

ORIGINAL ARTICLES

The Americleft Study: An Inter-Center Study of Treatment Outcomes for Patients With Unilateral Cleft Lip and Palate Part 1. Principles and Study Design

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., Kathleen Russell, D.D.S., M.Sc., Marilyn Cohen, B.A., Gunvor Semb, D.D.S., Ph.D., William Shaw, B.D.S., Ph.D.

Objective: The Americleft study is a North American initiative to undertake an intercenter outcome study for patients with repaired complete unilateral cleft lip and palate from five well-established North American cleft centers.

Design: Retrospective cohort study.

Setting: Five cleft palate centers in North America.

Methods: This is the first paper in a series of five that outlines the overall goals of the study and sets the basis for the clinical outcome studies that are reported in the following four papers. The five centers' samples and treatment protocols as well as the methods used for each study are reported. The challenges encountered and possible mechanisms to resolve them and reduce methodological error with intercenter studies are also reviewed.

KEY WORDS: *Americleft, cleft lip and palate, intercenter study, treatment outcome measures*

By the 1990s the array of surgical treatment approaches and primary infant management protocols for patients with

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. Dr. Russell is Associate Professor and Head, Division of Orthodontics, Dalhousie University, and Staff Orthodontist and Chair, Cleft Palate Team, IWK Health Centre, Halifax, Nova Scotia, Canada. Ms. Cohen is Administrative Director and Patient Care Coordinator, Regional Cleft-Craniofacial Program, Cooper University Hospital, Moorestown, New Jersey. Dr. Semb is Adjunct Professor, Dental Faculty, 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-Craniofacial Association meeting in Boulder, Colorado, April 2007; the 65th American Cleft Palate-Craniofacial Association meeting in Philadelphia, Pennsylvania, April 2008; and the 66th American Cleft Palate-Craniofacial Association meeting in Scottsdale, Arizona, 2009. In addition, portions were presented at the 8th European Craniofacial Congress, Bilbao, Spain, June 2007; Craniofacial Society of Great Britain and Ireland, Chester, United Kingdom, April 2008; the Japanese Cleft Palate Association, Hirsoshima, 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.

cleft lip and palate had become enormous, with little reliable information available upon which to base rational decisions to choose one method over another (Semb and Shaw, 1998). In a survey of 201 European cleft centers, 194 different surgical protocols were in use for the primary closure of a complete unilateral cleft lip and palate (CUCLP) alone (Shaw et al., 2000). With this number of approaches possible, the inadequate research methods of previous decades offered little chance to identify specific individual procedures that could be shown to be unequivocally superior to others (Spriestersbach et al., 1973; Roberts et al., 1991; Shaw et al., 1996; Semb and Shaw, 1998).

However, as described in a "State of the Art" paper by Long et al. (2000), an approach using rigorous comparisons of outcomes from different centers marked a significant change in direction from previous methodologies. In 1992, the landmark Eurocleft study (Asher-McDade et al., 1992;

The authors and Americleft group would like to acknowledge and thank the American Cleft Palate-Craniofacial Association and the Cleft Palate Foundation for their yearly 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. K. A. 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: 10.1597/09-180.1

Mars et al., 1992; Mølsted et al., 1992; Shaw et al., 1992a; Shaw et al., 1992b) clearly established the value of intercenter collaboration to carry out well-controlled assessments of treatment outcomes using retrospective data. Using records routinely collected for clinical diagnosis and treatment purposes, the study design and execution of the original Eurocleft investigation allowed for objective and unbiased detection of favorable versus unfavorable effects of varying primary infant management protocols for patients with CUCLP in the areas of dental arch relationship (Mars et al., 1992), craniofacial morphology (Mølsted et al., 1992), and nasolabial aesthetics (Asher-McDade et al., 1992). Most notably, the outcomes found most favorable to improve results in these key areas appeared to be in centers with simple surgical and orthodontic approaches (Shaw et al., 1992a; Shaw et al., 1992b). Though clearly not capable of identifying the individual procedures within a total treatment management protocol that are responsible for favorable or unfavorable outcomes, intercenter comparisons have become a key part of quality improvement programs, leading to international agreement on standardized documentation (Shaw et al., 2000).

Although the original Eurocleft project subsequently led to the networking of most cleft–craniofacial teams in Europe (Eurocleft Network) and to two randomized controlled trials of surgical and orthopedic interventions for primary infant management, similar progress among centers in North America has not materialized. North American centers have been unable to generate any significant momentum in inter-center, collaborative clinical research in spite of the large number of high-volume, well-organized centers and the organizational support available through the American Cleft Palate–Craniofacial Association. In 2004, the World Health Organization (WHO, 2004, pp. 46-47) devoted an entire report to promote international collaborative research and found that “... unlike the European experience, in which the original study generated a groundswell of support, extension of the clinical research approach throughout European centres and led to the establishment of Eurocran, Scandcleft and strong financial support from government and nongovernmental sources, the experience in the US has been the opposite.” Not only have there been no other US centers to join in the European collaborations, but attempts to duplicate the approach within the US between centers have been limited to isolated collaborations yielding very limited information, and failing to lead to expansion of interest to include additional centers.

The reasons for this failure are complex. They include a lack of funding, lack of research training, and lack of time and interest on the part of primary care providers to engage in research projects. Another critical problem in the United States is the current health care climate that tends to prioritize cost containment and access to care over quality of care. This in turn has created a tendency for decentralization of care provision. Although the existence of a large number of centers and individuals providing

treatment for craniofacial anomalies improves patients’ geographical accessibility to care, it simultaneously creates a fractionation of the study population, thereby reducing the probability of developing patient samples of adequate size to enable valid research. The entire landscape is further complicated by noncomparable patient populations, non-comparable treatment records, unquantifiable differences in operator skills, difficulties in letting go of biases, and an unwillingness to question our own protocols developed and used for years within individual centers. For intercenter collaboration to succeed, individual centers must be willing to entertain a degree of uncertainty about the effectiveness of their chosen protocols and a desire to discover ways to improve those protocols based on the experiences and outcomes of others. Finally, there remains a general lack of agreement among centers on minimal standards for reporting and recording outcomes, as well as cost and ethical concerns over taking records that cannot be clearly identified as essential for diagnosis and treatment. In an attempt to remedy this problem, in 2006 the American Cleft Palate–Craniofacial Association (ACPA) established a Task Force on Intercenter Collaboration as part of its Research Committee and provided support for the initiation of an Americleft study, based on the principles established in the original Eurocleft study. These basic principles are as follows.

(1) Understanding the Strength of Evidence Hierarchy

Numerous authors have emphasized the relative strengths of evidence obtained through various reporting methods and types of investigations (Roberts et al., 1991; Long and Deacon, 2008). The strongest evidence is clearly in the domain of the randomized control trial (RCT) because it best identifies the effects of specific features of treatment management protocols and control for bias. For cleft lip and palate treatment, such RCTs are difficult to execute due to the time, costs, and large samples required. Fortunately, though, the effect of these methodological challenges, confounding variables, and research biases can be minimized by study design. In the hierarchy of evidence, intercenter comparisons of outcomes are considered second strongest to RCTs, assuming appropriate interpretation of results and rigorous attempts to minimize bias (WHO, 2002).

(2) Controlling Bias

Bias has been identified by many as a primary weakness that limits the strength of evidence of retrospective studies. As emphasized in the 2002 WHO Report, “differences arising from the biases ... are likely to exceed actual differences attributable to the procedures” (p. 18). In this investigation, *susceptibility (case selection) bias* was controlled by the requirement for large, consecutively treated samples with documentation of CUCLP to ensure equivalency of the samples at the outset. *Follow-up bias* was

minimized by the requirement to account for any patients initially enrolled in the center for whom records were not available for the outcomes study (i.e., patients moving away, failing to return for team evaluations). Management of *analysis bias* was achieved through statistically determined minimal sample sizes, blinding all raters and examiners as to the source of the records; training and calibrating all raters and examiners; and statistically determining intrarater and interrater reliability. Only *proficiency bias* related to variation in operator skills could not be controlled and remains a potentially significant factor to explain the outcome differences identified.

(3) Using Standardized Clinical Records

One of the greatest obstacles to meaningful intercenter outcome comparisons is the inconsistency and noncomparability of the records that document the outcomes. Therefore, some agreement among participating centers on the minimal standards for record taking was critical to this study. It was critical that the records used for the intercenter comparison were those that were normally taken for treatment planning purposes. Given that mixed dentition orthodontic treatment is standard in most centers, dental models, photographs, and radiographs are generally taken by the orthodontists as a routine part of this treatment planning. Fortunately, as shown in the original Eurocleft, these records can be used simultaneously to evaluate dental arch relationships, skeletal morphology, and facial aesthetics outcomes resulting from prior interventions with appropriate ethics approval. Thus, these records and the outcomes they represent are the most immediate and simplest way to initiate an intercenter collaborative study.

KEY METHODOLOGICAL CONSIDERATIONS

In this study, there were seven main considerations that were established for the methodology. Again, these were largely attempts to mirror the methods used in the original Eurocleft study. Given the difficulties described above in making any progress whatsoever in establishing the interest in, and foundations for, successful intercenter comparisons, it was decided to take advantage of the successful Eurocleft experience as a model for Americleft. Therefore the following considerations were established as ideal targets for this study, with the understanding that, as with the Eurocleft experience, achieving them all would be a difficult undertaking. However, by setting these standards for the study at the outset, it was also hoped that any shortcomings encountered would be more clearly evident when considered in the context of what was ideal. Although five centers were included as the pilot group for this study, based on the varied ability of each center to meet all criteria for all of the outcome measures investigated, desired samples were nearly impossible to achieve for all centers for all outcome

measures. Nonetheless, it was felt of value to set these target methodological considerations at the outset of the study to mirror Eurocleft as much as possible and to use as a yardstick to help better understand the challenges of initiating a project of this nature. These key considerations were:

- (1) Sample sizes in the range of 30 to 40 were sought. This was based on the findings of the Eurocleft study, part 5 (Shaw et al., 1992b). The final conclusions in that study regarding sample sizes suggested that for three of the pertinent outcome measures (Goslon score of dental arch relationship, cephalometric maxillary prominence, and sagittal soft tissue relationships), five center comparisons based on a .5 significance level, 80% power, would need sample sizes of 30 to 40 to allow for detection of between 0.50 and 0.75 of a Goslon point and 3 to 4 degrees of soft tissue relationship difference. In the case of cephalometric maxillary prominence and to increase the sensitivity of difference detection for Goslon scores and soft tissue relationship, significantly larger sample sizes (97, 97, and 63, respectively) would be required. At the outset it was understood that sample sizes of that magnitude would not be feasible for most centers given the other methodological requirements of consecutiveness, complete nonsyndromic UCLP, and availability of records. Therefore, the 30 to 40 sample size was targeted to at least enable significant power for a few of the most sensitive and important outcome measures. Smaller samples were included in the study but would not represent a reliable and valid assessment of the outcome in question. Center A (see below) was only included in the dental arch relationship part of the study because that measure required only a sample of 16 for detection of a 1.0 Goslon point difference.
- (2) All patients in the study were white. In addition, samples had to be separated according to cleft type. Because CUCLP is the most frequently encountered type of cleft, CUCLP was chosen to facilitate achieving adequate sample sizes.
- (3) All unilateral cleft lip and palates had to be complete clefts and nonsyndromic; Simonart bands were permissible as part of a CUCLP as long as the skeletal (alveolar) cleft was complete. Furthermore, records had to be available to document and confirm the initial condition (i.e., initial entry chart notes, photos, or dental models).
- (4) Each center was required to provide evidence, usually through chart note entry dates, patient number, or patient birth date, that the samples consisted of consecutively enrolled patients. It was understood that although pure consecutiveness was desired, it might not be possible due to patients lost to long-term follow-up.
- (5) Records used for outcome assessments were unlinked to any cleft–craniofacial clinic or health center and

were prepared in such a way as to ensure blinding of raters and examiners. This was deemed especially important in intercenter comparisons of this nature due to the sensitivity of investigating favorable or unfavorable center protocols and the possibility of incorporating analyses bias.

- (6) Records describing the primary treatment protocol and number of operating surgeons involved in the primary surgeries had to be available. Furthermore, it was required that all primary surgeries were completed at each respective center.
- (7) All centers had to obtain patient permission and institutional review board (ethics) approval for use of the clinical records in an intercenter outcome assessment study. The study was designed to further ensure privacy of patient identity and information as well as the identity of the individual centers and surgeons/clinicians.

SUMMARY OF PARTICIPATING CENTERS AND THEIR PRIMARY PROTOCOLS

Five centers were chosen to participate in the Americleft study. The first criterion of the participating centers was to meet the above seven methodological conditions. Second, yet equally important, was the willingness and commitment of the individual cleft teams to openly enter into, without preconceived bias, an intercenter collaboration to determine the relative benefits of their protocols compared with those achieved by the other teams. All participating centers had well-established North American cleft teams with long histories of involvement in cleft lip and palate treatment that provided centralized, multidisciplinary care and used standardized protocols. In addition, the centers, identified as A through E in Table 1, were chosen to provide a wide range of protocols. The protocols for the individual centers are summarized in Table 1. Noteworthy within these tables is the inclusion of one center that used primary bone grafting (Center B), three centers that used variations of presurgical orthopedic treatment (Centers B, D, and E), and one center using two-stage palate repair (Center C). There was also a wide and representative range of lip and palate repair techniques. Finally, the number of surgeons involved at the various centers was small, especially compared with the number of surgeons reported for some centers in the original Eurocleft study (Shaw et al., 1992a). This is relevant due to the suggestion from the Eurocleft study of a surgical volume/skill effect on outcomes.

The total sample pool from the five centers was 172 patients, with individual sample sizes all reaching desirable levels except for Center A. For many patients complete records necessary for all three parts of this study were available. However, final sample sizes in parts 2, 3, and 4 were slightly less than that listed in Table 1 due to occasional missing or poor-quality dental models or

TABLE 1 Sample Characteristics and Treatment Protocols for the Study Centers

Treatment	Center A	Center B	Center C	Center D	Center E
Sample size and gender	18 (10 boys, 8 girls)	40 (28 boys, 12 girls)	41 (25 boys, 16 girls)	38 (29 boys, 9 girls)	35 (21 boys, 14 girls)
Average age (y:m)	9:4	8:6	8:10	9:1	9:2
Presurgical orthopedics	No	Yes*	No	Yes†	Yes‡
Lip repair	6–12 wk Millard or 5–6 m Delaire§	2–3 m Millard	3 m Tennison	3 m Millard	3–4 m Millard
Primary bone grafting	No	Yes 6–9 m	No	No	No
Hard palate repair	9–12 m Bardach or Delaire§	11–15 m hard palate Wardill-Kilner	12 m Vomer flap	12 m Wardill and Vomer flap	12–14 m Vomer flap/von Langenbeck
Soft palate repair	9–12 m Bardach or 5–6 m Delaire§	11–15 m IVP or some Furlow (1 surgeon)	18 m median suture with IVP	12 m Wardill and Vomer flap	12–14 m Veau pushback
Secondary bone grafting	6–7 y	8–9 y if needed	9 y	7–10 y	9–11 y
Nose/lip revisions	4–5 y	4–5 y	14–20 y	4–5 y	4–7 y
Surgeons	2	4	1	4	4

* Center B. The molding plate was started at 2 to 3 months and continued through primary bone grafting to the palate repair at 11 to 15 months. The molding plate was discontinued at the time of palate repair.

† Center D. Infant presurgical orthopedic treatment was done using a modified McNeil technique with extraoral traction. The orthopedic appliance was placed prior to lip repair at 3 months and discontinued once lip repair was done.

‡ Center E. Infant presurgical orthopedic treatment was done using a modified McNeil technique with extraoral traction. The orthopedic appliance was placed prior to lip repair at 3 to 4 months and continued until the time of palate repair at 12 to 14 months.

§ Those patients who received a 6- to 12-week Millard lip repair had a 9- to 12-month Bardach palate repair; whereas, those who received a 5- to 6-month Delaire lip and soft palate repair received a 9- to 12-month Delaire palate repair.

|| IVP = intravelar veloplasty.

radiographic or photographic records on individual patients. In addition, Center A was not included in parts 3 and 4 due lack of the necessary radiographs and photographs on all patients, further reducing the already suboptimal sample size.

Following identification of the participating centers, each center was given time to obtain ethics approval and prepare their samples. The initial outcomes to be assessed were identical to those reported in the initial Eurocleft study and included dental arch relationships using dental models and the Goslon rating system, skeletal and soft tissue craniofacial morphology using lateral cephalometric radiographs, and nasolabial aesthetics using standard frontal and lateral photographs. The methodology and results for each of these assessments will be reported in papers 2 through 4, and a summary paper will close the series as the fifth paper.

Acknowledgments. The authors and Americleft group would like to acknowledge Jennifer Gregory, Data Manager and Research Assistant, Lancaster Cleft Palate Clinic. We would also like to acknowledge Drs. Bryan Ruda, Mairaj Ahmed, and Brian Smith and Mr. J.B. Peterman, research assistants at the Lancaster Cleft Palate Clinic, and Dr. Mike Horst, Director of Research at Lancaster General Research Institute.

REFERENCES

- Asher-McDade C, Brattström V, Dahl E, McWilliam J, Mølsted K, Plint DA, Prah-Andersen B, Semb G, Shaw WC. 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.
- Long Jr RE, 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.
- Long Jr RE, Semb G, Shaw WC. State of the art—orthodontic treatment of the patient with complete clefts of the lip, alveolus and palate. The 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, Prah-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.
- Mølsted K, Asher-McDade C, Brattström V, Dahl E, Mars M, McWilliam J, Plint DA, Prah-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 2. Craniofacial form and soft tissue profile. *Cleft Palate Craniofac J.* 1992;29:398–404.
- Roberts CT, Semb G, Shaw WC. Strategies for the advancement of surgical methods in cleft lip and palate. *Cleft Palate Craniofac J.* 1991;28:141–149.
- Semb G, Shaw WC. Facial growth after different methods of surgical intervention in patients with cleft lip and palate. *Acta Odontol Scand.* 1998;56:352–355.
- Shaw WC, Asher-McDade C, Brattström V, Dahl E, McWilliam J, Mølsted K, Plint DA, Prah-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, Prah-Andersen B, Roberts C, et al. 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.
- 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; 1996:756–766.
- Shaw WC, Semb G, Nelson P, Brattström V, Mølsted K, Prah-Andersen B. *The Eurocleft Project 1996–2000*. Amsterdam: IOS Press; 2000.
- Spiersbach DC, Dickson DR, Fraser FC, Horowitz SL, McWilliams BJ, Paradise JL, Randall P. Clinical research in cleft lip and cleft palate: the state of the art. *Cleft Palate J.* 1973;10:113–165.
- World Health Organization (WHO). *Global Strategies to Reduce the Health-Care Burden of Craniofacial Anomalies. Report of WHO meetings on International Collaborative Research on Craniofacial Anomalies*. World Health Organization, Geneva Switzerland; 2002.
- World Health Organization (WHO). *Addressing the global challenges of craniofacial anomalies: Report of a WHO meeting on International Collaborative Research on Craniofacial Anomalies*. World Health Organization, Geneva Switzerland; 2004.