# The Americleft Study: An Inter-Center Study of Treatment Outcomes for Patients With Unilateral Cleft Lip and Palate Part 5. General Discussion and Conclusions

**Kathleen Russell, D.D.S., M.Sc., Ross E. Long, Jr., D.M.D., M.S., Ph.D., Ronald Hathaway, D.D.S., M.S., John Daskalogiannakis, D.D.S., M.Sc., F.R.C.D.(C.), Ana Mercado, D.M.D., Ph.D., Marilyn Cohen, B.A., Gunvor Semb, D.D.S., Ph.D., William Shaw, B.D.S., Ph.D.**

***Objective:*** **To summarize the Americleft study regarding treatment outcomes for patients with complete unilateral cleft lip and palate (CUCLP).**

***Setting:*** **Five cleft palate centers in North America.**

***Subjects:*** **One hundred sixty-nine subjects, between the ages of 6 years and 12 years, with repaired CUCLP who were consecutively treated at the five centers.**

***Methods:*** **Study consisted of model comparisons assessing maxillomandibular relationship using the GOSLON Yardstick (169 patients from all 5 centers), soft and hard tissue craniofacial morphologic comparisons using lateral cephalometric analyses (148 patients from four of the centers), and nasolabial esthetics assessments (125 patients from four of the centers).**

***Results:*** **Significant differences were found between the center with the best GOSLON scores and the remaining centers. These differences also corresponded to those found in the craniofacial morphologic cephalometric assessment. Sagittal maxillary prominence was found to be significantly better for the center with the best GOSLON scores, while no significant differences were seen among the centers for mandibular prominence, vertical dimensions, or dental inclinations. No differences were seen for nasolabial esthetics between the centers.**

***Conclusions:*** **Challenges experienced while undertaking the inter-center retrospective study are reviewed. Aspects of treatment that could potentially make the outcome of treatment less optimal included primary alveolar bone grafting and extensive treatment protocols. Differences in the outcomes identified between the centers were restricted to the maxilla, and no differences were identified for mandibular prominence, vertical dimensions, or dental inclinations.**

KEY WORDS: *Americleft, cephalometrics, craniofacial form, GOSLON, inter-center study, maxillary-mandibular relationships, model analysis, nasolabial esthetics, ratings, scores, skeletal relationship, treatment outcome measures*

Dr. Russell, is Associate Professor and Head, Division of Orthodontics, and Dalhousie University Staff Orthodontist and Chair, Cleft Palate Team, IWK Health Centre, Halifax, Nova Scotia, Canada. Dr. Long is Director, Lancaster Cleft Palate Clinic, Lancaster General Health, and Clinical Professor, Department of Surgery, Penn State College of Medicine, Lancaster, Pennsylvania. Dr. Hathaway is Medical Director, Craniofacial Center, Peyton Manning Children's Hospital at St. Vincent, Indianapolis, Indiana. Dr. Daskalogiannakis is Staff Orthodontist, SickKids Hospital, and Assistant Professor, Department of Orthodontics, University of Toronto, Toronto, Ontario, Canada. Dr. Mercado is Clinical Assistant Professor, Division of Orthodontics, College of Dentistry, The Ohio State University, Columbus, Ohio. Ms. Cohen is Administrative Director and Patient Care Coordinator, Regional Cleft-Craniofacial Program, Cooper University Hospital, Moorestown, New Jersey. Dr. Semb is Adjunct Professor, University of Oslo, and Affiliate, Oslo Cleft Team, Department of Plastic Surgery, Oslo, Norway, and Senior Lecturer in Craniofacial Anomalies, University of Manchester, Manchester, United Kingdom. Dr. Shaw is Professor of Orthodontics and Dentofacial Development, University of Manchester, Manchester, United Kingdom.

This article was presented, in part, at the 64th American Cleft Palate Association meeting in Boulder, Colorado, April 2007; the 65th American Cleft Palate Association meeting in Philadelphia, Pennsylvania, April 2008; and the 66th American Cleft Palate Association meeting in Scottsdale, Arizona, April, 2009. As well, portions were presented at the 8th European Craniofacial Congress, Bilbao, Spain, June 2007; the Craniofacial Society of Great Britain and Ireland, Chester, United Kingdom, April 2008; the Japanese Cleft Palate Association, Hiroshima, Japan, May 2008; Cleft 2008 - 5th International Congress of the World Cleft Palate Foundation, Dallas, Texas, September 2008; and the International Society of Craniofacial Surgery, XIII Biennial Congress, Oxford, United Kingdom, September 2009.

The authors and members of the Americleft group would like to acknowledge and thank the American Cleft Palate Association and the Cleft Palate Foundation for their annual generous financial support from 2006 to the present. Additional funding was provided by the Trout Family Trust, Lancaster, Pennsylvania, the Mellinger Medical Research Fund, Lancaster, Pennsylvania and the H.G. Barsumian Memorial Fund, Winston-Salem, North Carolina.

Submitted September 2009; Accepted April 2010.

Address correspondence to: Dr. Russell, Division of Orthodontics, Faculty of Dentistry, Room 5164, 5981 University Avenue, Dalhousie University, Halifax, Nova Scotia, Canada, B3H 3J5. E-mail Kathy.Russell@Dal.ca.

DOI: 10.1597/09-187.1

This is the final part in the initial Americleft study. Parts 1 through 4 presented the important background information that lead to the development of the project and the execution of the first three measures of outcomes of primary surgical protocols for children with complete unilateral cleft lip and palate (CUCLP). The outcome measures selected (dental arch relationship, craniofacial morphology, and nasolabial appearance) were chosen to follow along with the 1992 Eurocleft study, after which this study was modeled. In addition, dental study models, lateral cephalometric radiographs, and frontal and lateral facial photographs are probably the most commonly collected records by most cleft palate teams, and for which reliable and valid outcome measures have been developed. This section will review the salient conclusions reached as they apply to the use of comparative inter-center outcome assessments in advancing our knowledge and informing evidence-based treatment decisions.

## Prerequisite Conditions for Successful Outcome Studies

### 1. Group Dynamics and Participant Cultures

The success of an inter-center study depends a great deal on the commitment of the study investigators to the goals, conditions, and parameters of the study. The investigators have to be willing to: (1) enter into the project with a culture of willingness to entertain a degree of "uncertainty" regarding the true effectiveness of their individual protocols, (2) detach themselves from any personal connection to the results from their center, (3) let their treatment results be evaluated blindly and objectively by a group of their peers, and (4) accept the results as presented and consider explanations that may suggest the possibility of equally good or better outcomes with protocols used by other centers. Furthermore, the participants should be willing at the start of the study to consider making changes to their treatment protocol for the benefit of patient care in the case of overwhelming evidence to support such recommendations. Without this mindset at the beginning of such a study like Americleft, the result is bound for failure. This study started with the commitment of representatives from six established North American cleft centers to the objective assessment and comparison of overall treatment outcomes. Evidence to support the commitment of our team of investigators includes seeing the study through all the challenges to the end result and the actual changes made to one of the center's protocols based on the results presented through this series of papers.

### 2. Availability of Standardized Records

Once centers have committed to an inter-center research initiative, a second and significant challenge is the identification of standardized records to be used in the analyses that all centers have in common. The retrospective nature of the Americleft study was fraught with this challenge. Although all centers thought they had been taking adequate records as part of their treatment protocols, variations were found in the age and status of treatment when models, radiographs, and photographs had been taken; the appearance and type of models and images; and the availability of sufficient, appropriate records at the time the study was undertaken. At the initial planning meeting for Americleft, the investigators were "positive" that they had the required sample size of 40 consecutively treated cases with adequate records. Once available records were reviewed, one original center did not have sufficient records to be part of the study, reducing the original six centers to five; one had approximately 50% of the sample required; and one required an extra year to gather the required sample size. This challenge to gather an acceptable set of records further becomes extremely clear when, as described in Part 4, the number of facial profile images taken from the affected, cleft side was reviewed and found to be too low for statistical analyses in an inter-center study. It was found that the side from which the photographs were taken was routinely the right side following orthodontic convention and not related to the location of the cleft. This becomes even more astounding when the centers involved represent some of the largest and well-established cleft centers in North America.

### 3. Sample Characteristics

#### *a. Sample Sizes*

In this study, following along with the sample size requirements determined in the similar Eurocleft study, sample sizes between 30 and 40 were desired. To underscore the challenge that sample size represents, Shaw et al. (2005) noted the time frames necessary to reach sample sizes adequate to meet statistical requirements based on power analysis. Statistically, for a center treating 30 new patients with CUCLP per year, it would take 12 years to gather the required number of cases to enter into a five-center inter-center study. If only six new cases are seen per year, it would require 63 years to have a sufficient sample for a five-center study and 42 years for a two-center study; this represents more years than the career of an extremely dedicated researcher-clinician with the best of foresight! As a result of the large number of required cases and the challenge to accumulate these cases for retrospective inter-center studies, sample sizes are often less than optimal. In these situations with sample sizes less than optimal, the need to interpret the results with discretion and not over interpret differences between centers is emphasized. As well, a causative effect from specific aspects of individual treatment protocols cannot be made, but rather general aspects of treatment and relative treatment outcomes can be compared.

### *b. Sample Age Range*

A second characteristic of the sample is the age at the time of assessment. The mixed dentition age was chosen in this study to be consistent with the Eurocleft approach and allow for some cross-comparisons between studies. The importance of this relates to two factors. First, as mentioned above, the time frame required to accumulate large enough samples for valid comparisons of primary surgical protocols strongly suggests the need to carry out an assessment at the earliest age possible but with sufficient growth-time for a valid assessment of the effects of the primary protocols. An assessment completed too early would not likely provide a measure of the long-term growth, occlusion, or appearance outcomes to be expected from various primary protocols. Postponing assessment until the completion of growth and treatment would allow for the possibility of decades of use of injudicious procedures before detecting problems, and add multiple confounding treatment variables beyond the original primary protocols. Of note, the follow-up Eurocleft study in 2005 (Shaw et al., 2005) did indicate a strong, but yet not perfect, predictive value to the mixed dentition outcome measures that were used in that and this study.

In addition to the use of the 9-year-old records, there has been growing support for the use of 5-year-old records to assess dental arch relationships, craniofacial morphology, and nasolabial appearance. Many of the Americleft participants have initiated routine, complete 5-year-old assessments to be used for audits of primary protocol outcomes, as well as initial records for planning possible late primary dentition or early mixed dentition orthodontic treatment.

### *c. Sample Consecutiveness*

A final critical characteristic of the samples was the ability to document and confirm the consecutive enrollment of patients in the samples from the individual centers. Although well-established, large centers with centralized record keeping were chosen in this study, many centers struggled to confirm and document consecutiveness when required to do so. In addition, due to the nature of treating patients in a mobile society, absolute true consecutiveness was recognized as impossible to achieve. Between patients lost to follow-up, occasional poor quality of records, inability to locate records, and lapses in the record-taking protocol, all centers realized a certain attrition rate for some patients in their consecutive samples. However, all participants agreed that random exclusions for those reasons were occasionally unavoidable but that no patients in the consecutively-treated list would be excluded for any other reason. A final important lesson all participants learned from this experience was that in spite of the assumed precise procedures in place to ensure proper enrollment, follow-up, and record-taking, their individual centers all had improvements to make. Some centers have now established a data or records manager on the team to ensure better adherence to protocol.

## Outcome Measures and Methods

Knowing that the sample sizes were a limiting factor in our Americleft study, efforts were made to increase the ability and accuracy to detect differences with the chosen methodology. Intra- and inter-examiner reliability testing was done for all ratings in this study, including the GOSLON model, cephalometric, and nasolabial esthetics ratings. The reliabilities ranged from moderate to very good. Knowing that the reliability of our raters was equal or better than similar studies reported in the literature, the Americleft team was satisfied that any error from rating reliabilities was minimized and acceptable.

As well, the number of multiple comparisons for treatment outcomes was minimized. Model comparisons between centers were made using the comprehensive GOSLON Yardstick rating system encompassing many model characteristics such as overjet, overbite, and arch constriction. Using a validated inclusive outcome measure such as the GOSLON Yardstick eliminates type 1 errors that may result when numerous individual measures from models (such as overjet, overbite, crossbites, and many more) are each measured and statistically compared. When multiple measures were compared for the GOSLON, cephalometric, and nasolabial analyses, the Tukey-Kramer pairwise comparison analysis for multiple comparisons was used in addition to analysis of variance (ANOVA) to reduce analysis error from the multiple repeated testing and give further credibility to any reported statistical results.

## Lessons From the Project as a Whole

Of greatest significance in the interpretation of the results reported in the preceding four parts is to emphasize the nature of retrospective clinical audits of treatment outcomes in comparison to true clinical research. With the many converging primary protocol variables (such as extent of original deformity; surgical procedures, techniques, and timing; operator skill; and ancillary infant orthopedic procedures) possibly affecting a given outcome, it is impossible to establish a cause-and-effect relationship between any given protocol feature and outcome (Long and Deacon, 2008). However, as stated by Shaw et al. (1996), "the existence of significant disparities in outcomes of the overall treatment process provides a basis for speculating about the possible cause(s)…" In this regard, the comparison of results in this study to the original Eurocleft study provide the basis for serious consideration of favorable versus unfavorable outcomes that seem to be increasingly associated with specific infant protocol features.

### 1. Dental Arch Relationships

The results of the five-center comparison using the GOSLON Yardstick demonstrated significant differences between centers. Center B was the only center with a majority of GOSLON scores of 4 and 5 resulting in a worse mean score than the other four centers. The protocol for Center B employed primary alveolar bone grafting with infant orthopedics as well as secondary surgical revisions completed prior to the time of the records used in this study. As mentioned above, it is not possible to completely delineate or assign causality of less favorable results to primary bone grafting, orthopedics, or additional surgeries. It is noteworthy that the Eurocleft study also found that centers that employed primary alveolar bone grafting and presurgical orthopedics were unable to demonstrate results that were any better (and to the contrary actually less favorable) than centers using much simpler and less burdensome protocols.

These results both from this study and the Eurocleft study would suggest that centers might want to reconsider the use of these procedures and reexamine extensive complicated protocols. In fact, in the follow-up Eurocleft study (Shaw et al., 2005) it was reported that centers originally using presurgical orthopedics and primary bone grafting had discontinued use of those treatments in their infant management protocols following the results of the outcome comparison. Likewise, Center B in the current Americleft study has discontinued use of primary bone grafting in its infant management protocol, due to the failure to detect any measurable benefits and the possibility of detrimental effects on maxillary development. In both the Eurocleft and the current Americleft studies, it is noteworthy that the most favorable dental arch relationship outcomes were obtained in centers with the simplest surgical protocols and without use of presurgical orthopedics and primary alveolar repair. Also noteworthy is that in the two best-rated Eurocleft Centers (B and E) and the highest rated Americleft Center (C), anterior palate repair included the use of a vomer flap, a procedure that has occasionally been questioned due to concerns over possible growth disturbances that clearly were not overwhelmingly demonstrated here.

Finally, results of this study with GOSLON scores between 2.6 and 3.7 corresponded well to those results of a study by Johnson et al. (2000) in West Australia and also of the Eurocleft project where GOSLON scores between 2.47 and 3.46 were reported. Noted in common, are the possible similar ethnic and genetic backgrounds leading to underlying Class II craniofacial morphology (Mars et al., 1992; Johnson et al., 2000). In contrast were the poorer GOSLON scores that were reported for Japanese (Susami et al., 2006) and Malay (Zreaqat et al., 2009) children who had a naturally occurring higher prevalence of underlying Class III skeletal relationships (Woon et al., 1989; Lew et al., 1994). A naturally occurring higher incidence of underlying skeletal Class III relationship may lead to a higher proportion of GOSLON scores of 4 and 5 that would not necessarily be related to treatment protocols. The evidence would support the fact that the underlying skeletal relationship cannot be ignored and must be considered when drawing conclusions regarding any effects from treatment on outcomes.

### 2. Craniofacial Form

One criticism that is often made of the GOSLON Yardstick is using inter-arch dental relationships to judge the outcome of various treatment protocols with respect to craniofacial morphology and the likely need for future orthognathic surgery. Furthermore, a two-dimensional cephalogram for patients with CUCLP can also mask true maxillary deficiency through erroneous identification of the cleft side maxillary position from superimposition of the noncleft side of the maxilla. These were concerns identified by the Americleft team, so the research methodology included the comparison between the GOSLON scores and the craniofacial morphologic assessments. When comparing the results from Part 2 (dental arch relationships) to those from Part 3 (craniofacial form), a good correlation can be seen between the GOSLON scores using the dental casts and maxillary prominence from the cephalometric analyses. The cephalometric comparison enabled us to further elucidate the reasons for the differences in GOSLON scores and attribute them to poorer maxillary prominence and not to mandibular, vertical, or dental components. These similar results from both the model scores and the cephalometric measures created a degree of confidence that the results from using the GOSLON Yardstick accurately predict maxillary prominence and future potential need for maxillary orthognathic surgery.

### 3. Nasolabial Esthetics

The challenges to objectively assess esthetics have been well established in the literature. The results of this study did not identify any differences between the nasolabial esthetics outcomes among the four centers. This result could reflect the fact either that all centers did, in deed, have similar nasolabial esthetics or that the method used could not distinguish between esthetic outcomes. Although statistical differences were not seen, the subjective nature of esthetics creates the possibility for a remaining clinical difference. Also, differences identified statistically cannot automatically be assumed to be clinically significant and must be reviewed with caution when assessing the final outcomes, especially with regard to esthetics. The subjective and clinical nature of these assessments is a challenge evident still in the literature. Clearly much work remains in order to improve the reliability and validity of nasolabial esthetic outcome assessments, including standardization of the series, orientation and views of the images taken, as well

as consideration for incorporating three-dimensional imaging.

### 4. The Dilemma of Operator Skill

A complex consideration that applies to all of the outcomes reported in this series relates to the issue of the skill, experience, and proficiency of the surgeons on whose shoulders the responsibility for primary management of infants with CUCLP falls. Proficiency bias is clearly a factor that cannot be controlled through clinical audits of outcomes such as this. In the original Eurocleft study, operator volume was identified as a potential feature of interest that related more favorable outcomes with high-volume operators (<6 surgeons/center) and less favorable outcomes with low-volume operators (>6 surgeons/center). In this Americleft study, operator volume, and hence experience, would seem to be less of a feature of interest since all centers had four or fewer surgeons involved with their entire sample. All centers also used standardized protocols that were another feature mentioned in the Eurocleft study as being associated with more favorable outcomes. However, documentation of individual operator skill remains elusive and therefore a source of proficiency bias that is uncontrolled in a retrospective audit such as this.

### 5. The Extended Benefits of Inter-center Collaborative Outcome Studies

The Americleft project was motivated by, and specifically designed to mirror the 1992 Eurocleft study. The latter had already proven itself to be a valuable milestone in raising awareness of the shortcomings of our previous research efforts, as well as demonstrating to us all the value of well-conceived and well-executed retrospective comparative outcome audits using common clinical records. The Eurocleft study also represented a method to raise the awareness of evidence-based practice decisions and a vehicle through which we could evaluate and change, if indicated, our treatment protocols and philosophies to improve clinical care for our patients with clefts. Therefore, the design of this Americleft study, by intention, was to attempt to replicate the Eurocleft experience in North America. In that sense, the study design (samples, methods, statistics, outcome measures, and analysis) was very straightforward and intentionally nonoriginal. This clearly provided us with a shortcut to execute the study, from inception to publication within 3 years, as opposed to many years longer for the original Eurocleft study. This process was further facilitated through the use of advisors from the Eurocleft project who had already been through the rigors of setting up an inter-center outcome study.

Of even greater benefit, however, is the demonstration, through our Americleft findings, of the ability to cross-compare outcome data, not only between centers within a particular study, but between centers from different but similarly designed and executed studies. Reference has been made repeatedly throughout this series to the similarity of the findings and conclusions between Americleft and Eurocleft. Our ability to make these valid comparisons is directly due to the fact that the samples, records, and methods used were all identical. All ratings and measures were done in the context of the same standard, uniform reference scales, and measures (ie, GOSLON Yardstick, standardized cephalometric measures, Asher-McDade nasolabial reference photos). For example, the GOSLON Yardstick reference models used in Americleft were the same sets of models used in 1992 for Eurocleft. The use of raters who participated in both the original Eurocleft study as well as the current Americleft study also provided a common link between studies to ensure consistency of calibration and ratings among examiners.

With this attempt at consistency to design these studies, it thus becomes possible to draw reasonable assumptions about comparing an outcome achieved by any given Americleft center and its protocol to the same outcome from any given Eurocleft center and its protocol. In fact, in 1999, prior to this Americleft study, the dental models and cephalometric radiographs from Americleft Center C sample were actually taken to the Eurocleft Good Practice Archive in Manchester, U.K., and added to and blinded with the original Eurocleft sample. They were then compared to the six Eurocleft centers using the GOSLON Yardstick, and standard cephalometric analysis was reported in Part 3 in this report. The favorable dental arch relation for Center C found in this Americleft study (2.6) was nearly identical to the score obtained when rated with the Eurocleft sample (2.3) (Vargas, unpublished data, 1999). This 1999 comparison of the first American center to be assessed with the Eurocleft sample also demonstrated the favorable comparability of Center C with the two highest rated Eurocleft Centers, B (2.2) and E (2.3). Nearly identical conclusions were reached regarding the association of more favorable outcomes with standard, simpler, and less burdensome protocols without presurgical infant orthopedics, without primary bone grafting, and fewer surgeons. This shows significant support for the ability to compare results across different studies adhering to the same rigid methods and the benefits of these inter-center collaborative studies. The opportunity to reliably compare results across similar studies exponentially increases our ability to acquire valuable data concerning treatment outcomes over a much larger number of centers with a wide range of treatment approaches.

## Summary

The success of the Americleft study has resulted from the dedication and commitment of the original investigators, the support from their individual centers, and the financial support granted by various organizations to facilitate this research. Over 3 years, the group completed the model

assessment using the GOSLON Yardstick, the craniofacial morphologic study using conventional cephalometric analyses, correlations between the GOSLON scores and cephalometric analyses, and nasolabial esthetics assessments, as well as publication of the results. In the absence of prospective clinical trials, the inter-center retrospective cohort approach demonstrates the capability of discerning dental arch and craniofacial morphologic relationships that constitute a basis for evaluating favorable and unfavorable maxillary growth and nasolabial esthetics for patients with CUCLP. The Americleft study in particular represents a successful attempt to fulfill the recommendation of the 2002 report of the World Health Organization on "Global strategies to reduce the health-care burden of craniofacial anomalies" (WHO, 2002), which stated that: "…most likely at this point, the most promising avenue…may still lie in the original Eurocleft approach. With a core of interested and experienced clinicians, operating at high volume centers, and willing to agree on records, outcome measures of significance, and research protocols, and additionally with the possible guidance from those involved in the successful Eurocleft, Scandcleft and Eurocran programs, it might still be possible to initiate a major inter-center collaborative research effort (in North America)."

Future areas of interest for the Americleft inter-center study include evaluation of the burden of care of various treatment protocols, secondary alveolar bone grafting, comparison of speech outcomes, psychosocial assessment by patients and families, patients' evaluation of the treatment they received, nasoalveolar molding, and the long-term follow-up assessment of those factors assessed in this initial Americleft study.

## References


Johnson N, Williams AC, Singer S, Southall P, Atack N, Sandy JR. Dentoalveolar relations in children born with a unilateral cleft lip and palate (UCLP) in Western Australia. *Cleft Palate Craniofac J*. 2000;37:12–16.

Lew KK. Cephalometric ideals in Chinese, Malay and Indian ethnic groups. *Asian J Aesthet Dent*. 1994;2:35–38.

Long RE Jr, Deacon SA. Assessment of orthodontic outcomes in patients with clefts. In: Losee JE, Kirschner RE, eds. *Comprehensive Cleft Care*. New York: McGraw-Hill Medical; 2008:1061–1077.

Mars M, Asher-McDade C, Brattström V, Dahl E, McWilliam J, Mølsted K, Plint DA, Prahl-Andersen B, Semb G, Shaw WC, et al. 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.

Shaw WC, Roberts CT, Semb G. Evaluating treatment alternatives: measurement and design. In: Turvey TA, Vig KWL, Fonseca RJ, eds. *Facial clefts and craniosynostosis. Principles and management*. Philadelphia: WB Saunders Company; 1996:756–766.

Shaw WC, Brattström V, Mølsted K, Prahl-Andersen B, Roberts CT. The Eurocleft study: intercenter study of the treatment outcome in patients with complete cleft lip and palate. Part 5: discussion and conclusions. *Cleft Palate Craniofac J*. 2005;42:93–98.

Susami T, Ogihara Y, Matsuzaki M, Sakiyama M, Takato T, Shaw WC, Semb G. Assessment of dental arch relationships in Japanese patients with unilateral cleft lip and palate. *Cleft Palate Craniofac J*. 2006;43:96–102.

Woon KC, Thong YL, Abdul Kadir R. Permanent dentition occlusion in Chinese, Indian and Malay groups in Malaysia. *Aust Orthod J*. 1989;11:45–48.

World Health Organization. Global strategies to reduce the health-care burden of craniofacial anomalies. Report of WHO meetings on International Collaborative Research on Craniofacial Anomalies. WHO Human Genetics Programme; 2002.

Zreaqat M, Hassan R, Halim AS. Dentoalveolar relationships of Malay children with unilateral cleft lip and palate. *Cleft Palate Craniofac J*. 2009;46:326–330.