ORIGINAL ARTICLES

A Multicenter Outcomes Assessment of Five-Year-Old Patients With Unilateral Cleft Lip and Palate

Wanda Flinn, D.D.S., M.S., Ross E. Long, Jr., D.M.D., M.S., Ph.D., Giovanna Garattini, M.D., D.D.S., Ph.D., Gunvor Semb, D.D.S., Ph.D.

Objective: **Compare 5-year-old dental arch relationships of patients from three centers with differing primary protocols.**

Design: **Retrospective study of treatment outcomes using blinded evaluation of dental study casts.**

Setting: **Three major cleft-craniofacial centers; one (center A) is a free-standing institution, and two (centers B and C) are university hospitals.**

Patients: **118 (A = 41; B = 33; C = 44) consecutively treated 5-year-old patients with complete, nonsyndromic unilateral cleft lip and palate.**

Interventions: **Centers A and C completed primary repair without presurgical orthopedics by 18 months (center A in three surgeries and center C in two surgeries). Center B used passive presurgical orthopedics with lip/soft palate repair at 6 months and gingivo-alveoloplasty/hard palate repair at 18 to 36 months.**

Main Outcome Measure: **Averaged ratings of dental casts using the 5-year yardstick were computed for each patient. The Wilcoxon two-sample test was used to compare means; a chi-square test was used to compare distributions.**

Results: **Intra- and interexaminer reliability tests showed excellent reliability ($>$.90). Mean scores were not significantly different. Distribution of scores differed significantly. Center A had the highest percentage of good scores and the lowest percentage of poor scores (72% versus 6.5%), followed by center B (63% versus 6.6%) and center C (59% versus 16.3%).**

Conclusions: **Centers A and B had comparable scores and completely different protocols in surgical technique, timing, sequencing, and nonuse/use of appliances. Center C's results were slightly lower than those of 1 and 3, but the center had the protocol with the least burden of treatment (only two surgeries, without use of appliances).**

KEY WORDS: *dental casts, 5-year yardstick, intercenter, outcomes*

In spite of recent advances, the outcomes resulting from the use of different treatment protocols for primary management of patients with complete unilateral clefts (UCLP) may vary considerably. This has, in part, resulted in a lack of standardization in cleft treatment protocols observed worldwide. One reason for this problem is that despite numerous pertinent reports in the literature, controversial issues such as timing of palatal closure and use of presurgical orthopedics have been debated hotly for decades (Pruzansky, 1964). This is attributable to the fact that many treatment options are based on anecdotal case reports, case series, or poorly controlled research, and there is (unfortunately) little valid scientific information upon which to make evidence-based decisions. Although prospective randomized clinical trials likely would provide the best answers to such issues, ethical considerations render many of these studies difficult to design and execute. Consequently, intercenter retrospective studies have become valuable alternatives in bringing to light some of the common elements of treatment protocols shared by centers that produce higher per-

Dr. Flinn is affiliated with the Lancaster Cleft Palate Clinic, Lancaster, Pennsylvania. Dr. Long is Director, Lancaster Cleft Palate Clinic, Lancaster, Pennsylvania. Dr. Garattini is Professor of Orthodontics, University degli Studi di Milano, Milan, Italy. Dr. Semb is affiliated with Oslo Cleft Team, University of Oslo, Oslo, Norway, and is Senior Lecturer in Craniofacial Anomalies, University Dental Hospital of Manchester, Manchester, U.K.

This paper was submitted in partial fulfillment of the requirements for the master's degree in orthodontics, Temple University Dental School, Department of Orthodontics, Philadelphia, Pennsylvania. The data presented herein were presented at the 59th annual meeting of the American Cleft Palate-Craniofacial Association, Seattle, Washington, May 2002.

Submitted June 2004; Accepted February 2005.

Address correspondence to: Dr. Ross E. Long, Jr., Director, Lancaster Cleft Palate Clinic, 223 North Lime Street, Lancaster, PA 17602. E-mail rlong@supernet.com.

centages of patients with superior results. Conversely, some treatment elements shared by clinics that produce relatively poorer results also have been identified. Intercenter studies allow direct comparisons of various treatment outcomes with less risk of sampling bias as long as consecutive cases are compared. In addition, large sample sizes can be acquired, thereby increasing validity and allowing for subgrouping.

Early attempts at intercenter outcome comparisons were carried out by Bishara (1974) and Dahl et al. (1981), who evaluated the facial growth and occlusal results of different primary surgical procedures used at two and three different centers, respectively. In the study by Dahl et al. (1981), significant differences in crossbite incidence led the authors to conclude that the palate surgery used in the center with the highest incidence of crossbite was more traumatic. Also, it was suggested that the surgeon's individual skill may be as important as the particular technique chosen. This points to one of the major shortcomings of intercenter studies that attempt to relate a particular outcome to a specific step or technique in the primary protocol, namely that there is no way to account for the variability of surgical skill. Obviously, other confounding variables such as ethnic differences inherent in the sample populations being studied also complicate the attempt to relate outcomes to protocols.

In an intercenter study by Ross (1987a, 1987b, 1987c, 1987d, 1987e, 1987f, 1987g), 1600 cephalometric radiographs from males with complete UCLP were examined to discern the effects of surgical and orthopedic treatment on facial growth. Among other things, Ross concluded that simple treatment protocols produced the most favorable results, and similar to the suggestion made by Dahl et al., that surgical expertise was found to be a major determinant of overall success. Presurgical orthopedics was found to provide no long-term benefits.

Despite the obvious significance of this large study, the fact that its design was limited to cephalometric analysis made the analysis of many occlusal relationships impossible. These dental and arch-form factors often play an important role in treatment considerations.

In contrast, Mars et al. (1987) used dental casts and a new form of arch relation analysis called the Goslon Yardstick to compare outcomes between various clinics. The rating process is based on the Goslon reference models, which are divided into five groups. Depending on the amount of maxillary protrusion present, and to a lesser extent on transverse and vertical variables, the groups are ordered from the best arch relationships to the worst. ''Ones'' are considered the best, and conversely, ''fours'' and ''fives'' are considered severe enough to likely require surgical maxillary advancement during end-stage treatment. The dental casts to be studied are compared with these reference groups and are assigned a score. Simpler means of arch assessment have been suggested, such as crossbite evaluation (Huddart and Bodenham, 1972) and incisal overjet measurement (Morris et al., 1994); these techniques, however, are not as sensitive and do not predict facial morphology outcomes as accurately as the Goslon Yardstick. The Yardstick is a practical means of evaluating malocclusion severity and associated treatment difficulty, and was used in the Mars et al. (1987) study to compare outcomes between a sample from Oslo and two samples from Greater Ormond Street (only one of which received presurgical orthopedics). Although the Oslo ratings were superior to those of Greater Ormond Street, no significant difference was found between the two subgroups of the latter. Presurgical orthopedics was therefore reported as having no major effect in this study.

The Eurocleft study, published in 1992 as a series of five papers (Shaw et al., 1992a, 1992b; Mølsted et al., 1992; Mars et al., 1992; Asher-McDade et al., 1992) expanded the scope of intercenter research by comparing treatment outcomes of 8- to 10-year-olds with UCLP from six European cleft centers using cephalometric radiographs, dental casts evaluated with the Goslon Yarsdstick, and nasolabial photographs to evaluate craniofacial form, arch relationships, and nasolabial appearance. Most notably, the value of this landmark series of studies was the shift away from attempting to attribute outcomes to a specific aspect/procedure/technique of a center's protocol, to looking at the protocols as a whole and searching for characteristics of the protocols that were associated with better or poorer outcomes. In this fashion, clinics that used primary bone grafts or extraoral strapping were associated with poorer results, as were those that had multiple, low-volume operators. In contrast, the clinics that received the highest ratings were characterized by few surgeons and simpler primary surgical protocols. Use of presurgical orthopedics was not associated with any improvement in outcomes.

Two intercenter studies conducted by Atack et al. (1997, 1998) used a Goslon Yardstick-like index for 5-year-olds that the authors developed in hopes that outcomes of various treatment modalities might be evaluated at an even earlier age than was assessed in the Eurocleft study. The ''5-year-olds' index'' of Atack and colleagues was used for a two-center study comparing UCLP casts; it revealed a statistically significant disparity between the two clinics. This suggested that differences in surgical outcome may indeed be detectable by 5 years of age. This is meaningful, for in this manner earlier discovery and subsequent termination of potentially harmful procedures could be accomplished.

Although these studies represent a significant advancement in our approach to determining factors associated with improved outcomes, the number of factors available to consider are limited by the number of centers that have participated in this type of outcome assessment. Expansion of this approach to include as many additional centers as possible would help further to validate or to refute the conclusions reached based on the small number of initial centers involved, and perhaps elucidate other characteristics of better or worse outcomes. The purpose of this study was to identify any differences that exist between 5-year-old outcomes from three different centers and to identify the characteristics of the individual protocols used that relate to any possible differences in outcomes.

TABLE 1 Treatment Protocols of the Three Centers

	Center A	*Center B*	*Center C*
Birth		presugical orthopedics (passive plate)	
1 mo			
2 mo			
3 mo	lip repair (triangular flap)		lip closure (Millard), vomerplasty
6 mo		lip/nose surgery, soft palate closure	
12 mo	hard palate closure (vomer flap)		
18 mo	soft palate repair (median suture repair)	hard palate closure and gingivo-alveoloplasty often delayed up to 3 years of age	soft palate closure (modified Von Langenbeck)
2 yr			
3 yr			

MATERIALS AND METHODS

The study was designed as an intercenter comparison using the 5-year yardstick developed by Atack et al. (1997) to rate 118 study models of 5- to 7-year-old UCLP patients from three major, high-volume cleft/craniofacial centers (one in the United States, two in Europe). The primary surgical protocols for these three centers are provided in Table 1. Plaster casts from consecutively treated male and female UCLP patients from the three centers were evaluated. All subjects were nonsyndromic and were born with complete UCLP. Confirmation of these initial inclusion criteria required availability of pretreatment infant records. The subjects were chiefly in the primary dentition. Although patients with permanent upper incisors were excluded according to the protocol for the 5-year yardstick (Atack et al. 1997), none of the patients from centers B and C and just 2 patients of 43 from center A were excluded. Each patient had all relevant treatment performed at the same respective center. Table 2 shows the breakdown of patients by center and gender. Sample sizes of 25 to 30 per center had been determined in the Eurocleft study as providing necessary power for the detection of significant differences.

Study models from each of the 118 subjects were duplicated and were trimmed in maximum intercuspation according to the same specifications. Numbers were randomly assigned to each model, and no other form of identification was visible. Three trained and calibrated assessors rated the casts from 1 to 5 according to the 5-year yardstick groupings, with no discussion allowed (Table 3). The 5-year yardstick referent models, consisting of two different sets of casts representing each rating group, were displayed and were available for comparison with the casts that were being scored. The following day, the models were rearranged randomly to reduce the chance of memory bias. The three assessors then rated all of the casts again, giving each set of models a total of six individual scores over two sessions, which were averaged subsequently. The average for each center was then computed, as was the standard deviation of the average, the standard deviation of the sample, and the upper and lower limits for the average at a 95% confidence level. To compare each pair of centers, the *t* statistic, number of degrees of freedom, and corresponding *p* values were calculated using the Wilcoxon two-sample test. Score distributions were analyzed with the chi-square test. The Fleiss and Cohen weighted-kappa analysis (1973) was employed to determine interassessor and intra-assessor reliability.

TABLE 2 Sample

	Mean Age (yr)	*Age Range (yr)*	*Boys*	*Girls*	*n*
Center A	5.3	5.0 to 6.0	23	18	41
Center B	5.3	4.5 to 7.8	24	9	33
Center C	5.9	4.9 to 7.9	31	13	44

TABLE 3 Scoring Guidelines for the 5-Year-Old Yardstick (Atack et al., 1997)

Group	*Description*	*Prognosis*
1	+ overjet (normal or retroclined incisors); no crossbites or openbites; good upper arch form and palatal vault anatomy	excellent
2	+ overjet (normal or proclined incisors); unilateral crossbite / crossbite tendency	good
3	edge-to-edge bite (normal or proclined incisors) or slight negative overjet (retroclined incisors); unilateral crossbite	fair
4	negative overjet (normal or proclined incisors); unilateral crossbite ± bilateral crossbite tendency	poor
5	negative overjet (proclined incisors); bilateral crossbite	very poor

RESULTS

For both interassessor and intra-assessor reliability calculations, perfect agreement between two compared scores received 1 mark, whereas an error of one category between two compared scores received .05 mark. None of the rated models had more than a one-category discrepancy. All of the interassessor scores were high, with all but one falling in the ''very good agreement'' category of .81 and higher (Table 4). Intraexaminer scores (comparing assessors' first-day ratings to their own second-day ratings) were even higher, with all scores well above .90, considered to be in ''excellent agreement'' with the Fleiss and Cohen method (Table 5).

The mean scores of the three clinics were as follows: center A, 2.0; center B, 2.1; and center C, 2.4 (Table 6). Utilizing the

TABLE 4 Interassessor Kappa Scores

Session	*Assessor*	*Weighted Kappa*	*Standard Error*	*95% Confidence Intervals*
Day one	G.S./W.S.	0.797	0.028	0.742 to 0.851
	D.B./G.S.	0.880	0.026	0.830 to 0.930
	W.S./D.B.	0.843	0.026	0.792 to 0.894
Day two	G.S./W.S.	0.847	0.027	0.794 to 0.900
	D.B./G.S.	0.881	0.026	0.831 to 0.932
	W.S./D.B.	0.891	0.024	0.845 to 0.937

TABLE 5 Intra-assessor Kappa Scores

Assessor	*Weighted Kappa*	*Standard Error*	*95% Confidence Intervals*
G.S.	0.965	0.013	0.940 to 0.990
W.S.	0.937	0.015	0.907 to 0.967
D.B.	0.949	0.016	0.917 to 0.980

Wilcoxon two-sample test (Table 7), these differences were not found to be statistically significant at the 95% confidence level ($p > .05$). More significant differences were found, however, when the centers were compared by their distributions rather than by their means (Table 8, Fig. 1). Center A had 72.0% of its scores in either category 1 or 2. Center B had 63.6% of its scores fall into these groups, whereas center C had 59.1%. On the other end of the yardstick spectrum, centers A, B, and C recorded 6.5%, 6.6%, and 16.3% respectively, in either category 4 or 5. Figure 1 illustrates the differences in the distributions throughout all five categories. Although similar in that the majority of cases for all centers fell within the first three categories, the small differences in distributions category-to-category resulted in rejection of the null hypothesis of independence, with chi-square values of $p = .0013$ for center A versus center B, $p = .0002$ for center A versus center C, and $p = .0001$ for center B versus center C.

Discussion

A wide range of cleft management protocols currently exists throughout the world. Some issues, such as timing of palate closure and use of presurgical orthopedics, have been debated for many years without satisfactory resolution. Two practical realities have retarded progress on this front. One is the fact that due to small sample sizes and obvious ethical issues, prospective randomized clinical trials are extremely difficult. In addition, due to the slow and subtle effects of various procedures on dentofacial growth and development, it can take up to 10 years before changes resulting from specific treatment regimens are considered true reflections of future morphology. This delay in outcome assessment, however, has recently come under scrutiny.

By demonstrating statistically significant intercenter differences in clinical presentation by the age of 5 years, Atack et al. (1997, 1998) proved that harmful or unnecessary treatment modalities could be detected within 5 years of implementation. In addition, the use of a child's 5-year-old clinical presentation as a predictor of outcome for treatment planning purposes became a possibility. For research purposes, evaluation at 5 years of age is highly desirable, because the effects of primary surgery can be seen more purely. This is due to the fact that confounding variables such as orthodontics or secondary surgeries usually have not been performed yet. In addition, because a child's genetic predilection toward one skeletal pattern or another is not fully realized until after puberty, genetic influences also are less of a factor. This opportunity to procure a relatively pure sample is especially important, because in the absence of prospective trials, well-designed and controlled outcome assessment studies have become one of the most preferred alternatives.

Thus, the 5-year yardstick has become a potentially useful tool in intercenter outcome studies. This study model index has provided researchers with a robust and expeditious means of evaluating cases in the primary and early mixed dentition. In this study, 118 casts from three centers were rated using the 5-year yardstick. The mean yardstick scores, as well as their distributions, then were analyzed to discern any significant differences in outcome. Evaluation of the means alone produced no significant findings. The three centers had average scores ranging from 2.0 to 2.4. This is not surprising, inasmuch as at least two of the three centers have been shown to produce consistently good results in previous studies using the Goslon Yardstick at age 9 (Vargas, 2002). Analysis of the scoring distributions, however, revealed statistically significant differences between all of the centers. In this regard, it is tempting to apply the same assumptions of the Goslon Yardstick at age 9. In particular, in the Goslon ratings, it is assumed that orthodontic treatment of categories 1 and 2 can be carried out without orthognathic surgery, whereas those in categories 4 and 5 would be more than likely to require surgical assistance in end-stage orthodontic treatment due to the severity of the underlying skeletal dysplasia (Mars et al. 1987, 1992). Examination of the three centers in this study does reveal a distribution

TABLE 6 Intercenter Comparison of Yardstick Means and Standard Deviations

	Center A	*Center B*	*Center C*
Average	2.0463	2.1181	2.4181
Standard deviation of sample	0.9542	0.8636	1.0653
Sample size	41	33	44
Standard deviation of average	0.1490	0.1503	0.1606
Lower 95% confidence limit	1.7542	1.8235	2.1033
Upper 95% confidence limit	2.3384	2.4128	2.7329

TABLE 7 Intercenter Mean Pairwise Analysis of Yardstick Scores, Demonstrating No Significant Differences Between Means of Centers A, B, and C

	A versus B	*A versus C*	*B versus C*
p value	0.7353	0.0934	0.1768
t value	0.3394	1.6971	1.3637
df	71	83	75

TABLE 8 Numerical and Percentage Distributions of 5-Year-Old Yardstick Scores*

Score	*Center A (No.)*	*Center A (%)*	*Center B (No.)*	*Center B (%)*	*Center C (No.)*	*Center C (%)*
1	82	33.33	60	30.30	57	21.59
2	95	38.62	66	33.33	99	37.50
3	53	21.55	59	29.80	65	24.62
4	8	3.25	13	6.57	28	10.61
5	8	3.25	0	0	15	5.68

* Number values represent individual scores. Each model received a total of six individual scores.

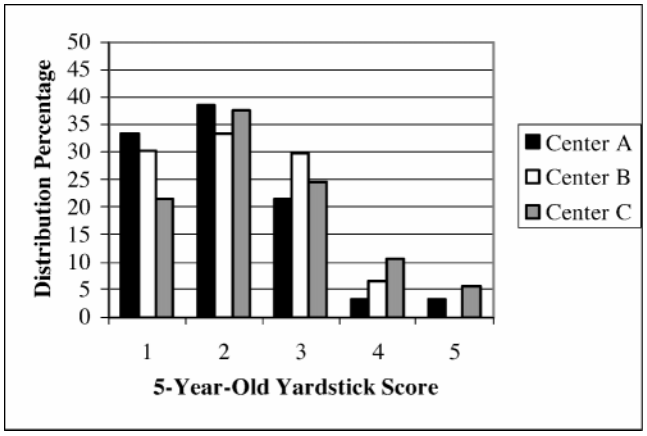

FIGURE 1 Comparison of yardstick score distributions for centers A, B, and C.

difference at both ends of the 5-year yardstick scale, which may have similar clinical implications. The distribution of cases in categories 1 and 2 was 72.0%, 63.6%, and 59.1% for centers A, B, and C, respectively. Conversely, the distribution of cases in categories 4 and 5 was 6.5%, 6.6%, and 16.3% for centers A, B, and C, respectively. However, it should be emphasized that to date, there are no longitudinal published reports confirming or refuting these assumptions of treatment needs in the permanent dentition.

The main value of the early, 5-year yardstick evaluation remains the detection of potentially harmful primary surgery protocols, and clearly, the majority of treatment outcomes for all centers in this study were in the fair to excellent category.

Because all three centers had different treatment protocols (Table 1), it is useful to examine their characteristics in light of the outcomes found. Center C had a markedly earlier primary repair of the hard palate at 3 months of age. In contrast, center A did not complete hard palate closure until about 12 months of age, and center B delayed hard palate closure even longer (up to 36 months). The role of hard palate surgery in scar tissue formation and subsequent maxillary growth inhibition has been debated for many years. Indeed, Graber (1954) proposed that scar tissue is the single most important growth-altering influence affecting patients with clefts. In his 1954 study, it was shown that patients who underwent palatal closure at younger than 6 months of age exhibited more severe maxillary aberrations than did those patients who had their repairs performed later. A recent review of the literature over the past 50 years also has demonstrated general agreement about the effects of palatal scar tissue (Kuijpers-Jagtman and Long, 2000).

Another factor that is known to play an important role in the quality of treatment results is the experience of the surgeons performing the palatal closures. In the landmark Eurocleft studies cited above, it was found that the centers with the best skeletal and dental results were those with caseloads distributed among the least number of surgeons. In the present study, all three centers had low numbers of surgeons who performed the primary repairs. Specifically, centers A and B each had only one, and center C had three. Given these similar numbers, it is possible that differences in surgical skills may have played a partial role in the observed outcome differences, but overall experience was not likely to be a factor because all centers had high-volume operators.

Of the three centers, center B was the only one that employed presurgical orthopedics (PSOT). A passive plate was used, and acrylic was removed every 3 weeks to allow narrowing of the cleft. Advocates of various PSOT techniques have put forth many persuasive arguments regarding the numerous benefits of such treatment during early infancy. However, there is also a larger body of evidence that has failed to demonstrate any measurable long-term benefit to PSOT (Prahl et al., 1997; Long et al., 2000). Interestingly, in this study, although center B's outcomes were indeed good, they were no better than the scores received by center A, which did not use PSOT. The burden of treatment, on the other hand, was obviously greater for those patients who underwent PSOT. This finding of lack of additional benefit paired with an increased burden of treatment is similar to the findings of the Eurocleft study. Using a similar benefit-to-burden analysis, whereas center A seemed to have better outcomes than center C, the primary surgical protocol for center A involved three surgeries (lip repair followed by two-stage palate repair) versus two surgeries for center C (lip and primary palate repair followed by one-stage repair of the posterior palate). Thus, the slightly better results came at the expense of an extra surgery for the patients. In that sense, the treatment burden of center C was superior to the other centers, requiring the fewest surgical interventions and no PSOT.

In general, the results of this study confirm that very desirable outcomes are possible with the use of very different primary protocols. It also highlights the benefits of attempting early outcome assessments of primary surgical management and the value of intercenter comparisons to be used in conjunction with randomized control trials to understand the effects of the treatments used for patients with clefts.

CONCLUSIONS

The distribution of the 5-year-old yardstick scores differed significantly between three high-volume centers with differing primary surgical protocols for patients with complete UCLP. However, the majority of cases from all centers were rated in the good to excellent categories.

The average scores for each center, using the 5-year yardstick assessment of dental models, were not significantly different.

The center with the largest percentage of cases in the poorer categories also had the least burden of care with fewest surgeries and no presurgical orthopedic intervention.

The use of PSOT in one center was associated with excellent outcomes, but the results were not better than those of the center whose protocol did not include that procedure.

Intercenter outcome studies represent a valuable tool in monitoring the relative success of the wide range of treatment

options available for the primary management of patients with clefts.

References

Asher-McDade C, Brattström V, Dahl E, McWilliam J, Mølsted K, Plint DA, Prahl-Andersen B, Semb G, Shaw WC, The RPS. A six-center international study of treatment outcome in patients with clefts of the lip and palate: part 4. Assessment of nasolabial appearance. *Cleft Palate Craniofac J.* 1992;29: 409–412.

Atack N, Hathorn I, Mars M, Sandy J. Study models of 5-year-old children as predictors of surgical outcome in unilateral cleft lip and palate. *Eur J Orthod.* 1997;19:165–170.

Atack NE, Hathorn I, Dowell, T, Sandy J, Semb G, Leach A. Early detection of differences in surgical outcome for cleft lip and palate. *Br J Orthod.* 1998;25:181–185.

Bishara SE. Comparisons of the effect of two palatoplasties on facial and dental relations. *Cleft Palate J.* 1974;11:261–271.

Dahl E, Hanusardottir B, Bergland O. A comparison of occlusions in two groups of children whose clefts were repaired by three different surgical procedures. *Cleft Palate J.* 1981;17:122–127.

Fleiss JL, Cohen J. The equivalence of weighted kappa and the intraclass correlation coefficient as a measure of reliability. *Educ Psychol Meas.* 1973; 33:613–619.

Graber TM. The congenital cleft palate deformity. *JADA.* 1954;48:375–395.

Huddart AG, Bodenham RS. The evaluation of arch form and occlusion in unilateral cleft palate subjects. *Cleft Palate J.* 1972;9:194–209.

Kuijpers-Jagtman AM, Long RE Jr. The influence of surgery and orthopedic treatment on maxillofacial growth and maxillary arch development in patients treated for orofacial clefts. *Cleft Palate Craniofac J.* 2000;37:527–537.

Long RE Jr, Semb G, Shaw WC. Orthodontic treatment of the patient with complete clefts of the lip, alveolus, and palate: lessons of the past 60 years. *Cleft Palate Craniofac J.* 2000;37:533.

Mars M, Asher-McDade C, Brattström V, Dahl E, McWilliam J, Mølsted K, Plint DA, Prahl-Andersen B, Semb G, Shaw WC, The RPS. A six-center international study of treatment outcome in patients with clefts of the lip and palate: part 3. dental arch relationships. *Cleft Palate Craniofac J.* 1992; 29:405–408.

Mars M, Plint DA, Houston WJ, Bergland O, Semb G. The Goslon Yardstick: a new system of assessing dental arch relationships in children with unilateral clefts of the lip and palate. *Cleft Palate J Jr.* 1987;24:314–322.

Mølsted K, Asher-McDade C, Brattström V, Dahl E, Mars M, McWilliam J, Plint DA, Prahl-Andersen B, Semb G, Shaw WC, The RPS. A six-center international study of treatment outcome in patients with clefts of the lip and palate: part 2. Craniofacial form and soft tissue profile. *Cleft Palate Craniofac J.* 1992;29:398–404.

Morris T, Roberts C, Shaw WC. Incisal overjet as an outcome measure in unilateral cleft lip and palate management. *Cleft Palate Craniofac J.* 1994; 31:142–145.

Prahl C, Kuijpers-Jagtman AM, Prahl-Andersen B. Effect of presurgical infant orthopedics on maxillary arch dimensions. In: Lee ST, Huang M, eds. *Transactions 8th International Congress on Cleft Palate and Related Craniofacial Anomalies.* Singapore: Stamford Press; 1997:1026.

Pruzansky S. Presurgical orthopedics and bone grafting for infants with cleft lip and palate: a dissent. *Cleft Palate J.* 1964;1:411–418.

Ross RB. Treatment variables affecting facial growth in complete unilateral cleft lip and palate. Part 1: treatment affecting growth. *Cleft Palate J.* 1987a; 24:5–23.

Ross RB. Treatment variables affecting facial growth in complete unilateral cleft lip and palate. Part 2: presurgical orthopedics. *Cleft Palate J.* 1987b; 24:24–32.

Ross RB. Treatment variables affecting facial growth in complete unilateral cleft lip and palate. Part 3: alveolus repair and bone grafting. *Cleft Palate J.* 1987c;24:33–44.

Ross RB. Treatment variables affecting facial growth in complete unilateral cleft lip and palate. Part 4: repair of the cleft lip. *Cleft Palate J.* 1987d;24: 45–53.

Ross RB. Treatment variables affecting facial growth in complete unilateral cleft lip and palate. Part 5: timing of palate repair. *Cleft Palate J.* 1987e;24: 54–63.

Ross RB. Treatment variables affecting facial growth in complete unilateral cleft lip and palate. Part 6: techniques of palate repair. *Cleft Palate J.* 1987f; 24:64–70.

Ross RB. Treatment variables affecting facial growth in complete unilateral cleft lip and palate. Part 7: an overview of treatment and facial growth. *Cleft Palate J.* 1987g;24:71–77.

Shaw WC, Asher-McDade C, Brattström V, Dahl E, McWilliam J, Mølsted K, Plint DA, Prahl-Andersen B, Semb G, The RPS. A six-center international study of treatment outcome in patients with clefts of the lip and palate: part 1. Principles and study design. *Cleft Palate Craniofac J.* 1992a;29:393–397.

Shaw WC, Dahl E, Asher-McDade C, Brattström V, Mars M, McWilliam J, Mølsted K, Plint DA, Prahl-Andersen B, Roberts C, Semb G, The RPS. A six-center international study of treatment outcome in patients with clefts of the lip and palate: part 5. General discussion and conclusions. *Cleft Palate Craniofac J.* 1992b;29:413–418.

Vargas A. *Comparison of Outcomes in UCLP Mixed Dentition Patients From 7 Different Centers.* Philadelphia: Temple University School of Dentistry, Department of Orthodontics; 2002. Unpublished master's thesis.