Predictors of Velopharyngeal Insufficiency After Surgical Advancement of the Maxilla among Young Adults with Cleft Palate

Sara Kinter

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2021

Reading Committee:

Noel Weiss, Chair

Kathy Chapman

Srinivas Susarla

Program Authorized to Offer Degree:

Epidemiology

University of Washington

Abstract

Predictors of Velopharyngeal Insufficiency After Surgical Advancement of the Maxilla among Young Adults with Cleft Palate

Sara Kinter

Chair of the Supervisory Committee:

Noel Weiss, MD, DrPH

Epidemiology

Introduction: Maxillary hypoplasia refers to undergrowth of the upper jaw and results in malocclusion of varying severity. While the prevalence of severe maxillary hypoplasia in the general population of the United States is 0.3%, the prevalence among patients with a cleft palate with or without cleft lip (CP/L) is more than 40%. The timeline of care for children with CP/L requires frequent follow-up and multiple surgeries throughout their childhood. When maxillary hypoplasia is severe, surgical maxillary advancement at the conclusion of craniofacial growth, typically between 17-21 years of age, is required to achieve functional occlusion. However, anterior movement of the maxilla also changes the relationship between the soft palate and the posterior pharyngeal wall, introducing risk of velopharyngeal insufficiency (VPI), a speech disorder that interferes with speech acceptability and if significant, speech intelligibility.

At present, there are two types of surgery commonly done to treat maxillary hypoplasia in children with a cleft palate: distraction osteogenesis (DO) and traditional maxillary

advancement (TMA). The results of prior studies that compare the prevalence of VPI in children after undergoing each type of surgery suggest a lower prevalence in patients who underwent DO, but the difference was small and was based on a relatively small number of patients.

The goal of this dissertation was: to assess in a larger sample the relation between type of surgery for maxillary hypoplasia and the presence of post-operative VPI; and to investigate different anatomic and speech characteristics in these patients that are associated with risk of developing VPI.

Methods: We analyzed data from a cohort study that utilized data from two internal clinical databases as well as the electronic medical record. The first analysis aimed to compare the risk of post-operative VPI between two different surgical approaches to maxillary advancement: distraction osteogenesis and traditional maxillary advancement. The second evaluated type of cleft, magnitude of maxillary advancement, borderline pre-operative VPI, and pre-operative velopharyngeal ratio as predictors of developing VPI after maxillary advancement. Identification of potential confounders was accomplished through a combination of *a priori* clinical knowledge as well as a data-drive approach.

Results: The prevalence of post-operative VPI among those who underwent distraction osteogenesis was 40.5%. After adjusting for type of cleft and magnitude of overjet, this was 47 percent (95% CI for the PR = 0.72-3.05) higher than the corresponding prevalence of VPI in the TMA group. A modest excess of post-operative VPI was also observed in the DO group when those with a unilateral cleft lip and palate were analyzed separately (adjusted PR=1.44, 95% CI 0.56 – 3.67). Neither history of Furlow palatoplasty nor history of any VPI surgery appeared to bear on the size of the association between maxillary surgery type and risk of post-operative VPI. The increased prevalence of VPI among those undergoing distraction osteogenesis was largely confined to patients with a pre-operative VP ratio of less than 0.8.

The prevalence of VPI after either type of maxillary advancement surgery was not associated with the magnitude of maxillary deficiency (adjusted PR 1.01 per 1 mm increase in

the distance between A-point and the y-axis, 95% CI 0.97 – 1.03). Conditional on syndrome status, those with borderline pre-operative VPI were more likely to present with post-operative VPI than those with normal pre-operative velopharyngeal function (PR=1.57, 95% CI 0.93 – 2.67). There was a suggestion that patients with bilateral and unilateral CLP were more likely to have post-operative VPI compared to those with a cleft palate only (PR 1.59, 95% CI 0.64 – 3.96; PR 1.67, 95% CI 0.66 – 4.27, respectively).

Conclusions: Our results suggest modestly poorer velopharyngeal function among patients who underwent DO compared to those who underwent TMA, after accounting for differences in type of cleft and magnitude of overjet. However, our results are not in accord with earlier findings. To date, however, all analyses on this issue are limited by statistical imprecision and additional contributions to the body of evidence are warranted.

Table of Contents

List of Figures.. ix

List of Tables..... xi

Chapter 1: Background and Aims

1.1 Timeline of Care for Children with Cleft palate +/- Cleft Lip..... 1

1.2 Anatomy and Physiology of Velopharyngeal Function..... 3

1.3 Sources of Increased Surgical Risk in Patients with CP/L..... 5

1.3.1 Skeletal Considerations..... 7

1.3.2 Occlusal Considerations..... 9

1.3.3 Vascular Considerations..... 10

1.3.4 Soft Tissue Considerations..... 11

1.3.5 Cranial Nerve Considerations..... 12

1.3.6 Velopharyngeal Function..... 13

1.4 Specific Aims..... 14

Chapter 2: Approaches to Supporting Internal Validity of the Study

2.1 Outcome Measurement Error: Assessment & Approach Rationale..... 16

2.1.1 Methods of Reliability of Measurement Assessment Within & Across Evaluators..... 16

2.1.2 Validity & Reliability Results..... 19

2.1.3 Discussion & Conclusion..... 21

2.2 Rationale for Complete-Case Analysis..... 24

2.3 Confounder Selection Rationale..... 30

Chapter 3: Maxillary Advancement & Velopharyngeal Function: A Comparison of Surgical Approaches

3.1 Introduction..... 36

3.2 Methods..... 37
3.2.1 Participants..... 37
3.2.2 Surgical Characteristics..... 38
3.2.3 Lateral Cephalometric Measurement..... 38
3.2.4 Perceptual Evaluation of Velopharyngeal Function..... 41
3.2.5 Statistical Analysis..... 42
3.3 Results..... 44
3.3.1 Reliability Results..... 46
3.3.2 Modified Poisson Regression with Binary Outcome..... 47
3.3.3 Multiple Linear Regression with Categorical Outcome..... 51
3.3.4 Sensitivity Analyses..... 53
3.4 Discussion & Conclusion..... 54
Chapter 4: Speech & Anatomic Predictors of Deterioration in Velopharyngeal Function after Maxillary Advancement
4.1 Introduction..... 58
4.1.1 Magnitude of Maxillary Advancement..... 58
4.1.2 Morphology of the Velopharyngeal Port at Rest..... 59
4.1.3 Pre-Operative VPI Status..... 61
4.1.4 Type of Cleft..... 61
4.2 Methods..... 63
4.2.1 Participants..... 63
4.2.2 Surgical Characteristics..... 64
4.2.3 Lateral Cephalometric Measurement..... 64
4.2.4 Perceptual Evaluation of Velopharyngeal Function..... 66
4.2.5 Statistical Analysis..... 67

4.3 Results....69
4.3.1 Reliability Results....71
4.3.2 Evaluation of Predictors....72
4.4 Discussion & Conclusion....74
References....81
Appendix....91

List of Figures

Chapter 1

Figure 1.1	Timeline of care for children with cleft palate +/- cleft lip	2
Figure 1.2	A three-quarter view of normal velopharyngeal anatomy	4
Figure 1.3	Normal and cleft palate levator veli palatini insertion	5
Figure 1.4	Pre- versus post-op change on lateral cephalometric radiographs	5
Figure 1.5	Pterygomaxillary disjunction	6
Figure 1.6	Skeletal characteristics prior to maxillary advancement in a patient with CP/L.	8
Figure 1.7	Vascularity of the maxilla	10
Figure 1.8	Innervation of the maxilla	12

Chapter 2

Figure 2.1	Potential predictors of outcome missingness	25
Figure 2.2	Associations between relevant covariates, missingness, and outcome	27
Figure 2.3	Scatterplot of age at surgery and VPI severity stratified by surgery type	28
Figure 2.4	Associations between missingness, exposure, outcome, and other covariates as identified within the data	29
Figure 2.5	DAG illustrating exposure, outcome, and covariate relationships as informed by clinical knowledge	30
Figure 2.6	DAG depicting proposed relationships leading to race as a potential confounder between exposure → outcome association	33
Figure 2.7	DAG illustrating paths between exposure and outcome blocked by adjustment for magnitude of advancement and cleft type	34

Chapter 3

Figure 3.1 Cephalometric landmarks, linear, and angular measures 40

Chapter 4

Figure 4.1 Cephalometric landmarks, linear, and angular measures 65

List of Tables

Chapter 2

Table 2.1	Severity definitions for perceptual speech ratings	17
Table 2.2	Subject-specific differences in intra-rater hypernasality by SLP	19
Table 2.3	Subject-specific differences in intra-rater VPI severity by SLP	20
Table 2.4	Subject-specific differences in intra-rater hyponasality by SLP	20
Table 2.5	Mean severity of each speech measure stratified by SLP	21
Table 2.6	Inter-rater reliability	21
Table 2.7	Intra-rater reliability	21
Table 2.8	Comparison of SCH & CAPS-A-AM severity rating definitions	24
Table 2.9	Proportion of missingness by covariate	26
Table 2.10	Correlation between age at surgery & post-op hypernasality stratified by surgery type	28
Table 2.11	Study population characteristics identified as potential confounders, stratified by surgery type	31

Chapter 3

Table 3.1	Pre-operative patient characteristics stratified by type of surgery	44
Table 3.2	Crude and adjusted relative risk of VPI/hypernasality comparing DO and TMA	48
Table 3.3	Association between VPI/hypernasality and surgery type, within strata of prior Furlow, prior VPI surgery, and pre-operative VP ratio	50
Table 3.4	Comparison of mean post-operative VPI/hypernasality severity rating between DO & TMA	52
Table 3.5	Observed versus predicted outcome prevalence based on assumption of 10% of unobserved patients having hypernasality	53

Chapter 4

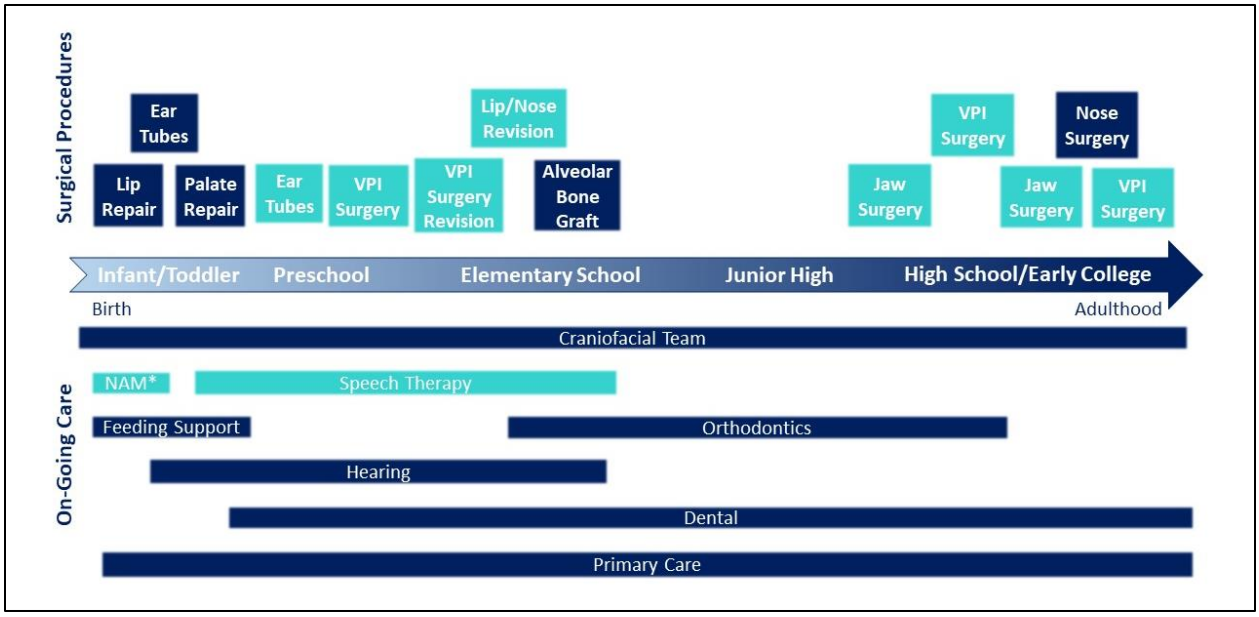
Table 4.1	Participant characteristics, stratified by presence versus absence of outcome	69
Table 4.2	Crude and adjusted VPI prevalence ratios for each predictor of interest	72
Table 4.3	Prevalence ratios comparing proportion of those with post-operative VPI by type of cleft, with secondary cleft palate as the referent category	73
Table 4.4	Proportion of participants lost to follow-up, stratified by higher versus lower risk groups	75

Chapter 1: Background and Aims

1.1 Timeline of Care in Patients with Cleft Palate +/- Cleft Lip

Care for children with cleft palate with or without cleft lip (CP/L) requires frequent follow-up and multiple surgeries throughout the course of their childhood. Figure 1.1 illustrates surgical and on-going care needs typical for children with CP/L. Items in light blue become part of the care timeline if needed, suggesting improvement of outcomes from earlier surgeries have the potential to reduce the frequency with which subsequent procedures and appointments are necessary for children with CP/L. A large body of literature suggests that quality of life (QoL) in children with CP/L is influenced by multiple factors, including appearance, speech, and surgical burden. Some studies report an inverse correlation between QoL and burden of care in multiple domains (physical health, psychological, social relationships, and environment), with QoL scores decreasing as burden of care increases [1]. Other studies report that responses on QoL measures can differ based on cleft type, and that young children with CP/L score lower on measures of self-esteem than their unaffected peers, suggesting that functional as well as aesthetic differences are predictors of a patient's perceived QoL [2,3]. Literature evaluating psychosocial functioning in children with surgical congenital conditions has suggested that both multiple hospitalizations and traumatic treatment procedures may predict with long-term psychosocial functioning [4].

Figure 1.1: Timeline of care for children with cleft palate +/- cleft lip[+].



[+] Dark blue boxes occur in most patients with CP/L; light blue boxes occur if needed.
[*] NAM = nasoalveolar molding

By adolescence, 30-50% of patients with CP/L develop maxillary hypoplasia requiring surgical advancement of the maxilla through LeFort I osteotomy [5-8]. Maxillary hypoplasia is often considered aesthetically undesirable and can interfere with tongue placement for speech sound production, impair dental occlusion needed for adequate chewing, and inhibit nasal airflow for breathing. Recovery from this surgery is prolonged and occurs at a time during late adolescence when patients are looking ahead to high school graduation and beyond. Maxillary advancement also carries risk, including introduction of hypernasality and nasal air emission during speech sound production, a condition known as velopharyngeal insufficiency (VPI). Between 10-40% of these adolescents have already had one or more surgeries to correct this condition [9]. For others, they have never experienced the social challenges of not being understood. QoL in patients with CP/L and VPI is often rated lower when compared to patients with CP/L who do not have VPI [10-12]. VPI generally requires surgical management to correct, and VPI-specific QoL measures indicate an improvement in QoL following VPI surgery [12].

Maxillary advancement is typically considered the "last" procedure in a long list of surgeries

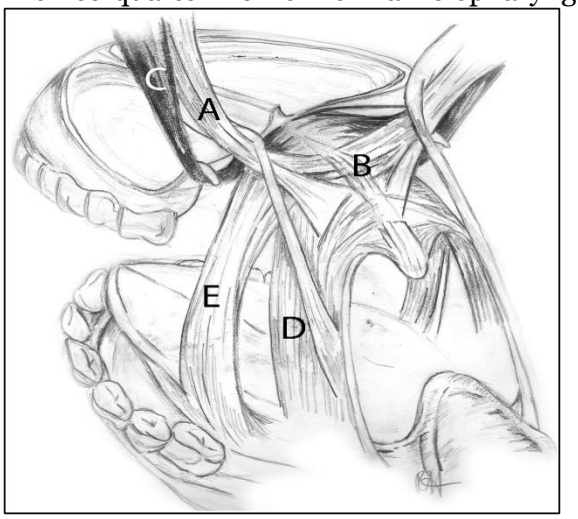
for patients with CP/L. Deterioration in intelligible speech and the need for additional surgery in young adults is not inconsequential. Therefore, surgeons have developed different surgical techniques for maxillary advancement aimed at minimizing risk of VPI without interfering with aesthetic and occlusal outcome. Traditional maxillary advancement surgery (TMA) entails LeFort I osteotomy to allow anterior movement through separation of the maxilla from adjoining bony structures. The maxilla is then fixed in an anterior position using permanent internal hardware to maintain the new position. Distraction osteogenesis (DO) is a newer technique developed in response to the surgical complexity present in patients with CP/L. DO is more likely to be chosen for patients requiring a large magnitude of advancement, as a reduced likelihood of relapse using this technique has been reported [13]. DO involves LeFort I osteotomy and therefore surgical manipulation similar to TMA, but hardware is placed that applies increasing anterior force over time [14]. These distractors are then removed after a consolidation phase of several weeks. It has been suggested that risk of VPI after DO is lower than after TMA due to the increased period of time allowed to achieve optimal advancement [15-17]. Despite wide use of DO, only a single small RCT has directly compared speech outcomes between DO and TMA[15,18]. Strengthening the level of evidence for VPI risk after DO and TMA will not only enhance the validity of pre-operative counseling, but will further delineate how providers, patients and families weigh risk of VPI when balancing surgical risks versus benefits.

1.2 Anatomy & Physiology of Velopharyngeal Function

The velopharyngeal port is comprised of the soft palate (or velum), lateral pharyngeal walls and posterior pharyngeal wall. Superior and posterior elevation of the soft palate, medial movement of the lateral pharyngeal walls, and anterior motion of the posterior pharyngeal wall occur to create complete closure of the velopharyngeal port and separation of the oral from the nasal cavity, a movement necessary for normal pressure consonant production and balanced oral-nasal resonance. The soft palate is the structure most involved in velopharyngeal closure

and the levator veli palatini is the primary muscle responsible for superior and posterior elevation (labeled "A" in Figure 1.2) [19].

Figure 1.2: A three-quarter view of normal velopharyngeal anatomy [19].



A = levator veli palatini; B = musculus uvulae;
C = tensor veli palataini; D = palatopharyngeus;
E = palatoglossus
Image from: Kotlarek & Perry, 2018, Perspectives of the ASHA Special Interest Groups.

The levator is a paired muscle with the right and left bundle arising from the temporal bone. Muscle fibers then course inferiorly, anteriorly, and medially and finally change degree of inclination and enter the soft palate at the junction of the middle and posterior thirds to meet its paired partner at midline [20]. This path results in what is known as the "levator sling" (Figures 1.2 and 1.3). Contraction of the levator sling results in superior and posterior elevation of the soft palate to allow contact with the posterior pharyngeal wall necessary for speech and swallowing. In an unrepaired cleft palate, the levator inserts into the posterior nasal spine anteriorly rather than medially within the soft palate (highlighted in red in Figure 1.3). Sometimes the levator sling is successfully recreated with surgery and other times it remains in an atypical position that is not optimal for speech in many people. A primary goal of initial palate repair is to restore anatomical midline insertion of the levator to optimize VP function. When VPI persists after initial palate repair, there are several additional surgical options for management.

Figure 1.3: Normal and cleft palate levator veli palatini insertion [21].

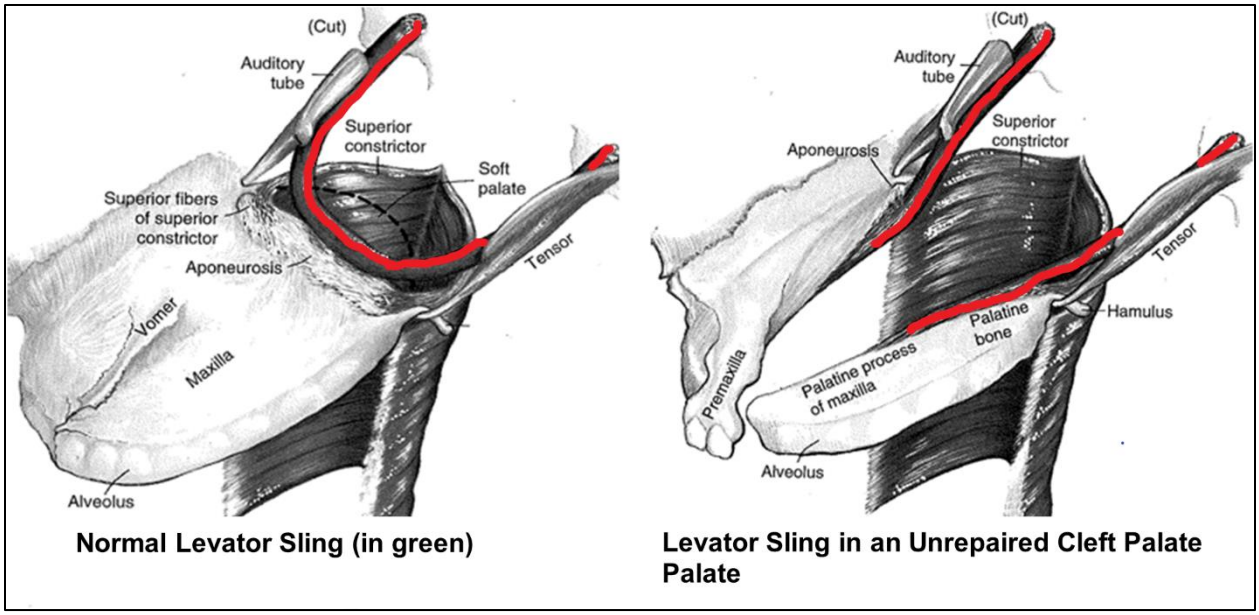
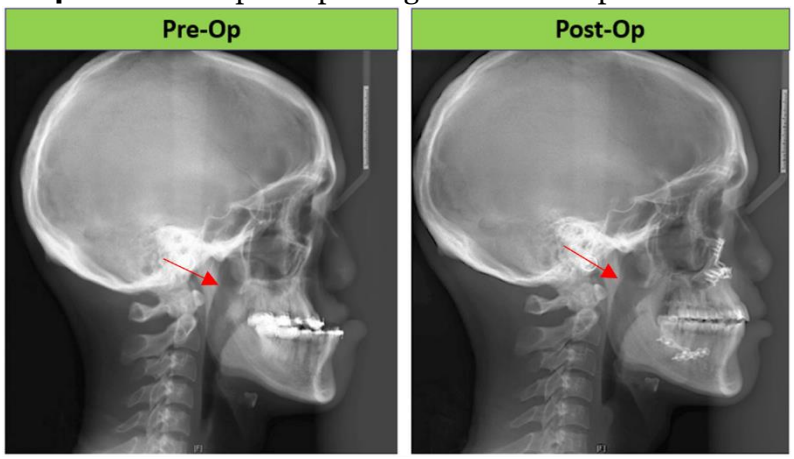


Image from Peterson-Falzone et al, 2006. *Clinician's guide to treating cleft palate speech*.

1.3 Sources of Increased Surgical Risk in Patients with CP/L

Surgical movement of the maxilla requires an orthognathic procedure designed to move the upper jaw (maxilla) forward, known as LeFort I osteotomy (LFI). LFI is a common procedure with minimal risk among patients with maxillary hypoplasia that do not have CP/L. Figure 1.4 illustrates occlusal changes with maxillary surgery. Note it is also apparent that the distance from the posterior pharyngeal wall to the posterior edge of the maxilla has increased (red arrows), resulting in increased size of the VP port.

Figure 1.4: Pre- versus post-op change on lateral cephalometric radiographs.



"Osteotomy" simply refers to the surgical cutting or removing of bone. As described in the literature, TMA involves an incision on the inside surface of the upper lip, where the lip meets bone (upper buccal sulcus), to gain access to the underlying maxilla. Following this incision, osteotomies of the lateral maxillary buttress, anterior wall of the maxilla and the lateral nasal wall are performed [22,23]. The primary goal of these osteotomies is complete separation of the maxilla from the pterygoid process ("pterygomaxillary disjunction", as seen in Figure 1.5).

Figure 1.5: Pterygomaxillary disjunction

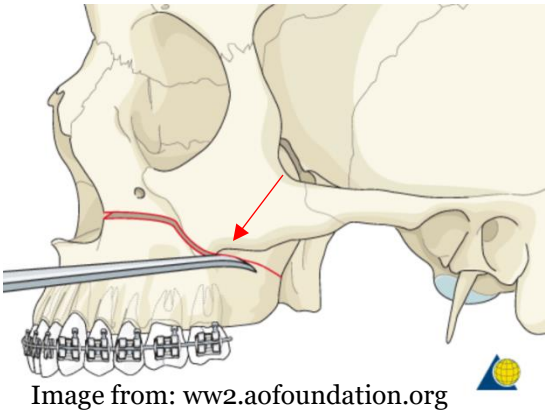


Image from: ww2.aofoundation.org

Once osteotomies are complete, "down-fracture" of the maxilla is then initiated by applying firm manual pressure to the alveolar arch to complete pterygomaxillary disjunction to loosen any remaining soft tissue attachments between the maxilla and pterygoid process. Complete passivity of maxillary segments is necessary to minimize oppositional force and to prevent immediate relapse once the maxilla has been placed into its new anterior position [24]. Alves et al summarize LFI as being based on the two primary principles described above: (1) complete pterygomaxillary disjunction and (2) down-fracture of the maxilla [25]. Complete disjunction and successful down-fracture maximize maxillary mobility for repositioning and minimize common risks associated with this procedure. Despite the invasive nature, this procedure is considered reliable and safe among individuals without a history of CP/L. The probability of primary risks associated with the procedure, relapse, and surgical site infection, are less than 10% and nearly absent risk for changes in velopharyngeal function [13].

LFI for maxillary advancement is now a common procedure used to manage maxillary hypoplasia in patients with CP/L with as many as 50% of those with unilateral CLP (UCLP) and 65% of those with bilateral CLP (BCLP) benefitting from maxillary advancement [24]. However, there are several anatomical considerations among adolescents with CP/L that make traditional LFI a more technically challenging and less reliable procedure compared to adolescents without CP/L. Underlying skeletal differences, scarring of soft tissue, absent incisors within the cleft defect, residual oronasal fistula, atypical vascularity of the maxilla, increased severity of maxillary hypoplasia and differences in velopharyngeal anatomy all contribute to increased risk of complications among adolescents with CP/L undergoing LFI [13,26,27].

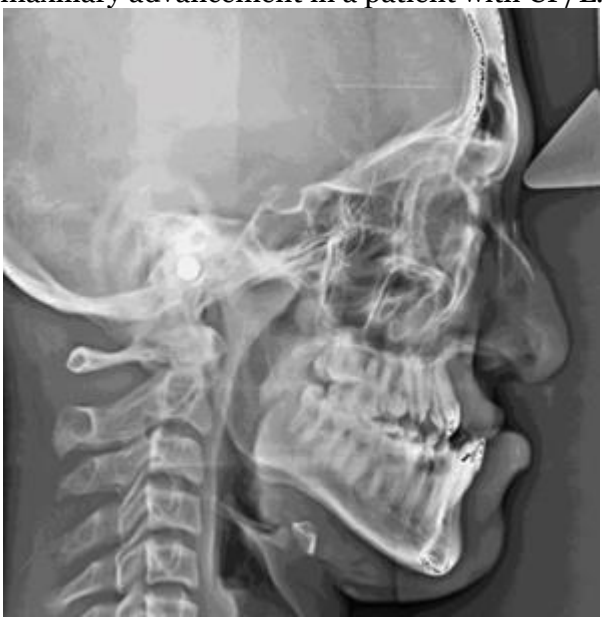
Anatomical disparities inherent in individuals with CP/L have led to modifications in surgical technique for patients with CP/L undergoing maxillary advancement surgery [23]. The most significant of these is development of DO, the technique currently preferred if a large degree of anterior movement is required [14,16,28]. For DO, LFI is completed similarly to that described above, but rather than rigid fixation, hardware is applied that allows progressive advancement of the maxilla over a period of several days, followed by a consolidation phase [29]. Modifications to the LFI osteotomy procedure itself have included strategic placement of soft tissue incisions that allow for increased access and visibility of the underlying bony maxilla as well as maintaining vascularity. Over-correction in the horizontal and transverse dimensions is also recommended to compensate for the increased risk and degree of relapse in patients with CP/L. Structural variation in patients with CP/L undergoing maxillary advancement influencing these changes to surgical technique and resulting in increased complications are described in more detail below.

1.3.1 Skeletal Considerations

A primary challenge of maxillary advancement among adolescents with CP/L is the presence of a maxilla that is deficient in all three dimensions: vertical, horizontal, and

transverse[22]. Figure 1.6 illustrates these skeletal characteristics in a patient with CP/L.

Figure 1.6: Skeletal characteristics prior to maxillary advancement in a patient with CP/L.



Specific characteristics and severity of these deficiencies vary by type of cleft. In patients with unilateral cleft lip and palate (UCLP), one side of the maxilla is often hypoplastic and displaced superiorly, posteriorly, and medially necessitating these changes to surgical technique and resulting in a maxillary midline deviation towards the side of the cleft [22,27].

With bilateral cleft lip and palate (BCLP), there is typically a very narrow maxilla with bilateral posterior crossbite due to the posterior alveolar segments being collapsed medially. Additionally, the premaxilla can be protrusive and either superiorly-inferiorly or palatally-superiorly displaced. Vertical maxillary deficiency (height of the maxilla) is difficult to address due to the degree of deficiency, reduced amount of regional soft tissue and need for revascularization of grafted bone. Prior successful alveolar bone grafting (recommended to be completed around ten years of age in children with CP/L) can mitigate these deficiencies. For patients with adequate alveolar continuity across the cleft from prior alveolar bone grafting, the maxilla can generally be advanced in a single piece. If there is deficient bone within the cleft,

segmentation of the maxilla may be needed (two pieces for UCLP or three segments for BCLP) to achieve optimal post-operative results [27].

Skeletal challenges described above have been indicated as a factor in increased risk of relapse of anterior maxillary movement achieved with LFI in patients with CP/L. Relapse risk among individuals with CP/L is 20-25% after one year, more than twice that of patients without CP/L who underwent the same procedure [22]. Other studies have shown that this relapse occurs in both the horizontal and vertical dimensions [30]. The amount of anterior movement needed (horizontal dimension) has traditionally been identified as a predictor of post-operative relapse[31-33]. However, not all studies agree on this point. Other studies have identified additional risk factors for relapse that may not be related to the degree of anterior movement required [23,26]. DO has been identified as a way to mitigate relapse risk when a large degree (over 10 mm) of movement is required, although the literature is overall equivocal on whether DO truly results in decreased relapse risk compared to traditional LFI in patients with CP/L, primarily due to a dearth of high quality evidence comparing both techniques between comparable populations [29].

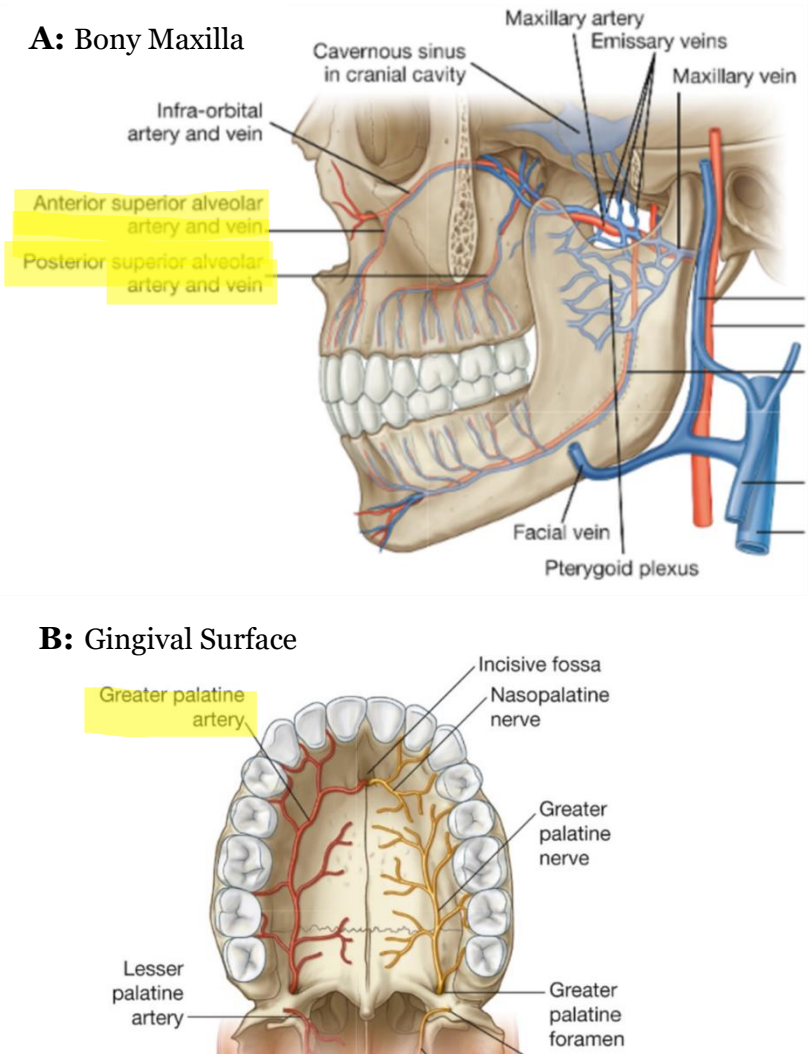
1.3.2 Occlusal Considerations

Completion of presurgical orthodontics is paramount in maximizing success of LFI among patients with CP/L and compensates for three common issues: (1) agenesis of lateral incisor(s), (2) residual alveolar clefts and/or residual oronasal fistula and (3) a multi-dimensional deficient maxilla [22]. Optimal alignment of teeth prior to maxillary advancement can improve stability of the post-operative outcome as well as prevent intra-operative complications. For example, if segmentation of the maxilla is anticipated, adequate positioning of the upper incisors will prevent their interference with optimal alignment of the dental arches during surgery [22]. Dental extractions prior to orthognathic surgery may also be required to reduce the amount of skeletal manipulation needed for functional occlusion [24].

1.3.3 Vascular Considerations

Normal vascularity of the maxillary region involves three arteries: the greater palatine, the anterior superior alveolar, and the posterior superior alveolar arteries (see labels highlighted in yellow in Figure 1.7) [34]. The greater palatine artery is considered the primary blood supply to the hard palate (maxilla) and supplies palatal gingiva. The posterior superior alveolar artery provides blood supply to the molar and premolar teeth in the maxillary arch. The anterior superior alveolar artery supplies the maxillary incisors and canines.

Figure 1.7: Vascularity of the maxilla



In patients with CP/L, several surgeries have already been completed to address the cleft prior to maxillary advancement. Details of these earlier surgeries are often unavailable and vascular supply to different regions of the maxilla are less predictable than in patients who have never had maxillary surgery. There has also been suggestion that the greater palatine artery in individuals with CP/L is smaller than those unaffected by CP/L, although it was unclear whether the observed difference was related to underlying congenital differences or as a consequence of scarring and tissue interference from multiple prior surgeries [35]. Additionally, because the posterior-superior alveolar vessels are necessarily transected with the initial incisions and during down-fracture required for successful LFI, preservation of the greater palatine artery is crucial in ensuring adequate blood supply to the maxilla post-operatively [36]. Injury to the greater palatine artery can significantly decrease vascular supply to the maxillary structures, introducing increased risk of post-operative necrosis. These discrepancies are likely even more pronounced in patients with CP/L, particularly in those with BCLP. Some authors suggest preservation of the midline mucosa over the premaxilla in the initial incision to preserve adequate blood supply post-operatively [23].

1.3.4 Soft Tissue Considerations

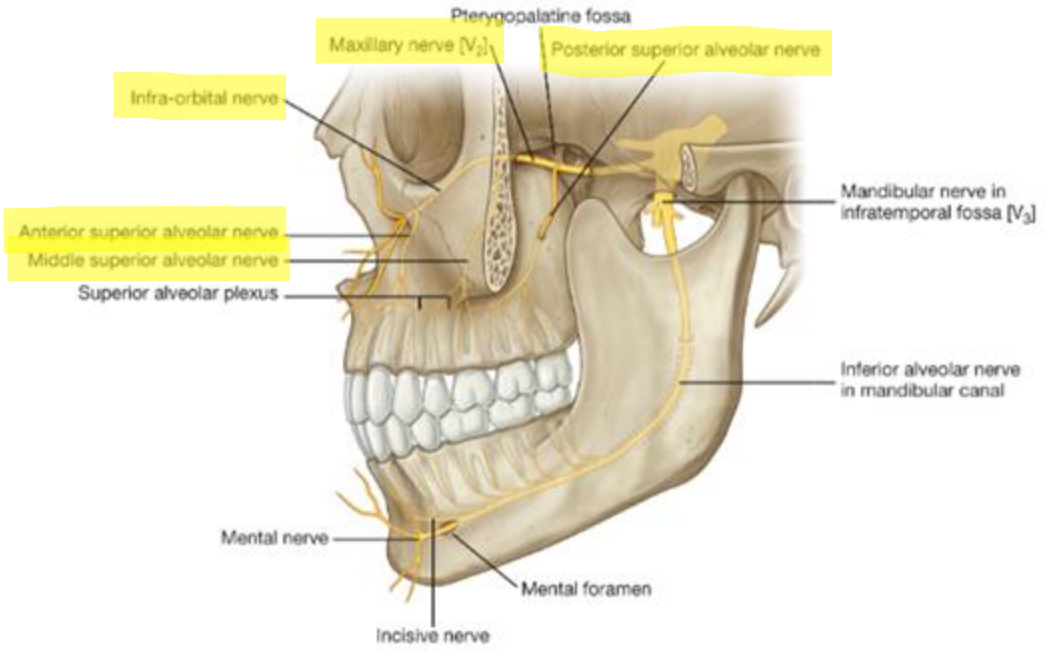
Roy et al report three soft tissue findings that can increase the risk of relapse post-operatively [22]. These include retraction due to scar tissue, upper lip tightness and inadequate mobilization of the bony segment. All three of these issues can result in forces in opposition to the newly anteriorly displaced maxilla that can lead to loss of some (or all) of the anterior movement gained during surgery. The need to consider the integrity of the soft tissue within the surgical site is related to the presence of scar formation and contracture created following prior surgeries to the hard palate. Reduced soft tissue malleability creates additional challenges during the down-fracture phase of the surgery that can lead to post-operative complications. As previously noted, passivity of maxillary movement is paramount for surgical success. During the

down-fracture phase, however, significant scarring of the soft tissue restricts maxillary mobility even after osteotomy is complete and can make complete pterygomaxillary disjunction very challenging [24]. A prolonged period of traction during the down-fracture phase may be required to achieve complete passivity of maxillary mobilization, sometimes for up to 20-30 minutes in patients needing large advancements (> 10mm) [22]. Increased force and stretching required for traction may in turn lead to additional soft tissue injury.

1.3.5 Cranial Nerve Considerations

Normal innervation of the maxilla and mandible is depicted in Figure 1.9 [34]. Alves et al

Figure 1.8: Innervation of the maxilla



completed a systematic review of the literature on cranial nerve injuries among all patients undergoing LFI and found that cranial nerve injuries can occur, but most often resolve on their own within a few months after surgery [13,25,37]. Prevalence of cranial nerve injuries is higher among patients with CP/L, however. Because the maxillary branch of the trigeminal nerve provides sensory innervation to the maxilla and its branches are directly in the path of surgical manipulation required for LFI, the most common cranial nerve complication post-operatively is reduced sensation in the roof of the mouth, upper lip, and cheek (maxillary hyposthesia). As can be seen in the figure above, the maxillary nerve becomes the infra-orbital nerve then further branches into the anterior and middle superior alveolar nerves, all of which are at risk of injury during LFI due to their location [25]. Injuries to these nerves are most often secondary to excessive retraction, compression and stretching of the nerve during mobilization of the maxilla as well as due to the mucosal incision needed for surgical approach [13,38]. Alves et al also found a handful of single case reports involving injuries to additional cranial nerves [25]. Although very rare, there is potentially significant morbidity associated with these injuries (i.e. optic nerve injury can lead to permanent blindness compared to those described above for the maxillary branch of the trigeminal nerve). Injuries to these other nerves are most often related to a fracture that occurs during separation of the maxilla from the pterygoid process that then extends either towards the orbit or skull base.

1.3.6 Velopharyngeal Function

Because of anatomical variations in the musculature necessary for velopharyngeal closure among patients with CP/L, there is increased risk of VPI in patients with CP/L when compared to adolescents without CP/L undergoing the same procedure. Indication of a significant difference in risk between the two groups is extrapolated from the absence of literature reporting VPI among non-CP/L patients undergoing maxillary advancement compared to the many articles going back 40 years reporting changes in velopharyngeal

function after maxillary advancement among patients with CP/L. Despite the known anatomical relationship between the bony maxilla and velopharyngeal musculature, very limited literature is available regarding specific changes to the velopharyngeal musculature. Chen et al investigated the angular changes within the levator veli palatini after DO using 3-dimensional computed tomography (3D CT) volume-rendered images [39]. In a small series of patients, they found that a deterioration in velopharyngeal function was associated with a reduction in the angle of soft palate.

Depending on the integrity of the underlying musculature as well as the relationship between the length of the soft palate and the distance to the posterior pharyngeal wall, it is physiologically possible that complete closure of the soft palate with the posterior pharyngeal wall will not be possible following maxillary advancement. While many papers have been published investigating changes in velopharyngeal function after maxillary advancement, there is very little evidence for predictors of these changes. An increase in depth of the velopharyngeal port has been reported following LFI, although the correlation between these changes and presence of VPI was not measured [40]. Reports comparing risk of VPI in DO versus TMA have been contradictory with one reporting no difference between the groups and another reporting "improved outcomes" among individuals undergoing DO [15,16,41]. A recent paper reported significant differences in soft palate mobility pre- versus post-maxillary advancement for both DO and TMA that were positively correlated with a decline in velopharyngeal function [17].

Among individuals undergoing LFI, anatomical differences between individuals with CP/L and those without are significant and lead to increased risk of negative post-operative sequelae. There are many anatomic details that may contribute to different morbidities after maxillary advancement. Understanding these details is vital to successfully designing studies that will illuminate changes needed to curtail these morbidities and bolster information available for adequately informed consent.

1.4 Specific Aims

Since January 2007, patients with CLP undergoing DO or TMA through Seattle Children's Craniofacial Center (N=236) have routinely had pre- and post-operative lateral cephalograms and speech measures collected and stored in a database. This study seeks to: (1) examine whether DO is associated with less deterioration in velopharyngeal function compared to TMA; and (2) assess whether pre-operative cephalometric measures, history of VPI surgery and/or perceptual speech findings predict presence of VPI post-operatively. Findings of this study have the potential to contribute to strengthening the existing body of literature's level of evidence as well as to lay a stronger foundation for design of future studies.

Chapter 2: Approaches to Supporting Internal Validity of the Study

Threats to internal validity include measurement error, selection bias, and confounding. While these threats can be mitigated to some degree with appropriate statistical analyses, decisions made during the study design phase have the greatest potential to influence internal validity.

2.1 Outcome Measurement Error: Assessment & Approach Rationale

Outcome measures used in analysis include severity ratings of hypernasality, VPI, and hyponasality made by speech-language pathologists (SLPs) at the time of clinical evaluation. Over the study period, a total of six SLPs were involved in assigning ratings that were later stored in the clinical database. One of these SLPs contributed ratings over the full study period and two others contribute ratings for more than half the study period. The remaining three SLPs each contributed ratings for two to three years of the study period, although their patient volumes are low. These values, stored in a clinical database, are used for analysis of the exposure-outcome association in this study. Likert scales are used to rate each outcome. Hypernasality and VPI severity each have five levels and hyponasality has three levels of severity. Because quantification of level cut-offs for ordered variables can differ across observers, the degree to which misclassification of outcome is present must be assessed [42].

2.1.1 Methods of Reliability of Measurement Assessment Within & Across Evaluators

Fifty audio recordings of patients obtained during pre- or post-operative speech evaluation for patients of appropriate age for maxillary advancement were used. Recordings were obtained using a Sony Digital ICD Recorder and included syllable and sentence repetition as well as counting 1-20 and 60-70. The speech sample used is available in Appendix A. Each audio sample was edited to eliminate all identifiable information. Spontaneous speech samples were not consistently available, and therefore were not included. Three SLPs specializing in the

evaluation of velopharyngeal function and with experience using these rating scales as defined in Table 2.1 rated each of the 50 recording for severity of hypernasality, velopharyngeal insufficiency (VPI) and hyponasality. Severity definitions of hyper- and hyponasality were adapted from the Cleft Audit Protocol-Augmented with Americleft Modifications (CAPS-A-AM), the only severity measurement protocol for speech characteristics specific to cleft palate rigorously evaluated for reliability and validity [43,44]. VPI severity is a measure that has been used for outcomes assessment at our institution for over 20 years, and accounts for both hypernasality secondary to a structural defect as well as the listener's perception of audible nasal air emission. Table 2.1 reflects definitions used for each level of severity. Each SLP then re-rated 15 randomly selected audios from the original 50 two weeks after initial ratings. Weighted kappa [45] and two-way mixed consistency, single-measures intra-class correlation coefficients[46] (ICC) determined intra- and inter-rater reliability, respectively. Interpretation of reliability coefficients as defined by Cicchetti et al were used [47]. All analyses were completed in R 4.0.3.

Table 2.1: Severity definitions for perceptual speech ratings.

Characteristic	***Definition***
VPI Severity	Degree to which HN & obligatory nasalization of pressure consonants (PCs) & nasal air emission (NAE) noticeable in speech
0 = None	No audible NAE and balanced resonance
1 = Minimal	Infrequent audible NAE in absence of HN OR assimilation nasality in absence of audible NAE
2 = Mild	Occasional audible NAE on 1-2 PCs and HN of high vowels /i, ɪ/
3 = Moderate	Audible NAE on 3-5 PCs and HN of high and mid-vowels /i/ and /u, o/; speech intelligibility begins to be impacted
4 = Severe	NAE audible on most/all PCs and HN on all vowels and voiced PCs; speech intelligibility clearly impacted by VPI
Hypernasality (HN)	Any abnormal increase in nasal resonance of voice during speech production; most easily perceived on vowels and voiced consonants
0 = None	Nasal resonance WNL for region
1 = Borderline	Some perceptible increase in nasal resonance
2 = Mild	Evident on closed vowels /i, ɪ, u/
3 = Moderate	Evident on both closed and open vowels /ae, a, etc./
4 = Severe	Evident on vowels and voiced consonants
Hyponasality	Abnormal reduction or absence in nasal resonance during speech; most easily perceived on nasal consonants
0 = None	Nasality WNL
1 = Mild	Nasal consonants partially denasalized
2 = Marked	Nasal consonants /m, n, ŋ/frequently denasalized

Use of weighted kappa for ordered categorical variables, as is used for intra-rater reliability measures in this study, are relatively standard when comparing two measurements [45]. Because three raters were used for inter-rater reliability assessment, a different calculation of the reliability coefficient was needed. The variant of ICC used in this study was selected through consideration of four factors [48,49]:

1. *Model of agreement:* In a one-way model, raters are randomly selected from a population of raters for each subject. Design of reliability assessment for this study, however, involved the same raters rating all 50 audio files (a "fully-crossed" design). Therefore, a two-way model was selected.
2. *Type of agreement:* "Good inter-rater reliability" for perceptual speech outcomes is defined by strong agreement when assigning rank order of severity. Reliability assessment was designed to determine degree of correlation in outcome measurement across raters, thus consistency rather than absolute agreement was selected.
3. *Unit of agreement:* Single-measure ICCs are used when a subset of subjects is rated by multiple raters. Although all 50 audio recordings were rated by all three SLPs, it was not the average of these ratings for each subject used in analysis, as is the case when reporting average-measure ICCs. In the interest of generalization of reliability results to the study population, the more conservative single unit of measurement was used.
4. *Effect of agreement:* Mixed effects were used, as raters were not randomly sampled from the population, therefore reliability coefficients needed to account for both within and between rater variability. Of note, selecting the agreement effect does not change the calculation of the reliability coefficient itself, but does influence inference that may be drawn based on these coefficients.

2.1.2 Validity & Reliability Results

Tables 2.2-2.4 illustrate subject-specific differences in severity between first and second ratings for each SLP. If a difference between first and second ratings existed, differences were typically within one severity level and never differed by more than two levels for any of the speech measures.

Table 2.2: Subject-specific differences in intra-rater hypernasality by SLP * (N=15).

Subject	SLP 1		SLP 2		SLP 3		Mean Difference
	A	B	A	B	A	B	
1	0	0	0	0	0	0	0
2	0	0	1	1	1	0	0.333
3	0	0	0	0	0	0	0
4	0	0	0	2	0	2	1.333
5	0	0	0	0	0	0	0
6	3	3	3	3	2	3	0.333
7	2	2	2	0	3	2	1
8	2	2	1	1	0	0	0
9	0	0	0	0	1	0	0.333
10	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0
12	0	0	0	0	1	1	0
13	4	3	3	3	3	2	0.667
14	1	0	0	0	0	0	0.333
15	1	1	1	0	0	1	0.667

* 0-4 point Likert scale with levels as defined in Table 2

Table 2.3: Subject-specific differences in intra-rater VPI severity by SLP* (N=15).

Subject	SLP 1		SLP 2		SLP 3		Mean Difference
	A	B	A	B	A	B	
1	0	1	0	0	0	1	0.667
2	0	1	1	1	1	0	0.667
3	0	0	0	0	0	0	0
4	0	0	0	2	0	2	1.333
5	0	0	0	0	0	0	0
6	3	3	3	3	2	3	0.333
7	3	3	2	1	3	3	0.333
8	2	2	2	1	0	0	0.333
9	0	1	0	0	1	0	0.667
10	1	0	1	0	0	1	1.3
11	0	0	0	1	0	0	0.333
12	0	0	0	0	1	1	0
13	4	3	4	3	3	2	1
14	1	0	0	0	0	0	0.333
15	1	1	1	1	1	1	0

*0-4 point Likert scale with levels as defined in Table 2.1.

Table 2.4: Subject-specific differences in intra-rater hyponasality by SLP * (N=15).

Subject	SLP 1		SLP 2		SLP 3		Mean Difference
	A	B	A	B	A	B	
1	0	1	1	0	0	0	0.667
2	1	0	0	0	0	0	0.333
3	0	0	1	1	0	0	0
4	0	0	1	1	0	0	0
5	0	1	0	1	0	0	0.667
6	0	0	0	0	0	0	0
7	0	0	1	1	0	0	0
8	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0
10	1	1	1	1	1	1	0
11	2	2	2	1	1	1	0.333
12	1	0	0	0	0	0	0.333
13	0	0	0	0	0	0	0
14	0	1	0	1	0	0	0.667
15	0	0	0	0	0	0	0

*0-2 point Likert scale with levels as defined in Table 2.1.

Mean severity rating across all 50 subjects stratified by SLP are shown in Table 2.5. With one exception (VPI severity for SLP 3), all mean differences are below 1.0.

Table 2.5: Mean severity of each speech measure stratified by SLP* (N = 50).

	SLP 1	SLP 2	SLP 3
Hypernasality	0.88 (1.21)	0.9 (1.13)	0.92 (1.18)
VPI Severity	0.96 (1.34)	0.98 (1.24)	1.08 (1.21)
Hyponasality	0.26 (0.53)	0.34 (0.56)	0.04 (0.20)

*Likert scales as defined in Table 2.1. All values presented as mean (standard deviation).

Correlation coefficients for inter- and intra-rater reliability revealed "good" to "excellent" reliability for both hypernasality and VPI severity [47]. For hyponasality, however, inter- and intra-rater reliability fell in the "fair" range of agreement, with the exception of SLP 3 who demonstrated perfect intra-rater agreement. Intra-rater reliability results are presented in Table 2.6 and intra-rater reliability is presented in Table 2.7.

Table 2.6: Inter-rater reliability (N = 50)

	ICC*	CI
Hypernasality	0.752	0.639-0.841
VPI Severity	0.768	0.661-0.852
Hyponasality	0.5	0.334-0.654

*Two-way mixed consistency single-measures ICC with three raters.

Table 2.7: Intra-rater reliability (N=15)

	Speech Measure	Kappa*
SLP 1	Hypernasality	0.953
	VPI Severity	0.87
	Hyponasality	0.545
SLP 2	Hypernasality	0.738
	VPI Severity	0.768
	Hyponasality	0.577
SLP 3	Hypernasality	0.686
	VPI Severity	0.702
	Hyponasality	1

*Weighted kappa

2.1.3 Discussion & Conclusion

The relatively small differences in means between raters for both hypernasality and VPI severity suggest limited systematic bias in the measurement of these outcome variables. The presence of small differences is likely the result of the fact that each SLP providing the ratings had at least three years of experience specializing in velopharyngeal dysfunction, minimizing the

probability of overt misdiagnosis. Experienced raters of vocal quality and phonetic distinctions have been shown to have higher reliability than listeners inexperienced with vocal quality [50-54] and hypernasality [44,55]. Additionally, speech characteristics including disordered vocal quality and non-orally articulated pressure consonants shown to interfere with perception of resonance were essentially absent in the audio samples used for reliability analysis in this study [56-58]. Agreement in hyponasality measurement both within and across raters was in just the "fair" range (*ICC* = 0.5 for inter-rater and K_w between 0.55 and 1.0 for intra-rater reliability) [47]. Of note, "perfect" reliability for SLP 3 was based on just two indications of hyponasality present across the 15 samples. Therefore, this finding is likely an artifact of very small sample size than it is an indication of very strong intra-rater reliability for hyponasality. Differences in mean hyponasality severity were also larger when compared to those for both hypernasality and VPI severity, suggesting a greater degree of bias may exist in hyponasality measurement. Lower reliability in hyponasality measurement, however, is consistent with other literature as there are inherent differences in the speech characteristics used to define severity of hyponasality [44,59].

Lower reliability of hyponasality measurement could be attributed to several factors, including low prevalence of hyponasality across audio recordings, limited opportunity to perceive hyponasality within samples, and fewer levels of measurement. In the 50 audio recordings, hyponasality was indicated as present in between just two and eight samples. In contrast, VPI and/or hypernasality was present in between 21 and 29 samples. Similarly, the proportion of speech sounds available to measure hyponasality within a speech sample is low compared to indications of hypernasality or VPI severity. Of a total of 171 consonant productions in each audio sample, 80% of them provide information regarding presence of VPI while just 19% inform presence of hyponasality. Because the value of kappa decreases as prevalence of the measured characteristic reaches zero, it is not a surprising finding that hyponasality is not measured as reliably as characteristics occurring more frequently within the study population [60]. Finally, with just three levels of hyponasality severity versus five for hypernasality and VPI

severity, hyponasality measurement is more prone to smaller correlation coefficients, as weighted kappa tends to be larger when more levels of severity measurement are used [61].

While agreement for hypernasality and VPI severity based on audio recordings is good to excellent, how this measure generalizes to clinical ratings used in analysis warrants further discussion. All three SLPs available for clinic ratings contributed a large proportion of study population ratings. Clinical ratings do have some advantage over audio recordings, as they are based on an overall longer speech sample and studies support stronger reliability in longer samples [62,63]. Furthermore, all items present in audio recordings were also present in the speech sample used to determine severity ratings during clinical evaluation, as they make up our standard clinical protocol for perceptual evaluation of velopharyngeal function. Conversely, clinical ratings may contain listener bias and not all SLPs providing clinical ratings were available to demonstrate reliability for audio samples. At the time clinical ratings were made, the SLP did not have knowledge of the interest in comparing surgery types and was not consistently aware of which type of LFI had been completed.

One may express concerns for whether hypernasality & VPI severity definitions changed over time, particularly when the team chose to model their ratings after the CAPS-A-AM in 2013. Comparisons between SCH and CAPS-A-AM rating definitions are shown in Table 2.8. Fortunately, definitions for hypernasality did not appreciably change and earlier hyponasality measures were easily merged with CAPS-A-AM rating definitions. VPI severity is not present in CAPS-A-AM ratings, thus definition for this variable did not concretely change over time.

Table 2.8: Comparison of SCH & CAPS-A-AM severity rating definitions[+].

Speech Rating	***SCH Definition***	**CAPS-A-AM Definition**
Hypernasality	Excessive nasal energy perceived during production of vowels & voiced consonants	Any abnormal increase in nasal resonance of voice during speech production; most easily perceived on vowels and voiced consonants
0 -None	Balanced oral-nasal resonance	Nasal resonance WNL for region
1-Minimal	Occasional perception of excessive nasal energy on vowels; includes assimilation nasality	Some perceptible increase in nasal resonance
2-Mild	Present on high vowels /i, ɪ/, others may not detect	Evident on closed vowels /i, ɪ, u/
3-Moderate	Consistently present on high & mid-vowels; may be occasionally present on voiced consonants	Evident on closed & open vowels /ae, a, etc./
4-Severe	Present on all vowels & voiced consonants	Evident on vowels & voiced consonants
Hyponasality	Presence of intra-oral pressure build-up for nasal consonants /m, n, ŋ/ a reduction in nasal resonance of vowels	Abnormal reduction or absence in nasal resonance during speech production; most easily perceived on nasal consonants
0-None	Balanced oral-nasal resonance	Nasality WNL
1- Mild	Occasional denasalization of nasal consonants /m, n, ŋ/; vowels balanced	Nasal consonants partially denasalized
2-Marked	Consistent denasalization of nasal consonants /m, n, ŋ/ & increased oral versus nasal energy on vowels.	Nasal consonants frequently denasalized → prevalence of denasalized consonants /m, n, ŋ/

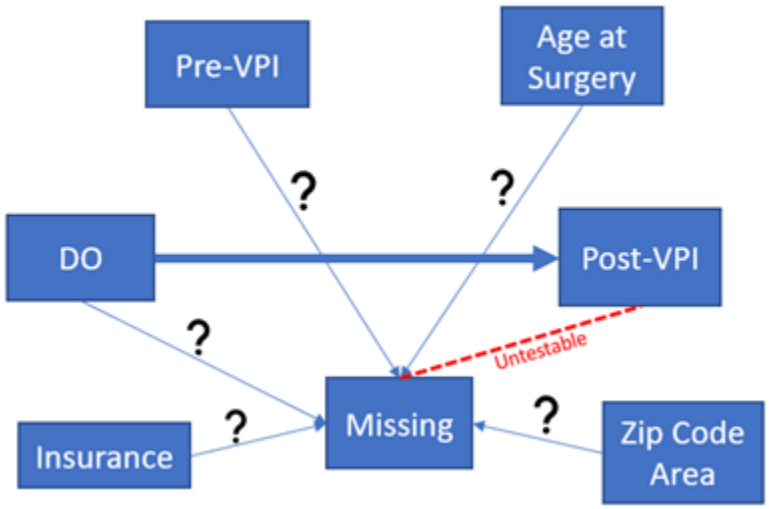
[+] NAE=nasal air emission; HN=hypernasality; PCs=pressure consonants; WNL=normal limits

Based on 50 audio recordings, reliability of primary outcomes hypernasality and VPI severity was good to excellent among experienced SLPs. Reliability for hyponasality was comparable to that reported elsewhere in the literature. While there are differences between speech samples obtained during clinical evaluation and those used for reliability analyses, degree of potential measurement error is considered acceptable and not likely to lead to an important degree of bias in analyses comparing outcomes between surgery types.

2.2 Rationale for Complete-Case Analysis

Early in the study design phase of this project, concern for the large proportion of subjects missing outcome measurement (32%) led to considerations for how this missingness may interfere with observed results and what techniques may be employed to mitigate bias. In the presence of loss to follow-up, selection bias arises when there are systematic differences between those who return for follow-up and those who do not, potentially leading to different exposure → outcome associations between the two follow-up groups [64]. Initially, inverse probability weighting (IPW) was proposed to address selection bias resulting from participants who did not return for post-operative speech evaluation (at which outcome is ascertained) [65]. IPW is a statistical approach that upweights individuals with a low probability of being selected into the study population, thereby allowing inference to extend to the whole study population rather than just those who returned for follow-up. In the current study, it was anticipated that pre-operative VPI status, distance from the hospital, insurance status and age at surgery were predictors of a subject's returning for follow-up. Figure 2.1 illustrates potential associations of predictors of missingness in the context of the primary exposure → outcome association of interest.

Figure 2.1: Potential predictors of outcome missingness



The above model was based on a priori knowledge of the patient population and factors that could reasonably interfere with a patient's ability to return to the hospital for post-operative evaluations. However, IPW is only as effective as the accuracy of assumptions made about factors that predict loss to follow-up. Therefore, a data-driven approach was employed to further evaluate predictors of missingness in the dataset.

Table 2.9 reflects the proportion of missingness within strata of each covariate with the potential for predicting absence of follow-up data. Additional covariates not shown here did not demonstrate an association with probability of missing follow-up.

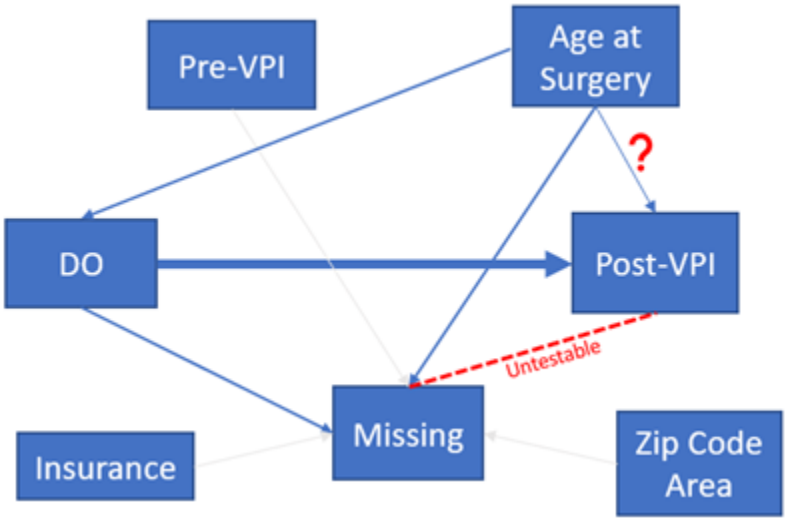
Table 2.9: Proportion of missingness by covariate.

	Missing	Not Missing	Total
Surgery Type			
DO	10 (16.4)	51 (83.6)	61
TMA	65 (37.1)	110 (62.9)	175
Zip Codes			
Eastern WA	6 (25)	18 (75)	24
Out of State	6 (28.6)	15 (71.4)	21
Western WA	61 (32.6)	126 (67.4)	187
Pre-Op Hypernasality			
None	41 (27.9)	106 (72.1)	147
Min	3 (16.7)	15 (83.3)	18
Mild	6 (20.7)	23 (79.3)	29
Mod	4 (33.3)	8 (66.7)	12
Severe	0 (0)	3 (100)	3
Insurance Status			
Financial Assist	5 (50)	5 (50)	10
Private	36 (27.9)	93 (72.1)	129
Public	30 (37.5)	50 (62.5)	80
Pre-Op Hyponasality			
None	52 (26.5)	144 (73.5)	196
Mild	0 (0)	10 (100)	10
Marked	2 (66.7)	1 (33.3)	3
Age at Surgery (Yrs)			
< 17	12 (15.2)	67 (84.8)	79
17-19	41 (37.3)	69 (62.7)	110
> 19	21 (44.7)	26 (55.3)	47

At first glance, there appears to be an imbalance of missingness for mild and marked hyponasality. However, these numbers are very small and there is no physiologic reason to explain why hyponasality would predict whether a patient returned for follow-up. Proportion of missingness was relatively balanced within levels of pre-op hypernasality, zip code area, and insurance status. Imbalances in outcome missingness are present within levels of surgery type and age at surgery. Proportion of missingness is more than two times greater among those who underwent TMA compared to DO. Similarly, the proportion of those missing outcome measurement in the oldest age group is nearly three times that present in the youngest age category. Because surgery type, the primary exposure of interest, is associated with missingness and the association between outcome and missing is untestable, it is possible that conditioning on those with follow-up data would lead to biased results. To further evaluate this concern, additional exploration of the association between type of surgery, age at surgery and outcome is necessary.

Descriptive analysis revealed that those undergoing DO were younger than those undergoing TMA (mean age 15.1 and 18.9 years, respectively). An association between age and surgery type potentially opens a backdoor path between exposure → outcome through age at surgery if there is a direct effect between age at surgery and the prevalence of post-op VPI (see Figure 2.2).

Figure 2.2: Associations between relevant covariates, missingness, and outcome.



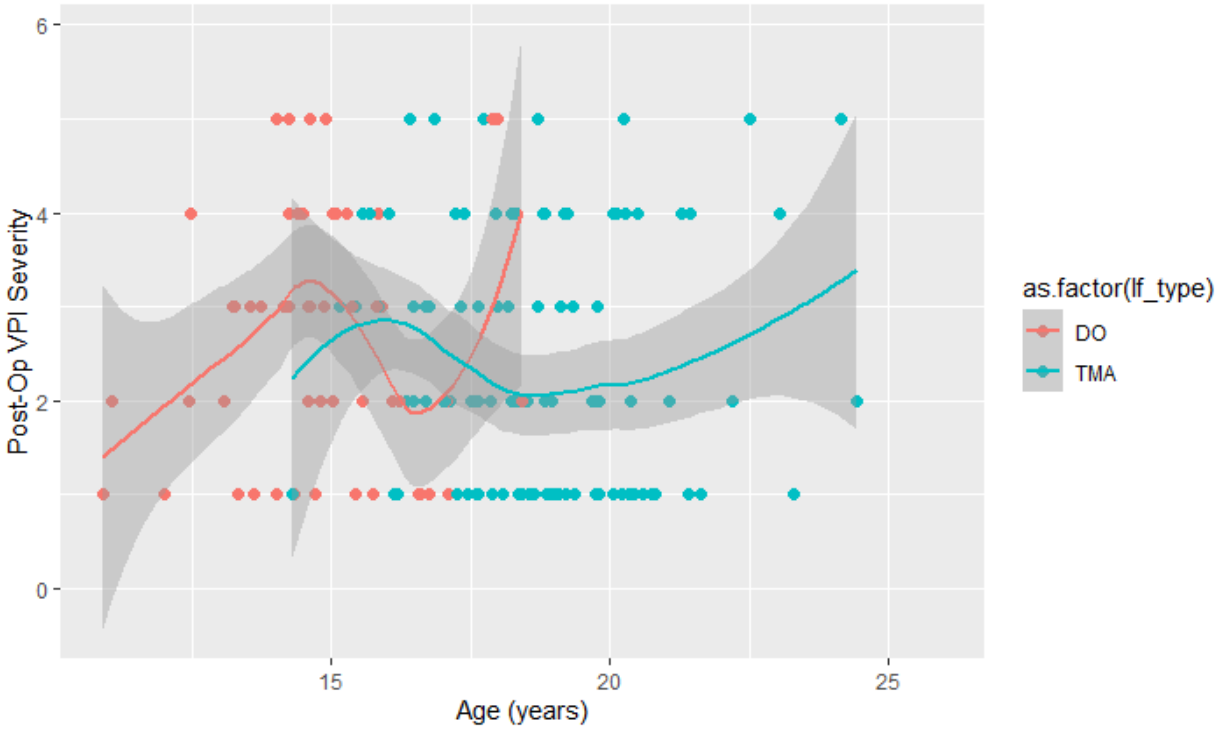
If present, this bias could not be adequately controlled for due to the relationship between age at surgery and missing outcome. Therefore, determination of whether a residual direct effect of age at surgery on outcome exists within strata each surgery type through both correlation coefficients and data visualization strategies. Results of Pearson's product-moment correlation analyses revealed absence of a linear correlation between age at surgery and post-op hypernasality (Table 2.10).

Table 2.10: Correlation between age at surgery & post-op hypernasality stratified by surgery type.

	Correlation	95% CI	p-value
DO	0.089	-0.191 - 0.356	0.535
TMA	-0.014	-0.202 - 0.175	0.886

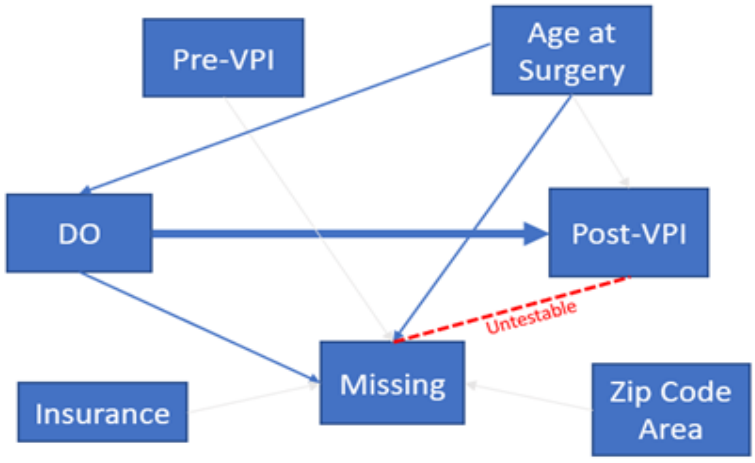
A scatterplot of age at surgery on outcome (Figure 2.3) provided further support for absence of a relationship between age and outcome within surgery types. Similar results were obtained when outcome was defined as post-operative hypernasality.

Figure 2.3: Scatterplot of age at surgery and VPI severity stratified by surgery type.



As can be seen from the LOWESS (Locally Weighted Scatterplot Smoothing) curves, a predictable trend of different outcome dependent on age at surgery is not present within either surgery group. Taken together, the lack of correlation as well as absence of a pattern of age-outcome association for either DO or TMA suggest no direct effect of age at surgery on outcome and relationships between exposure, outcome and covariates are as seen in Figure 2.4.

Figure 2.4: Associations between missingness, exposure, outcome, and other covariates as identified within the data.

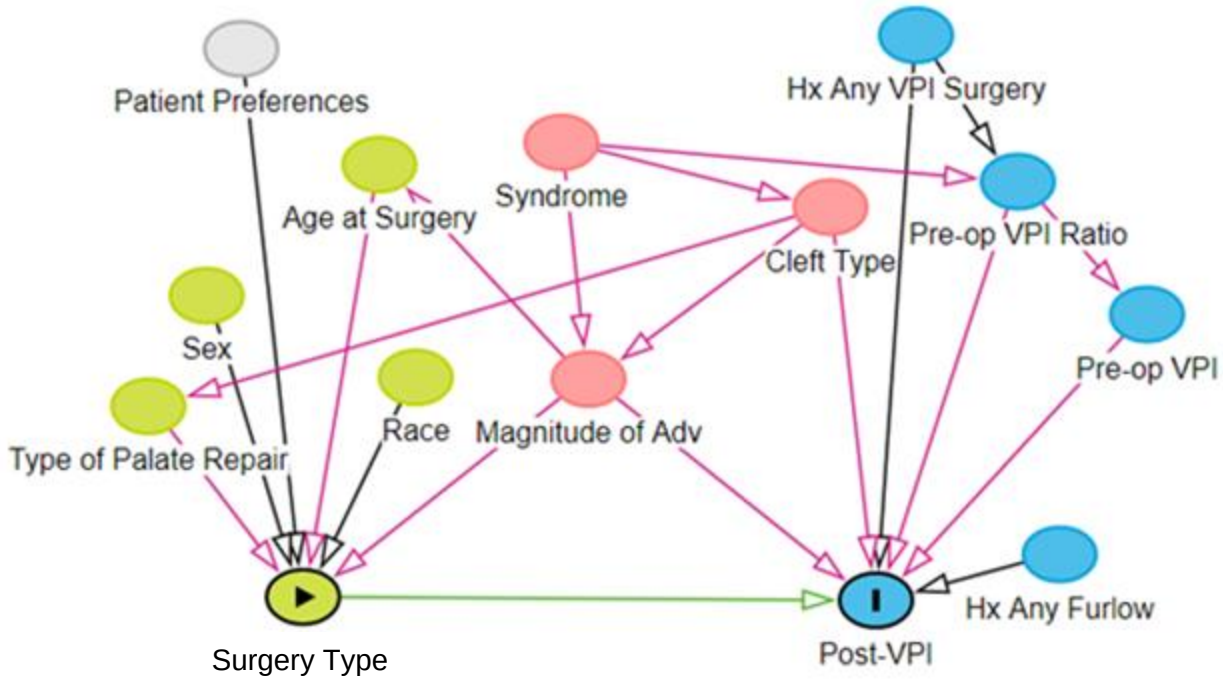


The purpose of an analysis using IPW in this study would have been to reduce selection bias created by conditioning only on subjects that contribute outcome data. However, selection bias requires both exposure and outcome to affect the probability of follow-up. Because only surgery type and age were identified as predictors of missing outcome and no direct effect of age on outcome was identified, IPW is not indicated. Furthermore, as can be seen in Figure 2.4, absence of a direct effect between age and outcome removes the backdoor path between exposure and outcome through age, leaving no relationship between a factor predictive of missingness and outcome and suggesting that selection bias is not a strong concern for this cohort. Given relationships among variables described above, assuming outcome data are missing completely at random (MCAR) is reasonable [66,67]. In the presence of MCAR outcome and absence of bias related to selection into the study, a complete-case analysis with adjustment for relevant covariates is a valid approach for primary exposure-outcome analyses in this study.

2.3 Confounder Selection Rationale

An observed association between exposure and outcome is confounded when there is an imbalance in the prevalence of a third variable between the exposed and unexposed [68]. When these imbalances are not controlled for, either through study design or data analysis, the observed exposure→outcome association is distorted by the effects of this third variable, a confounder. Only variables related in some way to both exposure and outcome should be considered as potential confounders. Additionally, a variable is not a potential confounder if it is on the causal path between the primary exposure and outcome of interest. In this study, potential confounders were identified through a combination of *a priori* clinical knowledge as well as assessment of statistical associations between variables in the data [69]. Figure 2.5 is a Directed Acyclic Graph (DAG) illustrating presumed relationships between variables in this dataset based only on clinical knowledge [70,71].

Figure 2.5: DAG illustrating exposure, outcome, and covariate relationships as informed by clinical knowledge*.



* Key: green = direct association with exposure; blue = direct association with outcome; red = directly or indirectly associated with exposure and outcome; red arrow = biased paths between exposure and outcome; gray = unmeasured

Intra-operative surgical characteristics (i.e. concurrent mandibular surgery, multi-segmental osteotomy, and alveolar bone graft) are not included in the diagram, as we are interested in the total effect of exposure on outcome and they are consequences of surgery type rather than independently informing surgery performed. Demographic variables (i.e. zip code area and insurance status) are also not included because there is no plausible biologic explanation for their association with outcome. Of the variables shown in Figure 2.5, magnitude of advancement, age at surgery, presence of a syndrome, cleft type, type of palate repair, pre-op VPI ratio and presence of pre-op VPI all have the potential to confound the relation between surgery type and prevalence of post-op VPI. As discussed in Chapter 2.2, age at surgery is not associated with outcome and will therefore no longer be considered as a potential confounder.

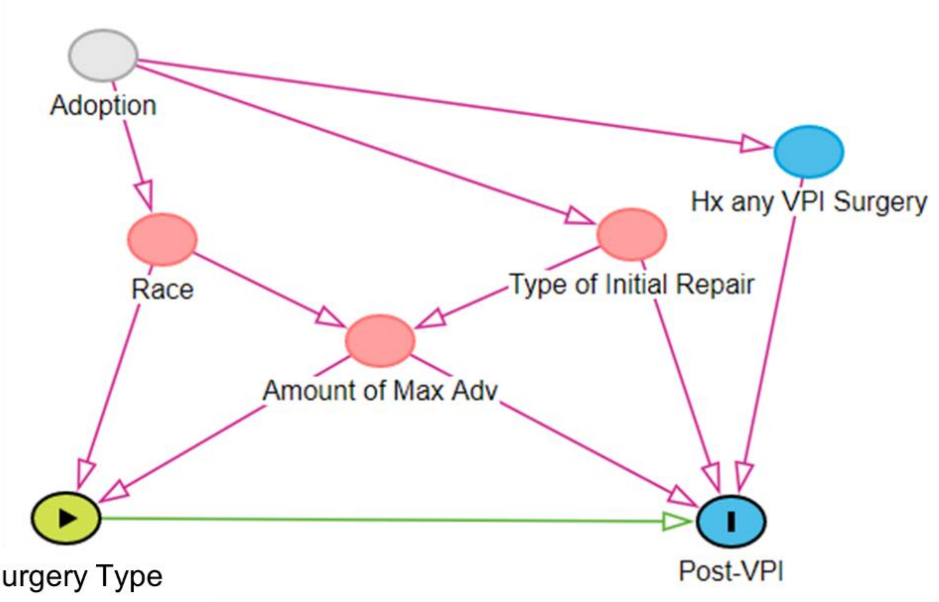
Further exploration of the remaining potential confounder relationships was achieved through descriptive data exploration. The proportion of each level of additional covariates between children undergoing each of the two types of surgery is reflected in Table 2.11.

Table 2.11: Study population characteristics identified as potential confounders, stratified by surgery type.

	DO n = 61	**TMA** n = 176
Demographics		
Male	43 (70.5)	96 (54.9)
Race		
Asian	20 (33.9)	31 (18.5)
Other	13 (22.0)	54 (32.1)
White	26 (44.1)	83 (49.4)
Clinical Characteristics		
Cleft Type		
BCLP	25 (41.0)	44 (25.1)
CPO	3 (4.9)	30 (17.1)
SMCP	0 (0.0)	12 (6.9)
UCLP	33 (54.1)	89 (50.9)
Type of Initial Palate Repair		
Furlow	0 (0.0)	4 (2.3)
Intravelar Veloplasty	35 (57.4)	105 (60.0)
Other	0 (0.0)	9 (5.1)
Unknown	26 (42.6)	57 (32.6)
History of Any VPI Surgery	24 (39.3)	71 (41.0)
Pre-Op VPI Severity		
None	33 (57.9)	94 (61.8)
Minimal	13 (22.8)	32 (21.1)
Mild	5 (8.8)	18 (11.8)
Moderate	3 (5.3)	7 (4.6)
Severe	3 (5.3)	1 (0.7)
Pre-op Hypernasality Severity		
None	42 (73.7)	105 (69.1)
Minimal	5 (8.8)	13 (8.6)
Mild	6 (10.5)	23 (15.1)
Moderate	2 (3.5)	10 (6.6)
Severe	2 (3.5)	1 (0.7)
Syndrome Present	6 (9.8)	40 (22.9)
History of Primary/Secondary Furlow	18 (29.5)	56 (32.4)
Pre-Operative Cephalometrics		
Overjet (mm)	-11.27 (4.88)	-3.43 (4.87)
ANB (degrees)	-8.31 (4.40)	-3.57 (4.64)

Based on observed imbalances, the following variables were considered further as potential confounders: sex, race, cleft type, presence of a syndrome, and magnitude of maxillary advancement (as defined by millimeters of overjet on pre-operative cephalometric radiograph). While it is feasible that sex at birth could influence surgery type due to differences in aesthetic acceptability between males and females, there is no reasonable explanation for sex to be associated with prevalence of VPI after maxillary advancement. Imbalance between surgery groups with respect to race was due to the fact that almost twice the proportion of the DO group were Asian compared to the TMA group. While it is possible that Asian descent may be associated with outcome through cephalometric variables due to known differences in craniofacial norms between races [72], it is unknown how these differences may be present within the population of individuals with CP/L. An additional consideration includes the fact that there is a high percentage of internationally adopted patients served by the Seattle Children's Craniofacial Center. Many of these patients had palate repair in their country of birth at an older age than is typical in the United States and with surgeons who may or may not specialize in craniofacial surgery. Differences in their early care may result in increased history of VPI surgery as well as greater maxillary deficiency by their teen years, thereby establishing a potential relationship between race and both exposure and outcome in this study (Figure 2.6).

Figure 2.6: DAG depicting proposed relationships leading to race as a potential confounder between exposure→outcome association.

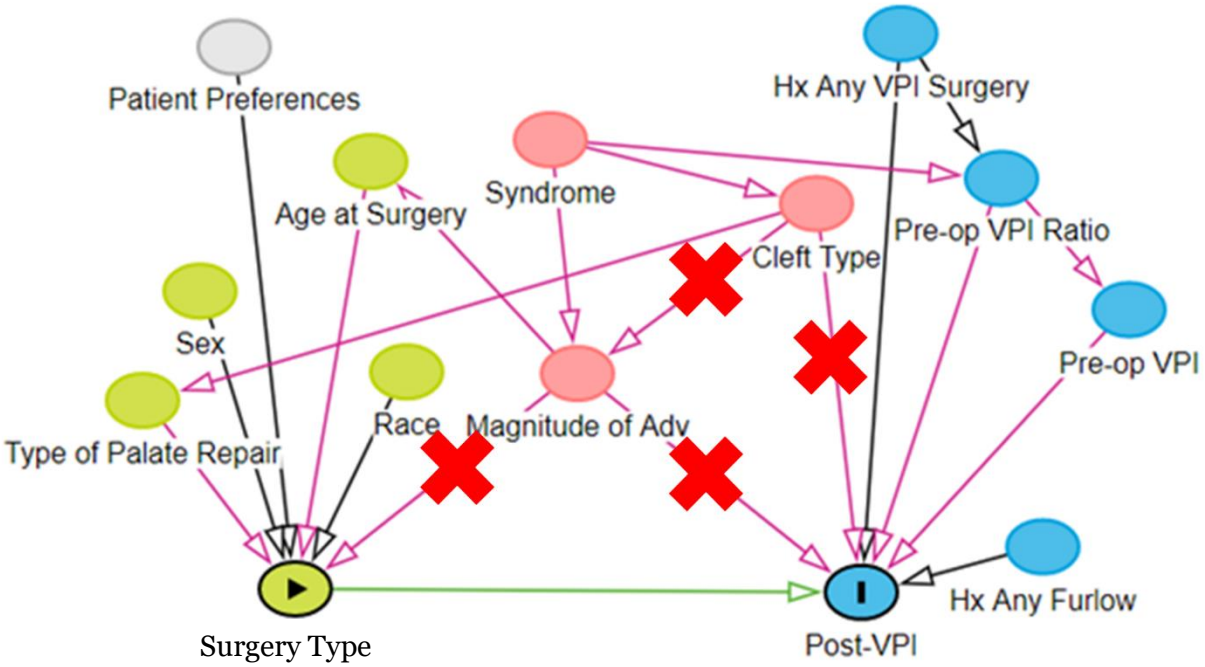


However, it is unknown what proportion of those of Asian descent in this cohort were adopted and a direct association between race and type of palate repair would be tenuous, at best. Similarly, the imbalance in proportion of race between exposed and unexposed is not large in the other two race categories, within which there is also likely a degree of heterogeneity. Adjusting for amount of maxillary advancement will block one path from exposure→outcome through race but will not block the path from exposure→race→adoption→VPI surgery→outcome. Given the unclear associations discussed as well as unmeasured adoption status, race was not adjusted for in the main analysis.

Syndrome status, cleft type, and amount of maxillary advancement were the remaining variables in consideration as confounders. As seen in Figure 2.5, the only path from syndrome status to surgery type is through amount of maxillary advancement. Since magnitude of maxillary advancement has a clear direct relationship between both exposure and outcome (and is not on the causal pathway), it is a clear confounder that requires adjustment in any analyses testing the association between type of maxillary advancement and prevalence of post-operative

VPI. Adjusting for amount of maxillary advancement will then block the path between exposure and outcome that runs through syndrome status, and so additional adjustment by syndrome status is not indicated. Adjustment for amount of maxillary advancement also blocks one potential path between exposure and outcome that runs through cleft type. It does not however, block the path from exposure→type of repair→cleft type→outcome. Because there is not an imbalance of type of repair between surgery groups, and because type of initial palate repair is a not a well-measured variable, main analyses should also be adjusted for type of cleft. Figure 2.7 illustrates the pathways remaining open and blocked based on the discussion above.

Figure 2.7: DAG illustrating paths between exposure and outcome blocked by adjustment for magnitude of advancement and cleft type.



Despite careful confounder selection methods, residual confounding is always a concern. In this study, race and syndrome status are the most plausible sources of residual confounding. To address this possibility, sensitivity analyses that include race and syndrome status have been run to determine whether inclusion of these variables has an important impact on the observed association between surgery type and prevalence of post-operative VPI.

Chapter 3: Maxillary Advancement and Velopharyngeal Function: A Comparison of Surgical Approaches

3. 1 Introduction

LeFort I osteotomy with maxillary advancement is a common surgical procedure used to correct maxillary hypoplasia. Maxillary hypoplasia interferes with dental occlusion, nasal airflow, chewing, speech production, and may contribute to the development of obstructive sleep apnea. Traditional maxillary advancement (TMA) is considered a reliable and safe procedure among individuals without cleft palate with or without cleft lip (CP/L) [13,22,27]. Individuals with CP/L however, present with multi-dimensional maxillary deficiency, increased severity of hypoplasia, and prior surgical history, making TMA a riskier and more technically complex procedure with a relatively greater chance of negative post-operative sequelae. Consequently, maxillary advancement achieved through distraction osteogenesis (DO) was introduced [14]. DO also involves LeFort I osteotomy, but distractors attached to the maxilla assert anterior force over a period of time rather than displacement occurring within a single procedure. For those with severe maxillary deficiency, there is evidence to suggest that DO can provide a more stable occlusal results long-term [15]. Because anatomical changes occur over time, some providers suggest that DO also could reduce the risk of post-op velopharyngeal insufficiency (VPI) compared to TMA in individuals with CP/L. However, to date this question has not been well addressed. The purpose of the present study is to compare risk of post-operative VPI between patients with CP/L undergoing DO and those with CP/L undergoing TMA, as well as to identify patient characteristics that may modify the effect of surgery on risk of post-operative VPI.

Individuals with CP/L who have more severe maxillary hypoplasia are relatively more likely to undergo DO than other patients with CP/L [15,30]. Similarly, it is plausible that those requiring a greater magnitude of advancement are at increased risk of post-operative VPI [16,41].

Therefore, a valid assessment of the impact of DO and TMA on post-operative velopharyngeal function necessitates a design or an analysis that permits a comparison between groups requiring a similar amount of advancement. A randomized controlled trial (RCT) of individuals requiring a moderate amount of advancement (4-10 mm) [15] observed no difference in post-op VPI risk between surgical approaches. However, outcome data were available for less than half of their original cohort and it was possible the study was not adequately powered to observe a difference. Additionally, systematic differences between those who did and did not return for speech follow-up was not considered. A retrospective cohort study of patients with CP/L and severe maxillary hypoplasia (>= 10 mm) observed that those undergoing DO had only half the prevalence of post-operative VPI compared to patients undergoing TMA [16]. Further statistical analysis for VPI outcome was not performed.

Maxillary advancement surgery is often considered the end of the timeline of care and occurs at a time of great transition in a young adult's life. Therefore, there is great interest for both providers, patients, and patients' families in choosing a surgical technique that fulfills the surgery's physiological and aesthetic purpose long-term while simultaneously minimizing risk of adverse functional changes that lead to need for further surgical and/or clinical management. The current study provides a direct comparison of risk of postoperative VPI between DO and TMA in a large cohort of patients with CP/L, thereby enhancing available evidence when considering risk of changes in velopharyngeal function following surgical advancement of the maxilla.

3.2 Methods

3.2.1 Participants

For this retrospective cohort study, eligible participants were identified from a clinical database that includes all patients of the Seattle Children's Craniofacial Center. All patients with CP/L who underwent LeFort I osteotomy for maxillary advancement between January 2007 and

June 2019 were considered eligible. For inclusion, patients had to have either an isolated or syndromic secondary palatal cleft of any type with or without a cleft lip. Patients with cleft lip/alveolus only were excluded, as were patients with repeat LeFort I osteotomy procedures or those with LeFort I osteotomy procedures not involving anterior movement of the maxilla. Demographics, surgical history, and medical diagnoses were obtained from the Craniofacial Center's clinical database (CRA DB) and verified in the electronic medical record. All participants underwent a pre-operative orthodontic evaluation, standard surgical evaluation with cephalometric radiographs, and pre-/post-operative perceptual evaluation of velopharyngeal function per clinical protocol. In the analysis, patients were divided into two groups by type of procedure, either DO or TMA. All surgeons performing the procedures were board-certified in either plastic and reconstructive surgery, oral-maxillofacial surgery, or both, and specialized in the management of children with craniofacial differences. This study was approved through expedited review by the Institutional Review Board at the University of Washington (IRB #STUDY00008156). Because all data utilized in the study were present in the electronic medical record or in internal clinical databases, written consent/assent was waived.

3.2.2 Surgical Characteristics

All participants underwent LeFort I osteotomy with pterygomaxillary disjunction and down-fracture, as described in the literature [22,24,27]. Type of surgical approach was identified directly from operative reports. Additional pertinent details pertaining to the surgery were also recorded. As indicated by presurgical evaluation, concurrent procedures including mandibular osteotomy, alveolar bone graft, and pharyngeal flap release were recorded. Need for multi-dimensional movement, as indicated by division of the maxilla into two or three segments prior to advancement, was also noted. Additional traction techniques were applied to ensure complete maxillary mobilization, as necessary. For TMA, the newly mobile maxilla was placed into a prefabricated acrylic surgical splint to obtain intermaxillary fixation then wired together with

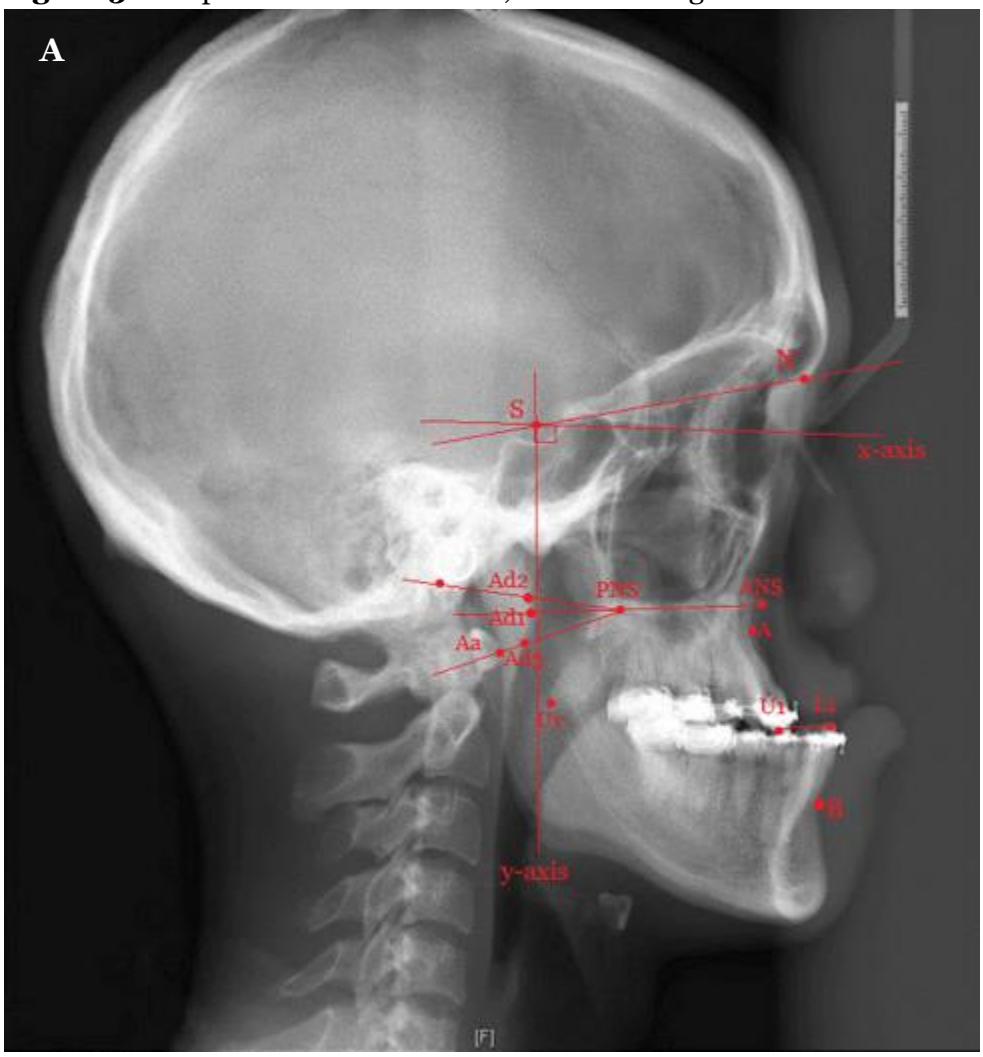
26-gauge wire loops. The new maxillary position was then stabilized with titanium miniplates secured with screws.

For DO, LeFort I osteotomy was performed as previously described. Following maxillary mobilization, however, the maxilla was secured into the splint in its baseline position rather than the desired anterior position. Rigid external distractors (RED) were then secured to the temporal bone with six screws, three bilaterally. The splint and RED device were then secured together with two 24-gauge wires. After a three-day latency period, activation of distraction was initiated at 1 mm per day until a class I occlusion relationship was achieved, typically between 7-15 days. Distractors were then removed in a second procedure following a three- to six-week consolidation phase.

3.2.3 Lateral Cephalometric Measurement

A lateral cephalometric radiograph was completed pre-operatively and serially post-operatively, per protocol. All lateral radiographs were taken using the same cephalostat with patient aligned to Frankfurt horizontal, molars in occlusion, and with the lips and soft palate at rest. For analysis, the radiograph closest to the date of surgery pre-operatively and to the date of perceptual velopharyngeal function evaluation post-operatively were used. Two orthodontists with the Craniofacial Center completed digitized tracings of pre- and post-operative radiographs for each participant with Dolphin Imaging Premium 11.95 software. Magnification differences were accounted for in tracings. For reference, the horizontal axis was established by dropping a line seven degrees from the center of sella turcica (S) to nasion (N) [18,73,74]. The vertical axis was then identified by a line drawn perpendicular to the x-axis through S. Documented landmarks are illustrated in Figure 3.1A. Definitions of linear and angular measurements using these landmarks are summarized in Figure 3.1B.

Figure 3.1: Cephalometric landmarks, linear and angular measures *.



* ANS = anterior nasal spine, PNS = posterior nasal spine, Uv = tip of uvula, A = deepest point on anterior contour of maxilla, S = center of sella turcica, N = nasion, Ba = Basion, Aa = atlas, Ad1 = intersection of palatal plane and posterior pharyngeal wall, Ad2 = intersection of Ba-PNS line and PPW, Ad3 = intersection of line Aa-PNS and PPW, U1 = upper incisor tip, L1 = lower incisor tip, B = deepest point on anterior contour of mandible

Figure 3.1B:

Linear Measures (mm)	Definition
Vertical maxillary deficiency	Linear distance from A point along line drawn perpendicular to x-axis
Anterior-posterior maxillary deficiency	Linear distance from A point along line drawn perpendicular to y-axis
Soft palate length	Linear distance from PNS to Uv
Nasopharyngeal depth	1. Linear distance from PNS to Ad1 2. Linear distance from PNS to Ad2 3. Linear distance from PNS to Ad3

Overjet (Predicted magnitude of advancement)	Anterior-posterior projection of upper beyond lower incisor tips as measured parallel to occlusal plane relative to upper & lower incisor tips
Angular Measures (degrees)	**Definition**
Cranial base angle	Angle formed by lines S-N and S-Ba
Palatal plane angle	Angle formed by ANS-PNS plane where it intersects x-axis
ANB	Angle formed by A point-N-B point; measures anterior-posterior discrepancy between maxilla and mandible
SNA	Angle formed by S-N-A point
Velar inclination angle	Angle formed by ANS-PNS and PNS-Uv; angulation of soft to hard palate

In addition to measures described in Figure 3.1, velopharyngeal need ratio (VP ratio), as defined by nasopharyngeal depth in the numerator and velar length in the denominator [75,76], was calculated. Ten percent of radiographs were traced by both orthodontists to evaluate inter-rater reliability. Each orthodontist also retraced ten percent of radiographs to evaluate intra-rater reliability.

3.2.4 Perceptual Evaluation of Velopharyngeal Function

Pre- and postoperative speech data were obtained through a second clinical database (VPI DB) for each eligible participant identified through the CRA DB. Perceptual evaluation of velopharyngeal function was performed by one of six speech-language pathologists (SLPs) specializing in the diagnosis and management of velopharyngeal dysfunction, although three of them contributed ratings for less than three years of the 12-year study period. Per clinical protocol, preoperative evaluations occurred within one year of maxillary advancement surgery and post-operative evaluations around three months after surgery. For each patient, ratings of VPI severity, hypernasality and hyponasality were stored in the VPI clinical database at the time of evaluation. Ratings were based on overall clinical impression, and each evaluation included syllable and word repetition, pressure consonant loaded sentence repetition, counting 1-20 and 60-70 and a spontaneous speech sample. Both VPI severity and hypernasality were rated on a 5-point severity scale and hyponasality on a 3-point severity scale with definitions as described in

Table 2.1. To establish interrater reliability with this scale, 50 audio recordings of syllable repetition, sentence repetition and counting 1-20 and 60-70 (Appendix A) of patients of similar age and diagnosis were rated using the above criteria by three of the six SLPs contributing ratings over the study period. Recordings were obtained using a Sony Digital ICD Recorder in a quiet exam room. Each of the three SLPs re-rated 20% of audio recordings two weeks after initial ratings to measure intra-rater reliability.

3.2.5 Statistical Analysis

The primary predictor was surgery type (DO or TMA) as indicated in the operative report. The primary outcomes of interest were post-operative VPI and hypernasality. These were defined in two ways: (1) as a binary variable with ratings 2-4 indicating presence of post-op VPI/hypernasality and 0-1 indicating absence and, (2) as the original 5-level severity scale. For the binary outcome, relative risk regression using the Poisson likelihood function [77] were run to estimate the difference between DO and TMA on the prevalence of post-operative VPI/hypernasality. Patients presenting with pre-operative VPI/hypernasality were excluded from analyses with binary outcome definition. For categorical outcomes, multivariable linear regression was used to investigate the difference in mean post-operative VPI/hypernasality among those who underwent DO relative to those who underwent TMA. For all regression analyses, sandwich robust standard errors were used to calculate confidence intervals. Relevant confounders identified through a combined *a priori* clinical knowledge and data driven approach included cleft type and magnitude of advancement needed, and all results included adjustment by these variables [69,70]. For adjustment purposes, magnitude of advancement was defined as the linear distance between the between the upper and lower incisal tips ("overjet"). Bias resulting from the failure of some patients to return for follow-up was not a strong concern, as no relationship between a characteristic predictive of missingness and outcome was

identified. Therefore, it was assumed that outcome data were missing completely at random (MCAR), and complete-case analysis was considered a valid approach [78].

Sensitivity analyses were also run to determine whether additional covariates meaningfully impacted results. When race, syndrome, pre-operative VPI/hypernasality severity, and time between speech evaluation and date of surgery (in weeks) were included as covariates in the adjusted model, there were no observed changes in exposure-outcome relationship. The primary model as described above was also re-run using different definitions of magnitude of advancement (ANB angle and linear distance from A-point to a point perpendicular to the y-axis). Primary analysis was also re-run after removing those with requiring multi-dimensional osteotomy to examine effect of multi-dimensional maxillary deficiency on the association between surgery type and post-operative VPI/hypernasality. Effect modification by any history of Furlow palatoplasty (as initial palate repair or for VPI management), any history of VPI surgery, and by whether the pre-operative VP ratio was above or below 0.8 was also assessed by including an interaction term for the effect modification variable of interest and type of surgery in the primary model. An indicator variable for whether the VP ratio was above or below the 0.8 cut-off was created through dichotomization of the continuous VP ratio variable. A cut-off of 0.8 was determined based on available evidence of the threshold that is consistent with presence of VPI [79,80].

Two-way mixed consistency single-measures intra-class correlation coefficients (ICCs) were calculated for all cephalometric measurements, for both inter- and intra-rater reliability [46]. For perceptual speech ratings, weighted kappa [45] and two-way mixed consistency, single-measures ICCs were calculated for intra- and inter-rater reliability, respectively. Interpretation of reliability coefficients as defined by Cicchetti et al were used [47]. Independent samples t-tests were used to evaluate consistency of measurements within and between orthodontists contributing cephalometric tracings. All analyses were completed in R version 4.0.3.

3.3 Results

Children who underwent a LeFort I osteotomy (LFO) that did not involve anterior displacement of the maxilla were excluded from the analysis. There remained a total of 237 patients who underwent LFO for maxillary advancement over the 12-year study period. These patients were divided into two groups, those who underwent DO (N = 61) and those who underwent TMA (N = 176) to achieve desired occlusal outcome. Table 3.1 summarizes patient characteristics by type of maxillary advancement surgery. Pre-operative cephalometric radiographs completed more than 100 weeks before surgery were considered missing. Two additional patients had pre-operative cephalometric radiographs at 78 and 64 weeks before surgery. These patients remained in the study population, as both were more than 21 years of age at the time of surgery, thus mitigating concern for significant growth between these timepoints. After these considerations, median time between pre-operative cephalometric radiographs and date of surgery was 2.93 weeks (IQR 2.14-8.04). Median time between pre-operative speech evaluation and surgery date was 21.6 weeks (IQR 3.00-76.6). Absence of interim VPI surgery was confirmed for all participants. Speech evaluations completed more than two years before surgery were considered missing.

Table 3.1: Pre-operative patient characteristics stratified by type of surgery [+].

	DO n = 61	**TMA** n = 176
Demographics		
Age at Surgery (years)	15.10 (1.93)	18.86 (2.03)
Male	43 (70.5)	96 (54.9)
Race		
Asian	20 (33.9)	31 (18.5)
Other	13 (22.0)	54 (32.1)
White	26 (44.1)	83 (49.4)
Area of Residence		
Eastern WA	7 (11.5)	17 (9.9)
Out of State	3 (4.9)	18 (10.5)
Western WA	51 (83.6)	136 (79.5)
Type of Insurance		

Financial assistance only	1 (1.8)	9 (5.5)
Private	34 (60.7)	95 (58.3)
Public	21 (37.5)	59 (36.2)
Receiving Financial Assistance	7 (11.5)	38 (21.7)
Clinical Characteristics		
Cleft Type		
BCLP	25 (41.0)	44 (25.1)
CPO	3 (4.9)	30 (17.1)
SMCP	0 (0.0)	12 (6.9)
UCLP	33 (54.1)	89 (50.9)
Syndrome Present	6 (9.8)	40 (22.9)
Type of Initial Palate Repair		
Furlow	0 (0.0)	4 (2.3)
Intravelar Veloplasty	35 (57.4)	105 (60.0)
Other	0 (0.0)	9 (5.1)
Unknown	26 (42.6)	57 (32.6)
VPI Surgical History		
Furlow	18 (29.5)	54 (31.2)
Sphincter pharyngoplasty	5 (8.2)	23 (13.3)
Pharyngeal flap	4 (6.6)	11 (6.4)
Other	2 (3.3)	9 (5.2)
VPI Severity*		
None	33 (57.9)	94 (61.8)
Minimal	13 (22.8)	32 (21.1)
Mild	5 (8.8)	18 (11.8)
Moderate	3 (5.3)	7 (4.6)
Severe	3 (5.3)	1 (0.7)
Hypernasality Severity*		
None	42 (73.7)	105 (69.1)
Minimal	5 (8.8)	13 (8.6)
Mild	6 (10.5)	23 (15.1)
Moderate	2 (3.5)	10 (6.6)
Severe	2 (3.5)	1 (0.7)
Hyponasality*		
None	55 (96.5)	141 (92.8)
Mild	2 (3.5)	8 (5.3)
Marked	0 (0.0)	3 (2.0)
Additional Surgical Characteristics		
Alveolar Bone Graft	2 (3.3)	103 (58.9)
Mandibular Setback	0 (0.0)	35 (20.0)
Mandibular Advancement	0 (0.0)	50 (28.6)
Multi-Segmental Osteotomy	2 (3.3)	40 (22.9)
Pre-Operative Cephalometrics**		
Cranial Base Angle (degrees)	130.6 (6.18)	131.68 (6.16)

ANB (degrees)	-8.31 (4.40)	-3.57 (4.64)
Velar Length	26.34 (7.07)	27.10 (6.91)
Nasopharyngeal Depth (per Ad1)	19.82 (5.57)	22.27 (5.50)
Maxillary Length (PNS-A point)	39.41 (8.32)	40.41 (6.14)
Overjet (mm)***	-11.27 (4.88)	-3.43 (4.87)
VP Ratio (per Ad1)	0.78 (0.25)	0.86 (0.25)

+ Continuous variables as mean (SD); categorical/binary variables as N(%)
* 4 missing in DO group, 24 in TMA group
**4 missing in DO group, 38 in TMA group
***5 missing in DO, 42 missing in TMA group

Those undergoing DO tended to be younger, were more likely to be male and on average had a greater degree of negative overjet than those undergoing TMA. Magnitude of planned advancement ("overjet") in the DO group ranged from -21.8 to 2.9 millimeters, while the range in the TMA group was -13.9 to 18.7 millimeters. The large range in the TMA group was due to one patient with significant bimaxillary hypoplasia, thus anterior displacement of the maxilla was required despite the large magnitude of positive overjet. After removing the 18.7 outlier in the TMA group, maximum overjet was 8.6 millimeters. The proportion of those identifying as Asian was nearly two times as great in the DO group compared to the TMA group, while the proportion of white or other race was similar between surgery groups. Those in the TMA group were more likely to have a syndrome diagnosis and included a larger proportion of participants with a cleft of the secondary palate than those in the DO group. The proportion of BCLP was greater among those undergoing DO while the proportion of UCLP was relatively balanced between surgery groups. Prevalence of VPI or hypernasality prior to surgery did not meaningfully differ between DO and TMA groups.

3.3.1 Reliability Results

Two-way mixed consistency single-measures intra-class correlation coefficient (ICC) analyses revealed good to excellent inter- and intra-rater reliability for all cephalometric measures used in analyses (ICCs ranging from 0.75 to 0.99). ICCs for inter-rater reliability of

VPI and hypernasality were 0.77 and 0.75, respectively, also revealing good agreement across the three raters. Weighted kappa values revealed good to excellent intra-rater reliability for each SLP, with values ranging from 0.69 to 0.95 for VPI and hypernasality.

3.3.2 Relative Risk Regression with Binary Outcome

With relative risk regression using the Poisson likelihood function, we compared the prevalence of any post-operative VPI and hypernasality between the DO and TMA groups. Because we were interested in the association between surgery type and VPI/hypernasality outcome, participants presenting with pre-operative VPI/hypernasality (rating 2-4) were excluded from this analysis. The prevalence ratios (PR) of VPI and hypernasality were estimated across the whole cohort as well as within subgroups defined by type of cleft. Models including the whole cohort were adjusted for cleft diagnosis and magnitude of planned advancement (overjet), while subgroup models were adjusted just for magnitude of planned advancement. Table 3.2 summarizes these results.

Table 3.2: Crude and adjusted relative risk of VPI/hypernasality comparing DO and TMA [+].

		Outcome Prevalence	Crude PR	Adjusted PR[*]	95% CI
Whole Cohort[**]			*VPI*		
	DO N = 42	17 (40.5)	1.49	1.47	0.72 - 3.05
	TMA N = 86	24 (27.9)			
			Hypernasality		
	DO N = 42	19 (45.2)	1.83	1.64	0.79 - 3.40
	TMA N = 81	20 (24.7)			
UCLP[***]			*VPI*		
	DO N = 20	9 (45.0)	1.63	1.44	0.56-3.67
	TMA N = 48	13 (27.1)			
			Hypernasality		
	DO N = 27	10 (47.6)	2.1	1.93	0.74-5.05
	TMA N = 72	10 (22.2)			

+ Prevalence N(%); PR = prevalence ratio; CI = confidence interval
[*] Adjusted for overjet & cleft type; subgroup models adjusted for overjet
[**] VPI: 9 missing in DO, 64 missing in TMA groups;
Hypernasality: 9 missing in DO, 62 missing in TMA group
[***] 6 missing in DO, 28 in TMA groups

Of patients who underwent DO, 40.5 percent had VPI postoperatively. Adjusting for type of cleft and magnitude of overjet, this was 47 percent (95% CI for the PR = 0.72-3.05) higher than the corresponding prevalence of VPI in the TMA group. The prevalence of post-operative hypernasality among those who underwent DO was 27.5 percent. This was 64 percent (95% CI for the PR = 0.793-3.40) higher than the corresponding prevalence in the TMA group, holding the other two covariates constant. However, all these differences were plausibly the result of chance given no true difference between the two surgical techniques on the post-operative prevalence of VPI or hypernasality.

With respect to VPI, the modest excess among patients in the DO group observed in the entire patient population was also seen when the data for children with UCLP were analyzed separately (Table 3.2). The association between type of surgery and speech outcome among subgroups defined by patients with CPO or BCLP could not be examined with any statistical precision due to small sample sizes.

Neither a history of Furlow palatoplasty (as initial repair or as VPI surgery) nor a history of any VPI surgery appeared to bear on the size of the association between maxillary surgery type on risk of post-operative VPI or hypernasality (Table 3.3). However, the increased prevalence of VPI and hypernasality associated with having undergone DO was largely confined to patients who had a pre-operative VP ratio less than 0.8.

Table 3.3: Association between VPI/hypernasality and surgery type, within strata of prior Furlow, prior VPI surgery, and pre-operative VP ratio.[+]

		Outcome PR, DO vs TMA	Ratio of Ratios	95% CI
Any Furlow*			*VPI*[****]	
	+ Furlow N = 34	1.51	1.12	0.31 - 3.98
	- Furlow N = 92	1.35		
			Hypernasality	
	+ Furlow N = 34	1.53	1.02	0.28 - 3.45
	- Furlow N = 92	1.51		
Any VPI Surgery**			*VPI*	
	+ Surgery N = 50	1.42	1.01	0.33 - 3.10
	- Surgery N = 76	1.39		
			Hypernasality	
	+ Surgery N = 60	1.43	0.93	0.31 - 2.73
	- Surgery N = 76	1.55		
Pre-Op VP Ratio***			*VPI*	
	>= 0.8 N = 47	0.91	0.45	0.12 - 1.70
	< 0.8 N = 67	2.01		
			Hypernasality	
	>= 0.8 N = 46	1.42	0.69	0.24 - 1.98
	< 0.8 N = 67	2.05		

+ PR = prevalence ratio; CI = confidence interval

* 22 missing outcome in +Furlow and 50 missing in -Furlow groups

** 25 missing outcome in +VPI surgery and 47 in -VPI surgery groups

*** 29 missing in >= 0.8 and 27 missing in <0.8 group

**** VPI & hypernasality ratings based on 5-point Likert scales

***** All models adjusted for overjet and type of cleft

3.3.3 Multiple Linear Regression with Categorical Outcome

Multiple linear regression models were used to compare the difference in mean post-operative VPI/hypernasality rating between surgical approaches. For these analyses, all cohort members with post-operative speech evaluations independent of pre-operative VPI/hypernasality status were included. Table 3.3 summarizes these results. After adjusting for type of cleft and magnitude of overjet, mean post-operative VPI severity was 2.69 in the DO group, an average of 0.40 points (95% CI for mean difference = -0.18 – 0.98) higher than the corresponding mean severity rating in the TMA group. The level of post-operative hypernasality also was greater in the patients who underwent DO (mean difference = 0.31, 95% CI = -0.34 – 0.95), though the latter observation was well within the limits of chance given no true difference between the two procedures. The mean difference in severity of VPI and hypernasality in relation to type of surgery that was observed overall was mirrored in the separate results for patients with BCLP and UCLP (Table 3.4). Mean differences in post-operative VPI and hypernasality for the CPO and SMCP groups were not calculated due to very sample size.

Table 3.4: Comparison of mean post-operative VPI/hypernasality severity rating between DO & TMA+.

		Mean Outcome Severity	Mean Difference	
			Unadjusted	Adjusted*
Whole Cohort**			*VPI* *****	
	DO N = 51	2.69	0.33 (-0.11 - 0.78)	0.40 (-0.18 - 0.98)
	TMA N = 107	2.36		
			Hypernasality	
	DO N = 51	1.55	0.30 (-0.18 - 0.77)	0.31 (-0.34 - 0.95)
	TMA N = 107	1.25		
BCLP***			*VPI*	
	DO N = 25	2.62	0.31 (-0.34 - 0.95)	0.29 (-0.85 - 1.43)
	TMA N = 23	2.48		
			Hypernasality	
	DO N = 25	1.43	0.17 (-0.72 - 1.05)	-0.09 (-1.37 - 1.35)
	TMA N = 23	1.26		
UCLP****			*VPI*	
	DO N = 27	2.81	0.62 (0.02 - 1.22)	0.57 (-0.11 - 1.26)
	TMA N = 56	2.19		
			Hypernasality	
	DO N = 27	1.64	0.54 (-0.09 - 1.17)	0.57 (-0.11 - 1.26)
	TMA N = 56	1.1		

+ Mean difference represented as mean (95% CI); CI = confidence interval

* Whole cohort adjusted for overjet & cleft type; subgroup models adjusted for overjet

** 10 missing in DO, 69 missing in TMA groups for VPI & hypernasality

*** 4 missing in DO, 22 in TMA groups for VPI & hypernasality

**** 6 missing in DO, 33 in TMA groups for VPI & hypernasality

***** VPI ratings based on severity scale of 1-5 and hypernasality ratings on a 0-4 severity scale

3.3.4 Sensitivity Analyses

When patients requiring multi-segmental osteotomy in addition to maxillary advancement were excluded from the analysis (2 in the DO group and 40 in the TMA group), the PR for VPI comparing DO and TMA increased to 1.89 (95% CI for the PR = 0.86 – 4.15) from the value of 1.47 among the whole group of patients. With respect to post-operative hypernasality, excluding patients requiring multi-dimensional expansion, the prevalence of post-operative hypernasality was twice as high in the DO group compared to TMA (95% CI for the PR = 0.94 – 4.24), an increase in the PR observed when analyzing the whole cohort (PR=1.64, 95% CI = 0.79 – 3.40).

Because of the differential proportion of missingness of outcome between the two surgery groups, sensitivity analyses were also run to simulate the effect different proportions of outcome among those missing would have had on observed exposure-outcome relationship. It is often assumed that patients who have no speech concerns have a lower likelihood of returning for follow-up, and thus would have a lower likelihood of having been assessed for the presence of post-operative VPI and hypernasality. Because a relatively larger proportion of patients who had undergone TMA had not received a post-operative speech evaluation, it is plausible that the prevalence of decline in velopharyngeal function in this group was spuriously high. If it is assumed that 10% of those with unobserved outcome truly did have VPI or hypernasality, this would increase the number of participants with presence of outcome by one (10% of 9 with unobserved outcome) in the DO group and by six (10% of 62 with unobserved outcome) in the TMA group.

Table 3.5: Observed versus predicted outcome prevalence based on assumption of 10% of unobserved patients having hypernasality.

	Observed N	Predicted N with 10%	Proportion Missing	Unweighted Prevalence	Weighted Prevalence
DO	19	20	17.6%	39.2%	31.0%
TMA	20	26	43.7%	18.3%	9.78%

If the assumption holds that only 10% of those with unobserved outcome truly had VPI or hypernasality, and accounting for the differences in proportion of unobserved outcome in the two surgery groups, the proportion of outcome in both groups would decrease, but the magnitude of decreased prevalence would be greater in the TMA group (as seen in Table 3.5). If true, then the PRs based on observed outcome may be biased towards the null, as a higher PR would be observed if the weighted prevalences above are an accurate reflection of the proportion of outcome in the total cohort of those with DO and TMA. While the values in the table above reflect calculations for hypernasality, similar trends would be observed for post-operative VPI.

3.4 Discussion & Conclusion

Our results suggest a modest increase in the prevalence of VPI and hypernasality in patients who have undergone DO relative to that among patients who have undergone TMA, after adjusting for relevant confounders (PR for VPI = 1.47, 95% CI 0.72 – 3.05, PR for hypernasality = 1.64, 95% CI 0.79 – 3.40). Similarly, patients who underwent DO had a slightly higher mean post-operative VPI rating than patients who underwent TMA (mean severity 2.69 versus 2.36, respectively, 95% CI for the adjusted mean difference -0.18 – 0.98). However, in the absence of any true difference in VPI and hypernasality prevalence following each type of surgery, the observed differences quite plausibly could have been the result of chance.

Prior history of either Furlow palatoplasty or any VPI surgery did not modify observed associations. In patients with a VP ratio of < 0.8 on pre-operative cephalometric radiographs, DO was associated with a two-fold increase in risk of post-operative VPI and hypernasality (PR for VPI = 2.01, 95% CI 0.79 – 5.10; PR for hypernasality = 2.05, 95% CI 0.81 – 5.80). Though it would be reasonable to attribute this increase in risk among those with smaller pre-operative VP ratio to a need for greater magnitude of anterior movement, these results were already adjusted for overjet. Nonetheless, it is possible that adjustment for a more accurate measure of the magnitude of maxillary deficiency --such as the distance from A-point to the line perpendicular

to the x-axis at sella turcica – would have led to a PR closer to the null. A modest increase in prevalence of VPI and hypernasality was also seen in patients requiring only anterior-posterior advancement. Sensitivity analyses did not suggest obvious residual confounding.

An important limitation of this study is the absence of complete data on all study participants. Both differential return for follow-up between the two surgery groups, as well as incomplete measurement of the primary confounder, magnitude of advancement, threaten the validity of our analyses to some degree. With respect to missing outcome assessments, it is plausible that those with no speech concerns were less likely to return for follow-up. Because relatively more patients who underwent TMA did not receive a post-operative speech evaluation, the observed small differences in risk of post-operative VPI or hypernasality between surgical techniques may in truth be larger. Missing measurement of the magnitude of maxillary advancement was typically due to the absence or poor quality of the pre-operative lateral cephalometric radiograph in the EMR. An upgrade in x-ray technology after 2012 meant that x-rays classified as untraceable were more likely to occur earlier in the study period. Because cephalometric data were missing to a similar degree in patients undergoing each type of surgery, the magnitude of bias from this source is likely to be small. Missing data in both the outcome and primary confounder variables also reduced the number of complete cases available for analysis, diminishing our sample size to a degree that we were unable to reliably identify a true small difference between surgery groups.

Our approach to reliability measures for VPI and hypernasality ratings also warrants mention. In the presence of imperfect reliability, errors in the assessment of post-operative VPI and hypernasality could have occurred. However, the degree of misclassification would not be expected to differ between surgery groups, and so any misclassification would likely be non-differential.

In a trial of patients with CP/L requiring 4-10 mm of advancement, 47 patients were randomized to receive either DO (N = 22) or TMA (N = 25) [15]. Post-operative data on speech

were available for only ten subjects in the DO group and 11 subjects in the TMA group. There was no difference in change in hypernasality between surgery groups at either three- or 12-months post-operatively ($\chi^2 = 2.2$, $p > 0.05$; $\chi^2 = 0.39$, $p > 0.05$, respectively). Comparable results were obtained for other measures of post-operative velopharyngeal function analyzed in their study, including nasalance values, nasoendoscopic findings, and nasal air emission. A cohort study of patients with moderate to significant maxillary hypoplasia (> 10 mm of advancement required) compared prevalence of post-operative VPI between 20 patients who underwent DO and 11 who underwent TMA, defining presence of VPI as a score of greater than three on the Pittsburgh Weighted Speech Scale (PWSS) [16,81]. Post-operatively, 82 percent of patients who received TMA and 45 percent of patients who received DO met the criteria for VPI (p-value for t-test of the difference was < 0.05, value of t-statistic not reported). The prevalence ratio was 0.55, suggesting that receipt of DO was associated with a 45 percent *decrease* in risk of post-operative VPI compared to TMA in their study population. While overjet was identified as the best variable to use for adjustment of confounding by indication in the current study, overjet may result in misclassification with respect to the actual anterior maxillary movement required among patients requiring bimaxillary surgery. In the two prior studies, patients requiring concurrent mandibular surgery were excluded, whereas in the current study those requiring concurrent mandibular surgery were present only in the TMA group. Therefore, our results may reflect some degree of differential misclassification of the confounder, contributing to the observed increased risk of VPI among DO compared to TMA.

Overall, there is insufficient evidence to determine whether the type of maxillary advancement surgery influences the risk of impaired velopharyngeal function. Similarly, it is unknown whether a particular type of maxillary advancement surgery has a deleterious impact on velopharyngeal function in just some subgroups of patients. Pre-operative VP ratio and need for multi-dimensional expansion are suggested as possible characteristics that warrant further study in this regard.

In conclusion, our results suggest modestly poorer velopharyngeal function among patients who had undergone DO compared to those who had undergone TMA, controlling for differences in type of cleft and magnitude of planned maxillary advancement. However, the results of earlier (quite small) studies are not in accord with this. Additional contributions to the body of evidence bearing on the issue are warranted.

Chapter 4: Predictors of Decline in Velopharyngeal Function Following Maxillary Advancement

4.1 Introduction

Because of both congenital and iatrogenic differences in their velopharyngeal anatomy, patients with cleft palate with or without cleft lip (CP/L) are at a high risk of developing velopharyngeal insufficiency (VPI) after surgical correction of maxillary hypoplasia [82-85]. Individuals with CP/L typically have an extensive surgical history, and risk of VPI following maxillary advancement surgery is a significant enough concern that anticipatory guidance around this risk is often a component of pre-operative evaluation and counseling. However, at present there is limited information regarding pre-operative anatomic and perceptual speech characteristics that predict deterioration in velopharyngeal function following maxillary advancement surgery in patients with CP/L.

4.1.1 Magnitude of Maxillary Advancement

Increasing risk of VPI with increasing magnitude of maxillary movement necessary to achieve functional occlusion is a plausible hypothesis, and this hypothesis is supported in some studies [41,86-88]. Other investigators report contradictory findings [40,89-93]. All of these studies, however, defined the amount of advancement as the *change in* A-point based on superimposition of pre- and post-operative lateral cephalograms. As complications of maxillary advancement surgery inherent in individuals with CP/L suggest [22,26], the predicted magnitude of advancement that informs type of procedure performed and whether concurrent mandibular surgery is indicated may not reflect the amount of sagittal movement achieved post-operatively. In the absence of evidence confirming a strong correlation between predicted and actual millimeters of advancement achieved among patients with CP/L, the validity of comparing location of pre- versus post-operative reference points as a predictor of VPI is unknown.

Even in studies that measured advancement from pre-operative cephalometric radiographs or surgical planning models, identification of *planned* advancement as a predictor of VPI outcome is unclear. A retrospective cohort study of 26 patients with CP/L defined predicted anterior movement of the soft palate from surgical planning models and found that there was no appreciable increase in the presence of post-operative VPI associated with a one-millimeter increase in soft palate advancement (OR = 0.98, 95% CI not reported) [94]. However, it is possible that defining the predictor of interest as a continuous variable missed a true association present when comparing two groups that differ by a larger magnitude of maxillary advancement. Additionally, the authors did not report how closely their predicted and actual advancement measurements correlated with each other. A recent descriptive study measured planned maxillary advancement from pre-operative cone-beam computed tomography (CBCT) scans, although reference points used in this measurement were not reported and post-operative CBCT scans were not available to test correlation between planned and actual magnitude of advancement [95]. The investigators observed no association between their definition of planned advancement and VPI outcome. Small sample size, however, limited the ability of this study to detect an association. In summary, the predictive value of anticipated amount of maxillary advancement, as identified during pre-operative planning, on the postoperative prevalence of VPI remains uncertain.

4.1.2 Morphology of the Velopharyngeal Port at Rest

Static dimensions of the velopharyngeal port at rest have also been suggested as predictors of decline in velopharyngeal sufficiency after maxillary advancement surgery in patients with CP/L. Short pre-operative velar length as identified on lateral cephalometric radiographs has been associated with the need for pharyngeal flap surgery as well as increased prevalence of hypernasality post-operatively [75,93,96]. A recent study utilizing measurements obtained via CBCT scans rather than lateral cephalometric radiographs, however, did not

observe velar length to be predictive of VPI after maxillary advancement in individuals with CP/L [17]. It was possible that the relatively greater reliability of soft tissue measurements on CBCT scans reduced the amount of measurement error that may have been present in soft tissue measurements based on lateral cephalograms [97].

The magnitude of the depth of the nasopharynx relative to velar length ("VP ratio") has been investigated in cross-sectional studies as association with velopharyngeal function. A ratio of 1.0 is reported as the threshold between normal and insufficient velopharyngeal function in some studies [75], and in others a value of around 0.8 is suggested [76,98]. In a cohort of eight patients with CP/L, investigators reported mean pre-operative VP ratio to be 0.75 (SD 0.12). The two patients with pre-operative VPI had VP ratios of 1.0. After maxillary advancement, the mean VP ratio among those without pre-operative VPI was 0.8 (SD 0.20) while the VP ratio among the two with pre-operative VPI increased to 1.2 [75]. Based on these findings, authors suggested that the VP ratio present after maxillary advancement can be predicted by adding the number of millimeters of anticipated advancement to the numerator and adding 0.4 times the millimeters of advancement to the denominator. The association between this "corrected" VP ratio as a predictor of VPI after maxillary advancement was later investigated in an observational study of 24 patients with BCLP and UCLP [91]. When velopharyngeal function as determined via pressure-flow measurements of velopharyngeal gap size [99] was used as an outcome definition, eight patients in the cohort were considered likely to develop post-operative VPI. Of these eight, however, just three developed VPI after maxillary advancement and one patient that had a predicted VP ratio of less than 1.0 also developed post-operative VPI. Of note, the mean maxillary advancement for both studies was 4.0 with the largest advancement under 10.0 millimeters. Whether the predicted VP ratio calculated based on pre-operative measurements correlated well with the actual VP ratio present on post-operative evaluation was not reported. Because of advances in surgical technique, the three-dimensional complexity of both the maxillary deficiency as well as the function of the velopharyngeal port, the accuracy of this

calculation to predict VPI is unclear. Studies evaluating the magnitude of the VP ratio present before maxillary surgery as an independent predictor of post-operative VPI have not been completed to date. Other morphological aspects of the velopharyngeal port, including velar inclination angle (angle formed by the lines drawn from ANS-PNS and from PNS to Uv) [40,75] and a short maxilla (linear distance between A-point to PNS) [89], have also been considered, but studies of these aspects have been small and have had limited statistical power.

4.1.3 Pre-Operative VPI Status

Pre-operative status of velopharyngeal function has also been identified as a possible predictor, although how this status is defined has varied across studies [81,94,100-102]. Visualization of a small gap on nasendoscopy has been associated with VPI in some studies [94,103], but not in others [104]. It has also been suggested that pre-operative perceptual ratings of VPI severity are associated with post-operative VPI outcome. In a study of 26 individuals, patients identified as "at-risk" for VPI after maxillary advancement based on presence of any degree of hypernasality pre-operatively (n = 9) all presented post-operative VPI [94]. Just two of 16 participants identified as "not at risk" based on pre-operative perceptual evaluation presented with VPI post-operatively. In a large retrospective cohort study [101], the proportion of those with a decline in velopharyngeal function after maxillary advancement was just 4% in the group that presented with no VPI pre-operatively (n = 71 with VPI rating = 0). In contrast, half of the 18 that demonstrated borderline pre-operative VPI (VPI rating = 1) experienced a deterioration in velopharyngeal function after maxillary advancement, with VPI severity ranging from mild to severe. Of those presenting with mild or moderate VPI pre-operatively (n = 11), 45% were rated as having more severe VPI post-operatively.

4.1.4 Type of Cleft

While investigators have reported differences in airway and pharyngeal dimensions between different types of CP/L [30,105], type of cleft as an independent predictor of VPI after maxillary advancement has not been extensively studied. One study found that the proportion of those with deterioration of velopharyngeal function was higher among those with bilateral cleft lip and palate (BCLP) than those with unilateral cleft lip and palate (UCLP) (50% versus 24% or 33% in TMA and DO groups, respectively), although these differences were within the limits of chance given no true difference [41]. In another study, 9% and 17% of individuals with UCLP (n=53) and CPO (n=23), respectively, demonstrated a decline in velopharyngeal function compared to nearly 30% of those with BCLP (n=24) [101]. While differences in risk of post-operative VPI after maxillary advancement by type of cleft is a plausible variable to investigate as an outcome predictor, there are many differences in early surgical care dependent on type of cleft. Therefore, it is possible that cleft type does not have a direct influence on VPI outcome after maxillary advancement, but that the influence of cleft type on VPI outcome is mediated through early surgical history.

Identification of speech and cephalometric characteristics present pre-operatively that put a patient at increased risk of VPI after maxillary advancement requires accounting for other variables that may be directly or indirectly related to that characteristic and VPI outcome. For example, it is plausible that palate length is associated with either type of initial palate repair and/or history of VPI surgery. Similarly, patients with both a shorter soft palate and need for a large amount of anterior movement could reasonably be at greater risk for post-operative VPI than those with only one of these characteristics. The complexity of potential relationships between past surgical history, diagnosis, cephalometric measurements, and perceptual speech findings has not been adequately investigated to date.

Maxillary advancement surgery is often considered the end of the timeline of care and occurs at a time of great transition in a young adult's life. Therefore, informed clinical decision-making that includes information regarding pre-operative characteristics put a patient at increased risk of VPI and need for further surgical and/or clinical management is of great interest to providers, patients, and their families. The current study uses multivariable regression analysis to not only directly compare individuals presenting with presence or absence of potential predictors of worsening velopharyngeal function after maxillary advancement, but to also evaluate whether combinations of these characteristics within individuals further increase or mitigate risk in individuals with CP/L undergoing surgical maxillary advancement.

4.2 Methods

4.2.1 Participants

For this retrospective cohort study, eligible participants were identified from a clinical database that includes all patients of the Seattle Children's Craniofacial Center. All patients with CP/L who underwent LeFort I osteotomy for maxillary advancement between January 2007 and June 2019 were considered eligible. For inclusion, patients had to have either an isolated or syndromic secondary palatal cleft with or without a cleft lip. Patients with cleft lip/alveolus only were excluded, as were patients with repeat LeFort I osteotomy procedures or those with LeFort I osteotomy procedures not involving anterior movement of the maxilla. Demographics, surgical history, and medical diagnoses were obtained from the Craniofacial Center's clinical database (CRA DB) and verified in the electronic medical record. All participants underwent a pre-operative orthodontic evaluation, standard surgical evaluation with cephalometric radiographs, and pre-/post-operative perceptual evaluation of velopharyngeal function per clinical protocol. All surgeons performing the procedures were board-certified in either plastic and reconstructive surgery, oral-maxillofacial surgery, or both, and specialized in the management of children with craniofacial differences. This study was approved through expedited review by the Institutional Review Board at the University of Washington (IRB #STUDY00008156). Because all data

utilized in the study were present in the electronic medical record or in internal clinical databases, written consent/assent was waived.

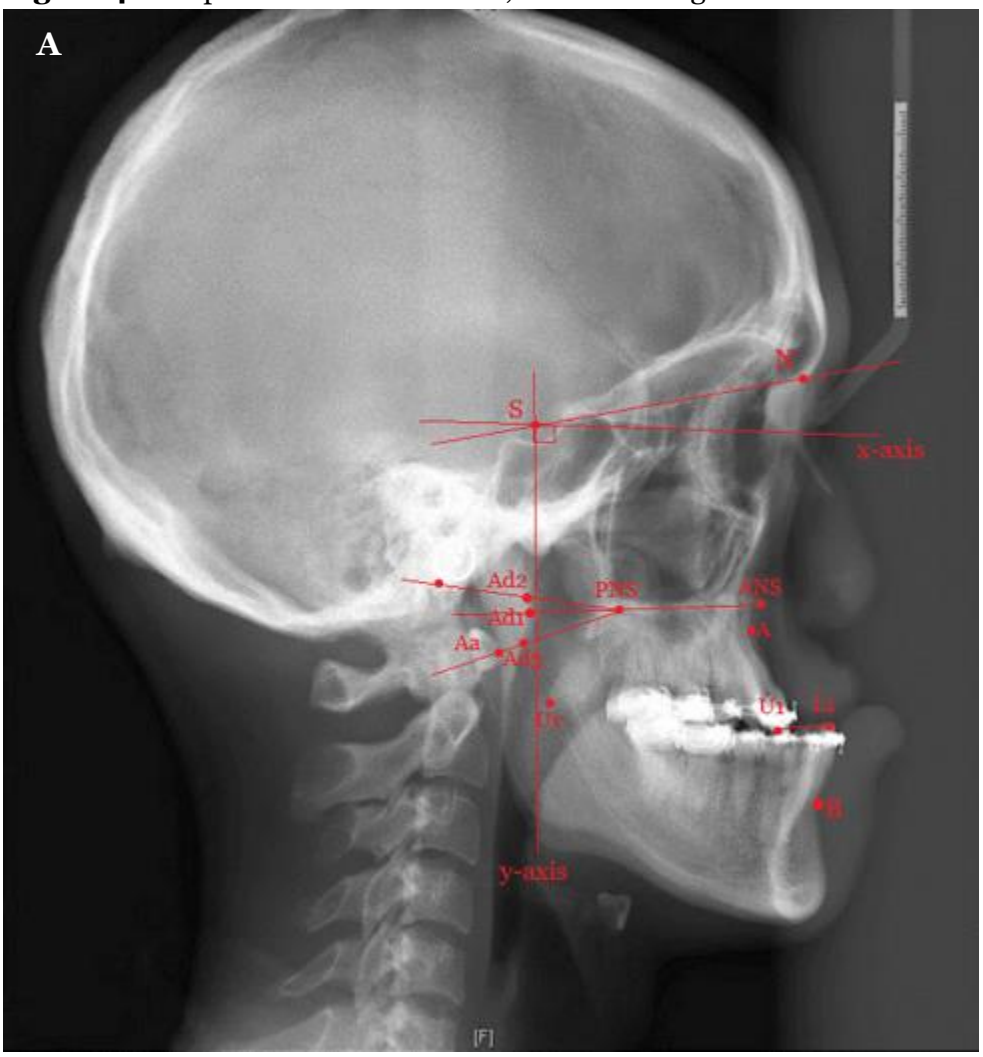
4.2.2 Surgical Characteristics

All participants underwent LeFort I osteotomy with pterygomaxillary disjunction and down-fracture, as described in the literature [22,24,27]. Additional traction techniques were applied to ensure complete maxillary mobilization, as necessary. Pertinent surgical details including type of surgical approach, need for further division of the maxilla, and concurrent procedures performed (i.e., pharyngeal flap release, mandibular osteotomy, and alveolar bone graft) were identified directly from operative reports. All clinical decisions relating to surgical technique were made as indicated by presurgical evaluation.

4.2.3 Lateral Cephalometric Measurement

A lateral cephalometric radiograph was completed pre-operatively and serially post-operatively, per protocol. All lateral radiographs were taken using the same cephalostat with patient aligned to Frankfort horizontal, molars in occlusion, and with the lips and soft palate at rest. For analysis, the radiograph closest to the date of surgery pre-operatively and to the date of perceptual velopharyngeal function evaluation post-operatively were used. Two orthodontists with the Craniofacial Center completed digitized tracings of pre- and post-operative radiographs for each participant with Dolphin Imaging Premium 11.95 software. Magnification differences were accounted for in tracings. For reference, the horizontal axis was established by dropping a line seven degrees from the center of sella turcica (S) to nasion (N) [18,73,74]. The vertical axis was then identified by a line drawn perpendicular to the x-axis through S. Documented landmarks are illustrated in Figure 4.1A. Definitions of linear and angular measurements using these landmarks are summarized in Figure 4.1B.

Figure 4.1: Cephalometric landmarks, linear and angular measures *.



* ANS = anterior nasal spine, PNS = posterior nasal spine, Uv = tip of uvula, A = deepest point on anterior contour of maxilla, S = center of sella turcica, N = nasion, Ba = Basion, Aa = atlas, Ad1 = intersection of palatal plane and posterior pharyngeal wall, Ad2 = intersection of Ba-PNS line and PPW, Ad3 = intersection of line Aa-PNS and PPW, U1 = upper incisor tip, L1 = lower incisor tip, B = deepest point on anterior contour of mandible

Figure 4.1B:

Linear Measures (mm)	Definition
Vertical maxillary deficiency	Linear distance from A point along line drawn perpendicular to x-axis
Anterior-posterior maxillary deficiency	Linear distance from A point along line drawn perpendicular to y-axis
Soft palate length	Linear distance from PNS to Uv
Nasopharyngeal depth	1. Linear distance from PNS to Ad1 2. Linear distance from PNS to Ad2 3. Linear distance from PNS to Ad3

Overjet	Anterior-posterior projection of upper beyond lower incisor tips as measured parallel to occlusal plane relative to upper & lower incisor tips
Angular Measures (degrees)	**Definition**
Cranial base angle	Angle formed by lines S-N and S-Ba
Palatal plane angle	Angle formed by ANS-PNS plane where it intersects x-axis
ANB	Angle formed by A point-N-B point; measures anterior-posterior discrepancy between maxilla and mandible
SNA	Angle formed by S-N-A point
Velar inclination angle	Angle formed by ANS-PNS and PNS-Uv; angulation of soft to hard palate

In addition to measures described in Figure 4.1, velopharyngeal need ratio, as defined by the nasopharyngeal depth in the numerator and velar length in the denominator [75], was calculated for all three definitions of nasopharyngeal depth. Ten percent of radiographs were traced by both orthodontists to evaluate inter-rater reliability. Each orthodontist also retraced ten percent of radiographs to evaluate intra-rater reliability.

4.2.4 Perceptual Evaluation of Velopharyngeal Function

Pre- and postoperative speech data were obtained through a second clinical database (VPI DB) for each eligible participant identified through the CRA DB. Perceptual evaluation of velopharyngeal function was performed by one of six speech-language pathologists (SLPs) specializing in the diagnosis and management of velopharyngeal dysfunction, although three SLPs contributed ratings for less than three years of the 12-year study period. Per clinical protocol, preoperative evaluations occurred within one year of maxillary advancement surgery and post-operative evaluations approximately three months after surgery. For each patient, ratings of VPI severity, hypernasality and hyponasality were stored in the VPI DB at the time of evaluation. Ratings were based on overall clinical impression, and each evaluation included syllable and word repetition, pressure consonant loaded sentence repetition, counting 1-20 and 60-70 and a spontaneous speech sample. VPI severity and hypernasality were rated on 5-point severity scales and hyponasality on a 3-point severity scale utilizing definitions described in

Table 2.1. To establish interrater reliability with this scale, 50 audio recordings of syllable repetition, sentence repetition and counting 1-20 and 60-70 (Appendix A) of patients of similar age and diagnosis were rated using the above criteria by three of the six SLPs contributing ratings over the study period to establish inter-rater reliability. Recordings were obtained using a Sony Digital ICD Recorder in a quiet exam room. Each of the three SLPs re-rated 20% of audio recordings two weeks after initial ratings to measure intra-rater reliability.

4.2.5 Statistical Analysis

Variables investigated as potential predictors of post-operative VPI included the following: (1) magnitude of maxillary hypoplasia, as defined by the linear distance from A-point to a point perpendicular to the y-axis on pre-operative cephalometric radiographs, (2) type of cleft (BCLP, UCLP, or CPO/SMCP), (3) presence of borderline VPI, as defined by a rating of one ("minimal VPI") on pre-operative speech evaluation, and (4) the ratio of nasopharyngeal depth with Ad1 as the posterior landmark) to velar length, termed "VP ratio". The primary outcome of interest was presence versus absence of post-operative VPI, with presence of VPI defined as a rating of 2-4 and absence defined as a rating of 0 or 1 on post-operative perceptual speech evaluation. Relative risk regression using the Poisson likelihood function [77] was run to investigate the ability of each exposure of interest to predict presence of VPI after maxillary advancement surgery. Patients presenting with pre-operative VPI rated at 2 or higher were excluded from analyses. Sandwich robust standard errors were used to calculate confidence intervals. Relevant confounders identified through a combined *a priori* clinical knowledge and data driven approach were identified for each exposure-outcome association separately [69,70]. Models investigating the relationship between magnitude of maxillary hypoplasia and prevalence of post-operative VPI were adjusted for type of cleft and age at time of surgery. When investigating cleft type as a primary predictor, dummy variables were used to allow comparison of VPI prevalence between types of cleft with CPO/SMCP identified as the referent category. The

association between cleft type and outcome was adjusted for presence versus absence of a syndrome. Whether this association was mediated through the magnitude of maxillary hypoplasia was also investigated. Models testing whether the presence of borderline VPI pre-operatively predicted post-operative VPI were adjusted only for syndrome status. Finally, models investigating the relationship between pre-operative VP ratio and outcome were adjusted for prior VPI surgery of any type. Bias resulting from the failure of some patients to return for follow-up was not a strong concern, as no relationship between a characteristic predictive of missingness and outcome was identified. Therefore, it was assumed that outcome data were missing completely at random (MCAR), and complete-case analysis was considered a valid approach [78].

To determine if the results differed using alternative definitions of nasopharyngeal depth, models investigating VP ratio as a primary predictor were re-run using the Atlas and Basion as the posterior landmark for nasopharyngeal depth. For investigation of magnitude of maxillary hypoplasia as a primary predictor of post-operative VPI, analyses were re-run by using both the ANB angle and overjet as the predictor definition to determine if different exposure definitions had any bearing on prevalence of post-operative VPI.

Two-way mixed consistency single-measures intra-class correlation coefficients (ICCs) were calculated for inter- and intra-rater reliability of all cephalometric measurements [46]. For perceptual speech ratings, weighted kappa [45] and two-way mixed consistency, single-measures ICCs were calculated for intra- and inter-rater reliability, respectively. Interpretation of reliability coefficients as defined by Cicchetti et al were used [47]. Independent samples t-tests were used to evaluate consistency of measurements within and between orthodontists contributing cephalometric tracings. All analyses were completed in R version 4.0.3.

4.3 Results

After excluding all patients who had pre-operative VPI (defined as a severity rating of 2-4), a total of 200 patients underwent LeFort I osteotomy for maxillary advancement over the 12-year study period. Seventy-three of these did not return for post-operative speech evaluation, leaving 127 patients available for analysis. A little over 30 percent of these patients presented with VPI post-operatively (severity rating between 2-4). Table 4.1 summarizes the distribution of demographic, clinical, surgical, and cephalometric characteristics of patients analyzed in the study, stratified by presence versus absence of post-operative VPI. Pre-operative cephalometric radiographs completed more than 100 weeks before surgery were considered missing. Two additional patients had pre-operative cephalometric radiographs at 78 and 64 weeks before surgery. These patients remained in the study population, as both were more than 21 years of age at the time of surgery, thus mitigating concern for significant growth between these timepoints. After these considerations, median time between pre-operative cephalometric radiographs and date of surgery was 2.93 weeks (IQR 2.14-8.04). Median time between pre-operative speech evaluation and surgery date was 21.6 weeks (IQR 3.00-76.6). Absence of interim VPI surgery was confirmed for all participants.

Table 4.1 Participant characteristics, stratified by presence versus absence of outcome (VPI).

	Post-Operative VPI*			
	Presence N = 41		Absence N = 86	
Demographics				
Age at surgery	16.70	(2.49)	17.70	(2.57)
Proportion male	25	(61.0)	54	(62.8)
*Race***				
Asian	9	(22.0)	16	(18.6)
Other	14	(34.1)	30	(34.9)
White	17	(41.5)	36	(41.9)
Residence by Zip Code				
Eastern WA	5	(12.2)	7	(8.1)
Out of State	3	(7.3)	11	(12.8)
Western WA	33	(80.5)	67	(77.9)
*Type of Medical Insurance****				

Financial assistance only	2	(4.9)	3	(3.5)
Private	23	(56.1)	47	(54.7)
Public	12	(29.3)	30	(34.9)
Clinical Characteristics				
Type of Cleft				
BCLP	12	(29.3)	24	(27.9)
CPO	7	(17.1)	11	(12.8)
SMCP	0	(0.0)	6	(7.0)
UCLP	22	(53.7)	45	(52.3)
Presence of a Syndrome	10	(24.4)	15	(17.4)
Borderline Pre-Op VPI ****	15	(36.6)	20	(23.3)
*Pre-Op Hyponasality*****				
None	35	(85.4)	71	(82.6)
Mild	2	(4.9)	7	(8.1)
Marked	0	(0.0)	1	(1.2)
Post-Op Hyponasality				
None	41	(100.0)	78	(90.7)
Mild	0	(0.0)	7	(8.1)
Marked	0	(0.0)	1	(1.2)
Surgical History				
Initial Palate Repair				
Furlow	1	(2.4)	1	(1.2)
IVV	25	(61.0)	54	(62.8)
None	0	(0.0)	5	(5.8)
Unknown	15	(36.6)	26	(30.2)
VPI Surgery				
Furlow	25	(29.1)	10	(24.4)
Sphincter pharyngoplasty	4	(9.8)	11	(12.8)
Pharyngeal Flap	3	(7.3)	5	(5.8)
Other	1	(2.4)	5	(5.8)
Type of Maxillary Advancement				
Distraction osteogenesis	17	(41.5)	24	(27.9)
Traditional advancement	24	(58.5)	62	(72.1)
*Multi-segmental osteotomy******	5	(12.2)	14	(16.3)
Pre-Operative Cephalometrics******				
Linear Measures (in mm)				
Vertical Deficiency	55.42	(10.53)	54.45	(9.80)
Horizontal Deficiency	51.63	(7.57)	51.72	(8.55)
Velar Length	25.98	(6.88)	27.17	(6.75)
Nasopharyngeal Depth				
Per palatal plane (Ad1)	20.71	(5.40)	21.26	(5.34)
Per Basion (Ad2)	20.99	(5.55)	21.37	(4.86)
Per Atlas (Ad3)	25.86	(6.49)	25.63	(5.32)
Maxillary Length	39.91	(7.05)	39.36	(7.27)

Overjet	-6.51	(6.09)	-5.78	(6.34)
Angular Measures (in degrees)				
Cranial Base Angle	129.63	(6.47)	131.73	(6.21)
ANB Angle	-5.38	(4.93)	-4.66	(5.31)
Velar Inclination Angle	126.24	(10.32)	128.31	(9.69)
Ratio Measures				
VP Ratio per Ad1	0.84	(0.28)	0.81	(0.26)
VP Ratio per Ad2	0.85	(0.29)	0.82	(0.25)
VP ratio per Ad3	1.05	(0.33)	0.99	(0.32)

· Continuous variables represented as mean (SD); categorical/binary as N (%)

* 73 participants did not return for post-operative evaluation

** 1 missing in +VPI group; 4 missing in no VPI group

*** 4 missing in +VPI group; 6 missing in no VPI group

**** 4 missing in +VPI group; 7 missing in no VPI group

***** 60 with multi-dimensional osteotomy did not return for follow-up

****** 3 missing in +VPI group; 11 missing in no VPI group

Demographic characteristics did not differ between those who had post-operative VPI and those who did not. UCLP was the most prevalent type of cleft; the distribution of types of cleft did not meaningfully differ between the two outcome groups. A modestly larger proportion of those with post-operative VPI had a syndrome as well as borderline pre-operative VPI (severity rating of 1). Of those who did not have post-operative VPI, more than 70 percent of these patients underwent traditional osteotomy rather than distraction. The distribution of traditional versus distraction surgical technique within the group presenting with VPI post-operatively was more balanced, although a larger proportion of those who underwent traditional advancement was present. Mean linear and angular cephalometric measurements were comparable between those who did and did not have VPI post-operatively.

4.3.1 Reliability Results

Two-way mixed consistency single-measures intra-class correlation coefficient (ICC) analyses revealed good to excellent inter- and intra-rater reliability for all cephalometric measures used in analyses (ICCs ranging from 0.75 to 0.99). ICCs for inter-rater reliability of VPI and hypernasality were 0.77 and 0.75, respectively, also revealing good agreement across

the three raters. Weighted kappa values revealed good to excellent intra-rater reliability for each SLP, with values ranging from 0.69 to 0.95 for VPI and hypernasality.

4.3.2 Evaluation of Predictors

Whether each exposure of interest predicted presence of post-operative VPI was assessed using relative risk regression using the Poisson likelihood function with sandwich robust standard errors. As previously indicated, adjustment for confounding was made for each exposure-outcome association individually. Table 4.2 summarizes these results.

Table 4.2: Crude and adjusted VPI prevalence ratios for each predictor of interest [+].

	Unadjusted		Adjusted	
Predictor of Interest	**PR**	**95% CI**	**PR**	**95% CI**
Magnitude of Maxillary Deficiency[*]				
A point to x-axis (mm)	1.01	0.98 - 1.03	1.00	0.97 - 1.03
Overjet (mm)	0.99	0.93 – 1.03	1.00	0.96 - 1.05
ANB (degrees)	0.98	0.95 – 1.03	1.00	0.95 - 1.05
Borderline pre-op VPI[**]	1.58	0.93 - 2.67	1.57	0.93 - 2.67
Pre-Op VP Ratio[***]				
Per PNS-Ad1 NP depth	1.18	0.45 - 3.10	1.19	0.45 - 3.15
Per PNS-Ad2 NP depth	1.31	0.51 - 3.42	1.33	0.51 - 3.46
Per PNS-Ad3 NP depth	1.35	0.66 - 2.76	1.33	0.66 - 2.77

[+] PR=prevalence ratio; CI=confidence interval; VP=velopharyngeal; NP=nasopharyngeal
[*] Adjusted for age and cleft type
[**] Adjusted for syndrome status
[***] Adjusted for history of any VPI surgery

The prevalence of VPI was not related to the magnitude of maxillary deficiency (adjusting for age at time of surgery and type of cleft, the PR associated with a one mm change was 1.01, 95% CI 0.97-1.03). There was no observed change in the exposure-outcome association when defining magnitude of maxillary deficiency by either overjet or the ANB angle. Of those presenting with borderline VPI pre-operatively, 36.6% presented with VPI post-operatively. After accounting for presence of a syndrome, this prevalence was 57% (95% CI for the PR 0.93 – 2.67) higher than that among patients with normal pre-operative velopharyngeal function (PR 1.57, 95% CI 0.93 – 2.67). When defining nasopharyngeal depth by the linear distance from PNS

to Ad1, patients with one-unit higher ratio of nasopharyngeal depth to soft palate length were slightly (PR 1.19, 95% CI 0.45 – 3.15) more likely to have post-operative VPI, after adjusting for history of any VPI surgery. A small increase in risk of post-operative VPI for patients with one-unit higher ratio of nasopharyngeal depth to soft palate length when using Ad2 or Ad3 as the posterior landmark for nasopharyngeal depth (PR for Ad2 1.33, 95% CI 0.51 – 3.46; PR for Ad3 1.33, 95% CI 0.66 – 2.77), holding other covariates constant. However, all observed differences were plausibly the result of chance given no true difference.

Relative risk regression was also used to test the association between type of cleft and prevalence of post-operative VPI. However, to allow comparison between types of clefts, dummy variables were created with CPO/SMCP as the referent category. Unadjusted and adjusted PRs for each type of cleft are summarized in Table 4.3.

Table 4.3: Prevalence ratios comparing proportion of those with post-operative VPI by type of cleft, with secondary cleft palate as the referent category.

		Unadjusted		**Adjusted***	
	Prevalence of VPI	**PR**	**95% CI**	**PR**	**95% CI**
CPO/SMCP (referent)**	7 (29.1)	(ref)	(ref)	(ref)	(ref)
BCLP	12 (33.3)	1.14	0.51 - 2.56	1.59	0.64 - 3.96
UCLP	22 (32.8)	1.13	0.54 - 2.35	1.67	0.66 - 4.27

+ Prevalence N(%); PR = prevalence ratio; CI = confidence interval

* Adjusted for presence of a syndrome

** Lost to follow-up: CPO/SMCP - 13; BCLP - 25; UCLP - 35

One third of patients with BCLP presented with post-operative VPI, a value 59% higher (95% CI for the PR 0.64 – 3.96) than the corresponding prevalence among those with CPO/SMCP after adjusting for presence of a syndromic cleft palate. The prevalence of post-operative VPI among those with UCLP was 32.8 percent and this was 67% higher (95% CI for the PR 0.66 – 4.27) than the prevalence of VPI among those with CPO/SMCP after adjusting for syndromic cleft palate.

However, these results are compatible with chance, given no true associations. To test whether the association between cleft type and post-operative VPI is mediated through the magnitude of maxillary deficiency, models were further adjusted for the linear distance between A point and the x-axis. Observed differences after adjusting for the direct versus total effect of cleft type of post-operative VPI did not meaningfully differ from those shown above (PR comparing BCLP to CPO/SMCP 1.52, 95% CI 0.61 – 3.81; PR comparing UCLP to CPO/SMCP 1.71, 95% CI 0.67 – 4.33), suggesting that cleft type does influence post-operative VPI beyond its impact on magnitude of maxillary deficiency.

4.4 Discussion & Conclusion

The magnitude of maxillary deficiency, as defined by the linear distance from A-point to the y-axis, was not associated with the prevalence of post-operative VPI after adjusting for age at time of surgery and type of cleft (PR associated with one mm change was 1.01, 95% CI 0.97-1.03). Re-defining magnitude of maxillary deficiency as either overjet or the ANB angle did not alter the observed null relation between exposure and outcome. Conditional on presence of syndromic versus non-syndromic cleft palate, there was a suggestion that patients with borderline VPI pre-operatively had an increased prevalence of post-operative VPI relative to those with normal velopharyngeal function prior to maxillary advancement (PR 1.57, 95% CI 0.93 – 2.67). A modest increase in prevalence of post-operative VPI was observed among patients with a one-unit higher VP ratio when using the point at which the palatal plane crosses the posterior pharyngeal wall as the posterior landmark for nasopharyngeal depth after adjusting for history of VPI surgery (PR associated with one unit change 1.19, 95% CI 0.45 – 3.15). This difference in risk increased to 33% when using alternative posterior landmarks to define nasopharyngeal depth (PR associated with one-unit change in VP ratio 1.33 for both, 95% CIs 0.51-3.46 and 0.66-2.77). Patients with BCLP and UCLP were 59 and 67 percent more likely to have post-operative VPI compared to those without a primary cleft palate after adjusting for

presence of a syndrome (95% CI for the PRs 0.64-3.96 and 0.66-4.27, respectively). Further adjustment for magnitude of maxillary deficiency did not meaningfully change the observed exposure-outcome association, providing some evidence that an association between cleft type and post-operative VPI exists beyond the effect acting through the magnitude of maxillary deficiency. All observed differences, however, were plausibly the result of chance given no true difference.

Incomplete data is an important limitation of this study. Seventy-three of the 200 patients undergoing maxillary advancement during the study period did not have a post-operative speech evaluation, although the proportion of missingness between exposure groups varied depending on the predictor of interest. To assess potential bias related to differential missingness of outcome assessment between exposure groups, the proportion of patients missing outcome were calculated. Continuous predictors were dichotomized at different cut-offs to create "exposed" versus "unexposed" groups for the purposes of illustrating degree of missingness in groups empirically considered at higher versus lower risk of VPI. These results are summarized in Table 4.4.

Table 4.4: Proportion of participants lost to follow-up, stratified by higher versus lower risk groups[+].

Predictor	Higher Risk	Lower Risk
Magnitude of Maxillary Deficiency		
A-point --> x-axis (< vs > mean)	29 (39.2)	27 (28.1)
A-point --> x-axis (< 1 SD below mean)	51 (33.5)	5 (27.8)
Overjet (> vs < mean)	10 (29.4)	44 (33.1)
Velopharyngeal Ratio (> vs < 0.8)	28 (36.8)	28 (29.7)
Borderline vs No Pre-Operative VPI	9 (20.5)	36 (30.8)
BCLP vs CPO/SMCP	25 (41.0)	13 (35.1)
UCLP vs CPO/SMCP	35 (34.3)	13 (35.1)

[+] Missingness proportions presented as N(%)

[*] Total missing in each group does not add up to 73 because of those missing both pre- and post-operative data.

While it is plausible that those who have no speech concerns are less likely to return for follow-up, the degree of missingness was similar between higher and lower risk groups for each predictor. Thus, the magnitude of bias related to missing outcome data is expected to be small.

Lack of complete data in the predictors was also present in this study. With regard to the magnitude of maxillary deficiency and VP ratio, values missing were due to either absence or poor quality of the pre-operative lateral cephalometric radiograph in the EMR. An upgrade in x-ray technology after 2012 meant that x-rays classified as untraceable were more likely to occur earlier in the study period. Because absence of a pre-operative cephalometric x-ray was unrelated to the outcome, bias related to missing cephalometric data is not of significant concern. Thirty-nine subjects did not have available pre-operative speech data. Just eleven of these, however, were ultimately included in the analysis, as the remainder did not return for follow-up speech evaluation. Presence of speech concerns did not appear to increase the likelihood of follow-up, as seven of the 11 without pre-operative speech evaluation did not present with post-operative VPI. Finally, incomplete data across exposure, outcome, and some covariates depleted the number of complete cases available for analysis, thereby diminishing our sample size to a degree that we were unable to reliably identify true small differences in risk of post-operative VPI between exposure groups.

Other studies investigating planned maxillary advancement as a predictor also reported lack of an association between magnitude of maxillary advancement and post-operative velopharyngeal function. In a case series of 26 patients with non-syndromic CP/L, the amount of planned advancement was measured at the second molar on pre-surgical models and the outcome was defined as presence versus absence of VPI post-operatively. With magnitude of advancement ranging from 5 to 17 mm, the unadjusted odds ratio (OR) associated with one millimeter change in the amount of advancement was 0.98 (95% CI not reported, p=0.86) [94]. Adjusting for presence of pre-operative VPI did not meaningfully change the observed association (OR 0.92, 95% CI 0.6 – 1.33). In a more recent study of 18 patients with CP/L

undergoing maxillary advancement, the range of planned maxillary advancement based on pre-operative cone beam computed tomography (CBCT) scans was 2.8 to 12 mm, although the landmarks used to define the amount of horizontal movement were not reported [95]. Post-operative outcome was the CAPS-A-AM 4-point severity rating of hypernasality [44]. Results of a chi-square analysis comparing the amount of maxillary advancement with post-operative hypernasality in this study were non-significant (chi-square statistic, CI or p-value not reported). In both of these studies, the total number of maxillary advancement surgeries versus the number that contributed outcome data was not reported and implications of incomplete outcome, if present, was not discussed.

In a series of 50 patients with CP/L who underwent maxillary advancement, 11 of 15 (73 percent) patients with borderline velopharyngeal closure on pre-operative nasopharyngoscopy had clinically significant VPI post-operatively [103]. While results of statistical analyses were not reported, this proportion is 30 percent higher than the proportion of those with borderline VPI presenting with post-operative VPI in the current study (15 of 35, 42.9 percent). This difference is likely related to the use of nasopharyngoscopy versus perceptual evaluation to define borderline status, as nasopharyngoscopy is more sensitive in identifying borderline function and a patient with a small gap visible on nasopharyngoscopy may not present with perceptual correlates of borderline VPI, as used in our study. In a recent study of 100 non-syndromic patients with CP/L, investigators used Spearman's correlations to assess the association between pre- and post-operative velopharyngeal function. Velopharyngeal function was defined as a 0-4 point severity rating of VPI [101]. Of the 18 patients with borderline function (rating of 1) before maxillary advancement, nine met the criteria for VPI after surgery (rating of 2-4), and an overall positive correlation between pre- and post-operative velopharyngeal function was reported (Spearman's rho, 95% CI and p-value not reported). The post-operative prevalence of VPI was 50 percent among patients with borderline VPI pre-operatively, far greater than the prevalence (4.2 percent) in those with normal pre-operative velopharyngeal function (PR = 11.9,

calculated directly from data available in paper). While the direction of this association is comparable with our results, the magnitude is considerably greater than the 58 percent increased risk present in our study population. It is possible that the definition of borderline velopharyngeal function in the current study (infrequent audible nasal air emission in the absence of hypernasality or assimilation nasality in the absence of audible nasal air emission) was less likely to classify a patient as having borderline function before surgery than the definition (insignificant, borderline, mild and occasional VPI detected by ears and/or nasometer) employed in the earlier paper.

The previous study also assessed the effect of type of cleft on changes in velopharyngeal function using chi-square analysis [101]. Using the same outcome definition as described in the previous paragraph, they reported that cleft type did not affect VPI outcome in their study ($p=0.75$, χ^2 statistic and 95% CI not reported). While the proportion of different types of cleft was comparable between studies, the prevalence of post-operative VPI was overall lower for UCLP, BCLP, and CPO/SMCP in the earlier study (10.1 percent, 16.7 percent, and 14.3 percent, respectively) compared to 32.8, 33.3, and 29.1 percent in our study. While authors noted missing post-operative evaluation as an exclusion criterion, the proportion of missingness within different types of clefts was not reported. In the current study, the proportion of missingness is comparable between each type of cleft (see Table 4.4), minimizing concern for potential bias related to differential missingness between exposure groups when calculating measures of relative risk. However, if those with no concerns regarding negative change in velopharyngeal function post-operatively are less likely to return for follow-up, it is plausible that the observed prevalence of post-operative VPI in our study population is overall larger than what would have been observed if all participants had returned for post-operative speech evaluation.

In the absence of studies measuring VP ratio prior to surgery and comparing measurements with post-operative outcome, it is difficult to compare our results to others.

Although one cannot rule-out chance as the source of our results if no true relationship exists, the modest difference in the magnitude of the effect may be of interest for future study. Use of the point at which the palatal plane crosses the posterior pharyngeal wall as the posterior landmark for the nasopharyngeal depth is common practice, as the height of velopharyngeal closure is typically at or near the level of the palatal plane. However, in patients with a repaired cleft palate, their level of attempted velopharyngeal closure may vary depending on the location of insertion of the levator veli palatini as well as scarring from prior surgeries restricting velopharyngeal movement. If nasopharyngeal depth from PNS-Ad1 is on average higher than the true level of attempted closure, this value may overestimate the posterior distance the velum needs to travel to achieve velopharyngeal closure, making the calculated VP ratio larger than if it was calculated using landmarks more closely representing the height of velopharyngeal closure. Additionally, the angle of the palatal plane may not be near parallel to horizontal in patients with a repaired CP/L, as is typical for patients without cleft palate. If the maxilla is rotated forward and down, Ad1 may occur quite high on the nasopharynx, increasing the distance from PNS to Ad1 in a way that does not reflect nasopharyngeal depth at the level of the velopharyngeal port. PNS-Ad2 and PNS-Ad3 are less influenced by changes aside from horizontal maxillary movement that occur as a consequence of LeFort I osteotomy and thus may be more valid as predictors of post-operative change related to maxillary advancement. Alternatively, a true association between VP ratio as measured on pre-operative lateral cephalometric radiographs and development of VPI after maxillary advancement may not exist given the complex three-dimensional changes to the velopharyngeal port that occur as a result of LeFort I osteotomy.

Identifying anatomic and functional characteristics that place a patient at increased risk of VPI after surgical management of maxillary hypoplasia is a challenge due to the heterogeneity of both the patient population in the presence of CP/L as well as the three-dimensional changes that occur as a result of surgery. Consequently, there is insufficient evidence to support specific

factors as predictive of post-operative VPI. The presence of borderline velopharyngeal function pre-operatively is suggested as an area for future study. Additionally, analyses evaluating whether observed associations are modified by or mediated through other variables will be important in expanding our understanding of these complex relationships.

In conclusion, our results suggest that patients with BCLP and UCLP are at modestly greater risk of developing VPI after maxillary advancement than patients with a CPO or SMCP with the same syndrome status. Patients with borderline velopharyngeal function on perceptual exam also appear to have somewhat greater risk of post-operative VPI. After accounting for the relationships between age and type of cleft on the exposure-outcome association, the magnitude of maxillary advancement did not appear associated with presence of post-operative VPI. Given the variability in the literature for reliable factors in predicting the presence of VPI after maxillary advancement, further contributions to the overall body of evidence are needed.

References

1. Beluci ML, Mondini, Cleide Carolina da Silva Demoro, Trettene ADS, Dantas RAS. Correlation between quality of life and burden of family caregivers of infants with cleft lip and palate. *Rev Esc Enferm USP*. 2019;53:e03432. https://www.openaire.eu/search/publication?articleId=od3056::1d465c7c1e843b03ef6c818eb84043a6. doi: 10.1590/s1980-220x2017047603432.

2. Kramer F, Gruber R, Fialka F, Sinikovic B, Schliephake H. Quality of life and family functioning in children with nonsyndromic orofacial clefts at preschool ages. *J Craniofac Surg*. 2008;19(3):580-7.

3. Nolte FM, Bos A, Prahl C. Quality of life among dutch children with a cleft lip and/or cleft palate: A follow-up study. *Cleft Palate Craniofac J*. 2019;56(8):1065-1071. https://www.narcis.nl/publication/RecordID/oai:dare.uva.nl:publications%2F1c47d01e-6ff1-4cef-8ee7-498d36daa2ae. doi: 10.1177/1055665619840220.

4. Diseth TH, Emblem R. Long-term psychosocial consequences of surgical congenital malformations. *Semin Pediatr Surg*. 2017;26(5):286-294. http://dx.doi.org/10.1053/j.sempedsurg.2017.09.009. doi: 10.1053/j.sempedsurg.2017.09.009.

5. Good PM, Mulliken JB, Padwa BL. Frequency of le fort I osteotomy after repaired cleft lip and palate or cleft palate. *Cleft Palate Craniofac J*. 2017;44(4):396-401. https://search.datacite.org/works/10.1597/06-075.1. doi: 10.1597/06-075.1.

6. Pereira V, Sell D, Tuomainen J. The impact of maxillary osteotomy on speech outcomes in cleft lip and palate: An evidence-based approach to evaluating the literature. *Cleft Palate Craniofac J*. 2013;50(1):25-39. https://journals.sagepub.com/doi/full/10.1597/11-116. doi: 10.1597/11-116.

7. Friede H, Lilja J, Lohmander A. Long-term, longitudinal follow-up of individuals with UCLP after the gothenburg primary early veloplasty and delayed hard palate closure protocol: Maxillofacial growth outcome. *Cleft Palate Craniofac J*. 2012;49(6):649-656.

8. Sitzman T, Mara C, Long R, et al. The americleft project: Burden of care from secondary surgery. *Plast Reconstr Surg Glob Open*. 2015;3(7):e442. https://www.ncbi.nlm.nih.gov/pubmed/26301131. doi: 10.1097/GOX.0000000000000415.

9. Sitzman TJ, Carle AC, Heaton PC, Helmrath MA, Britto MT. Five-fold variation among surgeons and hospitals in the use of secondary palate surgery. *Cleft Palate Craniofac J*. 2018;56(5):586-594. https://search.datacite.org/works/10.1177/1055665618799906. doi: 10.1177/1055665618799906.

10. Barr L, Thibeault SL, Muntz H, de Serres L. Quality of life in children with velopharyngeal insufficiency. *Arch Otolaryngol Head Neck Surg*. 2007;133(3):224-229. http://dx.doi.org/10.1001/archotol.133.3.224. doi: 10.1001/archotol.133.3.224.

11. Skirko JR, Weaver EM, Perkins J, Kinter S, Sie KCY. Modification and evaluation of a velopharyngeal insufficiency quality-of-life instrument. *Arch Otolaryngol Head Neck Surg*. 2012;138(10):929-935.

12. Skirko JR, Weaver EM, Perkins JA, et al. Change in quality of life with velopharyngeal insufficiency surgery. *Arch Otolaryngol Head Neck Surg*. 2015;153(5):857-864.

13. Moran I, Virdee S, Sharp I, Sulh J. Postoperative complications following LeFort I maxillary advancement surgery in cleft palate patients: A 5-year retrospective study. *Cleft Palate Craniofac J*. 2018;55(2):231-237.

14. Cohen SR, Burstein FD, Stewart MB, Rathburn M. Maxillary-midface distraction in children with cleft lip and palate: A preliminary report. *Plast Reconstr Surg*. 1997;99(5):1421-1428.

15. Chua HDP, Whitehill TL, Samman N, Cheung LK. Maxillary distraction versus orthognathic surgery in cleft lip and palate patients: Effects on speech and velopharyngeal function. *Int J Oral Maxillofac Surg*. 2010;39(7):633-640. https://search.datacite.org/works/10.1016/j.ijom.2010.03.011. doi: 10.1016/j.ijom.2010.03.011.

16. Kumar A, Gabbay JS, Nikjoo R, et al. Improved outcomes in cleft patients with severe maxillary deficiency after LeFort I internal distraction. *Plast Reconstr Surg*. 2006;117(5):1499-1509. https://www.ncbi.nlm.nih.gov/pubmed/16641719. doi: 10.1097/01.prs.0000206308.86089.86.

17. de Medeiros-Santana, Maria Natália Leite, Perry JL, Yaedú RYF, Trindade-Suedam IK, Yamashita RP. Predictors of velopharyngeal dysfunction in individuals with cleft palate following surgical maxillary advancement: Clinical and tomographic assessments. *Cleft Palate Craniofac J*. 2019;56(10):1314-1321. https://search.datacite.org/works/10.1177/1055665619852562. doi: 10.1177/1055665619852562.

18. Chanchareonsook N, Whitehill TL, Samman N. Speech outcome and velopharyngeal function in cleft palate: Comparison of LeFort I maxillary osteotomy and distraction osteogenesis—early results. *Cleft Palate Craniofac J*. 2007;44(1):23-32. https://journals.sagepub.com/doi/full/10.1597/05-003. doi: 10.1597/05-003.

19. Kotlarek KJ, Perry JL. Velopharyngeal anatomy and physiology. *Perspect ASHA Spec Interest Groups*. 2018;3(5):13-23. https://pubs.asha.org. doi: 10.1044/persp3.SIG5.13.

20. Perry J. Anatomy and physiology of the velopharyngeal mechanism. *Seminars in speech and language*. 2011;32(2):83. https://search.datacite.org/works/10.1055/s-0031-1277712. doi: 10.1055/s-0031-1277712.

21. Peterson-Falzone S, Trost-Cardamone JE, Karnell M, Hardin-Jones M. *The clinician's guide to treating cleft palate speech*. 2nd ed. St Louis: Elsevier Inc; 2017:320.

22. Roy AA, Rtshiladze MA, Stevens K, Phillips J. Orthognathic surgery for patients with cleft lip and palate. *Clin Plast Surg*. 2019;46(2):157-171. https://doi.org/10.1016/j.cps.2018.11.002. doi: 10.1016/j.cps.2018.11.002.

23. Posnick JC, Tompson B. Modification of the maxillary LeFort I osteotomy in cleft-orthognathic surgery: The unilateral cleft lip and palate deformity. *J Oral Maxillofac Surg*. 1992;50(7):666-675. https://search.datacite.org/works/10.1016/0278-2391(92)90092-e. doi: 10.1016/0278-2391(92)90092-e.

24. Phillips JH, Nish I, Daskalogiannakis J. Orthognathic surgery in cleft patients. *Plast Reconstr Surg*. 2012;129(3):535e-548e.

25. dos Santos Alves JM, de Freitas Alves BW, de Figueiredo Costa AC, Carneiro BGDS, de Sousa LM, Gondim DV. Cranial nerve injuries in LeFort I osteotomy: A systematic review. *Int J Oral Maxillofac Surg*. 2019;48(5):601-611. https://search.datacite.org/works/10.1016/j.ijom.2018.11.012. doi: 10.1016/j.ijom.2018.11.012.

26. Yamaguchi K, Lonic D, Lo L. Complications following orthognathic surgery for patients with cleft lip/palate: A systematic review. *J Formos Med Assoc*. 2016;115(4):269-277. https://search.datacite.org/works/10.1016/j.jfma.2015.10.009. doi: 10.1016/j.jfma.2015.10.009.

27. Posnick JC. *Orthognathic surgery: Principles & practice*. St Louis: Elsevier Inc; 2014.

28. Scolozzi P. Distraction osteogenesis in the management of severe maxillary hypoplasia in cleft lip and palate patients. *J Craniofac Surg*. 2008;19(5):1199-1214. https://www.ncbi.nlm.nih.gov/pubmed/18812842. doi: 10.1097/SCS.0b013e318184365d.

29. Austin SL, Mattick CR, Waterhouse PJ, Austin S. Distraction osteogenesis versus orthognathic surgery for the treatment of maxillary hypoplasia in cleft lip and palate patients: A systematic review. *Orthod Craniofac Res*. 2015;18:96-108.

30. Heliövaara A, Ranta R, Hukki J, Haapanen M. Cephalometric pharyngeal changes after le fort I osteotomy in patients with unilateral cleft lip and palate. *Scand J Plast Reconstr Surg Hand Surg*. 2002;60(3):141-145. http://www.tandfonline.com/doi/abs/10.1080/000163502753740142. doi: 10.1080/000163502753740142.

31. Dowling PA, Espeland L, Sandvik L, Mobarak KA, Hogevold HE. LeFort I maxillary advancement: 3-year stability and risk factors for relapse. *Am J Orthod Dentofacial Orthop*. 2005;128(5):560-567. http://dx.doi.org/10.1016/j.ajodo.2004.07.051. doi: 10.1016/j.ajodo.2004.07.051.

32. Hirano A, Suzuki H. Factors related to relapse after LeFort I maxillary advancement osteotomy in patients with cleft lip and palate. *Cleft Palate Craniofac J*. 2001;38(1):1-10. https://www.ncbi.nlm.nih.gov/pubmed/11204674. doi: 10.1597/1545-1569(2001)0382.0.CO;2.

33. Hochban W, Ganss C, Austermann KH. Long-term results after maxillary advancement in patients with clefts. *Cleft Palate Craniofac J*. 1993;30:237-243.

34. Drake RL, Vogl AW, Mitchell AW. Head and neck. In: *Gray's basic anatomy*. Second ed. Philadelphia, PA: Elsevier Inc; 2018:588-593.

35. Drommer R. Selective angiographic studies prior to LeFort I osteotomy in patients with cleft lip and palate. *J Maxillofac Surg*. 1979;7(4):264-70. http://www.ncbi.nlm.nih.gov/pubmed/292741.

36. Epker BN. Vascular considerations in orthognathic surgery: II. maxillary osteotomies. *Oral surgery, oral medicine, oral pathology*. 1984;57(5):473-478. http://dx.doi.org/10.1016/0030-4220(84)90302-5. doi: 10.1016/0030-4220(84)90302-5.

37. Posnick JC, Al-Qattan MM, Pron G. Facial sensibility in adolescents with and without clefts 1 year after undergoing le fort I osteotomy. *Plast Reconstr Surg*. 1994;94(3):431-435. https://www.ncbi.nlm.nih.gov/pubmed/8047593. doi: 10.1097/00006534-199409000-00002.

38. McLeod NMH, Bowe DC. Nerve injury associated with orthognathic surgery. part 3: Lingual, infraorbital, and optic nerves. *Br J Oral Maxillofac Surg*. 2016;54(4):372-375. https://www.clinicalkey.es/playcontent/1-s2.0-S0266435616000516. doi: 10.1016/j.bjoms.2016.01.028.

39. Chen PK, Por Y, Liou EJ, Chang FC. The effect of cleft maxillary distraction osteogenesis on the levator veli palatini and velopharyngeal function. *J Craniofac Surg*. 2015;26(3):687-690. https://www.ncbi.nlm.nih.gov/pubmed/25974774. doi: 10.1097/SCS.0000000000001586.

40. Wu Y, Wang X, Ma L, Li Z. Velopharyngeal configuration changes following LeFort I osteotomy with maxillary advancement in patients with cleft lip and palate: A cephalometric study. *Cleft Palate Craniofac J*. 2015;52(6):711-716. http://ir.bjmu.edu.cn/handle/400002259/126039. doi: 10.1597/14-146.1.

41. Chung J, Lim J, Park H, Yoo A, Kim S, Koo Y. Correlation between speech outcomes and the amount of maxillary advancement after orthognathic surgery (LeFort I conventional osteotomy and distraction osteogenesis) in patients with cleft lip and palate. *J Craniofac Surg*. 2019;00(00):1855-1858. https://www.ncbi.nlm.nih.gov/pubmed/31107383. doi: 10.1097/SCS.0000000000005623.

42. Scott SC, Goldberg MS, Mayo NE. Statistical assessment of ordinal outcomes in comparative studies. *J Clin Epidemiol*. 1997;50(1):45-55. https://www-sciencedirect-com.offcampus.lib.washington.edu/science/article/pii/S0895435696003125. Accessed Nov 11, 2020. doi: 10.1016/S0895-4356(96)00312-5.

43. John A, Sell D, Sweeney T, Harding-Bell A, Williams A. The cleft audit protocol for speech-augmented: A validated and reliable measure for auditing cleft speech. *Cleft Palate Craniofac J*. 2006;43(3):272-288. https://search.proquest.com/docview/204913826. doi: 10.1597/04-141R.1.

44. Chapman KL, Baylis A, Trost-Cardamone J, et al. The americleft speech project: A training and reliability study. *Cleft Palate Craniofac J*. 2016;53(1):93-108.

45. Cohen J. Weighted kappa: Nominal scale agreement with provision for scaled disagreement or partial credit. . 1968;70:213-220.

46. Hallgren KA. Computing inter-rater reliability for observational data: An overview and tutorial. *Tutor Quant Methods Psychol*. 2012;8(1):23-34. https://search.proquest.com/docview/1826558069. doi: 10.20982/tqmp.08.1.p023.

47. Cicchetti DV. Multiple comparison methods: Establishing guidelines for their valid application in neuropsychological research. *J Clin Exp Neuropsychol*. 1994;16(1):155-161. http://www.tandfonline.com/doi/abs/10.1080/01688639408402625. doi: 10.1080/01688639408402625.

48. Shrout PE, Fleiss JL. Intraclass correlations: Uses in assessing rater reliability. *Psychol Bull*. 1979;86(2):420-428. https://search.proquest.com/docview/614274515. doi: 10.1037/0033-2909.86.2.420.

49. McGraw KO, Wong SP. Forming inferences about some intraclass correlation coefficients. *Psychol Methods*. 1996;1(1):30-46. doi: 10.1037//1082-989X.1.1.30.

50. Kreiman J, Gerratt BR, Kempster GB, Erman A, Berke GS. Perceptual evaluation of voice quality: Review, tutorial, and a framework for future research. *J Speech Hear Res*. 1993;36(1):21-40. http://jslhr.asha.org/cgi/content/abstract/36/1/21. doi: 10.1044/jshr.3601.21.

51. Gerratt BR, Kreiman J, Antonanzas-Barroso N, Berke GS. Comparing internal and external standards in voice quality judgments. *J Speech Lang Hear Res*. 1993;36(1):14-20. http://jslhr.asha.org/cgi/content/abstract/36/1/14. doi: 10.1044/jshr.3601.14.

52. Helou LB, Solomon NP, Henry LR, Coppit GL, Howard RS, Stojadinovic A. The role of listener experience on consensus auditory-perceptual evaluation of voice (CAPE-V) ratings of postthyroidectomy voice. *Am J Speech Lang Pathol*. 2010;19(3):248-258. https://www.ncbi.nlm.nih.gov/pubmed/20484704. doi: 10.1044/1058-0360(2010/09-0012).

53. Munson B, Johnson JM, Edwards J. The role of experience in the perception of phonetic detail in children's speech: A comparison between speech-language pathologists and clinically untrained listeners. *Am J Speech Lang Pathol*. 2012;21(2):124-139. https://www.ncbi.nlm.nih.gov/pubmed/22230182. doi: 10.1044/1058-0360(2011/11-0009).

54. Gooch JL, Hardin-Jones M, Chapman KL, Trost-Cardamone JE, Sussman J. Reliability of listener transcriptions of compensatory misarticulations. *Cleft Palate Craniofac J*. 2001;38:59-67.

55. Lee A, Whitehill TL, Ciocca V. Effect of listener training on perceptual judgement of hypernasality. *Clin Linguist Phon*. 2009;23(5):319-334.

56. Henningsson G, Isberg A. A cineradiographic study of velopharyngeal movements for deviant versus nondeviant articulation. *Cleft Palate Craniofac J*. 1991;28(1):115-118. https://www.ncbi.nlm.nih.gov/pubmed/2004089. doi: 10.1597/1545-1569(1991)0282.3.CO;2.

57. Henningsson GE, Isberg AM. Velopharyngeal movement patterns in patients alternating between oral and glottal articulation: A clinical and cineradiographical study. *Cleft Palate J*. 1986;23(1):1-9.

58. Kataoka R. Quantitative evaluation of hypernasality - relation between spectral characteristics and perception of hypernasality. *J Jpn Cleft Palate Assoc*. 1988;13:204-216.

59. Castick S, Knight R, Sell D. Perceptual judgments of resonance, nasal airflow, understandability, and acceptability in speakers with cleft palate: Ordinal versus visual analogue scaling. *Cleft Palate Craniofac J*. 2017;54(1):19-31. https://journals.sagepub.com/doi/full/10.1597/15-164. doi: 10.1597/15-164.

60. Thompson WD, Walter SD. Variance and dissent. A reappraisal of the kappa coefficient. *J Clin Epidemiol*. 1988;41:949-958.

61. Brenner H, Kliebsch U. Dependence of weighted kappa coefficients on the number of categories. *Epidemiol*. 1996;7:199-202.

62. McNutt JC, Wicki L, Paulsen J. Judgments of phoneme errors under four modes of audio-visual presentation. *J Speech Lang Pathol Audiol*. 1991;15:37-52.

63. Stevens I, Daniloff RG. Trouble with /s/: A methodological study of factors affecting the judgment of misarticulated /s/. *J Comm Disord*. 1977;10:207-220.

64. Chatfield MD, Brayne CE, Matthews FE. A systematic literature review of attrition between waves in longitudinal studies in the elderly shows a consistent pattern of dropout between differing studies. *J Clin Epidemiol*. 2005;58(1). https://pubmed-ncbi-nlm-nih-gov.offcampus.lib.washington.edu/15649666/. Accessed Dec 8, 2020. doi: 10.1016/j.jclinepi.2004.05.006.

65. Hernan MA, Hernandez-Diaz S, Robins JM. A structural approach to selection bias : Epidemiology. *Epidemiol*. 2004;15(5):615-625. https://journals-lww-com.offcampus.lib.washington.edu/epidem/Fulltext/2004/09000/A_Structural_Approach_to_Selection_Bias.20.aspx. Accessed Dec 6, 2020. doi: 10.1097/01.ede.0000135174.63482.43.

66. Perkins NJ, Cole SR, Harel O, et al. Principled approaches to missing data in epidemiologic studies. *Am J Epidemiol*. 2018;187(3):568-575. https://www.ncbi.nlm.nih.gov/pubmed/29165572. doi: 10.1093/aje/kwx348.

67. Schafer JL, Graham JW. Missing data: Our view of the state of the art. *Psychol Methods*. 2002;7(2):147-177. https://www.ncbi.nlm.nih.gov/pubmed/12090408. doi: 10.1037//1082-989x.7.2.147.

68. Weiss NS, Koepsell TD. *Epidemiologic methods: Studying the occurrence of illness*. Second ed. New York: Oxford University Press USA - OSO; 2014. https://ebookcentral.proquest.com/lib/[SITE_ID]/detail.action?docID=1707881.

69. Hernan MA. Causal knowledge as a prerequisite for confounding evaluation: An application to birth defects epidemiology. *Am J Epidemiol*. 2002;155(2):176-184. https://search.datacite.org/works/10.1093/aje/155.2.176. doi: 10.1093/aje/155.2.176.

70. Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. *Epidemiol*. 1999;10(1):37-48. https://www.jstor.org/stable/3702180. doi: 10.1097/00001648-199901000-00008.

71. Textor J, van der Zander B, Gilthorpe MS, Liśkiewicz M, Ellison G. Robust causal inference using directed acyclic graphs: The R package 'dagitty'. *International Journal of Epidemiology*. 2016;45(6):1887-1894. https://doi.org/10.1093/ije/dyw341. Accessed May 15, 2021. doi: 10.1093/ije/dyw341.

72. Gu Y, McNamara J, James A, Sigler LM, Baccetti T. Comparison of craniofacial characteristics of typical chinese and caucasian young adults. *Eur J Orthod*. 2011;33(2):205-211. https://doi.org/10.1093/ejo/cjq054. Accessed Dec 11, 2020. doi: 10.1093/ejo/cjq054.

73. Cheung LK, Samman N, Hui E, Tideman H. The 3-dimensional stability of maxillary osteotomies in cleft palate patients with residual alveolar clefts. *The British Journal of Oral & Maxillofacial Surgery*. 1994;32(1). https://pubmed-ncbi-nlm-nih-gov.offcampus.lib.washington.edu/8136344/. Accessed Feb 27, 2021. doi: 10.1016/0266-4356(94)90163-5.

74. Quejada JG, Bell WH, Kawamura H, Zhang X. Skeletal stability after inferior maxillary repositioning. *Int J Adult Orthodon Orthognath Surg*. 1987;2(2):67-74. https://www.ncbi.nlm.nih.gov/pubmed/3295077.

75. Schendel SA, Oeschlaeger M, Wolford LM, Epker BN. Velopharyngeal anatomy and maxillary advancement. *J Maxillofac Surg*. 1979;7(116-124).

76. Simpson RK, Austin AA. A cephalometric investigation of velar stretch. *The Cleft palate journal*. 1972;9:341-51. https://pubmed-ncbi-nlm-nih-gov.offcampus.lib.washington.edu/4509238/. Accessed Jan 22, 2021.

77. Zou G. A modified poisson regression approach to prospective studies with binary data. *Am J Epidemiol*. 2004;159(7):702-706. https://pubmed-ncbi-nlm-nih-gov.offcampus.lib.washington.edu/15033648/. Accessed Feb 16, 2021. doi: 10.1093/aje/kwh090.

78. Daniel RM, Kenward MG, Cousens SN, De Stavola BL. Using causal diagrams to guide analysis in missing data problems. *Stat Methods Med Res*. 2011;21(3). https://journals-sagepub-com.offcampus.lib.washington.edu/doi/10.1177/0962280210394469?url_ver=Z39.88-2003&rfr_id=ori%3Arid%3Acrossref.org&rfr_dat=cr_pub++0pubmed. Accessed Sep 12, 2020.

79. Satoh K, Wada T, Tachimura T, Fukuda J. Velar ascent and morphological factors affecting velopharyngeal function in patients with cleft palate and noncleft controls: A cephalometric study. *Int J Oral Maxillofac Surg*. 2005;34(2):122-126. https://www-sciencedirect-com.offcampus.lib.washington.edu/science/article/pii/S0901502704001535. Accessed Mar 14, 2021. doi: 10.1016/j.ijom.2004.05.002.

80. Lu Y, Shi B, Zheng Q, Xiao W, Li S. Analysis of velopharyngeal morphology in adults with velopharyngeal incompetence after surgery of a cleft palate. *Ann Plast Surg*. 2006;57(1):50-54. https://journals-lww-com.offcampus.lib.washington.edu/annalsplasticsurgery/Fulltext/2006/07000/Analysis_of_Velopharyngeal_Morphology_in_Adults.10.aspx. Accessed Mar 14, 2021. doi: 10.1097/01.sap.0000208937.05684.38.

81. Janulewicz J, Costello BJ, Buckley MJ, Ford MD, Close J, Gassner R. The effects of LeFort I osteotomies on velopharyngeal and speech functions in cleft patients. *J Oral Maxillofac Surg*. 2004;62(3):308-314. https://www-sciencedirect-com.offcampus.lib.washington.edu/science/article/pii/S0278239103011182. Accessed Jan 18, 2021. doi: 10.1016/j.joms.2003.08.014.

82. Pereira VJ, Sell D, Tuomainen J. Effect of maxillary osteotomy on speech in cleft lip and palate: Percetpual outcomes of velopharyngeal function. *Int J Lang Comm Disord*. 2013;48(6):640-650.

83. Kim S, Kim J, Moon J, Lee K, Choi KY, Cho BC. Perceptual speech assessment after maxillary advancement osteotomy in patients with a repaired cleft lip and palate. *Arch Plast Surg*. 2012;39(3):198-202. https://www.ncbi.nlm.nih.gov/pubmed/22783526. doi: 10.5999/aps.2012.39.3.198.

84. Kudo K, Takagi R, Kodama Y, Terao E, Asahito T, Saito I. Evaluation of speech and morphological changes after maxillary advancement for patients with velopharyngeal insufficiency due to repaired cleft palate using a nasometer and lateral cephalogram. *J Oral Maxillofac Surg Med Pathol*. 2014;26(1):22-27. http://www.sciencedirect.com/science/article/pii/S2212555813001269. Accessed Jan 22, 2021. doi: 10.1016/j.ajoms.2013.07.006.

85. Saleh E, Saleh J, Beauchemin G, El-Jalbout R, Borsuk DE. Velopharyngeal space assessment in patients undergoing le fort 1 maxillary advancement. *Plast Reconstr Surg Glob Open*. 2020;8(11):e3232. https://pubmed-ncbi-nlm-nih-gov.offcampus.lib.washington.edu/33299700/. Accessed Jan 22, 2021. doi: 10.1097/GOX.0000000000003232.

86. Taha M, Elsheikh YM. Velopharyngeal changes after maxillary distraction in cleft patients using a rigid external distraction device: A retrospective study. *Angle Orthod*. 2016;86(6):962-968. https://www.ncbi.nlm.nih.gov/pubmed/27007755. doi: 10.2319/011216-33.1.

87. Ko EW, Figueroa AA, Guyette TW, Polley JW, Law WR. Velopharyngeal changes after maxillary advancement in cleft patients with distraction osteogenesis using a rigid external distraction device: A 1-year cephalometric follow-up. *J Craniofac Surg*. 1999;10(4):312-320.

88. Maegawa J, Sells RK, David DJ. Speech changes after maxillary advancement in 40 cleft lip and palate patients. *J CRANIOFAC SURGERY*. 1998;9(2):177-82; discussion 183. doi: 10.1097/00001665-199803000-00017.

89. Harjunpaa R, Alaluusua S, Leikola J, Heliovaara A. LeFort I osteotomy in cleft patients: Maxillary advancement and velopharyngeal function. *J Craniomaxillofac Surg*. 2019;47:1868-1874.

90. Harada K, Ishii Y, Ishii M, Imaizumi H, Mibu M, Omura K. Effect of maxillary distraction osteogenesis on velopharyngeal function: A pilot study. *Oral surgery, oral medicine, oral pathology, oral radiology and endodontics*. 2002;93(5):538-543. http://dx.doi.org/10.1067/moe.2002.123827. doi: 10.1067/moe.2002.123827.

91. Watzke I, R D, Turvey TA, Warren DW. Alterations in velopharyngeal function after maxillary advancement in cleft palate patients. *Journal of oral and maxillofacial surgery : official journal of the American Association of Oral and Maxillofacial Surgeons*. 1990;48(7). https://pubmed-ncbi-nlm-nih-gov.offcampus.lib.washington.edu/2358944/. Accessed Jan 2, 2021. doi: 10.1016/0278-2391(90)90050-c.

92. Smedberg E, Neovius E, Lohmander A. Impact of maxillary advancement on speech and velopharyngeal function in patients with cleft lip and palate. *Cleft Palate Craniofac J*. 2014;51(3):334-343. https://journals.sagepub.com/doi/full/10.1597/12-304. doi: 10.1597/12-304.

93. Impieri D, Tønseth K, Hide Ø, Brinck E, Høgevold H, Filip C. Impact of orthognathic surgery on velopharyngeal function by evaluating speech and cephalometric radiographs. *J Plast Reconstr Aesthet Surg*. 2018;71(12):1786-1795. https://pubmed-ncbi-nlm-nih-gov.offcampus.lib.washington.edu/30172730/. Accessed Jan 30, 2021. doi: 10.1016/j.bjps.2018.07.018.

94. Phillips JH, Klaiman P, Delorey R, MacDonald DB. Predictors of velopharyngeal insufficiency in cleft palate orthognathic surgery. *Plast Reconstr Surg*. 2005;115(3):681-686. https://search.datacite.org/works/10.1097/01.prs.0000152433.29134.79. doi: 10.1097/01.prs.0000152433.29134.79.

95. Schultz KP, Braun TL, Hernandez C, et al. Speech outcomes after LeFort I advancement among cleft lip and palate patients. *Ann Plast Surg*. 2019;82:174-179.

96. McComb RW, Marrinan EM, Nuss RC, LaBrie RA, Mulliken JB, Padwa BL. Predictors of velopharyngeal insufficiency after LeFort I maxillary advancement in patients with cleft palate. *J Oral Maxillofac Surg*. 2011;69(8):2226-2232. https://search.datacite.org/works/10.1016/j.joms.2011.02.142. doi: 10.1016/j.joms.2011.02.142.

97. Raja Lakshmi C, Ayesha Thabusum D, Bhavana SM. An innovative approach to evaluate the morphological patterns of soft palate in oral submucous fibrosis patients: A digital cephalometric study. *International Journal of Chronic Diseases*. 2016;2016:5428581-6. https://dx.doi.org/10.1155/2016/5428581. doi: 10.1155/2016/5428581.

98. Subtelny JD. A cephalometric study of the growth of the soft palate. *Plastic and reconstructive surgery (1946)*. 1957;19(1):49-62. https://www.ncbi.nlm.nih.gov/pubmed/13419567.

99. Warren DW, Dubois AB. A pressure-flow technique for measuring velopharyngeal orifice area during continuous speech. *Cleft Palate J*. 1964;16. https://pubmed-ncbi-nlm-nih-gov.offcampus.lib.washington.edu/14116541/. Accessed May 15, 2021.

100. Epker BN, Wolford LM. Middle-third facial osteotomies: Their use in the correction of congenital dentofacial and craniofacial deformities. *J Oral Surg*. 1976;34(4):324-342. https://www.ncbi.nlm.nih.gov/pubmed/1062535.

101. Alaluusua S, Turunen L, Saarikko A, Geneid A, Leikola J, Heliövaara A. The effects of le fort I osteotomy on velopharyngeal function in cleft patients. *J Craniomaxillofac Surg*. 2019;47(2):239-244. https://www-sciencedirect-com.offcampus.lib.washington.edu/science/article/pii/S1010518218306188. Accessed Jan 18, 2021. doi: 10.1016/j.jcms.2018.11.016.

102. Trindade IEK, Yamashita RP, Suguimoto RM, Mazzottini R, Trindade AS. Effects of orthognathic surgery on speech and breathing of subjects with cleft lip and palate: Acoustic and aerodynamic assessment. *Cleft Palate Craniofac J*. 2017;40(1):54-64.

https://search.datacite.org/works/10.1597/1545-1569_2003_040_0054_eoosos_2.0.co_2. doi: 10.1597/1545-1569_2003_040_0054_eoosos_2.0.co_2.

103. Witzel MA. The effects of le fort I osteotomy with maxillary movement on articulation, resonance, and velopharyngeal function [commentary]. . 1989;26:199-200.

104. Pereira VJ, Tuomainen J, Hay N, Mars M, Suchak A, Sell DA. Identifying predictors of acquired velopharyngeal insufficiency in cleft lip and palate following maxillary osteotomy using multiple regression analyses. *J Craniofac Surg*. 2020;31(8):2260-2266. https://pubmed-ncbi-nlm-nih-gov.offcampus.lib.washington.edu/33136867/. Accessed Jan 22, 2021. doi: 10.1097/SCS.0000000000006775.

105. Yatabe-Ioshida MS, Campos LD, Yaedu RY, Trindade-Suedam IK. Upper airway 3D changes of patients with cleft lip and palate after orthognathic surgery. *Cleft Palate Craniofac J*. 2019;56(3):314-320.

Appendices

A. Speech sample used in audios reviewed for inter- and intra-rater reliability:
   - Mamama
   - Papapa
   - Kakaka
   - Sasasa
   - Mommy made lemon jam
   - Pull the baby buggy
   - Take Teddy to Daddy
   - Give Kate the cake
   - Sissy sees the sky
   - Jim and Charlie chew gum
   - Count 1-20
   - Count 60-70