The Goslon Yardstick: A New System of Assessing Dental Arch Relationships in Children with Unilateral Clefts of the Lip and Palate

Michael Mars, F.D.S., D. Orth.
Dennis A. Plint, F.D.S., D. Orth.
William J.B. Houston, Ph.D., F.D.S.
Olav Bergland, D.D.S., Dr. Odont.
Gunvor Semb, D.D.S.

The Goslon Yardstick is a clinical tool that allows categorization of the dental relationships in the late mixed and or early permanent dentition stage into five discrete categories. Cases are allocated to these categories on a value judgment basis by reference to the anchor groups of the Goslon Yardstick. The categorization was sufficiently sensitive to distinguish the treatment results at different centers in this study. It is proposed that the Goslon Yardstick should facilitate cross-center studies.

In order to evaluate and compare the results of different approaches to the early management of the child with a cleft of the lip and palate, it is essential to have a reliable method of assessing dental arch relationships.

Conventional methods of scoring arch relationships and dental irregularities have a number of deficiencies, even when applied to routine orthodontic problems, and they are not directly applicable to malocclusions in children with clefts of the lip and palate. Several classifications have been designed specifically for this purpose. Pruzansky and Aduss (1964) and Matthews et al (1970) proposed methods based essentially on the presence and extent of crossbites, but these did not account for a number of clinically important variables, such as open bites. In addition, reliability is often low. Huddart and Bodenham (1972) described a system in which a score was allocated to each upper tooth according to its occlusion with the lower arch. This was an improvement on previous methods, but as with all such scoring systems, it is possible that the overall score does not accurately represent the severity of the malocclusion: mild generalized irregularity may yield a higher score than a more severe but localized anomaly.

Objectives of the Present Study

The present study was designed to categorize malocclusions in patients with unilateral clefts of lip and palate in a way that would represent the severity of the malocclusion and the difficulty of correcting it. The method had to be simple to use and highly reliable even when used by different observers. The classification was to be used to compare the long-term results of different approaches to the early treatment of children with clefts of lip and palate (e.g., the effectiveness of presurgical orthopaedic treatment and of different surgical procedures).

The stages in the investigation that are described here are:
1. Development of the yardstick
2. Application of the yardstick
3. Testing the yardstick
4. Comparison of different groups using the yardstick.

The initial study has been confined to children with unilateral clefts of the lip and palate, and these are at the early permanent dentition stage because this is the age at which skeletal and occlusal problems are clearly manifest and at which definitive orthodontic and surgical treatment are usually planned.

Development of the Yardstick

Three of the authors proposed the clinical features that they considered most important in characterizing malocclusions in the early perma-
dentition stages of children with unilateral clefts of lip and palate. These were as follows:

1. **Anteroposterior Arch Relationships.** Severe Class III incisor relationships were obviously least satisfactory, and in general it was felt that a Class II division 1 relationship in the early permanent dentition, though rare, was most favorable for subsequent orthodontic correction. Preexisting dentoalveolar compensation in the presence of a reverse overjet was not considered to be favorable, since it limits the possible orthodontic correction of the incisor malrelationship; thus, it was included in the evaluation.

2. **Vertical Labial Segment Relationships.** A deep overbite was preferable to a reduced overbite which, in turn, was considered a better situation than an open bite. Since overclosure tends to exaggerate the Class III tendency, it is included in the anteroposterior assessment.

3. **Transverse Relationships.** Canine crossbites of the smaller segment were considered worse than molar crossbites, which might be clinically acceptable. The degree of transverse arch narrowness, rather than the number of teeth in crossbite, is the critical factor.

Anteroposterior relationships were considered to be of greatest clinical importance, and generalized crowding and irregularity were agreed to be relatively unimportant. The yardstick was to be based clinically on the features that pose the greatest difficulties in treatment.

In order to test the application of these subjective criteria, it was decided to investigate the extent of agreement between different orthodontists in ordering a series of study models of 30 cases in the early permanent dentition. These were selected from the files at The Hospital for Sick Children at Great Ormond Street, London and were chosen to represent the full range of results. The patients had not at that stage received orthodontic treatment except for correction of reverse overjets in the early mixed dentition.

The models were ranked subjectively by four experienced orthodontists working independently and without further consultation after the initial discussion of the criteria outlined above. The ranking was repeated after an interval of one week and the reliability was measured by Spearman’s rank correlation coefficient (Table 1). Bias was tested using the Wilcoxon matched-pairs signed rank test. Intra- and interexaminer agreement was high, indicating a satisfactory degree of reliability.

After the models had been ranked, it became apparent that the cases could readily be separated into five groups, which then formed the basis of the Yardstick. Since the intention at this stage was to produce clear examples of the different groups, it was decided to exclude eight borderline cases, leaving 22 cases as the Goslon Yardstick.

These were distributed as follows:

- Group 1—excellent: 2 cases
- Group 2—good: 7 cases
- Group 3—fair: 3 cases
- Group 4—poor: 4 cases
- Group 5—very poor: 6 cases

In general terms, groups 1 and 2 have occlusions that require either straightforward orthodontic treatment or none at all. Group 3 require complex orthodontic treatment to correct the Class III malocclusion and possibly other arch malrelationships, but a good result can be anticipated. Cases in group 4 are at the limits of orthodontic treatment without orthognathic surgery to correct skeletal malrelationships, and if facial growth is unfavorable, orthognathic surgery will be required. Cases in group 5 require orthognathic surgery to correct skeletal malrela-

<table>
<thead>
<tr>
<th>Examiner</th>
<th>DP</th>
<th>MM</th>
<th>MC</th>
<th>JK</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spearman Rank Correlation ($r$)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DP</td>
<td>0.94</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MM</td>
<td>0.95</td>
<td>0.94</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MC</td>
<td>0.94</td>
<td>0.95</td>
<td>0.96</td>
<td></td>
</tr>
<tr>
<td>JK</td>
<td>0.97</td>
<td>0.96</td>
<td>0.98</td>
<td>0.97</td>
</tr>
<tr>
<td>Wilcoxon Signed Ranks Test (Wilcoxon T/Number of non-zero differences)*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DP</td>
<td>108.5/16</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MM</td>
<td>187.0/28</td>
<td>123.0/22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MC</td>
<td>159.0/27</td>
<td>221.0/30</td>
<td>179.0/27</td>
<td></td>
</tr>
<tr>
<td>JK</td>
<td>169.0/29</td>
<td>170.0/27</td>
<td>183.0/27</td>
<td>174.0/26</td>
</tr>
</tbody>
</table>

* $P>0.30$ for all comparisons.
tionships if there is to be any prospect of obtaining satisfactory occlusal relationships.

Representative models from these groups, together with the corresponding facial photographs and lateral skull radiographs are shown in Figures 1-5; these records are of patients with the teeth held lightly in occlusion but without overclosure. The cases used to derive the Yardstick are used only for reference and are not included in the subsequent analyses reported here.

FIGURE 1  A representative case from Goslon Group 1
FIGURE 2  A representative case from Goslon Group 2
FIGURE 3  A representative case from Goslon Group 3
FIGURE 4 A representative case from Goslon Group 4
FIGURE 5  A representative case from Goslon Group 5
The Yardstick can be applied reliably only by those trained in its use with the original cases and who have undertaken the categorization of a calibration series.

**APPLICATION OF THE YARDSTICK**

The instructions given to the assessor are as follows:

**Stage 1: Anteroposterior Assessment**

The overjet is examined first. If for example there is a reverse overjet of 3 to 5 mm, this indicates that the case might belong to group 4. However, if there is already dentoalveolar compensation with marked proclination of the upper incisors and retroclination of the lower incisors indicating that the overjet underestimates the severity of the case, a higher category should be considered. For example, 4+ might then be provisionally allocated at this stage. On the other hand, if the inclinations of the incisors or if over-closure of the mandible exaggerates the severity of the reverse overjet, this should be taken into account and a less severe category than might originally have been considered, may be appropriate. The anteroposterior relationships of the buccal segments are not of importance in determining the grouping of a case.

**Stage 2: Vertical Assessment**

Favorable vertical features (i.e., deep overbite) do not indicate a modification of the provisional category except in borderline cases. A reduced overbite or anterior openbite suggests a higher grouping. For example, a case placed at the borderline between groups 3 and 4 on the anteroposterior assessment, but with a deep overbite might be confirmed as belonging to group 3. On the other hand, a case provisionally grouped a 3 but with an anterior openbite would probably be transferred to group 4 at this stage.

**Stage 3: Transverse Assessment**

A normal transverse relationship or a crossbite that can be treated orthodontically does not indicate a change of group. Marked narrowing of the upper arch with bilateral crossbite could indicate a more severe category for a case already at the upper limits of a group for other reasons.

In practice, the majority of cases can quickly and confidently be allocated to the appropriate group by the trained examiner.

**TESTING OF THE YARDSTICK**

Fifty-five consecutive sets of study models of children in the early permanent dentition with unilateral complete clefts of lip and palate were retrieved from the files of the Oslo Cleft Lip and Palate Clinic. No cases were excluded after the initial selection. They had study models in the early permanent dentition (mean age = 12 years, 1 month; SD = 1 year, 6 months). No orthodontic treatment other than correction of anterior crossbites in the early mixed dentition had been undertaken at that stage.

Four assessors employed the Goslon Yardstick to independently categorize the 55 models on 2 separate occasions. No discussion was allowed, but the Yardstick models were used for reference. Bias was not statistically significant (Wilcoxon's test) and reliability judged by Spearman's rank correlation coefficient was high (Table 2). It was concluded that the Yardstick was sufficiently reliable for general use.

---

**TABLE 2 Reliability Between and Within Examiners: Oslo Material**

<table>
<thead>
<tr>
<th>Examiner</th>
<th>OB</th>
<th>GS</th>
<th>DP</th>
<th>MM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spearman Rank Correlation ($r_s$)</td>
<td>0.90</td>
<td>0.87</td>
<td>0.91</td>
<td>0.95</td>
</tr>
<tr>
<td>OB</td>
<td>0.84</td>
<td>0.90</td>
<td>0.93</td>
<td></td>
</tr>
<tr>
<td>GS</td>
<td>0.90</td>
<td>0.93</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DP</td>
<td>0.93</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MM</td>
<td>0.89</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Wilcoxon Signed Ranks Test (Wilcoxon T/Number of non-zero differences)*

<table>
<thead>
<tr>
<th>Examiner</th>
<th>OB</th>
<th>GS</th>
<th>DP</th>
<th>MM</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.0/1</td>
<td>36.5/14</td>
<td>2.0/3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>39.0/14</td>
<td>45.5/18</td>
<td>10.0/9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17.0/10</td>
<td>20.0/15</td>
<td>36.0/12</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*In no case did a difference between scores for a case exceed one category. p >0.05 for all comparisons.*
A COMPARISON OF DIFFERENT GROUPS USING THE YARDSTICK

Two further groups of records were selected from the files of Great Ormond Street: group A comprising children who had received presurgical orthopaedic treatment; and group B, those who had not received this treatment because they were treated at a time when this was not standard practice at the hospital. Case selection was as for the Oslo group, with which these cases were to be compared (Fig. 6). The records were of the early permanent dentition, and there had been no orthodontic treatment other than correction of lingually positioned upper incisors in the early mixed dentition. Patients in both groups had been treated by several different plastic surgeons using a variety of procedures. Group A was assessed by two London authors (Mars and Plint) and group B by two other experienced orthodontists. Intra- and interexaminer reliability was of the same order as on the previous occasion (r > 0.92) for all comparisons (Wilcoxon’s test, non significant). The results from these two groups were then compared with those from the Oslo group. A comparison of the results from the three groups is presented in Figure 6. Groups A and B from Great Ormond Street could not be distinguished statistically, but the Oslo group had a statistically significantly lower score (p < 0.001; Chi squared test).

It may be concluded that within the context of the Great Ormond Street cases, presurgical orthopaedic treatment had no major effect, for better or for worse. The Oslo group received no presurgical orthopaedic treatment, and so the superiority of the results must depend on other factors, probably the most important of these being the manner of surgical repair of the cleft.

DISCUSSION

The results of the present study demonstrate that the Goslon Yardstick is highly reliable and is capable of discriminating among the quality of results at different centers. Some may consider the categories rather coarse and the Yardstick incapable of distinguishing the finer differences in the severity of malocclusions. This is intentional and contributes to the reliability of the assessment. Fine discrimination is inappropriate when cases span the full range between categories 1 and 5; the most urgent questions relate to the factors that result in this spread. Clearly, if groups that span only categories 1 and 2 were to be compared, a finer subdivision would be required, but this could readily be superimposed upon the present structure by using a scoring system along the lines proposed by Huddart and Bodenham (1972).

REFERENCES