Correlation between the modified Huddart and Bodenham index and the GOSLON yardstick for assessing occlusal characteristics at 5 and 10 years of age in individuals born with unilateral cleft lip and palate

Yazeed Almuhizi1, Anette Leser2 and Marie Pegelow3

1Division of Orthodontics, Department of Dental Medicine, Security Forces Hospital, Riyadh, Saudi Arabia, 2Department of Orthodontics, Folktondvården Sörmland, Nyköping, Sweden and 3Division of Orthodontics and Jaw Orthopedics, Department of Dental Medicine, Karolinska Institute, Stockholm Craniofacial Team, Stockholm, Sweden

Correspondence to: Marie Pegelow, Division of Orthodontics, Department of Dental Medicine, Karolinska Institute, Box 4064, 141 04 Huddinge, Sweden. E-mail: marie.pegelow@ki.se

Summary

Background: The Great Ormond Street, London and Oslo, Norway (GOSLON) yardstick and the modified Huddart and Bodenham (MHB) index are two indices for evaluating study models to discern the effects of treatment in individuals born with unilateral cleft lip and palate (UCLP). GOSLON is a tool for ranking study models based on their sagittal, vertical, and transversal planes that requires calibration of the examiners. MHB is a tool for scoring the arch constriction between the jaws for each tooth pair, with scores ranging from +1 to −3.

Aim: Our study aimed to identify the degree of correlation between the GOSLON and MHB indices and to compare inter- and intra-examiner reliability.

Material: The study scored 107 study models of 76 children aged 5 (n = 50) and 10 (n = 57) years using the MHB and GOSLON indices.

Results: We found a very high intra- and inter-examiner reliability (>0.90) for the MHB and the GOSLON indices as well as a moderate to high negative correlation between the two. Among 5-year-olds, the relationship between the MHB and the GOSLON was linear, while among 10-year-olds, GOSLON Groups 2 and 4 occurred most frequently, with a corresponding wide range in MHB.

Conclusion: Standardization of maxillary arch constriction measurements following CLP treatment would help evaluation of treatment outcome and make comparability between studies possible. We found a moderate to high correlation between GOSLON and MHB. Both indices showed very good intra- and inter-examiner reliability. MHB is a more sensitive system compared to GOSLON, with a greater range of scoring (40 grades) than GOSLON (five categories).

Introduction

Management of cleft lip and palate (CLP), which is the most common congenital defect involving the face and jaws, requires a multidisciplinary approach for surgical and non-surgical treatment (1). Malformations of craniofacial growth specific to CLP can range from a midface deficiency and a constricted maxillary dental arch to congenitally missing and malformed teeth (2, 3).

The goal of surgical intervention to repair CLP is restoration of normal morphology and function without disrupting growth potential. As yet, no agreement on optimal technique (4) or timeframe
We wanted to evaluate...on 05 October 2017.

...by guest Downloaded from https://academic.oup.com/ejo/article-abstract/38/4/359/1748917/Correlation-between-the-modified-Huddart-and

...the mandible. The tooth scores are then summed: the lower the total score, the greater the maxillary arch constriction (11). Mossey et al. (11) modified it for use in the mixed dentition as well.

The aims of this study were to assess and compare the inter- and intra-examiner reliability of the GOSLON and MHB indices and to identify the degree of correlation between them.

**Hypotheses**

Our study had two hypotheses: 1. the reliability of the two indices will differ, and 2. the two indices will have no correlation.

**Materials and methods**

The material comprised 111 plaster study models from 76 children born between 1991 and 2005 who had been treated by the Stockholm Craniofacial Team; they were selected for the study by consecutive inclusion. All individuals had non-syndromic (NS) UCLP (diagnosis code in ICD10: Q37.5).

The study models available for our research had been made when the patients were aged 5 (57 models) and 10 years (50 models). The study excluded four individuals due to broken or missing models. Each model received a number from 1 to 107. All study models were of good quality and accurately trimmed. We included patients even if they had had orthodontic treatment before bone grafting, and we scored the models using both the GOSLON and the MHB.

Our study recruited two examiners (A and Y), by the time postgraduate dentists participating in the Orthodontic specialisation programme at Karolinska Institutet in Stockholm. Examiner calibration was done in a pilot study, where the 10 GOSLON reference models were used to analyse the study models of 20 patients. The calibration process also used software for GOSLON calibration (not yet in commercial use). Figures 1–5 show the GOSLON categories, while we present the instructions for administering the MHB.

The examiners scored the study models on four occasions, recording scores for the GOSLON on two occasions and the MHB on two occasions; the examiners also recorded cleft side and gender for each patient. The examiners scored the models independently. The washout period between scoring the models using the same index was 8 weeks. Examiners recorded results on a separate sheet of paper at each scoring session.

**Figure 1.** A representative model for Group 1 of the Great Ormond Street, London and Oslo, Norway yardstick that shows the most favourable skeletal relationship with a positive overjet and overbite. Cases in this category exhibit Angle class II Division 1 malocclusion.

**Figure 2.** A representative model for Group 2 of the Great Ormond Street, London and Oslo, Norway yardstick that shows a favourable skeletal relationship that can only be corrected with orthodontic treatment. Cases in this category exhibit Angle class I malocclusion.

**Figure 3.** A representative model for Group 3 of the Great Ormond Street, London and Oslo, Norway yardstick that shows an edge-to-edge occlusion that will need more complex orthodontic treatment, though results are likely to be good.
Modified Huddart and Bodenham scoring instructions: All maxillary teeth, except the lateral incisors, are scored according to their bucco-lingual relationship to the corresponding mandibular tooth. The scores are as follows: +1 = more than normal relation, 0 = normal relation, −1 = edge-to-edge relation; −2 = crossbite/inversion relation, −3 = more than crossbite/inversion relation. Individual scores are added together to give a score for each set of models. The more negative the score, the greater the maxillary arch constriction. Before the age of 6 years, the first permanent molars are not scored, even if they have erupted; thus, scores can range between −24 and +8 (on a 32-point scale). After the age of 6 years, first permanent molars are scored, if present; if absent, the midpoint of the maxillary alveolar ridge is used. In this case, scores can range from −30 to +10 (on a 40-point scale). Premolars should be scored as for primary molars. If a central incisor is missing, the other central incisor is used to score the missing one. If a canine is unerupted, its score is determined by the midpoint of the maxillary alveolar ridge. If a premolar is absent, a score is allocated equivalent to the adjacent premolar, if erupted. If no premolars are erupted, the score is determined by the midpoint of the maxillary alveolar ridge (12, 13).

The Ethics Committee in Stockholm approved this study [Daybook no. (Dnr) 2012/2182-31/4].

Statistics
Statistical analysis used IBM SPSS Statistics version 22.0 (IBM Corp., Armonk, New York, USA). To determine inter- and intra-examiner reliability, we used an intra-class correlation coefficient (ICC) for MHB, which is a continuous scale and has quantitative variables (14). A score of 1 would indicate perfect agreement. Statistical calculations are based on the total MHB score. We used Cohen’s weighted kappa (κ) statistics to analyse the reliability of the GOSLON, which is a categorical scale and has ordinal qualitative variables (15, 16). Strength of agreement of κ is as follows: <0.2 = Poor; 0.21–0.40 = Fair; 0.41–0.60 = Moderate; 0.61–0.80 = Good; 0.81–1.00 = Very good.

Spearman’s rank correlation coefficient and Kendall’s rank correlation coefficient revealed the degree of correlation between MHB scores and GOSLON scores. These tests assess how well the relationship between two variables can be described; perfect correlations in these tests have a score of 1 (17). The mean and median values for the GOSLON and the MHB were calculated from the means of four measurements (each examiner made two) on each model.

Results
In this study, most cases were boys, and in both genders, the left was most common on the left side (Table 1).

Inter- and intra-examiner reliability
The MHB index: Intra- and inter-examiner reliability were very good (Table 2).

The GOSLON yardstick: Both intra- and inter-examiner reliability were very good (Table 3) according to Altman (16). Table 4 shows a cross tabulation between the examiners.

Correlation
The negative correlation between the MHB and the GOSLON was moderate to high: Spearman’s rank correlation coefficient was −0.7, and Kendall’s rank correlation coefficient was −0.6 (Table 5).

Among 5-year-olds, the means for the GOSLON and the MHB were 3.04 and −6.86, respectively. In the 10-year-olds, the means were 2.96 for the GOSLON and −7.06 for the MHB (Table 6). Most cases in the 5-year age group were classified as Group 2, 3, or 4 on the GOSLON yardstick (Table 7), while in the 10-year age group, most were classified as Group 2 or 4 with only a few cases in Group 3 (Table 8).

The grouping of the continuous numerical scale of the MHB into the five categories of the GOSLON demonstrates that there is a trend from best to worst in a negative correlation. In 5-year-olds, GOSLON Groups 2 and 3 had an almost identical range on the MHB scale. GOSLON Group 4 showed a slightly lower value on the MHB, both in range and on average, however, the values were overlapping (Figure 6A). This group had one outlier (ID 41), a single measurement with a value 1.5 box lengths away from the box. GOSLON Group 1 contained only one patient (Figure 6A).

The 10-year-olds showed two GOSLON groups consistently occurring: these were Groups 2 and 4. To some extent, these groups showed overlapping values in the MHB. In GOSLON Group 2, the box was shorter and the values more clustered around the median. In Group 4, the box was taller and there was a larger range of values around the median. This shows less variance in Group 2 than in Group 4. Group 2 contained one outlier (ID 66; Figure 6B).

In both age groups, some patients who were classified in the same GOSLON group received different values in the MHB, and some with the same MHB value were classified into different GOSLON groups. For example, in the 5-year age group, ID 5 and ID 26 were both classified into GOSLON Group 2 but received MHB values that differed greatly. In contrast, ID 93 and ID 1, for example, received the same MHB value but were classified into GOSLON Groups 2 and 4, respectively (Figure 7A). Among 10-year-olds, three cases (IDs 107, 13, and 2) had almost the same MHB value but were classified into three different GOSLON groups. Two individuals (ID 75 and 99) were both categorized in GOSLON Group 4 but received MHB values that varied greatly (Figure 7B).

When mean values were not rounded up, the ranges in MHB scores for the cases classified in GOSLON Groups 2 and 3 were especially large; and the ranges were roughly the same in both groups. GOSLON Group 4 corresponded to a rather large range of MHB scores: from approximately −6 to −18. GOSLON Group 5 had...
only two patients and one of them had a mean MHB of −14, while the other had a mean MHB of −19. Overall, there seemed to be a linear relationship between MHB and GOSLON among 5-year-olds (Figure 8A). As described earlier, in the 10-year age group, GOSLON scores seemed to cluster into two groups rather than linearly; the two most frequently occurring GOSLON groups were Groups 2 and 4. The range of MHB scores for the cases classified in GOSLON Group 4 was large (between −5 and −19; Figure 8B).

**Discussion**

The evaluation of the reliability of the MHB scoring system showed that both examiners had very good intra- as well as inter-examiner reliability. ICC scores were above 0.95 for both examiners, while inter-examiner agreement was similarly high for both assessments.

**Table 1.** Cleft side and gender in the two groups of study models, made at 5 and 10 years of age, of children with congenital non-syndromic cleft lip and palate.

<table>
<thead>
<tr>
<th>Cleft side</th>
<th>Number of cases</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Right</td>
</tr>
<tr>
<td>Boys</td>
<td>47</td>
</tr>
<tr>
<td>Girls</td>
<td>29</td>
</tr>
</tbody>
</table>

**Table 2.** Inter-examiner reliability of measurements with the modified Huddart and Bodenham index. Df, degrees of freedom; ICC, intra-class correlation coefficient.

<table>
<thead>
<tr>
<th>Occasion</th>
<th>Lower bound</th>
<th>Upper bound</th>
<th>F-test true value</th>
<th>Df1</th>
<th>Df2</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st rating</td>
<td>0.98</td>
<td>0.97</td>
<td>0.99</td>
<td>100.3</td>
<td>106</td>
<td>106</td>
</tr>
<tr>
<td>2nd rating</td>
<td>0.98</td>
<td>0.98</td>
<td>0.99</td>
<td>118.7</td>
<td>106</td>
<td>106</td>
</tr>
</tbody>
</table>

**Table 3.** Inter- and intra-examiner reliability of measurements with the Great Ormond Street, London and Oslo, Norway yardstick.

<table>
<thead>
<tr>
<th>Weighted κ</th>
<th>Standard error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intra-examiner reliability</td>
<td>Examiner A: 0.91, 0.053</td>
</tr>
<tr>
<td>Inter-examiner reliability</td>
<td>Assessment 1: 0.95, 0.039</td>
</tr>
</tbody>
</table>

**Table 4.** Score of cross tabulation for the first assessment by the examiners with the Great Ormond Street, London and Oslo, Norway yardstick.

<table>
<thead>
<tr>
<th>GOSLON 1st assessment–Examiner Y</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>GOSLON 1st assessment–Examiner Y</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Examiners A</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Total</td>
<td>7</td>
<td>38</td>
<td>17</td>
<td>40</td>
<td>5</td>
<td>107</td>
</tr>
</tbody>
</table>

**Table 5.** Non-parametric correlations between the Great Ormond Street, London and Oslo, Norway yardstick and the modified Huddart and Bodenham index among 5- and 10-year-olds with congenital unilateral cleft lip and palate.

<table>
<thead>
<tr>
<th>Age group (years)</th>
<th>Number (N)</th>
<th>Kendall’s τ</th>
<th>Spearman’s ρ</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>57</td>
<td>−0.61</td>
<td>−0.74</td>
</tr>
<tr>
<td>10</td>
<td>50</td>
<td>−0.64</td>
<td>−0.77</td>
</tr>
</tbody>
</table>

**Table 6.** Means, medians, and SD of four measurements on models of 5- and 10-year-old children with congenital unilateral cleft lip and palate using both the Great Ormond Street, London and Norway, Oslo yardstick (GOSLON) and the modified Huddart and Bodenham (MHB) index.

<table>
<thead>
<tr>
<th>Age group (years)</th>
<th>Mean</th>
<th>Median</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-year age group (N = 57)</td>
<td>GOSLON: 3.04</td>
<td>3.00</td>
<td>0.94</td>
</tr>
<tr>
<td>MHB: −6.86</td>
<td>−6.75</td>
<td>4.61</td>
<td></td>
</tr>
<tr>
<td>10-year age group (N = 50)</td>
<td>GOSLON: 2.96</td>
<td>3.00</td>
<td>1.14</td>
</tr>
<tr>
<td>MHB: −7.06</td>
<td>−6.00</td>
<td>5.81</td>
<td></td>
</tr>
</tbody>
</table>

**Table 7.** Cross tabulation of the Great Ormond Street, London and Oslo, Norway yardstick (GOSLON) and the modified Huddart and Bodenham (MHB) index.

<table>
<thead>
<tr>
<th>MHB</th>
<th>GOSLON</th>
<th>Mean</th>
<th>SD</th>
<th>Min</th>
<th>Median</th>
<th>Max</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>−0.25</td>
<td>0</td>
<td>−0.25</td>
<td>−0.25</td>
<td>−0.25</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>−3.68</td>
<td>3.11</td>
<td>−11.25</td>
<td>−3.00</td>
<td>0.25</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>−5.69</td>
<td>2.91</td>
<td>−10.75</td>
<td>−6.25</td>
<td>−0.25</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>−10.36</td>
<td>3.29</td>
<td>−18.00</td>
<td>−10.00</td>
<td>−6.00</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>−16.63</td>
<td>3.71</td>
<td>−19.25</td>
<td>−16.63</td>
<td>−14.00</td>
<td>2</td>
<td></td>
</tr>
</tbody>
</table>

Descriptive statistics for the distribution of MHB and GOSLON in the 5-year age group.

**Table 8.** Cross tabulation of the Great Ormond Street, London and Norway, Oslo yardstick (GOSLON) and the modified Huddart and Bodenham (MHB) index.

<table>
<thead>
<tr>
<th>MHB</th>
<th>GOSLON</th>
<th>Mean</th>
<th>SD</th>
<th>Min</th>
<th>Median</th>
<th>Max</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>−1.00</td>
<td>3.72</td>
<td>−5.25</td>
<td>−0.50</td>
<td>2.25</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>−3.17</td>
<td>3.40</td>
<td>−9.25</td>
<td>−2.75</td>
<td>2.00</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>−6.81</td>
<td>3.78</td>
<td>−12.25</td>
<td>−5.75</td>
<td>−3.50</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>−10.71</td>
<td>4.43</td>
<td>−18.75</td>
<td>−10.75</td>
<td>−4.75</td>
<td>21</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>−18.25</td>
<td>1.77</td>
<td>−19.50</td>
<td>−18.25</td>
<td>−17.00</td>
<td>2</td>
<td></td>
</tr>
</tbody>
</table>

Descriptive statistics for the distribution of MHB and GOSLON in the 10-year age group.

These results are in line with the results reported in other studies (12, 18, 19).

GOSLON showed very good intra-examiner agreement, and the inter-examiner agreement was similarly high for both the first and second assessments. These results also agree with the results of other studies (9, 18, 20, 21). Different assessment methods generate different types
of data including various measures of reliability. As such, it was not possible to directly compare the reliability scores between the methods.

Our findings that CLP was more common in boys than in girls and that the cleft was most common on the left side (Table 1) agree with findings reported in other studies (3, 22, 23).

We used Spearman’s and Kendall’s rank correlation coefficients to determine how MHB scores correlate with GOSLON categories. These tests describe the relationship between two variables (17). In this study, Spearman’s and Kendall’s correlation coefficients for the MHB–GOSLON correlation were good. Thus, the correlation between MHB and GOSLON was moderate to high and negative: the higher the GOSLON category, the lower the MHB score (Table 5). Mossey et al. (11) and Gray and Mossey (12) also found a statistically significant correlation between GOSLON and MHB.

Among 5-year-olds, most cases were classified into GOSLON Groups 2, 3, and 4 (Table 7). GOSLON suggests categorizing occlusion with an edge-to-edge relationship in Group 3 (9). But because an edge-to-edge bite is normal in the primary dentition, Mars et al. (21) suggested that an edge-to-edge incisor relationship with a minimal crossbite tendency should be classified in GOSLON Group 2 when the patient is 5 years old. We found that GOSLON Groups 2 and 3 have almost identical ranges of MHB scores in the 5-year age group (Figure 6A), which is in line with what Mars et al. (21) found. Thus, merging Groups 2 and 3 when assessing the primary dentition of 5-year-olds could be recommended. Among 10-year-olds, most cases were classified as GOSLON Group 2 or 4 (Table 8).

Dental arch relationships in some patients improved over the 5-year period, perhaps due to a combination of favourable growth and successful surgery, or because of orthodontics. Other patients experienced negative occlusal development. In this study, some patients in the 10-year age group had undergone orthodontic treatment before bone grafting and this, of course, improved the dental arch relationship and affected the spread of the data. Since the aim of this study was to compare two indices, not to judge treatment outcome, we decided to include these patients.

Figure 6. (A) Box plot for the distribution of the modified Huddart and Bodenham (MHB) index subdivided by the Great Ormond Street, London and Oslo, Norway (GOSLON) yardstick among 5-year-olds. GOSLON uses the median value of the four measurements rounded up to the closest whole number (e.g. 2.5 → 3). For MHB, all values are the mean of all four measurements. (B) Box plot of the distribution of the MHB index, subdivided by the GOSLON yardstick among 10-year-olds. The GOSLON yardstick uses the median of the four measurements rounded up to the closest whole number (e.g. 2.5 → 3). For MHB, values are the means of the four measurements.

Figure 7. (A) Scatter plot identifying cases with unequal modified Huddart and Bodenham (MHB) measurements to GOSLON yardstick among 5-year-olds. GOSLON uses the median of the four measurements rounded up to the closest whole number (e.g. 2.5 → 3). (B) Scatter plot identifying cases with unequal MHB measurements to GOSLON yardstick among 10-year-olds. GOSLON uses the median of the four measurements rounded up to the closest whole number (e.g. 2.5 → 3).
Various national and international studies have used GOSLON, which is more often used than other indices, because it has been available since 1987. Several studies evaluating the dental arch relationship in individuals with UCLP have used GOSLON, including the six centre comparison study in England (10). Such efforts to improve the treatment outcome of patients with CLP in randomized control trials are promoted by WHO because these studies can reduce the health care burden of the patients (25).

The GOSLON groupings were sufficiently sensitive to distinguish between treatment outcomes at the six centres (10). One of the limitations of GOSLON is that the yardstick requires a substantial degree of professional judgment concerning possible orthodontic correction, which introduces an element of subjectivity. Its second limitation is that calibration, with its use of 10 reference models for comparison, is time-consuming, but essential to ensure accuracy (11).

To really study the effect of surgical treatment on midface growth, we would require a sample of 10-year-old individuals born with UCLP treated only with primary surgery. Gathering such a sample is likely to prove impossible as most 10-year-olds have already undergone orthodontic, restorative, and bone-grafting procedures that could obscure the effects of the primary surgery.

The greater understanding of the factors adversely affecting treatment outcome has improved treatment protocols. As a result, surgical outcomes are better, protocols are more standardized, and differences between cleft centres are decreased. Current differences in treatment outcome are thus more difficult to detect with the five group GOSLON yardstick than with the MHB scoring system which has a larger range of scores capable of discerning fine distinctions in the arches.

Noverraz et al. (24) found that GOSLON is useful for assessing dental arch relationships at all stages of development and is suitable for longitudinal research, though GOSLON was originally developed for use on study models of 10-year-old individuals with UCLP (9). However, using GOSLON to score both 5- and 10-year-old models helped reduce the number of indices to be compared, achieving accuracy with less sophisticated statistics. Earlier studies also found that the GOSLON required less time in study model assessment than the MHB (18). The present study did not evaluate testing time, but the consensus among the examiners was that the GOSLON required less time than the MHB.

Altalibi et al. (26) conducted a systematic review of these indices and concluded that GOSLON is used more frequently than any of the other indices, possibly due to its longer time in use. The MHB, however, equalled or outperformed all other indices in all of the WHO criteria for an ideal index. Altalibi et al. (26) recommend using the MHB to assess malocclusions related to all types of CLP at all ages. Standardization of the measurement facilitates international inter-centre studies and allows optimization of cleft treatment protocols. The MHB provides a sensitive and objective assessment of maxillary arch constriction, and can be used to score both deciduous and permanent teeth. The scoring system also minimizes the effect of random operator error since the cumulative score derives from eight or ten separate categorical assessments (27).

Clinicians always compare maxillary constriction with the width of the lower arch, but it is also important to compare cleft and non-cleft sides. The MHB makes such comparison easy, because by recording the number and nature of dental arch segment displacements and dental crossbite, the MHB provides more information on sites of occlusal discrepancy than the GOSLON.

The MHB is a continuous scale of the severity of arch constriction, while the GOSLON is an ordinal, categorical scale. The greater range of the MHB scoring scale (32 or 40 points) increases

A few patients who received the same GOSLON score in both age groups received varying MHB scores, while some who received the same MHB scores in both age groups were classified into different GOSLON groups (Figure 7A and 7B). A reasonable explanation for this disagreement could be that GOSLON and MHB measure different components. GOSLON considers dental compensations and the vertical plane, evaluating how the malocclusion will be treated. MHB, on the other hand, is not a treatment-related index and does not assess the vertical relationship. Noverraz et al. (24) pointed out that the early mixed dentition (around age 8) is the most difficult stage to classify into a correct GOSLON group. Other possible factors that might adversely affect the correlation between the indices are the element of subjectivity in classification or unsuccessful examiner calibration. In this study, the two examiners had similar clinical experience, but some subjectivity is always a risk.

![Figure 8](image-url)
sensitivity, allowing differentiation of severity within each GOSLON group. The MHB is also more versatile, as it can be applied to any cleft type at any age (27). As we did, Gray and Mossey (12) evaluated the use of MHB and concluded that MHB is a reliable indicator since intra- and inter-rater agreement were both excellent.

The MHB is easy to use since it does not require calibration and has no need for reference models. Orthodontic assistants in Sweden receive professional education about dental crossbite and could learn to score the models, which would improve the efficiency of clinicians providing cleft care (11, 27).

Conclusion

Standardization of the measure used to assess maxillary arch constriction following CLP treatment would help evaluation of treatment outcome to improve treatment quality and make comparability between studies possible.

Our two hypotheses were rejected: GOSLON and MHB showed moderate to high correlation, and both the MHB and the GOSLON showed very good intra- and inter-examiner reliability. The MHB is a more versatile and sensitive system than the GOSLON, with the moderate to high correlation, and both the MHB and the GOSLON between studies possible.

Conclusion outcome to improve treatment quality and make comparability between studies possible.

Our two hypotheses were rejected: GOSLON and MHB showed moderate to high correlation, and both the MHB and the GOSLON showed very good intra- and inter-examiner reliability. The MHB is a more versatile and sensitive system than the GOSLON, with the MHB providing a greater range of scoring (40 grades) and revealing more detailed information about the malocclusion compared to the five-category GOSLON yardstick.

References