td>
<td>83</td>
<td>12</td>
<td>17</td>
</tr>
<tr>
<td><strong>Ratio</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>71</td>
<td>97</td>
<td>97</td>
<td>70</td>
<td>3</td>
<td>30</td>
</tr>
<tr>
<td>Age groups</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3–5</td>
<td>64</td>
<td>100</td>
<td>100</td>
<td>47</td>
<td>0</td>
<td>53</td>
</tr>
<tr>
<td>6–8</td>
<td>75</td>
<td>100</td>
<td>100</td>
<td>77</td>
<td>0</td>
<td>23</td>
</tr>
<tr>
<td>9–12</td>
<td>78</td>
<td>91</td>
<td>88</td>
<td>83</td>
<td>12</td>
<td>17</td>
</tr>
<tr>
<td><strong>Area</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>75</td>
<td>94</td>
<td>95</td>
<td>72</td>
<td>5</td>
<td>28</td>
</tr>
<tr>
<td>Age groups</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3–5</td>
<td>77</td>
<td>100</td>
<td>100</td>
<td>59</td>
<td>0</td>
<td>41</td>
</tr>
<tr>
<td>6–8</td>
<td>70</td>
<td>94</td>
<td>93</td>
<td>73</td>
<td>7</td>
<td>27</td>
</tr>
<tr>
<td>9–12</td>
<td>78</td>
<td>91</td>
<td>88</td>
<td>83</td>
<td>12</td>
<td>17</td>
</tr>
</tbody>
</table>
groups, and differences in correlation were found among these groups, further studies on this topic are warranted due to their clinical importance.

The findings of the present study showed high values for specificity and positive predictive value, reflecting a reduced false positive rate. On the other hand, the sensitivity and negative predictive value showed a moderate diagnostic ability, reflecting a relatively high false negative rate. Among the four measurements methods the linear demonstrated the more homogeneous results (12% false positive and 22% false negative rate). Nevertheless, the combination of ratio measurement (3% false positive rate) and linear measurement (22% false negative rate) seems to be the best option, regardless of the age group.

Those measurements are easy to perform, are not time-consuming and do not need any computer assistance to calculate airway area. Our results demonstrate that the validity indices are different in each age group; therefore, we suggest that such numbers should be used only as a reference guide.

From a clinical point of view, proper interpretation of these results makes LCR an important tool for the orthodontist. When a patient, especially a child over 6 years of age, has a radiographic indication of nasopharynx obstruction by the adenoids, the risk for sleep-disordered breathing, including OSA, is high. On the other hand, in situations when the image does not suggest that the size of the adenoid tissue is increased, there is a chance of a false negative result. Therefore, regardless of the age group, if the patient has other clinical features of mouth breathing, referral for a thorough ENT assessment should be recommended.

The orthodontist should be aware of the diagnosis of respiratory obstruction, not only because of the dentofacial implications but also because of issues related to ethical and legal liability. Recognizing children with severe nasal impairment can be facilitated by using LCR, an assessment methodology used daily in an orthodontic office. Overlooking mouth breathing exposes children not only to an etiologic factor contributing to malocclusion, but also to OSA. From a legal standpoint, the high positive predictive value of LCR implies that the orthodontist should not underestimate the utility of radiographic visualization as a measurement of adenoid hypertrophy.

CONCLUSIONS

- The reliability of LCR methods for diagnosing adenoid hypertrophy varies among different age groups.
- LCR is a valid method for measuring adenoid hypertrophy in children from 6 to 12 years old.
- A diagnosis of adenoid hypertrophy based on LCR measurements should be made with caution in children with primary dentition (3–5 years old).
- When LCR suggests severe adenoid hypertrophy, there is a high chance that the nasopharynx is obstructed.
- The combination of linear and ratio measurements on LCR is a reliable screening tool to determine the need for an ENT evaluation.

REFERENCES


