The Americleft Project

Experiences and Recommendations for Establishing Successful Inter-center Collaborative Outcome Study

The Americleft Project
Table of Contents

Section 1: Introduction and Background

Background

Thanks for your willingness to consider becoming involved in Inter-center Collaborative Outcomes Research. By way of background, the following is a summary of the events leading to the establishment of the Americleft Project.

Although this has been a topic of interest for many years and in spite of many good people with good intentions, before 2006 centers in the US and Canada had not been as successful as those in Europe in establishing interest and commitment to inter-center collaborative outcome studies. A recent WHO report, "Addressing the Global Challenges of Craniofacial Anomalies," has emphasized the need for, and benefits of, this type of research, based on the successes and accomplishments of the Eurocleft and Eurocran projects. As stated in that report:

> "*Professionals entrusted with the provision of health care have an obligation to review the success of their practices and, where shortcomings are revealed, to take remedial action. Such efforts should constitute a continuous cycle, sometimes known as a 'clinical audit'…….. (which) is divided into evaluating the process of care (the way in which care is delivered) and the **outcomes** of care (what is achieved)……Audit of the treatment of clefts is a considerable challenge, because of the lengthy follow-up required, the complexity, subtlety and number of relevant outcomes and, above all, the relatively small number of cases. Inter-center collaboration still offers significant advantages, by providing insight into the processes and outcomes of treatment of comparable services elsewhere, the establishment of future goals and the exchange of clearly successful practices. …..Perhaps the greatest benefit of inter-center comparisons is the cooperative spirit that they foster and a gradual diminution of rivalry.*"

The same report included the following summary of a lack progress in the US and North America for stimulating inter-center collaborative research:

> ### *"Current Status of Inter-center Collaboration for Clinical Research in the US and North America"*
>
> "*Although the US is a leader in many areas of the management of patients with clefts and craniofacial anomalies (CFA), has many well-organized CFA teams, and the largest professional organization in the world in this field (American Cleft Palate-Craniofacial Association), there has been little significant momentum in the area of inter-center, collaborative, clinical research, especially compared to the more successful efforts of Eurocleft and Eurocran. .*
>
> *…….of current and ongoing research projects in North America, relatively few were statistically sound, unbiased inter-center assessment and comparison of clinical outcomes. Few, if any, were actual randomized control clinical trials. This would suggest that much of the ongoing clinical research currently underway in the North America, may continue to generate little useful information which would contribute to the establishment of sound evidence-based decision making in clinical care.*
>
> *As another indication of the opportunities available in the North America, Uhrich et al. reported that over 100 teams in the US have yearly new CFA patient caseloads of over*

50. The necessity of high volume centers and care providers in providing sample sizes adequate to conduct outcome audits and clinical trials in a time- and cost-effective manner has been well-established through the Eurocleft study (1992), and subsequent CSAG report in the UK.

Several attempts have been made at a number of different levels to take advantage of these clinical research opportunities. However, unlike the European efforts in which original study generated a groundswell of support and extension of the clinical research approach throughout European centers, and led to the establishment of Eurocran, Scandcleft and strong financial support from governmental sources and NGOs', the experience in North America has been the opposite.

The reasons for this failure are a reflection of problems and obstacles..... While the large number of centers and individuals providing treatment for CFA in North America 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 non-comparable patient populations, non-comparable treatment records, unquantifiable differences in operator skills, and difficulties in letting go of biases. Also, while collaborative research can be structured without violating patient privacy laws, the rigors of doing so are sufficient discouragement for many clinicians to participate. Finally, there remains a general lack of agreement between centers on minimal standards for reporting and recording outcomes, as well as cost and ethical concerns over taking records which cannot be clearly identified as essential for diagnosis and treatment purposes.

Potential remedies for these obstacles have been attempted in the North America. These represent "top-down" solutions where organizations or groups have attempted to facilitate standardization of recording treatment histories and outcome data, centralization of data through a registry, and networking between individuals and centers with collaborative research interests. To date, these initiatives have also either failed or met with limited success. Most notably, the Craniofacial Outcomes Registry (COR), an NIDCR funded project..... was discontinued due to lack of renewal funding.

Another resource which is still operational is the APCA Data Base. While potentially useful in identification of patient samples which might be appropriate for trials or outcome studies, the Data Base however, has not been accepted and utilized by all centers and requires some modification to include outcome measures useful for collaborative studies.

*In summary, although the desire, research talent, and patient samples, would all seem to be readily available in the North America, the failure to get centers to agree on something as basic as standardization of recording and reporting outcomes, as well as governmental hurdles and a serious lack of funding, have all resulted in a huge and ongoing missed opportunity**. It seems most likely at this point, that the most promising avenue to break out of this inertia, 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."***

Based on the concluding bolded statement, in 2006, the Executive Council of ACPA approved funds for the Research Education Committee to organize a pilot project which has become "Americleft", an ACPA Task Force. Six centers were identified to participate in the pilot project which resulted in 3 face-to-face meetings over the past two years, completion of initial comparisons of dental arch relationship outcomes, cephalometric morphology outcomes, nasolabial esthetic outcomes and the protocol for assessing speech outcomes. The results of these have all confirmed the value and benefits of well-controlled and well-designed inter-center outcome comparisons. Most importantly however, is the experience and insight that has been gained in the understanding the requirements, demands, and possible obstacles that must be overcome in order to participate successfully in such collaborative studies. The purpose of this Eye Opener and Study Guide is to provide potential participants with a "roadmap" to participation including requirements, recommendations and suggestions, all based on the experiences of our Eurocleft colleagues and the initial Americleft members. While these requirements are not meant to be restrictive or exclusive for centers wishing to participate, they are based on two fundamental overriding goals. First is the absolute necessity of protecting the privacy and confidentiality of the centers wishing to participate, as well as their patients. Second is the need to maintain the highest possible scientific standards. If the research is compromised, the value and validity of the outcome studies are lost. Therefore the continued success of Americleft hinges on the integrity and intellectual honesty of those choosing to participate. In addition, several requirements which were established with the selection of the initial Americleft participating centers are still applicable:

1. A center with team members experienced and focally interested in CLP, and with an interest in seeking knowledge about the relative merits of various primary protocols rather than having an unquestioning loyalty to particular procedures. While we all believe that the procedures we are doing are the best possible for our patients, involvement in collaborative outcome studies implies a degree of **<u>uncertainty</u>** about the true effectiveness of our individual protocols, the ability to question our own beliefs and to accept the possibility that there may be other equally good or better outcomes with protocols different from the one(s) used by our own team.
2. A high volume center with reasonably consistent protocol for primary management of its CLP patient population.
3. A center with the resources to support team representatives in dedicating the time, and absorbing the costs of the effort.
4. The availability of the necessary records (privacy protected of course), and ability to secure IRB approval from the parent institution.

It has been our experience that many Centers may have difficulty identifying existing samples of sufficient size which meet all inclusuion/exclusion criteria and for which adequate records already exist. For those that do, outcome studies can begin immediately. For those that don't, this study guide may help in establishing protocols for record-taking which can be used prospectively in planning for eventual inter-center outcome audits.

Section 2: Americleft Task Force Strategic Planning

The Task Force has now completed and presented the results of its initial project, the first multi-center, multiple outcome study in North America. This focused on orthodontically-related outcomes of primary infant management protocols for UCLP (dental arch relationship, skeletal morphology, and nasolabial esthetics). The intent of this strategic planning session was to define the short-term and longer-term goals and objectives to expand the activities of the Task Force in the context of a long term vision of the ultimate purpose of the project. Especially important in this regard was the inclusion of representatives from surgery, speech and psychology to establish the multi-disciplinary nature of the project. Within this multi-disciplinary framework, an ultimate vision of Americleft as a resource for centers to carry out internal audits of their own outcomes, to take part in inter-center collaborative outcome studies, and to engage in randomized control trials of critical treatment protocol features, was developed.

Shorter-term objectives were established, including completion of other orthodontic outcomes, recruitment of other centers with protocols containing features of interest to join in the multi-center outcome study, creation of a core surgical outcomes planning team, development of surveys to evaluate burden of care and patient/parent satisfaction issues, beta testing of a speech outcomes assessment tool, and establishment of minimum standards for recording outcomes in the various disciplines. A workbook containing recommendations and requirements for participation in the Americleft inter-center outcome comparisons was also developed and distributed at the ACPA meeting to ACPA members expressing an interest in project.

Finally, immediate goals and plans were made for the next Americleft meeting to be held at the Lancaster Cleft Palate Clinic in fall of this year.
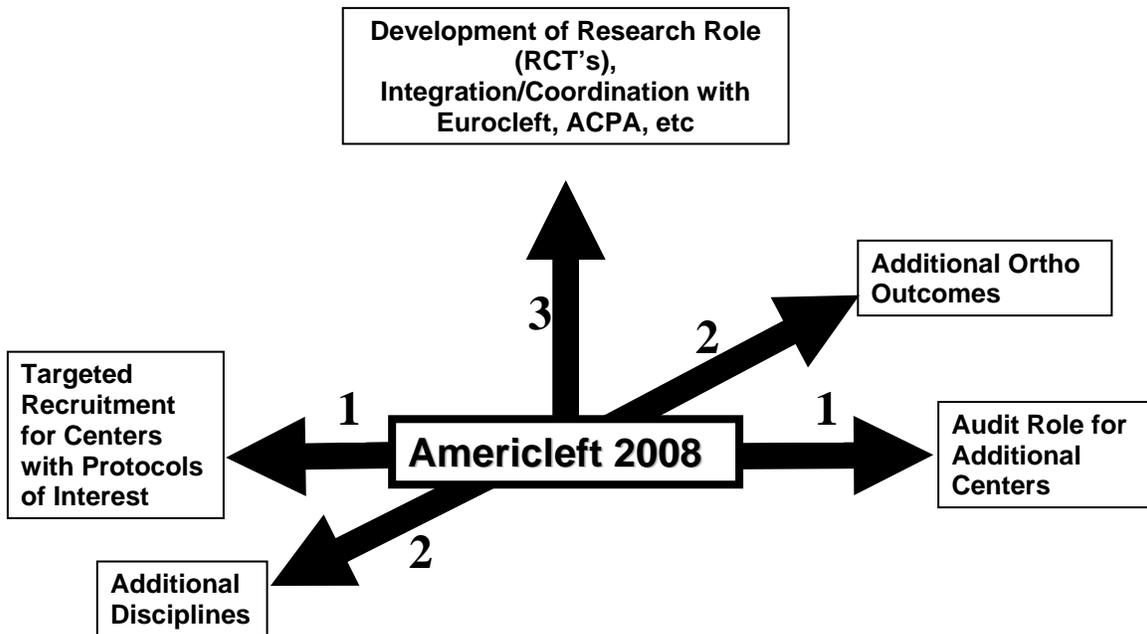
Details of the results of this meeting follow.

Purpose: At this stage, the best description of the purpose of Americleft is probably to base it on the charges to the Task Force from ACPA Executive Council…… "to establish the benefits of interdisciplinary team care". In that context, the more precise purpose could be "to demonstrate and document outcomes to be expected with team care, and to define the key features or characteristics of various team treatment procotcols and procedures that are associated with more or less favorable/desirable outcomes".

The Americleft Task Force will be working in conjunction with an additional Task Force on the Economics of Team Care, which is charged with the goal of documenting the financial benefits of team care. The common link between the Task Forces would seem to be to combine data on clinical outcomes (benefits of various treatments) with assessments of clinical burden of care and the associated financial burdens of care associated with those outcomes.

Shared Vision: Consensus on a "shared vision" among the participants was not fully articulated. However, discussion focused on the possible future directions Americleft (also considered "opportunities") could take at this juncture, as illustrated in the following diagram and described below.

Possible Future Directions (Opportunities??)

Development of Research Role (RCT's), Integration/Coordination with Eurocleft, ACPA, etc

3

Additional Ortho Outcomes

2

Targeted Recruitment for Centers with Protocols of Interest

1

Americleft 2008

1

Audit Role for Additional Centers

2

Additional Disciplines

Operationally, these directions for expansion of the project could be defined in 3 dimensions.

1. Lateral Expansion:
   a. To establish protocols and processes for ACPA member teams to participate in inter-center outcomes comparisons (research) or intra-center outcome audits.
   b. To establish an "Outcomes and Good Practice Archive" in North America for the purposes of inter-center comparisons and intra-center audits
   c. To identify and recruit centers with protocols which include features of interest to participate in inter-center outcomes comparisons
   d. To establish standards for records which would be required for comparisons and audits.
2. Depth Expansion:
   a. To complete orthodontically-related outcome comparisons
   b. To identify individuals and establish protocols and measures for inter-center outcome comparisons in speech, surgery and psychology.
3. Vertical Expansion:
   a. To participate in and contribute to the creation of a global network of similar initiatives with common standards of recording and reporting treatment outcomes

b. To use the results of its outcome comparisons to identify individuals, centers, treatment protocols, and samples for participation in randomized control trials. (see figure below)
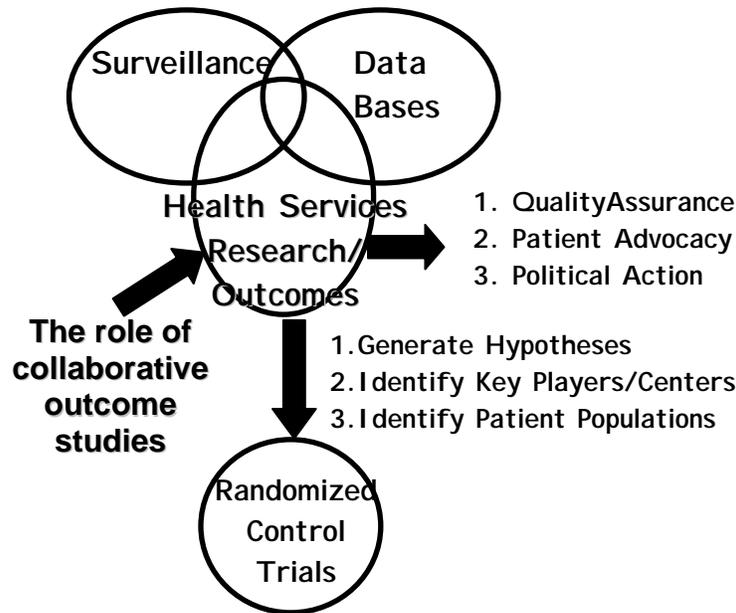
Based on the response to the Americleft Eye Opener and Panel at the Annual Meeting, and the ensuing discussion, the Task Force made the following assumptions:

1. There is a need, desire and benefit for Americleft expansion in all three directions.
2. The two main roles Americleft could fill would be as a resource for research (clinical outcomes, randomized control trials) and clinical audits.
3. The infra-structure and groundwork which result from inter-center outcome comparisons and which are necessary for vertical expansion into randomized control trials, are not yet in place.
4. Standards of record-taking and outcome recording are not currently adequate and accepted between centers to allow for significant lateral expansion and enrollment of additional centers.
5. The multi-disciplinary nature of CLP treatment, and the support of ACPA and CPF mandate immediate attempts at depth expansion in to other disciplines.
6. Through success of Americleft, ACPA may be able to utilize not only the data generated by Americleft initiatives, but also the standards and structure established in the process, as a product-line benefit to its members and teams.

Although definitive shared vision statements were not articulated at this session, in an attempt to summarize the preceding discussions and dialogue which took place at the meeting and distill them down to basic shared ideas, the current **Shared Vision** of the members might be expressed in the following statements:

1. Americleft will be the outcome registry and good practice archive for ACPA and its members.

2. Americleft will be the resource for ACPA member teams to participate in inter-center collaborative outcome comparisons or to assist in auditing their own treatment outcomes.
3. Americleft will be the lead organizing agency and primary resource for ACPA members and member teams to participate in and conduct randomized control trials.

Assuming acceptance of the purpose and shared vision as articulated above, a number of 3-5 year Goals and Objectives and 1-2 year Action Steps were discussed. Again, although final and formal documentation of those was never completed at this session, there was sufficient dialogue to allow for an attempt at structure and order. As a first attempt, the proposed Goals, Objectives and Action Steps have been placed in the Strategic Planning Template used by ACPA Executive Council in it strategic planning process. In following with the 3-dimensional approach to continuation and expansion of Americleft, the Goals expressed through discussion will be itemized according to the directions described.

GOAL 1: Orthodontically-related inter-center outcome comparisons will be completed for the original Americleft Centers with final determination of outcome measures to be used.

GOAL 2: Orthodontically-related inter-center outcome comparisons with additional Centers using protocols which include treatment procedures of particular interest, not yet included in previous comparisons, will be completed.

GOAL 3: Inter-center comparisons of 3-5 key outcomes of significance in other disciplines will be completed between select Centers with final determination as to outcome measures to be used.

GOAL 4: Minimal standards for record-taking in the main disciplines, necessary for participation in inter-center studies and clinical audits, will be established, based on findings and experiences with the initial participants.

GOAL 5: Protocols and procedures will be in place for the purpose of archiving the records and results of outcome comparisons, to provide a "Good Practice Reference Registry" against which ACPA member teams could benchmark their outcomes for internal audit purposes.

GOAL 6: A randomized control trial to investigate the key features of infant management protocols which were identified in initial outcome comparisons as being of greatest interest, potential benefit, or risk will be funded and initiated.

The forms to follow are an initial attempt to reduce each goal to shorter-term action steps which could be used to direct activities of the Americleft Task Force over the next 1-2 years.

Section 3: Methodological Considerations

The key to reliable and valid outcome audits lies in careful adherence to methodology. The fatal flaws in typical retrospective research are related to the inability to control the biases inherent in the use of retrospective records. With careful attention to the methodology used, these biases can be minimized. The details of the methods used for specific outcomes described in following sections are covered in the respective sections. Since we are at the beginning of this project our outcomes to date are very limited and based primarily on the Eurocleft project. As Americleft grows, additional outcomes will be included to eventual encompass the full range of disciplines involved in CLP care.

It is widely understood that primary infant management protocols have significant impact on future treatment needs across all disciplines ((lip and palate repairs (type/timing), infant alveolar repair (type/timing) and use/non-use of presurgical orthopedics)) In a survey done as part of the Eurocleft Project, Shaw et al (2000) found 194 different protocols used in 201 centers surveyed clearly demonstrating the lack of any agreement on most-favorable practices. Therefore, the most widely used outcome measures are generally those which are indicative of the effects of those primary procedures.

The key methodological considerations are:

1. For most of the outcomes described herein, power analyses have suggested the need for sample sizes in the 35-40 range. Samples smaller than that can be evaluated but may not represent a reliable and valid assessment of the outcome in question.

2. For most outcomes, samples must be separated according to cleft type. Since UCLP represents the most common cleft type, most outcome comparisons carried out to date are designed to evaluate that cleft group, since achieving adequate sample sizes is easier. However, outcome measures for all cleft groups are being developed and will be included in the project.

3. These patients must be complete, non-syndromic clefts although Simonart's bands are permissible. There should be records available to document and confirm the initial condition (initial entry chart notes, photos, dental casts, etc).

4. One of the most important methodological consideration in studies of this nature is the need to insure and document that the patients in the sample are consecutively enrolled at your center (usually through chart note entry dates, enrollment date, patient number, patient birth date, etc.). Without using a randomized control approach, consecutiveness and samples of sufficient size, are the only ways to have a reasonable assurance that samples being compared were equivalent at the outset of intervention, thereby reducing the chance of selection bias. Also, it is important to keep in mind that while pure consecutiveness is desired, it is understood that there will be gaps created by patients lost to follow-up. The intent is obviously to avoid selection bias (aka "cherry-picking") so that the sample evaluated truly is representative of the primary protocol outcome. Therefore participants in Americleft are expected to be able to provide proof of consecutiveness.

5. Another equally important methodological consideration is the use of measures, whenever possible and feasible, to insure blinding of those doing the outcome

assessment. This is especially important in inter-center comparisons because of the sensitivity of the issues raised by finding protocols which are more favorable or less favorable. It also is critical to guard against analysis bias since it would be normal to expect a team member to want to favor the results from his/her own center.

6. The necessary outcome records, itemized in subsequent sections, must be available on this consecutive series of patients.

7. Availability of primary treatment protocol records and number of operating surgeons involved in the primary surgeries must be available. It would be desirable to have the primary protocol for the center be the one used for the entire sample, although it is understood that some variations of that may have been used by the same or other surgeons at the center for some of the patients. Widely disparate protocols used by numerous surgeons at the same center would lessen our ability to detect differences which could be related to individual protocols. APPENDIX 1 provides the protocol table for the initial Americleft study and illustrates the information required on the sample.

8. Ability to obtain necessary patient permission and IRB approval for use of the clinical records in an outcome assessment. During the study, patient privacy will be protected with absence of any patient identifiers. Identification of the individual centers and surgeons would also be protected. Examples of IRB approval requests are included in each Section.

Section 4: Recommendations for Recording and Reporting Outcomes

One of the greatest obstacles to meaningful inter-center outcome comparisons is the inconsistency and non-comparability of the records that are taken to document those outcomes. This is usually due either to a lack of consistent protocols within the center for record-taking, a lack of oversight by the responsible team member(s) for records to be taken, and in some cases, an attempt to contain the costs of our treatments. Nonetheless, without some agreement between participating Centers on minimal standards for record-taking, inter-center outcomes research is impossible.

With this in mind, there are several key elements to establishing such standards. First, giving due consideration to the costs of records, it is critical that whenever possible, the records taken which could potentially be used for outcome audits, are those which would normally be taken for treatment planning purposes also. A perfect example of this is mixed dentition dental casts, photographs, and radiographs taken by the orthodontists as a routine part of treatment planning for mixed dentition intervention. While necessary for planning treatment, these records can simultaneously be utilized to evaluate dental arch relationship, skeletal morphology and facial esthetic outcomes resulting from prior interventions, if consistently taken and standardized, but without adding costs to the patient or the Center. This also lessens somewhat, the complications of getting approval for records taken for purely research purposes and improves chances of IRB approval for inter-center outcome comparisons and audits by reducing patient risks (eg no extra-ordinary X-ray exposure). Thus outcomes to be evaluated have been designed to able to be assessed using normal clinical records as much as possible.

Second, the minimal standards need to include records that are within the capabilities of the entire range of CLP Centers. If the records required are only available at the largest, best funded Centers, inter-center outcome comparisons only become possible for those select Centers. Records requiring expensive equipment and procedures or highly specialized personnel, are not included in the minimal standards being recommended. Clearly, Centers with greater resources can take additional records to those being recommended at their own choosing, but until such capabilities are available at all Centers, outcomes requiring those records would not be helpful for inter-center comparisons.

Third, once the minimal standards have been accepted by the Center, it is imperative that someone on the team be given responsibility for oversight of the process to insure that the records are, in fact, being taken.

With these in mind, Americleft is in the process of establishing its own standards of record-taking with a list of "minimally required" records and also additional "recommended records". As first step in the process, and also to bring Americleft in line with the Eurocleft project so that outcomes could be compared on a global scale, the following chart of minimal records which was established for European Centers, has also been adopted by Americleft. These will likely be modified in the future with possible additions

Eurocleft Consensus Recommendations on Timing of Minimum Records (X's). Additional Possibilities Indicated by (?)

1. Complete UCLP and BCLP

Timing	Models	Lateral Cephs	Photos	Speech	Audiometry	Patient/Parent Satisfaction
Primary Surgery	X		X			
3 years				X	X	
5/6 Years	X	?	X	X	X	
8-10 years	X	X	X	X	X	
18+ years	X	X	X	X		X

2. Cleft Palate Only

Timing	Models	Lateral Cephs	Photos	Speech	Audiometry	Patient/Parent Satisfaction
Primary Surgery	X		X			
3 years				X	X	
5/6 years	X		?	X	X	
8-10 years	?	?	?	?	?	
15-16 years	X	X	X	X	X	X

3. Cleft Lip Only

Timing	Models	Lateral Cephs	Photos	Speech	Audiometry	Patient/Parent Satisfaction
Primary Surgery	X		X			
3 years						
5/6 Years	X		X			
10 years						
18+ years			X			X

4. Alveolar Bone Grafting

Timing	Intra-oral x-ray	Photos
Just before Bone grafting	X	X
6 months After graft	X	
After canine Fully erupted	X	X

5. Pharyngoplasty

Timing	Speech sample
Just before operation	X
One year After operation	X

6. Orthognathic Surgery

Timing	Lateral ceph	Models
Just before operation	X	X
One year After operation	X	X

Section 5: Roadmap For Integration and Incorporation of Multi-disciplinary Outcomes

Both Eurocleft and Americleft got started by examining orthodontically-related outcomes of primary infant management protocols. The explanation for this lies in the previously mentioned usual availability of mixed dentition orthodontic records, taken for treatment planning purposes. The establishment of simple reliable and valid outcome measures which could be applied to these records then enabled their use for the dual purpose of treatment planning **and** outcome audit. However, the strategy of Americleft is not limited to just orthodontic or orthodontically-related outcomes, and obviously to be successful, must include the entire range of disciplines involved in cleft care. The following is a "roadmap" for incorporating the orthodontically-related outcomes of our Americleft project (circles 2 and 3) into a comprehensive, multidisciplinary outcome assessment (circle 4) with the ultimate goal of identifying those treatment, protocols and strategies that are best able to eliminate the stigma of cleft lip and palate. Strategies for the "roadmaps" to lead other disciplines to circle 4 are currently being developed (See Section 7). Also of importance to note is that because of the large number of treatment variables converging to produce a given outcome (circle 2), inter-center studies of this nature, by definition, are not intended to identify any specific feature of a protocol in a "cause and effect" fashion. As stated in the WHO report:

> **"…for primary cleft surgery, it is difficult, if not impossible to establish the key beneficial or harmful features of a specific treatment due to the invariably complex and arbitrary mix of surgical techniques, timing and sequence, ancillary procedures and surgical personnel"**

Section 6: Orthodontically-Related Outcomes of Primary Infant Protocols

a. <u>**Dental Arch Relationships Outcomes:**</u>

- <u>**Accomplishments and Findings to Date**</u>

The initial meeting of meeting of the "Americleft" project was held February 23-26, 2006 at the Lancaster Cleft Palate Clinic in Lancaster, Pennsylvania. Investigators representing six North American Centers were in attendance. It was here that the Goslon Yardstick (see below) method of scoring dental casts for the evaluation of dental arch relationships was proposed to the cohort of orthodontists. As mentioned in Section 2, initial target for sample sizes from each center was 40 patients with appropriate dental casts for statistical significances to be detected based on our power analysis. The material for this retrospective pilot study involved the pre-treatment diagnostic dental casts on patients who had already received their primary cleft lip and palate surgeries. Thus, consecutive records taken as part of the normal clinical treatment protocol for a phase of orthodontic care routinely done in the 7-10 year old age range collected by each center were used.

At this meeting only 3 of the 6 centers had been able to identify enough patients who meet all inclusion criteria (complete unilateral cleft lip and palate, non-syndromic, Caucasian, age range, no previous active orthodontic treatment, see below) and had appropriate dental models. Nonetheless, the results of the initial meeting were encouraging inasmuch as significant differences in outcomes were found between the two centers with the largest samples. The pilot results confirmed the results of the studies carried out previously in Europe that used the same rating method to identify differences in dental arch relationships that were possibly related to primary surgical outcomes. It was recognized that the data were incomplete due to the limited number of centers with sufficient data, but the potential importance of the results from 2 of the centers led to a presentation at the April 2007 meeting of the American Cleft Palate-Craniofacial Association. Since the protocols for infant treatment were completely different between these two centers, the results were a first step to identify more or less favorable approaches to initial treatment options.

The second meeting of the "Americleft" project was again held at the Lancaster Cleft Palate Clinic in March of 2007 and ratings were again performed on dental study casts representing six centers (the original three centers with adequate samples, plus another one of the original six centers participating in the first meeting who had been able to increase sample size adequately, plus two additional centers with interesting infant protocols but still with sub-optimal sample sizes). The data collected in March 2007 will be used in all future reporting versus the data obtained in 2006 as, with experience, the participants had better inter- and intra-rater reliability as shown by their high Kappa scores.

The results of the six center comparison for the parameter of dental arch relationships demonstrate significant differences between the highest and lowest scoring centers. Goslon average scores for the centers were 2.5, 2.6, 3.2, 3.3, 3.3 and 3.7. This was very similar to the Eurocleft experience that show the worst outcomes for centers employing primary alveolar bone grafting in their surgical protocol that was also associated with a threefold increase in

the likelihood of a patient requiring orthognathic procedures in later years. The prediction that patients might require orthognathic surgery comes from comparison with the results of the Eurocleft study that had an almost exact range and distribution of Goslon scores as the present Americleft study. In the Eurocleft study, a detection of a 0.5 Goslon scale point difference indicated a 20% difference in osteotomy rate (for samples of n=42, 5% probability and 80% power). These results will be presented at the 2008 meeting of the American Cleft Palate- Craniofacial Association.

Finally, a third interim meeting of the "Americleft" project was held at the Peyton Manning Children's Hospital Craniofacial Center in Indianapolis, Indiana in October of 2007. At this meeting, bilateral casts were scored from the original two reporting centers that now had adequate complete bilateral cleft sample sizes using the same criteria as the unilateral study. Similar results were demonstrated with the bilateral sample that received primary alveolar bone grafting having a significantly greater likelihood of requiring orthognathic surgery.

Based on these preliminary results and experiences, the following are the recommendations for future participation by additional centers in dental arch relationship studies. These comparisons represent the easiest way for centers to become involved in inter-center comparisons, since (1) dental study models are normally taken on a routine basis for any orthodontic intervention, especially in the mixed dentition, (2) they represent a non-invasive (low risk) procedure, (3) they allow for easy patient privacy protection, and (4) the Goslon Yardstick and other rating systems have been shown to be reliable, valid and simple outcome assessment methods which are easily mastered through brief training and calibration exercises.

- **Protocol for Dental Arch Relationship Comparisons**

1) *Example of Request Application for IRB approval* – For all aspects of inter-center comparisons, participating Centers must obtain IRB approval. An example of such a request that has been used successfully in Americleft is provided in APPENDIX 2. Please note that it includes a request to waive specific informed consent from the patients. Depending on the sample you may be using (current patients vs. historical records) and depending on the agreeability of your IRB, this may or may not be accepted and especially for more current or even prospectively gathered records, may not be suitable, so specific informed consent might be necessary for the outcome audit. Also, keep in mind that with our ability to carry out 5-year old assessments, depending on your Center's protocol 5-year study models may not fall under the category of those taken routinely for orthodontic treatment planning purposes, and therefore require special approval and informed consent for taking them and also for using them in such an inter-center comparison.

2) *Sample Considerations -* Various aspects of the inclusion criteria have been mentioned previously in Section 2 and in the preceding description of the Americleft accomplishments to date. To summarize, the following are the main inclusion criteria for samples to be satisfy the requirements for inter-center collaborative studies of dental arch relation outcomes in the 7-10 year old mixed dentition patient using the Goslon Yardstick

    ✓ Sample size approximately 40

✓ Complete, non-syndromic unilateral cleft lip and palate with no additional associated facial or dental malformations (expanding outcomes to include BCLP and CPO being developed, requiring stratification on those cleft types also)
✓ Consecutively enrolled (documented by patient number, charts, birth dates, etc)
✓ All primary treatment received at same center
✓ No additional orthodontic treatment between primary management and the date the dental study casts were taken
✓ Availability of total treatment history
✓ Availability of infant presurgical records to confirm complete skeletal clefts (study models, photographs, chart notes, and/or op notes.
✓ Availability of 9-year old dental casts (range 7-12) trimmed in occlusion (matching standard lateral cephalometric radiographs also desired to allow for concurrent evaluation of facial morphology outcomes)

3) *Rating Scales,*

✓ **The Goslon Yardstick.** This rating system for unilateral complete clefts is a valid and well-tested 5 point scale (1=excellent, 5=poor). It was used in the original Eurocleft study (CPCJ, 1992) and has been used extensively since then as many additional European centers collected samples and dental study models for outcome assessments. It is based on clinical features that simplify or complicate treatment and the "burden of care"(Mars M, Plint DA, Houston WJ, et al.: The Goslon Yardstick: a new system of assessing dental arch relationships in children with unilateral clefts of the lip and palate. *Cleft Palate J* 24:314, 1987). It is necessary to have the reference yardstick (the plaster casts) available for comparison with any given cast to be rated when conducting a study. A photographic representation of the discrete Goslon categories can be found in APPENDIX 3, but it is not intended to substitute for the original plaster casts that constitute the yardstick methodology.

All dental casts from all centers need to be prepared identically (see below) and randomized in their order of presentation to insure the records are blinded. The entire set of casts is rated twice by at least 3 experienced, calibrated raters to calculate percentage distribution of cases within each Goslon category and the mean Goslon score for each center. It is possible to add new cohort centers to the "Americleft" arch relationship study using this method with different, yet calibrated raters. However, to maintain a continuous link back to the original Americleft ratings, at least two of the raters will always be from the cohort of original Americleft raters. Inter- and intra- rater reliability testing is done with a weighted Kappa statistic. Means and standard deviations are calculated for each group and tested statistically using t-test (p<.05). Distribution of scores is tested using Chi-square and the Mann-Whitney U/Wilcoxin Rank Sum Test is used to test rank order.

The application of the "yardstick" has 3 determinants that influence the score given to each cast. The greatest influence is from the <u>antero-posterior assessment</u> or overjet. If there are dental compensations present such as proclinations of maxillary incisors or retroclination of mandibular incisors, the score may become the next higher or lower score, depending on the magnitude of the compensation. The second determinant is the <u>vertical assessment</u>. A deep overbite is preferable

to an openbite. Only in a borderline case, can a deep overbite influence the score to the next lower whole number indicating a better score. But, an openbite would likely raise the score to the next higher whole number indicating a poorer score. Finally, the third determinant is the transverse assessment of the arch relationships. Here, the transverse relationships infrequently influence to the Goslon score as this factor is weighted less than the others based on the assumption that many transverse relationships can be treated with orthodontic therapy. Severe narrowing of the arch might alter the score. The influence of the three determinants (antero-posterior, vertical, and transverse) is built into the Goslon Yardstick and this emphasizes the need to use the yardstick models as a reference for any calibrations or ratings that are done.

- ✓ **The Five Year Yardstick.** Given the success of the Goslon Yardstick in identifying more and less favorable dental arch relationship outcomes, a desire to do the same type of evaluation, but on younger patients, lead to the development of the 5-Year Yardstick (Atack NE, Hathorn IS, Semb, G, et al.: A new index for assessing surgical outcomes in unilateral cleft lip and palate subjects aged 5 – reproducibility and reliability. *Cleft Palate Craniofac J* 34:242, 1997). The same basic assessment methods described above, are used for the 5-Year ratings, but the reference dental casts are all primary dentition. This system is intended for use in the late primary dentition. With earlier identification of the protocols leading to the most favorable results, the ability for a Center to understand the key beneficial or harmful features of a protocol allow for adjustments to be made sooner. The reference models for the 5-Year Yardstick are provided in APPENDIX 4.

- ✓ **The Refined Bilateral (Bauru) Yardstick.** A new yardstick for rating dental arch relationships in BCLP patients in the mixed dentition stage has been developed and tested for reliability. There was a need for a different yardstick, as the Goslon was designed for children with UCLP, a different anatomical condition. The only outcome assessment available for BCLP was designed from the Goslon concept *(Ozawa, Soares, Santo, et al., 2005).* The newer generation of this Bauru Yardstick is known as the "Refined" Bauru Yardstick and is based on identical assessment steps as described for both the Goslon and the 5 Year Yardsticks, but with bilateral complete cleft lip and palate reference casts. It is a modification of the Bauru Yardstick to increase reliability. An initial set of reference models is available, but will be expanded.

The Refined Bauru Yardstick, like the Goslon, is a rating system with a valid and well tested 5 point scale (1=Excellent, 5=Poor). Unlike the Goslon, more attention is given to the A-P relationship of the apical bases of the premaxilla and mandible and also to the transverse dimension as a potential "modifier" of the score. A photographic representation of the discrete BCLP categories can be found in Appendix 5, but is not intended to substitute for the original casts that constitute the Refined Bauru Yardstick. Guidelines for scoring are as follows:

- Consider apical base relationship first
- Correct inclination of the incisors mentally (also consider excessive retroclination of lower incisors)
- Ignore crossbite of deciduous and permanent lateral incisors and/or canines

- Ignore edge to edge buccal cusp relationships
- If there is evidence of orthodontics, assume there was a crossbite pre-treatment (e.g. bands, teeth flared bucally or over expanded)

Score 1:
- Class I or Class II apical base relationship
- Positive overjet and overbite (no open bite)
- No crossbite
- Good arch form

Score 2:
- Class I or Class II apical base relationship
- Corrected incisors would be in positive overjet and overbite (or minimal open bite)
- May have crossbites or minor deviation in arch form
- If severe deviation in arch form or severe open bite: Score 3

Score 3:
- Edge to edge apical base relationship
- Corrected incisors would be edge to edge
- May have crossbites or major deviation in arch form

Score 4:
- Class III apical base relationship
- Corrected incisors would not be edge-to edge
- May have crossbites or major deviation in arch form

Score 5:
- Class III apical base relationship
- Corrected incisors would no touch lower incisors
- May have crossbite or poor arch form

All dental casts from all centers need to be prepared identically (see Appendix 6) and randomized in their order of presentation to insure the records are blinded. The entire set of casts is rated twice by at least 3 experienced, calibrated raters to calculate percentage distribution of cases within each Refined Bauru category and the mean score for each center. It is possible to add new cohort centers to the "Americleft" arch relationship study using this method with different, yet calibrated raters. However, to maintain a continuous link back to the original Americleft ratings, at least two of the raters will always be from the cohort of original Americleft raters. Inter- and intra-rater reliability testing is done with a weighted Kappa statistic. Means and standard deviations are calculated for each group and tested statistically using t-test (p<.05). Distribution of scores is tested using Chi-square and the Mann-Whitney U/Wilcoxin Rank Sum Test is used to test rank order.

A pilot study using this new rating system with the yardstick reference casts was carried out on samples from two of the original Americleft centers in October of 2007 and introduced at the 2008 meeting of the American Cleft Palate-Craniofacial Association. The guidelines for the preparation of the dental casts are in the index

and are the same as for the unilateral casts. Bilateral casts were scored from the original two reporting centers (N=35 and 37) which had adequate complete bilateral cleft sample sizes using the same criteria as the unilateral study. Mean intra-rater reliability was 0.935 and mean inter-rater reliability was 0.866, both higher than for the Goslon studies. Similar results (to the unilateral study) were demonstrated with the bilateral sample that received primary alveolar bone grafting having a significantly greater likelihood of requiring orthognathic surgery.

- ✓ **The Eurocran Yardstick.** A final dental model rating system which is a refinement of the original Goslon Yardstick and is based on a four point scale, but also places additional emphasis on maxillary arch form, is also being developed and tested. In Americleft we will continue to use the Goslon and 5 Year Yardsticks for now to enable out outcome studies to be comparable to the many other similar assessments done in Europe to this point.

4) *Dental Model Preparation* – In order to insure blinding of the models during the rating sessions, so that raters would be unable to determine the center(s) from which a given set of models originated, it is essential that the duplicated models all be prepared similarly from all centers, including type of stone used as well as trimming. In this regard, we have elected to follow the guidelines set in the Eurocleft Project. (It is understood that there will be variation in the model preparation from patient to patient and center to center, but significant deviations from the guidelines would require re-preparation of those models or exclusion from the study)

- ✓ Cast in vacuum mixed white stone
- ✓ Trimmed with a fine wheel to the standard heights and angles shown in APPENDIX 6
- ✓ Trimmed with heels parallel so that when models are place on their heels, teeth are in centric occlusion
- ✓ Finished with light sanding, but NOT soaped

5) *Model Rating procedures-*

- ✓ Once samples have been identified, and dental casts duplicated and prepared, currently all dental arch relationship outcomes are being carried out at the Lancaster Cleft Palate Clinic (LCPC), Lancaster, PA. The original Americleft sample is archived there, and samples from new Centers wishing to participate will be mixed in with select samples from the original Americleft study to insure a commonality between the original Americleft results and those from additional Centers joining the project. Dental casts would need to be shipped to LCPC in advance of a rating, so the LCPC's dedicated Data Manager and her team would have time to randomly number the casts, and mix/blind them with casts from other Centers already archived. LCPC Data Management team would be responsible also for randomly reassigning numbers between ratings, and for data entry and analysis. Web-based ratings which would eliminate the need for travel, are being explored but are currently not available.

- ✓ As stated above, ratings are done by at least 3 trained and calibrated raters on two separate occasions at least one day apart. Currently there are at least 8 experienced raters from the original Americleft team who have volunteered to meet as needed at the LCPC for additional ratings. A representative of a new

Center wishing to join is not required, and the dental casts can be rated by just a select group of the original Americleft team. However, it is STRONGLY recommended that a member from a Center sending dental casts to be rated actually travel to LCPC to participate in the process. Not only does that offer the opportunity for new participants to experience the positive benefits of these outcome comparisons with other Americleft members, but it also reduces the chances that findings might be attributed to bias against a newly participating Center if it had no representation on the panel. The training and calibration in the use of the Yardsticks has proven to be a straightforward and simple process taking only approximately one hour.

✓ Statistical analysis for intra- and inter-rater reliability, and tests for statistical significance are described above and will be carried out at the LCPC at the time of the ratings, so that new participating Centers will know the relative ranking of their dental arch relationship outcomes at the time of the study.

b. Cephalometric Outcomes of Skeletal Morphology

1) *Example of Request Application for IRB approval* – For all aspects of inter-center comparisons, participating Centers must obtain IRB approval. An example of such a request that has been used successfully in Americleft is provided in APPENDIX 7 which was an addendum to the IRB request for approval of the dental arch relationship part of the project (APPENDIX 2). Please note that it again includes a request to waive specific informed consent from the patients. Depending on the sample you may be using (current patients vs. historical records) and depending on the agreeability of your IRB, this may or may not be accepted and especially for more current or even prospectively gathered records, may not be suitable and specific informed consent might be necessary for the outcome audit. Also it is important to stress that if your protocol for mixed dentition orthodontic treatment planning includes routine taking of lateral cephalometric radiographs the use of those for outcome audit purposes and the fact that the radiographs can be totally void of any PHI for use in the audits, that you may be able to get approval without additional informed consent as long as parents had already given informed consent when they started orthodontic treatment planning and treatment. This may be especially true if the outcomes from an historical sample at your Center are the ones you will be investigating.

2) *Sample* - **Teams interested in participating should adhere to the following Sample Inclusion / Exclusion criteria**

✓ **Inclusion Criteria:**
   ● Caucasian subjects with a history of non-syndromic complete unilateral cleft lip and palate, (diagnosis confirmed by neonatal photographs, study models, and/or a clearly written preoperative description)
   ● Patients with Simonart's bands will be included, provided no hard tissue union exists
   ● Patients must have lateral cephalograms (with the teeth in occlusion) available at the approximate age of 9 years (range 7-11 years)
   ● Each subject has received all of his/her primary surgery and previous care in the Institution concerned.

- Consecutively treated cases are required

✓ **Exclusion Criteria:**
  - Patients with associated anomalies or syndromes.
  - Patients with incomplete clefts (other than a Simonart's band).
  - Patients who have had any (fixed- or removable-appliance) orthodontic treatment, maxillary expansion, headgear or face-mask therapy prior to taking the cephalometric radiographs
  - Patients that have undergone any orthognathic surgery or osteodistraction treatment prior to taking the cephalometric radiographs

3) *Descriptive Data -* As per the original Eurocleft Study (Shaw *et al.*, 1992), the following descriptive data will be collected for patients included within the investigation (APPENDIX 8):

✓ Date of birth
✓ Sex
✓ Side of cleft
✓ Presence of Simonart's band
✓ How the diagnosis was confirmed
✓ Date of lateral cephalogram
✓ Age at lateral cephalogram
✓ Age and date of alveolar bone grafting procedure, if performed
✓ Whether or not infant orthopedics was performed
✓ Other surgical procedures undergone
✓ Code(s) representing surgeon(s) who performed each procedure (in the case of multiple surgeons at the same Institution)

Additionally, a thorough description of each Center's surgical treatment protocol will be recorded including the technique/type of lip repair, the technique/type of palatal closure, the technique/type of alveolar bone graft (and whether primary or secondary) if performed, and the approximate age of the patient(s) at the time of operation (APPENDIX 1).

4) *Methods -* The investigator performing the assessment will be blind to these data (including the origin of the lateral cephalogram). Different investigators will perform the final evaluation of the numerical data for the synthesis of the discussion. If the records submitted are not in DICOM or JPEG format, a film scanner (Epson model #1680) will be used to convert them into JPEG format for cephalometric analysis. The radiographs from each CLP center will be digitized using Dentofacial Planner version 8.0 cephalometric software (http://www.dentofacial.com/). Sixteen hard tissue and twelve soft tissue landmarks will be used. Each cephalometric landmark will be identified twice, by two independent examiners. A number of hard-tissue and soft-tissue cephalometric variables per radiograph will be calculated (Tables I). The mean of the numerical outcomes per cephalometric measurement will be used for the intercenter cephalometric comparison. The linear distance Ba-N (Basion-Nasion) in mm will be used for size adjustment of all linear measurements. The cephalometric assessment will be performed at the

Division of Orthodontics, The Hospital for Sick Children, Toronto, Canada; and at the Department of Orthodontics, Faculty of Dentistry, University of Toronto, Toronto, Canada.

FIGURE 1: Reference landmarks on the lateral cephalometric tracing
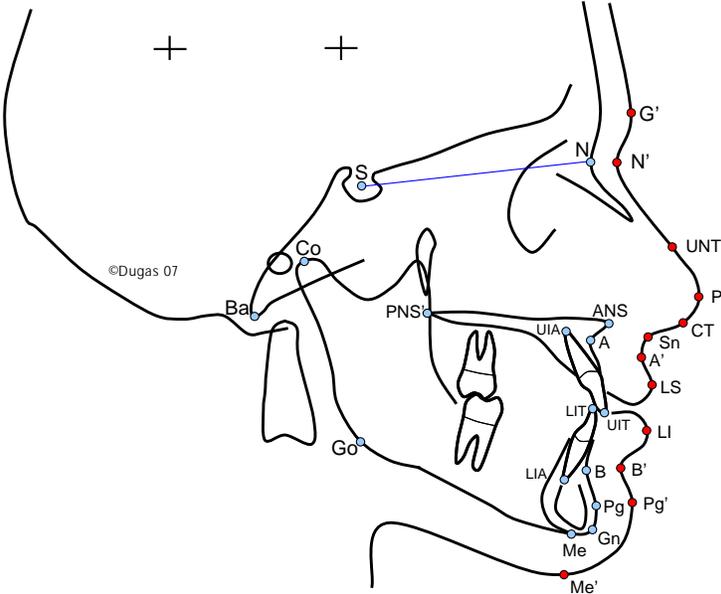
The following Table is the list of measurements that will be taken for comparison between centers. This list is designed to be consistent with the previous Eurocleft cephalometric studies, and also to provide a comprehensive assessment of craniofacial morphology with standard cephalometric measures that have been well established in orthodontics and for which there are norms available for comparison to unaffected controls.

TABLE I

Measurement	Evaluator Initials____
SNA (°)	
SNB (°)	
ANB (°)	
Ba-N-ANS (°)	
Ba-N-Pg (°)	
ANS-N-Pg (°)	
WITS appraisal (A$\perp$OP:B$\perp$OP) (mm)	
Ba-N (mm)	
PNS'-ANS (mm)	

Md length (Co-Gn) (mm)	
SN-MP (SN-GoGn) (˚)	
ANS-Me (mm)	
N-Me (mm)	
ANS-Me/N-Me (%)	
U1-PP (˚)	
L1-MP (L1-GoGn) (˚)	

5) *Statistical Analysis -* Statistical evaluation will be performed with *repeated-measures analysis of variance* comparing the *group means* for the different centers per measurement assessment and checking for an Institution effect, time effect, and time-Institution interaction. Variance terms will be included in the model to account for between-subject variation.

Section 7: Multi-disciplinary Outcomes of Primary Infant Protocols

<u>**Nasolabial Esthetic Ratings**</u>

1) *Example of Request Application for IRB approval*
   Refer to the information described in the Cephalometric Outcomes section. Additionally, it is important to state in the IRB proposal that patient names will be removed from all images and replaced with codes. Also, it should be stated that the eyes will be masked from all images to protect patients' identities not only during transit of records but also for the duration of the study.

2) *Sample considerations*
   Inclusion criteria:
   - Caucassian patients with non-syndromic complete unilateral or bilateral cleft lip and palate,
   - Patients who had complete orthodontic records taken in the mixed dentition, prior to any orthodontic movement of teeth (including maxillary expansion or incisor alignment). Patient may or may not have received primary and/or secondary alveolar bone grafting.
   - Each subject has received all primary surgery and previous care at the Institution concerned.
   - Cases must be consecutively treated patients.
   - Patients who had a complete set of of extra and intraoral photos: full face at rest, left and right profile pictures at rest.

   Exclusion criteria:
   - Non Caucassian patients, patients with associated syndromes, patients with incomplete clefts, cleft lip only, or cleft of secondary palate only.
   - Patients who did not have complete orthodontic records (radiographs, study models, and photos) taken in the mixed dentition.
   - Patients who had orthodontic treatment such as maxillary expansion or incisor alignment or orthopedic maxillary treatment (face mask, head gear, chin cup or functional appliances).
   - Patient who did not have a complete set of extra and intraoral photos: full face at rest, left and right profile pictures at rest. A patient with an incomplete set of quality photos should be excluded.
   - Patients whose photos or images are blurred, excessively dark or bright, or grainy (poor quality image).

3) *Photographic protocol*
   - Images that can be used include Polaroids, slides, photos, and digital images.
   - Photos should be taken at the same appointment but before alginate impressions are taken.
   - Use a single color, well-lit, non-textured background to take the photos. Remove eye-glasses, hats, nose jewelry, and tuck patient's hair behind ears.

- Full face frontal photo should be taken at repose (not smiling), without strain on the lip musculature. Attempt to line interpupillary plane parallel to the floor. Patient's head should be oriented at natural head position. If the camera has a single point flash, it should be oriented at either the right or left side of pt's head.
- Profile photos must be taken from **both** the right and left side of the patient's full face. Lips should be at rest. Head should be oriented at natural head position. Single point flash should be located on the same side as the patient's nose to prevent shadowing on facial outline.
- Each patient must have a complete set of quality photos. If a patient has one image in the set that is not of adequate quality, the patient must be excluded from the study.

4) *Coding of patients and descriptive data*
Each center should collect the descriptive data for each patient as described in the Cephalometric Outcomes section (please refer to previous pages), including the date at which the facial images were taken. Each center should disclose the surgical treatment protocol followed at that Institution. All patient names must be removed from photos and replaced by code numbers. Only codes for each patient (no names) should be used in the images and descriptive data sheets

5) *Scanning of Polaroids, photos and slides*
Scanning of Polaroids, photographic prints, and slides must be done at the centers of origin. Those should be scanned and saved as JPEG images with at least a 1400dpi resolution.

5) *Orienting the images and blocking the eyes*
- This applies both to scanned images and to digital photos.
- Frontal images – use a Photo-software like Adobe Photoshop to tilt the image so that an interpupillary plane is horizontally to the floor. This corrects for canting on the face due to posturing. Following this step, use a small white circle to block or cut out individually both irises of the eyes, while preserving the inner canthi of the eyes visible in the image. The Nasion area (between the eyes) should not be blocked.
- Profile images – use a small white triangle to block out the eye on each image. There is no need to tilt the profile images.
- Remember to save each image with the patient's code followed by the sufix ".jpeg".

6) *Storing of images for shipping*
- Burn all images (coded) into a compact disc and also include the data sheet listing each patient's descriptive data.
- A summary print out of all images in 1-2 pages is useful, but not mandatory.
- Ship the compact disc and any hard paper copies to:
    Dr. Ana M. Mercado
    Ohio State University College of Dentistry, Section of Orthodontics
    305 W. 12th Ave.
    Columbus, OH 43218

7) **Image cropping, subtraction of background, and re-sizing**
- This is done by DrMercado and staff at Ohio State University.
- Adobe Photoshop software is used to crop all images. The only areas to show will be the nasolabial area, innercanthus, nose bridge, nostrils, philtrum and upper lip.
- Any background shown on the profile images will be standardized to the same color.

- All images will be re-sized to scale into the same dimensions.

8) *Preparation of PowerPoint slides for rating*
   - Each PowerPoint slide will contain a patient's frontal and profile image.
   - A number will be assigned to each slide (patient) that is different from the original code (see figure below).
   - All slides will be grouped into a single PowerPoint file, stored into CD's, and distributed to raters for their evaluation.

Case #12

Example of a coded slide for rating. It has frontal and profile images for Case #12.

9) *Panel of raters and their responsibilities*
   - Raters can include but is not limited to orthodontists, plastic surgeons, oral surgeons, speech pathologists, lay persons, and parents of affected children.
   - Raters should agree to participate in a training and calibration session (about one hour long) with the principal investigator, to be described below.
   - Raters will receive one CD with all study subjects' slides and a second "reliability" CD with a smaller selection of subjects' slides randomly ordered and coded differently than those in the first CD.
   - Raters should agree to rate all study subjects' slides and also to re-rate a number of random "reliability" slides to determine intra-rater reliability testing.

10) *Rating scale of nasolabial outcomes*
   - Refer to the methodology described by Asher-McDade C. *et al.*, CPJ 1991, 28(4): 385-90.
   - Four features are rated: nasal form, nasal symmetry, vermilion border, and nasolabial profile.
   - Features are rated on a 1-5 scale:
     1 – Very good (for a patient with a cleft)
     2 – Good
     3 – Fair
     4 – Poor
     5 – Very poor

11) *Training and Calibration*
   - Raters will receive a brief training on the purpose of the study and the types of images that they will be evaluating. A series of PowerPoint slides of patients that are

not actual study subjects will be presented to the raters to familiarize them with the facial cropping and the layout of the images on the slides.

- A calibration session will be done with each rater or with a group of raters by showing them 20 slides and asking them to rate all slides. Ratings for each slide will be reviewed with the principal investigator.

11) *Reference images – Appendix 9*

- Raters will receive a printed color copy of images of different severities on the scale of 1-5, one page for each one of the facial features.
- These printed color copies are meant to be used by the raters as reference images or as a "yardstick" of the scale of severity for each nasolabial feature (developed by Katsaros et al., 2006). Raters may or may not use the yardstick.
- Reference images for each feature are included in Appendix 8.

12) *Recording of ratings*

- Raters will be given blank recording sheets. The following is an example of a table to circle the ratings from a single patient.

CASE # 1		Very Good	Good	Fair	Poor	Very Poor
	Nasal form	1	2	3	4	5
	Nose symmetry	1	2	3	4	5
	Vermillion border	1	2	3	4	5
	Nasolabial profile	1	2	3	4	5

13) *Statistical Analysis*

- The group means of the different centers for each nasolabial feature will be compared using analysis of variance. Weighed kappa statistics will be performed to evaluate inter-examiner and intra-examiner agreement.

Section 8: Bone Graft

Secondary Alveolar Bone Grafting **April 2009**

Secondary alveolar bone grafting (ABG) procedures vary between centres for many variables including the type of surgical procedure, the age of the patient, and the donor site. With the significant variation in treatment protocols and the unproven claims of superiority of certain procedures, there is an identified need for a controlled study to evaluate the outcomes of secondary ABG.

While there are many outcome parameters that are assessed and reported, within the AmeriCleft Study the goal was to have a reliable and reproducible objective outcome analysis that was 'simple' to use and meaningful for the outcome assessments. Factors in the assessment tool that have been considered include: time, a yardstick vs. true measures, and a method that is easy to apply, applicable in the mixed and permanent dentitions, statistically comparable, usable between centres, and appropriate for both retrospective and prospective studies. The possible use of the assessment tool via the internet was also a characteristic of the assessment outcome that was preferable.

Goals of a successful ABG that the method should capture include closure of vestibular and palatal oral-nasal fistulae, presence of bone for dental eruption, skeletal nasal base, adequate bone for the placement of implants, functional airway, reconstruct bony and muscular / soft tissue architecture. The AmeriCleft group determined that the amount and location of bone was to be assessed within the confines of the inter-radicular space from the CEJ to the apex of the adjacent teeth, including presence of incomplete bony bridging both vertically as well as laterally.

There are many different studies that have reported methods to assess the success of secondary ABG (Bergland 1986, Rosenstein 1997, Nightingale 2003, Kindelan 1997, Hynes 2003, Lilja 2000, Long 1995, Withrow 2002). The AmeriCleft group is proceeding to run a pilot test using the Withrow (Chelsea) Scale. It consists of an 8 point scale that accounts for bony bridging both at the apical and cervical aspects of the cleft site, measures both the amount and location of bone, a visual rating scale that appears to be easily implemented both retrospectively and prospectively and could be used as well via the internet for rating by different centres not required to be in the same location, uses periapical or occlusal radiographs that are routinely taken prior to and after ABG, and the reported Kappas are acceptable. Part of the pilot study for ABG will also include panoramic radiographs to assess the recreation of the nasal floor and the ability to rate this from the radiographs. The possible use of this method retrospectively will also depend on the availability of panoramic radiographs.

The patient information required to compare the ABG results would be: primary surgical procedures, age at ABG, previous surgical procedures including failed ABG, successful or failed fistula closure, surgical technique used, occurrence of expansion prior to ABG, and the presence of fistula at the time of surgery. Records to be assessed include periapical or occlusal radiographs 3-6 months pre-ABG surgery and at least 3 and preferably at least 6 months post-ABG.

Section 9: The Americleft Speech Outcomes Project
(Updated May 2014)

This section describes the speech protocol that is used to collect speech data. It involves the collection of high quality audio-video recordings of five- and six-year-old children with cleft palate with or without cleft lip.

Participant Characteristics

Inclusion criteria:
All children who had initial palate repair within a given center should be identified and recalled for the speech assessment and recordings. The entire cohort for the ages covered should be invited. Children adopted from foreign countries are eligible for inclusion, provided adoption occurred by age 30 months and they are speaking English as their predominant language. The number of children who attend out of the total eligible cohort should be listed. Those who are not available for speech recordings should be listed with reasons identified for non-attendance (e.g., failure to attend, moved away, illness, etc.).

Exclusion criteria:
Children with submucous cleft palate, diagnosed syndromic cleft palate with or without cleft lip, known cognitive impairment, documented generalized developmental delay or sensori-neural hearing loss should be excluded. All children should be speaking English as their primary language.

Methodology

Data Collection. The methodology is based on the Scandcleft methodology (Eurocran/ Scandcleft Protocol 2004; Lohmander et al., 2009), work done in the UK (John et al., 2006; Sell et al., 2009) and incorporates aspects of the Universal Parameters for Reporting Speech Outcome in Individuals with Cleft Palate (Henningsson et al. (2008).

Recordings should be made in a quiet room with the subject facing natural light if possible and the face and upper neck only framed in the picture. Ideally, there should be one consistent unbusy background. The stimuli should be presented by the speech-language pathologist (SLP), who should be seated directly facing the child. The camera should be at the child's eye level, so that the child is looking straight at the camera. An aerial view should be avoided. The target sentences should be produced following the SLP's model even if the child is able to read. In this way, the rate at which sentence elicitation takes place is controlled in order to facilitate subsequent transcription and analysis. There should be a stimulus presentation gap of two to three seconds between each sentence presentation. Recording equipment should be checked prior to each data collection session and following each session (using headphones) to make sure that a high quality sample has been recorded. The SLP will document the recording number (ID number) on the Background Information form developed for this project.

Equipment. It is not possible to recommend specific equipment given the continual changes in the market place but it is important to ensure equivalent performance to the descriptions provided below. Recordings are to be made on high quality digital video cameras using an external microphone. The clinician should also use headphones to check that sound has been recorded.

- **Camera.** A center's current equipment may be suitable. Digital video cameras should be used for ease of editing. The camera should have two separate ports, one for a high-quality external microphone and the other for headphones for monitoring of recordings.
- **Microphone.** Ideally, a stand-alone microphone is preferred to prevent extraneous noise from the child moving or fidgeting and to control the mic-to-mouth distance. To achieve this, a stand alone microphone can be placed on a stand 12 inches from the child, at the level of his/her mouth and placed to the side on a floor stand or tripod. If this is not possible then a lapel microphone can be used with attention to the issues described above. Remember to evaluate compatibility of all the components.

The Nature of the Speech Sample. As recommended by Henningsson et al. (2008) the speech sample should include conversational speech, sentence repetition and single word production. The sentence sample is controlled in the sense that sentence stimuli include only the target consonant being assessed and an adequate sampling of high and low vowels. The *American English Sentence Sample* (Trost-Cardamone, 2012), designed to capture cleft palate speech errors, is used in this project. In addition to the sentence repetition task, the speech sample collected includes conversational speech, automatic or rote speech (counting 1-20 and 60-70). The single word portion of *The Goldman-Fristoe 2 Test of Articulation* (GFTA-2) (Goldman & Fristoe, 2000) also is administered to provide a normative measure of consonant articulation in single words.

The order of data collection is as follows:
- Conversational speech: minimum length of 2 minutes
- Counting from 1-20 and 60-70
- Sentence repetition: American English Sentence Sample
- GFTA-2

The conversational speech sample should be 2-3 minutes of the child's speech and focused on questions where the answer is not easily predetermined. Yes or No questions should be avoided. For example, the following elicitation questions may be helpful and will help to standardize the stimulus questions:
- Tell me about your brothers and sisters.
- Tell me about your favorite movie.
- Tell me about your favorite TV program.
- Tell me about your favorite place to go.
- Tell me about what you like to do on your birthday.

o Tell me what you like to do on vacation.

   Background Information. For each child in the study, the SLP should complete the Background Information form, in conjunction with the parent(s) and others involved in the child's care. This form documents the child's first language, surgical history, history of past and current speech therapy intervention, and influencing factors.

Brief History of the Project and Accomplishments to Date

o **April 2009**: Several SLPs from different centers across North America who have collected data as part of this project participated in a half-day training and ratings calibration session (ACPA Phoenix, AZ) using a variety of patient recordings and practicing the Universal Parameters System (UPS) protocol. Subsequently, a decision was made to use the recently validated CAPS-A analysis framework rather than to attempt to both validate the UPS and use it for data analysis simultaneously.

o **2010**: A pilot study (which included 20 samples from three centers in North America) was done by a clinician from the UK, trained in using the CAPS-A analysis framework, to determine its compatibility with American English speech samples. The study consisted of independent ratings of resonance (*hyper*nasality, *hypo*nasality), nasal airflow (audible emission, turbulence), cleft-type speech characteristics, and developmental immaturities. In addition, DS and JTC also provided ratings of hypernasality on these same samples for purposes of inter-rater reliability.

o **February 2011**: SLPs participating in the project met at the University of Minneapolis Dental School for CAPS-A training, discussion and consensus ratings of video-recorded speech samples, and scoring practice, using the GOS.SP.ASS sentences used in the UK and Ireland. Baseline speech ratings on 10 British and Irish (UK/I) samples were rated prior to the meeting and were compared to immediate post-training ratings, and to a final set of perceptual speech ratings conducted one month later (March, 2011) to compute inter- and intra-rater reliability. Training in methodology was provided by Debbie Sell (UK), Triona Sweeney (Ireland), and Anne Harding-Bell, (UK), who are experienced in CAPS-A training and have led efforts in the UK in assessment of speech outcomes in individuals with clefts. Subsequently, a set of 10 North American (NA) speech samples was collected using the American English Sentence Sample. These samples were rated using the CAPS-A -Americleft Modification and inter- and intra-rater reliability measures were obtained.

o **August 2011**: A follow-up meeting was held in Minneapolis in to review and discuss the reliability data and to prepare pertinent submissions for the 2012 ACPA Annual Meeting. The results of both the UK and NA rater reliability studies were presented at the 2012 ACPA.

o **April, 2012:** The speech group met in San Jose, CA following the ACPA Annual meeting. Parameters were reviewed in relation to rater reliability measures obtained and modifications were made with regard to the rating protocol. The timeline and working groups were developed for preparing a manuscript on reliability outcomes. Potential presentations for the 2013 International Congress were discussed.

o **February 2013:** The speech group meeting was based in Houston with some of the group meeting face-2-face and others joining via teleconference. Methodology issues were discussed and clarified where necessary. KC reviewed the status of the NIH-NIDCR grant application and requirements for all SLP participants, including CITI training. Mechanisms for periodic listener calibration sessions were discussed and content for the 3 Americleft presentations at the International Congress was finalized.

o **October 2013:** The speech group met in Minneapolis with one member joining by teleconference. Topics discussed included: review of intra-and inter-reliability results; scoring protocol for the GFTA-2; changes to the global parameters of speech Acceptability and Intelligibility; decision to use Acceptability only, not rate intelligibility; rating recalibration and consensus listening; the plan for manuscript completion. A decision as made to rate hypernasality and acceptability using both EAI scaling and VAS.

o **March 2014:** KC and raters met for two days prior to the ACPA Annual Meeting in Indianapolis. A key purpose was to discuss and become familiar with the REDCAP system for documenting and storing ratings. A timeline for completing primary and secondary ratings was finalized. The raters transcribed several GFTA-2 samples for calibration practice and degree of "narrow" transcription so as not to hold children with clefts to a higher standard than the normed test prescribes. Raters also practiced rating hypernasality and acceptability using both EAI and VAS. The full group met for an hour on Sunday to discuss updates to the Background form and data collection. AB will start collecting data on "typical" 5-6 year olds who will serve as a control group (40/site).

Institutional Review Board (IRB) Documents.
IRB approval is required for data collection as part of the Americleft Project. Likewise, data collectors must complete CITI training. IRBs should be as broad as possible and should allow for analysis by listeners in other centers or listeners other than the SLPs collecting the data. The statement should be made that all data will be made anonymous. There is an exemplar IRB available on request. The University of Utah is the coordinating center for data analysis.

The Americleft Project Workbook is periodically updated and can be accessed as follows: http://www.acpa-cpf.org/research/americleft-study-guide.pdf.

References

Eurocran – ScandcleftProtocol: http://www.eurocran.org/content.asp?contentID=1507

Goldman R, Fristoe, M. (2000). *The Goldman-Fristoe Test of Articulation 2*. San Antonio, TX: Pearson.

Henningsson G, Kuehn DP, Sell D, Sweeney T, Trost-Cardamone, JE & Whitehill, TL. (2008). Universal parameters for reporting speech outcomes in individuals with cleft palate. *The Cleft Palate-Craniofacial Journal,* 45, 1-15.

John A, Sell D, Harding-Bell A, Sweeney T, Williams A. (2006). CAPS-A – A validated and reliable measure for auditing cleft speech. *The Cleft Palate-Craniofacial Journal,* 43, 272-288.

Lohmander A, Willadson E, Persson C, Henningsson G, Bowden B, Hutters B. (2009). Methodology for speech assessment in the Scandcleft project – An international randomized clinical trial on palatal surgery: experiences from a pilot study. *The Cleft Palate–Craniofacial Journal*, 46, 347-362.

Sell D, John A, Harding-Bell A, Sweeney T, Hegarty F. Freeman J. Cleft Audit Protocol for Speech (CAPS-A): A Comprehensive Training Package for Speech Analysis. (2009). *International Journal of Language and Communication Studies.* iFirst Article, 1-20. Downloaded from http://www.informalhealthcare.com. D10.1080/13682820802196815.

Trost-Cardamone, JE. (2007, revised 2012). American English Sentence Sample: A controlled sample for assessing cleft palate speech outcome. Poster presented at the Annual Meeting of the American Cleft Palate-Craniofacial Association, San Jose, CA.

Section 10: Discussion

a. <u>**Logistical Issues and Complications**</u>
  - *Costs, grants, other support* – At this point in time, the original Task Force on Inter-center Collaboration, since renamed as the Americleft Task Force, has received limited but valuable funding from ACPA as part of the budget for its Research Education Committee. In addition, local foundations in Lancaster and Indianapolis have enabled us to match the ACPA seed money. However, all participants have volunteered to pick up as much of their own costs as is necessary, once other sources are depleted. In the case of the dental casts, costs for duplication and shipping of dental casts to Lancaster have been considerable. All Centers have had to supply their own manpower at their respective sites, for sample identification, chart reviews, location of records, etc. Although ACPA support is to continue for this year, it is anticipated that additional outside sources of funding will be required. Unless found, then other expenses such as travel to Lancaster or one of the other primary sites, and lodging may have to be absorbed also.

  - *Sites for comparisons* – As mentioned above, currently it is felt that face to face meetings for the purpose of dental arch relationhip assessments is still far more beneficial and meaningful than other approaches. Lancaster is now serving as the site for archiving those records for future inter-center studies, and therefore is the primary location for carrying out these studies. Indianapolis will also be sharing this responsibility in the future. Toronto is the designated site for cephalometric comparisons, and Ohio State will serve as the site for Nasolabial ratings. Fortunately for these, actual travel to Toronto or OSU for these ratings may not be necessary inasmuch as Lancaster will serve as the coordinating center for these studies, and once cephalometric and photographic data are acquired, they can be forwarded to those sites for analysis. Sites for Speech Outcomes Assessments, Psychological and Surgical Studies have not yet been determined. Finally, it is anticipated that with the anticipated growth of Americleft, that these specific site assignments will likely change or expand.

  - *Most common obstacles/limitations* – The original Americleft team has found the following pitfalls to be the most common, some of which would preclude a Center's participation in these collaborative outcome studies.

    1) Inability to document consecutiveness of a sample with adequate sample size
    2) A record-taking protocol which did not include some of the necessary records for these outcome assessment
    3) Inconsistency of records which should have been taken, excluding some patients normally to be included, from the sample
    4) Incomplete documentation of treatment history
    5) Difficulties with IRB approval, patient consent, etc
    6) Lack of financial support available from the Center
    7) Lack of support from fellow team members for participation in inter-center outcome studies

b. **<u>Future Directions</u>**

- *Additional Outcome Measures of Interest -*

  1) Bone grafting
  2) Psychological
  3) Social
  4) Surgical sequelae (secondary revisions, VPI management, fistulas, etc)

- *Continued ACPA involvement, ownership(?)*

- *Expansion of project -*

  1) **<u>"Lateral"</u>** expansion to add more centers to participate in currently established outcome measures
  2) **<u>"Depth"</u>** expansion to add outcome measures and studies from other disciplines and development of similar converging roadmaps
  3) **<u>"Vertical"</u>** expansion to integrate Americleft with Eurocleft/Eurocran and develop strategies for international RCT's to include North American Centers

APPENDIX 1: Protocol Table

Treatment	Center A	Center B	Center C	Center D	Center E	Center F
Pre-surgical orthopedics	No	Yes	No	No	Yes	Yes
Lip repair	6 wks Millard or 6 mos Delaire	2-3 mos Millard	3 mos Tennison	3 mos Variable	7 wks Lip Adhesion; 7 mos Millard	3-4 mos Millard
Primary bone grafting	No	Yes 6-9 mos	No	No	No	No
Hard palate repair	9-12 mos Bardach or Delaire	11-15 mos Hard palate Wardill-Kilner	12 mos Vomer flap	12 mos Vomer flap	14 mos V-Y pushback	? mos Vomer flap/ Von Langenbeck
Soft palate repair	9-12 mos Bardach Or 6 mos Delaire	11-15- mos Furlow (1 surgeon) or IVV	18 mos Median suture with IVP	12 mos Von Langenback with IVP		? Veau pushback
Secondary bone grafting	6 yrs Delaire	8-9 yrs If needed	9 yrs	7-10 yrs	9 yrs	9-11 yrs
Surgeons	2	4	1	1	1	4
Sample Size	18	40	38	38	18	35
Avg Age	9:4	8:6	9:0	9:1	9:0	9:2

APPENDIX 2: Sample of IRB Application for Dental Arch Relationship Audit

Title: An Inter-center Comparison of Treatment Outcomes in Unilateral Cleft Lip and Palate.

Investigators:

Site:

Background: Desirable outcomes in the treatment of patients with cleft lip and palate can be measured in a number of different areas important to successful rehabilitation of the patient. These include intelligible speech, normalized facial esthetics, normal hearing and favorable facial and dental growth and development. In landmark inter-center comparative studies in Europe, called the "Eurocleft Project", all of these outcomes have been shown to be significantly related to the initial surgical protocol used for repair of the cleft in the infant, as well as patient volumes treated by the primary surgeons. The initial identification of primary protocols which produce favorable vs unfavorable outcomes has been started, although a recent survey of 201 European cleft palate centers revealed a total of 196 different primary protocols being used! Recently a similar initiative has been started in the North America, called "Americleft". This provides opportunity for additional attempts, through inter-center outcome comparisons, to examine outcomes from centers using different protocols, and especially to involve the North American centers in this international initiative.

 Of all the outcomes, the one that has the greatest impact is the subsequent development of the bones of the face, jaws and dental arches. Coincidentally, this outcome is also one which is most easily quantified and rated using non-invasive clinical records routinely gathered on patients for orthodontic diagnosis and treatment planning procedures. Plaster dental study casts, made from standard dental impressions are routinely taken in the 7-10 year age range for patients with cleft lip and palate in almost all centers. These are taken for the purpose of diagnosis and planning coordinated surgical and orthodontic treatment for bone grafting at this age and are considered as a valuable indicator of future treatment needs. As part of the routine necessary inclusion of these dental casts for treatment, no informed consent in addition to that obtained routinely for evaluation, diagnosis and treatment, is normally obtained for the dental study casts. A dental model rating system developed in England has been show to be a robust, valid and reliable method of differentiating between favorable vs unfavorable outcomes to that point in a patient's life, using future treatment needs as the index. Since the results observed and rated represent the outcomes of the particular primary surgical protocol used, the rating of these models becomes a method of quantifying favorable vs unfavorable infant management procedures.

Previous investigations have established well-defined inclusion criteria for such dental model rating studies, in order to reduce and eliminate sources of bias. These criteria include: (1) verification of initial condition being complete, non-syndromic unilateral cleft lip and palate; (2) verification of consecutively enrolled patients; (3) availability of dental study casts at the appropriate age; (4) confirmation of primary surgical procedures used, and numbers of primary surgeons involved; (5) verification of no other surgical or orthodontic treatment other than the primary procedures, up to the time of the dental casts; (6) verification of patient age at time of dental casts.

The _____Clinic has been a regional leader in the field of craniofacial anomalies and cleft lip and palate for the past _____ years. The availability of facial growth and treatment records on patients at this Center provides an opportunity for us to participate in the Americleft project. Intercenter collaborative outcome studies such as this have also recently been endorsed and supported by the World Health Organization. The possibility of expanding the number of participating teams and initiating similar collaborative efforts in North America has also been endorsed and supported by the American Cleft Palate-Craniofacial Association. As a result, to date, ___ centers have been identified which have patient samples which meet the inclusion criteria and are of sufficient size to allow for statistically valid comparisons of outcomes. We are seeking approval to become involved in this important project.

Research Design: A retrospective review of patient records is carried out to identify patients meeting the inclusion criteria. These records reviews are carried out by the professional staff members of this Center. The initial sample lists include only patient name and date of birth as PHI's in order for the sample selection to allow for determination of inclusion/exclusion, and once included, to determine surgeon of record, surgical protocol used, and age at the time dental models were taken. No dental model records are taken which are not already available as part of the patients' normal diagnostic and treatment procedures. Once dental models are identified for inclusion, duplicate models are made which remove all PHI from the dental casts themselves. The only identifiers used on the casts are a number assignment for the center from which they came, and a randomly assigned patient number. At this point, a data manager is assigned from professional staff, who is not part of the investigative team. The data manager is then responsible for generating a sample list which consists solely of the patients' randomly assigned numbers on the dental casts, the patient age at the time the dental casts were taken, a number corresponding to the surgeon of record, and a description of the surgical procedures used. At this point, all PHI becomes permanently de-linked from the dental casts and inaccessible to any of the investigators.

The numbered dental casts from all this Center and all others who will be participating in this outcome comparison are duplicated identically to blind the investigators/raters from the source of the records. The entire sample for all Centers is then randomly renumbered removing all indication of the center of origin.

The actual rating is carried out after a preliminary review of the rating system with the rating team, a training period and a calibration trial rating. All investigators for the collaborating centers are following the same procedures and have been informed as to the methods used to insure no risk to patients regarding record taking or PHI disclosure. Two separate ratings of the entire sample are carried out. The data gathered are given to the data manager who is then responsible for entering the data and statistically analyzing the results. Access to the computer used by the data manager is secure and password protected, even though the final outcomes assessment data will still contain no PHI. In addition, no photo or other reproduction of any of the patients' dental casts will ever appear in any presentation or publication which may result from this outcome assessment. Only group data are presented.

Once ratings and data analysis are complete, all duplicated dental study casts will be archived with no PHI attached. The data manager and the will be responsible for insuring adherence to the methods and procedures described above.

Finally, it is proposed that due to the nature of this outcomes assessment that the informed consent and HIPAA requirements be waived, for several reasons. First, due to the nature of this retrospective study, the majority of patients whose dental casts are included in the sample, have long since completed their treatment and have been dismissed from their respective centers,

thereby most likely making new contact with these patients both intrusive and possibly inconsistent. The likelihood that some patients meeting the inclusion criteria would be excluded from the study simply due to inability to contact them, would significantly reduce sample sizes to levels that would be statistically invalid. Second, the records being analyzed pose absolutely no risk to the patient since they have already been taken. Third, all dental casts were taken as part of the routine diagnostic and treatment planning procedures for all centers and considered as covered under the normal informed consent signed by all patient/parents. Since the purpose of the dental casts was to evaluate treatment needs, the rating system used in this study is simply a method to quantify those needs and allow for statistical analysis of group results. Therefore, with the assumption that treatment needs are directly and inversely related to outcomes, the only use of the dental casts in addition to that for which they were taken, is the statistical analysis of those treatment needs or outcomes, in the context of the primary surgical protocols used. Fourth, in the execution of the outcomes assessment, once the sample is identified, all PHI will be permanently de-linked from the dental casts used in the study, so there would also be absolutely no risk to patients of unintended PHI disclosure. Last, no copy, photo or other reproduction of any of the dental casts used would appear in presentation or publication, even though dental casts per se are not considered PHI.

Significance: The significance of this study lies in its potential value in the quest for information which would allow us to determine those primary infant management protocols which produce the most desirable outcomes. Since most centers providing care for patients with clefts use their own specific approaches, and since the number of different approaches used is overwhelming, and since it has been shown in Europe that not all approaches produce desirable results, it is incumbent on those of us treating these patients that we be willing to compare and scrutinize our results and outcomes methodically and scientifically in order to be able to make evidence-based decisions about treatment choices we offer to patients. Thus collaboration between centers is essential. Such collaboration as started in Europe in the 1990's has led to a rapid growth in our knowledge base as well as having laid the groundwork for more sophisticated investigations such as randomized control clinical trials and standards for recording and reporting outcomes, which offer even better chances to identify "good practices". "Americleft", the first of its kind in North America, has the potential to stimulate similar progress in the US and Canada, and thereby add substantially to the growing body of knowledge necessary to improve care for patients with cleft lip and palate.

APPENDIX 3: The Goslon Yardstick

Goslon Yardstick 1

Goslon Yardstick 2

3 W
3 W

Goslon Yardstick

3

4 D
4 D

Goslon Yardstick

4

Goslon Yardstick

5

APPENDIX 4: The Five-Year Yardstick

5-Year Yardstick
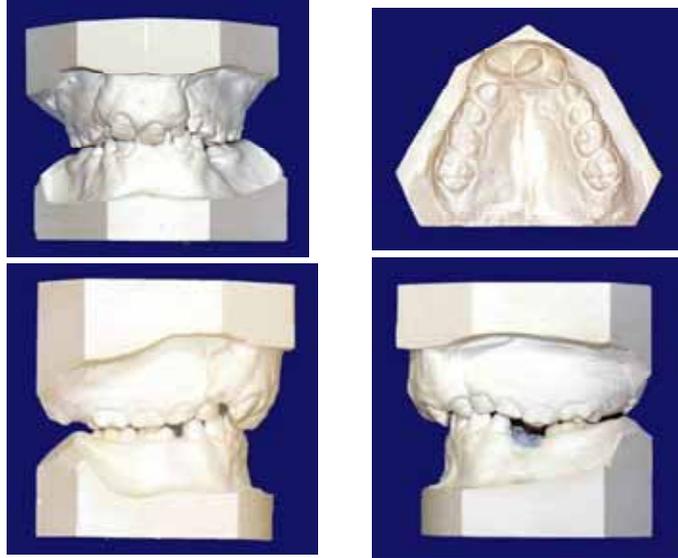# 1

5-Year Yardstick
2
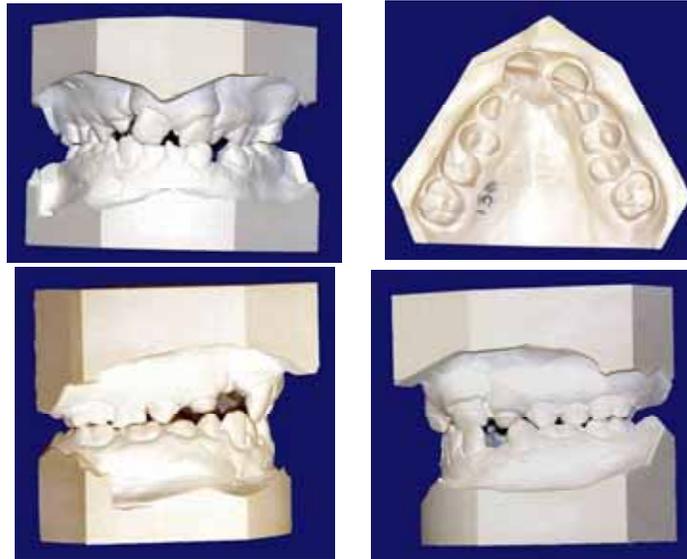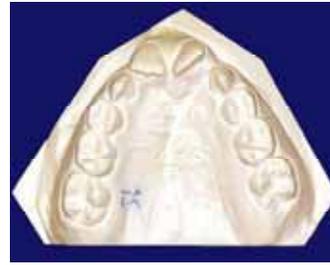
5-Year Yardstick
3

5-Year Yardstick
4

5-Year Yardstick
5

BCLP Yardstick 1

BCLP Yardstick 2

BCLP Yardstick 3
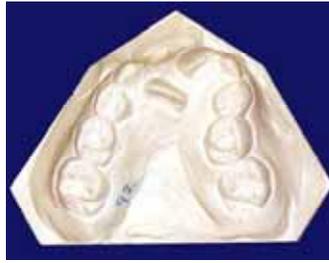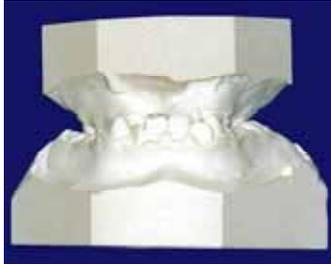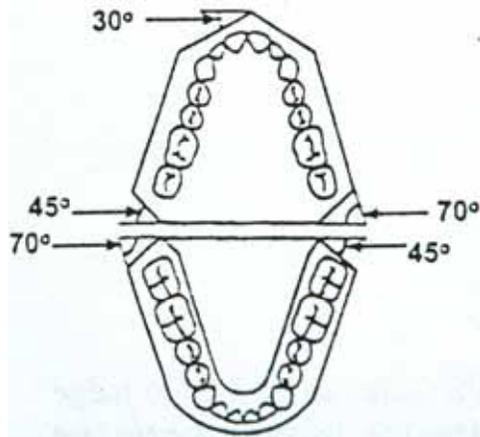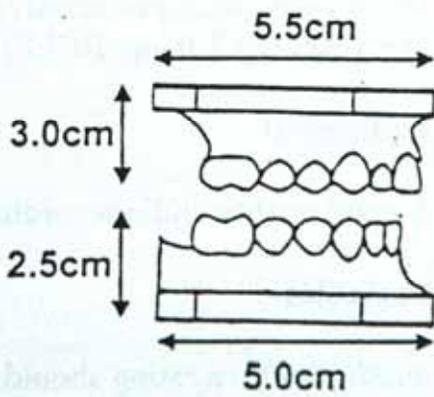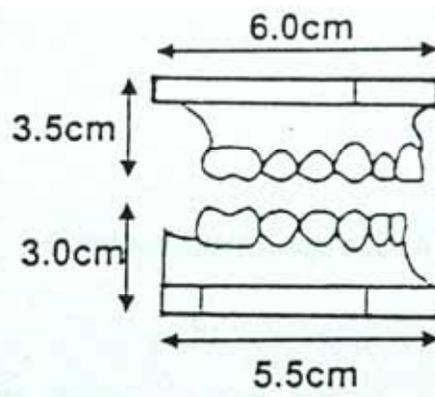

BCLP Yardstick 4

BCLP Yardstick 5

APPENDIX 6: Dental Cast Preparation

Dental casts' base angles

**Dental casts dimensions
For 5-year Yardstick**

**Dental casts dimensions
for Goslon Yardstick**

APPENDIX 7: Sample of IRB approval request for Lateral Cephalometric Study

Title: An Inter-center Comparison of Treatment Outcomes in Unilateral Cleft Lip and Palate.

Investigators:

Site:

Americleft Project Continuation Plans: The initial target for sample sizes is 40 patients determined by prior statistical power analysis to be necessary to identify statistically significant differences. "Enrollment" was not actual active recruitment of current patients, but the identification of historical treatment records (dental study casts) on patients who had already received their primary cleft lip and palate surgeries and had records taken as part of normal clinical treatment protocol to plan for a phase of orthodontic care routinely done in the 7-9 year old patient. As indicated, the intent of the study is not to prospectively monitor orthodontic treatment results, but to use the orthodontic treatment planning records (dental study casts and lateral cephalometric radiographs) to rate retrospectively the success of their primary surgeries in the context of dental occlusion and jaw relationship.

Since dental study casts only evaluate one aspect of outcomes resulting from various primary surgical protocols, the evidence collection needs to be expanded. Since dental study cast ratings can only <u>suggest</u> underlying jaw growth dysplasias, additional supportive evidence must be included. Fortunately, many of the same patients who have had dental study casts taken for clinical purposes of diagnosis and treatment planning, also routinely have had lateral cephalometric radiographs taken (a routine x-ray in orthodontics). Therefore, the current application calls for inclusion of measurements on those radiographs on those same patients taken at the same time as the dental study casts. Although radiography is considered more invasive than dental study casts, several points need to be emphasized: (1) the records have already been taken for clinical treatment purposes, (2) patients have already given consent for them to be taken as part of their consent for the treatment, (3) the measurements made on them are the same as those done for the diagnostic and treatment purposes, (4) all radiographs will be duplicated with all PHI being removed on the duplicates and only randomly assigned patient numbers remaining, and (5) the data derived will only be aggregate measurement data on the total samples for each center. For these reasons, it is proposed that since patients have already given their informed consent for the records to be taken, and the since outcome assessment and audit is a routine part of the treatment process, additional informed consent and HIPAA requirements for reseach purposes do not apply and can be wiaved.

The duplicates of the original radiographs will be made by scanning the radiographs and putting into digital form. On the digital copy, all PHI will be removed and the copies assigned new random numbers which will be generated by the Data Manager and kept on a password protected computer accessible only by the Data Manager. Radiographic images will be measured on each radiograph, and statistics on the group means, and standard deviations carried out. Aggregate group results from each center will be compared. No PHI or individual patient information or results would be available to the investigators. No copy, photo or other reproduction of any of the radiographs used would appear in any presentation or publication. Only aggregate data would be used.

Significance: Centers which have collaborated in inter-center outcome comparisons, both in the original Eurocleft study and the current Americleft Project, have been enthusiastic about this approach in contributing to evidence-based care, with use of retrospective treatment records rather that through a randomized control trial. However, due to the complex and multidisciplinary nature of cleft lip and palate treatment, outcomes more than just dental arch relationships need to be considered. The expansion of the study to include cephalometric radiographs adds a measure of facial skeletal morphology to the analysis which is not available through dental model ratings, and it does so without any additional risk to patients regarding the taking of the records or PHI disclosures. Through the gradual accumulation of outcome data from various additional assessment methods and the generation of a desire to improve treatments through collaborative efforts, this Americleft project has the potential for significant contributions to the field. The Project has been endorsed and supported by the American Cleft Palate-Craniofacial Association.

APPENDIX 8: Sample Descriptive Data Sheet For Lateral Ceph Study

CENTER CODE_____   TABLE IV: SUBJECT DATA

SHEET

Patient Code	DOB	Sex	Side of Cleft	S-Band	How was Diagnosis Confirmed	Date of Lateral Cephalogram	Age at Lateral Cephalogram	Alveolar Bone Graft		Early Orthopedics	
								Date	Age	Yes	No