The Cleft Palate-Craniofacial Journal

Effect of Maxillary Osteotomy on Speech in Cleft Lip and Palate: Instrumental Outcomes of Velopharyngeal Function

Journal:	*The Cleft Palate-Craniofacial Journal*
Manuscript ID	CPCJ-19-0349.R2
Manuscript Type:	Original Article
Keywords:	Orthognathic surgery, Velopharyngeal function, Resonance

Abstract:	Objective: To investigate the effect of maxillary osteotomy on velopharyngeal function in cleft lip and palate (CLP) using instrumental measures. Design: Prospective Participants: A consecutive series of 20 patients with CLP undergoing maxillary osteotomy by a single surgeon were seen at 0-3 months pre-surgery(T1), 3-months(T2) and 12-months(T3) post-surgery. Interventions: Nasalance was measured on the Nasometer II 6400. For videofluoroscopy and nasendoscopy data, visual perceptual ratings (VPRs) e.g. palatal lift angle (PLAn) and quantitative ratiometric measurements (QRMs) e.g. closure ratio (CRa), were made using a validated methodology and computer software. Reliability studies were undertaken for all instrumental measures. Main Outcome Measures: Repeated measures ANOVA (with time at 3 levels) for nasalance and each velar parameter. Planned comparisons across pairs of time points (T1-T2, T1-T3, T2-T3) including effect sizes. Results: A significant difference over time was found for nasalance (p =.001) and planned comparisons across pairs of time points were significant between T1-T2 (p = .008), T1-T3 (p = .002) but not between T2-T3 (p =.459) providing evidence that maxillary osteotomy can impact on nasalance adversely and that the changes seen are permanent and stable. There were also significant differences over time for PLAn (p =.012) and CRa (p =-.059) and planned comparisons for both velar parameters reflected similar findings to those of nasalance. Conclusions: Maxillary osteotomy can adversely affect velopharyngeal function in patients with CLP. The study provides evidence for a much earlier post-surgery review even as early as 3-months after surgery. Keywords cleft palate, velopharyngeal function, orthognathic surgery, speech

Title

Effect of Maxillary Osteotomy on Speech in Cleft Lip and Palate: Instrumental Outcomes of Velopharyngeal Function

Abstract

Objective: To investigate the effect of maxillary osteotomy on velopharyngeal function in cleft lip and palate (CLP) using instrumental measures.

Design: Prospective

Participants: A consecutive series of 20 patients with CLP undergoing maxillary osteotomy by a single surgeon were seen at 0-3 months pre-surgery(T1), 3-months(T2) and 12-months(T3) post-surgery.

Interventions: Nasalance was measured on the Nasometer II 6400. For videofluoroscopy and nasendoscopy data, visual perceptual ratings (VPRs) e.g. palatal lift angle (PLAn) and quantitative ratiometric measurements (QRMs) e.g. closure ratio (CRa), were made using a validated methodology and computer software. Reliability studies were undertaken for all instrumental measures.

Main Outcome Measures: Repeated measures ANOVA (with time at 3 levels) for nasalance and each velar parameter. Planned comparisons across pairs of time points (T1-T2, T1-T3, T2-T3) including effect sizes.

Results: A significant difference over time was found for nasalance ($p = .001$) and planned comparisons across pairs of time points were significant between T1-T2 ($p = .008$), T1-T3 ($p = .002$) but not between T2-T3 ($p = .459$) providing evidence that maxillary osteotomy can impact on nasalance adversely and that the changes seen are permanent and stable. There were also significant differences over time for PLAn ($p = .012$) and CRa ($p = -.059$) and planned comparisons for both velar parameters reflected similar findings to those of nasalance.

Conclusions: Maxillary osteotomy can adversely affect velopharyngeal function in patients with CLP. The study provides evidence for a much earlier post-surgery review even as early as 3-months after surgery.

Keywords

cleft palate, velopharyngeal function, orthognathic surgery, speech

Introduction

A well-known adverse sequela of cleft lip and palate (CLP) is abnormal facial growth which is reflected in a class III malocclusion resulting in maxillary retrusion. In individuals with CLP, maxillary retrusion becomes increasingly evident during the pubertal growth spurt (Ross, 1987; Semb, 1991) with reports that up to 50% of individuals with CLP require surgical correction of this facial deformity (e.g. Good, Mulliken and Padwa, 2007). The most commonly used surgical technique to correct this in CLP is a Le Fort I maxillary osteotomy with or without a mandibular setback (Cheung and Chua, 2006).

Maxillary osteotomy can adversely impact velopharyngeal function in individuals with CLP (e.g. Impieri et al., 2018). Velopharyngeal dysfunction (VPD) occurs when the velopharyngeal sphincter does not separate the oral from the nasal cavity efficiently for the production of (oral) pressure consonants. The well-known perceptual sequelae of VPD include hypernasality, nasal airflow errors, nasal/facial grimace, non-oral and passive cleft speech characteristics (CSCs) (Harding and Grunwell, 1988). Nasal airflow errors may also be due to an oronasal fistula (Mercer & Pigott, 2005). Passive CSCs include the cleft speech characteristics of weakened/nasalized consonants, nasal realizations of plosives or fricatives, and/**or** absent pressure consonants (Sell et al., 1994,1999; Grunwell & Sell, 2005). Compensatory articulation or non-oral CSCs such as glottal articulation (e.g. Trost, 1981) may reflect either ongoing or previous VPD (Lohmander et al., 2009).

The literature reports mixed evidence on velopharyngeal function following maxillary osteotomy. Some studies have found that surgery has a negative impact on velopharyngeal function resulting in acquired or increased hypernasality and/or nasal airflow errors (e.g. Trindade et al., 2003; Chua et al., 2010; Pereira et al., 2013a). Trindade et al. (2003) reported that 45% of their cohort showed increases in nasalance at 9-months post-surgery. Similarly, Chua et al. (2010) and Pereira et al. (2013a) reported a 36% and 31% (respectively) rate of acquired hypernasality based on perceptual assessment. Using two post-surgery time points, Pereira et al. (2013a) found that hypernasality, nasal turbulence and velopharyngeal composite score changed significantly immediately post-surgery (3-months) and the deterioration was maintained a year after surgery reflecting a stable speech outcome seen early on after surgery. Other studies, however, have reported no adverse impact on velopharyngeal function (e.g. McCarthy et al., 1979; Smedberg, Neovius & Lohmander, 2014). Smedberg et al. (2014) for example, found no impact of the surgery on nasalance and velopharyngeal function as measured by nasendoscopy and videoradiography.

Pereira, Sell and Tuomainen (2013b) undertook a systematic review of the speech osteotomy literature using a pre-determined framework including assignment of levels of evidence, calculation of effect sizes and post-hoc power. The authors surmised that such calculations allow for more objective comparisons of results across studies where there is variability in speech or study methodology and particularly where no statistically significant differences are found (Greenhalgh, 2001). In the field of CLP research where sample sizes tend to be small, there is a high probability of failing to reject a false null hypothesis (e.g. there is no impact of maxillary osteotomy on speech), known as a Type II error. The risk of a Type II error can be reduced by increasing the power of a study by means such as increasing sample size and/or the magnitude of an effect size. Post-hoc analyses uses the calculated effect and sample sizes to identify the power of the study. (Greenhalgh, 2001). From the initial 40 studies identified by Pereira et al (2013b), only seven met the inclusion criteria for description and discussion. The authors concluded that the results were conflicting for both resonance and nasalance, and that maxillary osteotomy may not have a true or clinical effect on velopharyngeal function when assessed using instrumentation. However, the authors also noted that the results were inconclusive as there were inherent study and speech methodological issues: some studies were retrospective in nature, did not report speech reliability studies, had small sample sizes and/or insufficient post-surgery follow-up.

In the assessment of velopharyngeal function, direct assessment is essential and the use of at least one instrumental measure is recommended (Dalston et al., 1988). The mainstay of such direct instrumental assessment of velopharyngeal function continues to centre around videofluoroscopy and nasendoscopy (Kuehn and Moller, 2000), with differences between the latter two methods in terms of analyses and measurement. In 1990, a multidisciplinary group of scientists, an 'International Working Group', published a paper proposing standards on the reporting of nasendoscopy and videofluoroscopy outcomes (Golding-Kushner et al., 1990), known as the Standardization method. For instance, for lateral view videofluoroscopy images, velum displacement can be measured ratiometrically from 0.0 (velum at rest) to 1.0 (complete closure). Reliability of the Standardization method, however, was not addressed in the proposal and continues to be ill-defined (e.g. Sell and Pereira, 2011). The authors described several drawbacks to the method stating that "mostly it has applicability when maximum movements occur in the midlines of the structures, at the same level and are symmetrical and consistent. Unfortunately, this is often not the case particularly in nasendoscopy, which probably affects the reliability of the method and brings into question its validity" (p. 155). Furthermore, the Standardization method does not address inherent issues with barrel distortion and effect of object-lens distance on magnification of nasendoscopic images (Gilleard et al., 2013).

A reliable and validated measurement system that allows for absolute and relative or ratiometric measures of velar parameters based on lateral videofluoroscopic images has been described (Birch, Sommerlad and Bhatt, 1994; Birch et al., 1999). This methodology involves the use of a computer software, Image Pro (Media Cybernetics) which allows for the manual measurement of point-to-point 2D and 3D line distances as well as angles, thus allowing for velar parameters such as closure ratio and velar extensibility to be measured reliably. The clinical and research utility of this measurement technique has been demonstrated in several studies (Sommerlad et al., 2002; Pereira, 2012). Gilleard (2008) and Gilleard et al. (2009) further extended this to include the measurement of *closure ratio* based on nasendoscopic images.

In spite of international standards and guidelines around the assessment of velopharyngeal function, Pereira et al. (2013b) found that less than half of the identified speech osteotomy studies in their review reported using an instrumental measure to assess velopharyngeal function. More recently, Smedberg et al. (2014) reported using both nasendoscopy and videoradiography in their speech osteotomy study. However, although reliability studies were undertaken, it was unclear if the study was undertaken retrospectively or prospectively. Additionally, their final sample size was reduced to N=9 with the exclusion of two participants due to association with a syndrome.

Nasometry (Pentax Medical) is another instrumental measure, an indirect method, for the assessment of velopharyngeal function. It is an internationally recognized tool for the acoustic measurement of nasality. The system produces the objective measure of nasalance based on a standard formula. Nasalance scores are highly language specific and so language specific norms are obligatory. The scores may also be inadvertently inflated by high proportion of high vowels and/or nasal consonants in the speech sample, as well as the presence of nasal airflow errors during speech. The relationship between perceptual ratings of nasality and nasalance scores has been found to rely on factors such as the speech sample (e.g. Sweeney and Sell, 2008). Notwithstanding, it continues to have both clinical and research utility (e.g. Kummer et al., 2012).

The aims of the current prospective study were to therefore address some of the study and speech methodological issues in the osteotomy literature in investigating the impact of maxillary osteotomy on velopharyngeal function in cleft lip and palate, using instrumental outcome measures. These included nasalance, an indirect acoustic measure of resonance and nasendoscopy and lateral videofluoroscopy, both direct assessments of velopharyngeal function during speech.

Materials and Methodology

This study was conducted according to the ethical principles of the World Medical Association Declaration of Helsinki. The study was reviewed and approved by the institutional review board (06NS08).

Participants

Twenty participants with a cleft palate +/- lip, representing a consecutive series of osteotomies, were recruited from a single regional cleft service. Written consent was obtained from all participants. All participants were native speakers of English and underwent maxillary osteotomy, with or without a mandibular setback, by a single surgeon. The mean age at surgery was 20;2 years (range = 18;1 – 30 years, SD = 2;6 years), with a gender distribution of 16 males to 4 females. None of the participants presented with hearing and/or learning difficulties rendering them unable to participate in any of the tasks. No participant had a known syndrome diagnosis. Table 1 shows participant details including cleft diagnosis and orthognathic surgery details. One participant (case 1) had a history of secondary velopharyngeal surgery. This participant had had a midline pharyngeal flap undertaken at age 7;9 years followed by a detachment of the flap and a muscle transfer pharyngoplasty/Orticochea pharyngoplasty (Orticochea, 1968) at 12;1 years. A group of normal controls (N=20), matched for age, was also recruited and seen for measurements of

Nasometry (Pereira, 2012). There were 10 females and 10 males and the mean age was 23;3 years (range = 19;8 - 26;0, sd = 2;11).

Baseline Measurements and Follow-up Time Points

There were three assessment time points: pre-surgery, 3-months and 12-months post-surgery. The most commonly reported post-surgery time points in the osteotomy literature is 6-months, followed by 3-months and then 12-months post-surgery (Pereira, 2013a). Three-months post-surgery was identified for use in the current study to capture early speech changes. By 12 months, the maxilla is considered to be relatively stable and thus a timeframe which can be justified for measuring outcomes after orthognathic surgery (Eurocran, 2003). Skeletal relapse of the maxilla which tends to occur within the first post-surgery year (Cheung and Chua, 2006) can have an impact on speech and velopharyngeal function post-surgery, and therefore justified the two post-surgery data points. Two (C12 and C18) of the 20 participants failed to attend the 3-month post- surgery appointments. There were no other missing datapoints.

Instrumental Measures

Nasality was assessed using the Nasometer II 6400 (*Kay Elemetrics, Pentax UK).* The speech sample consisted of 16 sentences grouped together according to whether they contained high pressure consonants, low pressure consonants or mixed consonants (Sweeney, 2000; Sweeney & Sell, 2007). The Nasometer was calibrated for each participant and the headset placed according to manufacturer's guidelines.

Lateral videofluoroscopy images were screened according to the standard clinical protocol. During the course of the study, the recording system was upgraded from Super VHS cassette recorder to a DVD system. An external hypercardiod condenser microphone (Rode NT 3) was used for all recordings. A head-alignment device with a calibration ring attached (Sommerlad, Rowland & Harland, 1994) was used during the procedure to reduce unwanted head movement, and to facilitate ratiometric measurements. The ring, located in the same midsagittal plane as the participant's head, was always screened at the same magnification in order to make quantitative ratiometric measurements. Lateral views were coned down to minimize radiation exposure and lasted for only 45 seconds per recording.

Nasendoscopic evaluation was undertaken by the first author. On-line recordings were undertaken for both nasendoscopy and lateral videofluoroscopy and speech sampling was based on the unit's clinical protocol and guidelines proposed by the International Working Group (Golding-Kushner et al., 1990).

Analyses and Coding

For lateral videofluoroscopy and nasendoscopy. two sets of analyses were undertaken: visual perceptual ratings (VPRs) and quantitative ratiometric measurements (QRMs) using validated methodology and software (Birch, Sommerlad & Bhatt, 1994; Birch et al. 1999; Gilleard, 2008; 2009). Data samples were edited to contain only production of the vowel /i/ with the soft palate at maximum closure and at rest for QRMs according to the methodology set out by Birch and colleagues (1994; 1999), whilst the full speech sample set was used for VPRs. All samples were randomized and re-labelled to prevent recognition of participant and/or time point.

Videofluoroscopy VPRs. Parameters identified were based on those used by the Unit as well as other published work (e.g. Golding-Kushner et al., 1990; Kummer, 2008). This included status of velopharyngeal closure (definitely adequate/probably adequate/borderline/probably inadequate/definitely inadequate), firmness of closure (touch type/firm/very firm/ and proportion of palate contacting the posterior pharyngeal wall (small/moderate/large). Ratings were undertaken independently by two experienced speech and language therapists (SLTs) in the field and consensus judgments for each parameter were made.

Videofluoroscopy QRMs. All videofluoroscopic images were converted to Audio Video Interleave (AVI) format using AVS Editor 4.2 (Online Media Technologies Ltd., 2010). Velar measurements were undertaken using Image Pro 6.3 (Media Cybernetics, 2009). Three velar parameters were identified for measurement based on the published and validated work by Birch et al. (1994;1999): *palatal lift angle*, *extensibility*, and *closure ratio*. A fourth velar parameter, *velar stretch* was identified for inclusion. Two experienced raters were convened to undertake these quantitative measurements. Both raters underwent direct training with the author of the methodology (Birch, M.). To identify the anatomical landmarks, raters used the slow-motion function in Image Pro on the videofluorosopic edits (production of the vowel /i/ at rest and with soft palate at maximum closure) (Figure 1a). The four velar parameters were measured based on standardized formulas. An example of measurements made with complete velopharyngeal closure is shown (Figure 1b).

Nasendoscopy and VPRs. Parameters identified were based on the unit's clinical protocol (Sell and Ma, 1990) as well as other published work (e.g. Golding-Kushner et al., 1990; Kummer, 2008). This included status of velopharyngeal closure (definitely adequate/probably adequate/borderline/probably inadequate/definitely inadequate), firmness of closure (touch type/firm/very firm) and velopharyngeal gap size (if any) (pinhole/small/moderate/large).

Nasendoscopy and QRMs. Nasendoscopic images were converted to AVI format and measurements were made using Image Pro 6.3 (Media Cybernetics, 2009). Ratiometric measurements of closure ratio were undertaken on production of /i/ with the soft palate at maximum closure and at rest (Gilleard, 2008). The rater manually traced the shape of the velopharyngeal gap if present, firstly, with the palate at maximum closure on production of /i/ (AREA 1) and secondly, with the palate at rest (AREA 2) (Figure 2). Closure ratio is calculated as AREA 1 divided by AREA 2 where a closure ratio of 1 indicates incomplete closure and no movement of the velopharyngeal sphincter and 0 indicates complete closure (Gilleard, 2009). Two raters were convened for this part of the study; rater 1 was a plastic surgery registrar and rater 2, a speech-language therapist

Reliability

For nominal-type data, the Kappa statistic (κ) was used and for interval data, inter- and intra-rater reliability were calculated using Pearson's correlation I, where a correlation coefficient of $r = \pm 0.3$ is weak/small, medium/moderate if $r = \pm 0.5$, and strong/large if $r = \pm 0.7$ (e.g. John Wiley & Sons, 2020). The correlation coefficient, however, reflects the association between two variables and does not consider the levels of categories in a rating scale. Hence, if one rater consistently rates one category above the second rater, the resultant coefficient will be 1, although both raters have not shown any or exact agreement (Pereira, 2012). Hence, an additional measure, percent agreement was also used. Percent of agreement between raters was calculated using two statistics: perfect agreement (Po) indicating that the two raters are rating on similar scalar points on a rating scale, and a less conservative agreement based on whether the raters agree to the precision of -1 to +1 scores (Po-1), where raters differ in their ratings by one scalar point on a rating scale. A conservative measure for agreement was adopted in that perfect agreement was deemed acceptable when percent agreement was equal to or more than 80%, and Po-1 was equal to or more than 90% (e.g. Lord & Corsello, 2005: p.738; Zarcone et al., 1991).

Nasalance. Test-retest reliability of the Nasometer with head gear change was undertaken, as there is evidence of test re-test variability in hypernasal speakers when the Nasometer headset is removed between testings (Watterson & Lewis, 2006). In our study, five participants were re-tested on the Nasometer at a post-surgery time point, either in two separate sessions in one day (am: test 1 and pm: test 2), or at the beginning (test 1) and at the end of the session (test 2). In both scenarios, the Nasometer headset was removed before the second recording. There was a minimum time lapse of 45 minutes between Test 1 and Test 2. The mean nasalance at Test 1 was 36.6% (sd = 11.8) and 34.6% (sd = 12.1) at Test 2 (mean difference = 5.6%, sd = 2.7, range = 3-10%). There was no statistically significant difference in test-re-test reliability $t(4) = 0.694$, $p = .526$.

Videofluoroscopy VPRs. Two experienced speech and language therapists were convened for the reliability studies. Ten samples were randomly identified and repeated for the calculation of intra-rater reliability. Inter-rater reliability for ratings based on /i/ ranged $r_s = .635$ to $\kappa = 0.815$, reflecting moderate to large agreement or correlations and percent agreement from 47.8% to 77.8% for Po, and from 88.9% to 98.4% for Po-1. Inter-rater reliability for ratings based on the full speech sample ranged from $\kappa = 0.475$ to $r_s = .781$, reflecting moderate to large agreement or correlations and percent agreement from 47.8% to 85.1% for Po and from 88.1% to 95.5% for Po – 1. Intra-rater reliability (rater 1) for ratings based on /i/ ranged from $\kappa = 0.545$ to 1.000, reflecting moderate to large agreement or correlations and percent agreement from 40%-80% for Po and at 100% for Po-1 for the range of parameters. Intra-rater reliability (rater 1) for ratings based on the full speech sample ranged from $\kappa = 0.500$ to to $r_s = .818$, reflecting moderate to large agreement or correlations and percent agreement from 40%-80% for Po and from 90%-100% for Po-1.

Videofluoroscopy QRMs. The same ten samples randomly selected and included as repeats in the intra-operator reliability visual perceptual ratings study were used in the inter-rater ratiometric reliability studies. Inter-rater reliability for the range of velar parameters ranged from r = .423 to r = .827 and intra-rater reliability for rater 1 ranged from r = .303 to r = .884, reflecting moderate to large correlations in the main.

Nasendoscopy VPRs. Ratings were undertaken independently and a consensus judgment for each parameter was made. Twelve samples (20%) were randomly identified and included in the dataset for the calculation of intra-rater reliability. Inter-rater reliability for ratings based on /i/ sample ranged from $\kappa = 0.776$ to $r_s = .896$, reflecting large correlations and percent agreement from 74.5% to 83% for Po and from 81.7% to 96.4% for Po-1. Inter-rater reliability for ratings based on the full speech sample ranged from $r_s = .819$ to $r_s = .898$, reflecting large correlations and percent agreement from 75.4% to 87.7% for Po and from 96.5% to 96.6% for Po-1. Intra-rater reliability (rater 1) based on /i/ ranged from $r_s = .893$ to $\kappa = 1.000$, reflecting large correlations and percent agreement from 75% to 100% for Po and 100% for Po-1. Intra-rater reliability for the full speech sample ranged from $\kappa = 0.676$ to $r_s = .888$, reflecting large correlations and percent agreement at 66.7% for Po and from 91.7% to 100% for Po-1.

Nasendoscopy QRMs. All measurements were undertaken independently. Intra-rater reliability of the main rater was r = .88 and inter-rater reliability was r = .94, both statistically significant at $p < .001$.

Skeletal Relapse

Standard lateral cephalometric radiographs were taken immediately post-surgery (PS_1, $\bar{x}$ = 13.4 days) and at the end of orthodontic treatment after surgery (PS_2, $\bar{x}$ = 14.8 months). Films were traced under standard conditions and the following points traced: centre of sella turcica (S), nasion (N) and deepest point on the anterior contour of the maxillary arch (A). An approximate Frankfort plane was constructed from the SN line and a perpendicular dropped through S and the horizontal position of point A (Hor A) was measured at right angles to this line. Almost 10% of the total sample available was randomly identified for use in the reliability studies. Two raters were convened where rater 1 was a Registrar in Orthodontics and rater 2 was a Consultant Orthodontist). Intra-rater (rater 1) and inter-rater reliability were r = 0.989 and r = 0.989 respectively, both statistically significant at $p < 0.001$, reflecting large correlations.

Results

All statistical analyses were undertaken using SPSS and effect sizes were calculated using G*Power (Faul et al., 2007) where $d \geq 0.2$ is small, $d \geq 0.5$ is medium and $d \geq 0.8$ is large (Cohen, 1992). Tests of normality and homogeneity of variance were undertaken. For within-subject comparisons across time (T1, T2 and T3), a repeated measures ANOVA was undertaken for interval data and a Friedman's test was used for ordinal data or when the normality assumption was violated. Planned comparisons across pairs of timepoints were undertaken using paired samples t-test for interval data and Wilcoxon Signed Rank Test for ordinal data or when the normality assumption was not met. Effect sizes were calculated and reported as they "indicate the strength of the association between two variables or the size of the normalized difference between the means, and enable objective comparisons to be made across studies (Egger et al., 1997; Coe, 2002), thus serving to facilitate the identification of best evidence" (Pereira et al., 2013b: p. 26-27). In addition to the two participants (C12 and C18) who failed to attend the 3-month post- surgery appointments, Case 1 was excluded from the videofluorosopic and nasendoscopic analyses in view of previous velopharyngeal surgery (pharyngeal flap in-situ).

Nasalance

The mean nasalance for the group at T1 was 23% (SD = 10.1, range = 7 – 44%), at T2, 31% (SD = 11.4, range = 11 – 55%) and at T3, 33% (SD = 11.2, range = 20 – 63%). The mean nasalance for normal speakers matched for age and gender (Pereira, 2012) was 25.9% (SD = 5.3, range = 17 – 37%). Using a repeated measures ANOVA, there was a statistically significant difference in nasalance over time, $F(2,34) = 8.020$, $p = .001$. Of the 16 cases who had normal nasalance at T1, six cases (37.5%) (C5, C12, C16, C17, C19 and C20) showed increases in nasalance post-surgery at T3 that was over a nasalance value of 31.2% (+1sd of normal nasalance mean) with an average increase of 18.6% (range = 7-41%, sd = 13.02). Individual nasalance scores across time are shown in Table 2. Planned comparisons across pairs of time points were statistically significant between T1-T2, $t(17) = -2.991$, $p = .008$ and between T1-T3, t(19) = -3.694, p = .002 but not between T2-T3. Effect sizes approached large between T1-T2 (d = 0.726) and large between T1-T3 (d = 0.894), suggesting a true effect of the surgery on nasalance in CLP. Together with a non-significant difference, the small effect size between T2-T3 (d = 0.166) suggests that the change in nasalance found at 3-months post-surgery is stable and permanent.

Lateral Videofluoroscopy

VPRs. Bivariate correlations were undertaken between ratings made on /i/ and ratings made on the full speech sample set across the range of velar parameters. Of the 20 possible correlations, 13 (65%) were statistically significant and 16 (80%) were moderate to large correlations (Table 3). A full speech sample set consists of a range of consonant types and includes samples across the linguistic hierarchy. As such, it has better face validity and is of clinical interest. Statistical results based on the full speech sample set are therefore reported here. There were no statistically significant differences for *presence or absence of a velopharyngeal defect (VP)*, Cochran's Q = 2.800, p = .247, *size of VP defect*, $\chi^2(2)$ = 2.800, p = .247, *adequacy of VP closure*, $\chi^2(2)$ = 1.902, p = .386, *firmness of closure*, $\chi^2(2)$ = 1.188, p = .552 and *proportion of palate contacting the posterior pharyngeal wall (PPW)*, $\chi^2(2)$ = 3.250, p = .197 (Figure 4.14). The only statistically significant difference over time was for *presence or absence of Passavant's Ridge* (Cochran's Q = 9.600, p = .008). Planned comparisons across pairs of time points (T1-T2, T1-T3 and T2-T3) showed no significant differences for any of the parameters.

QRMs. Repeated measures ANOVA showed significant main effects of time for *palatal lift angle*, F(2,30) = 6.362, p = .012, *closure ratio*, F(2,30) = 7.723, p = .002. Planned comparisons across pairs of time points were in the main, significant for all three velar parameters between T1-T2 and T1-T3 but not for T2-T3 (Table 4). Effect sizes were medium to large between T1-T2 and T1-T3 and less than small between T2-T3 for all three parameters (Table 4).

Nasendoscopy

VPRs. As with videofluoroscopy, ratings on /i/ were significantly correlated with ratings on the full speech sample for eight of the 12 velar parameters across the three time points and as such, results based on the full speech sample are reported here. There were no statistically significant differences over time for *presence/absence of a velopharyngeal(VP) defect,* $\chi^2(2) = 1.190$, p = .551, *size of VP gap*, $\chi^2(2) = 2.000$, p = .368, *adequacy of VP closure*, $\chi^2(2) = 1.111$, p = .574, and *firmness of closure against PPW*, $\chi^2(2)= 1.636$, p = .441. Planned comparisons across pairs of time points also showed no significant differences for any of the parameters.

QRMs. There was no significant change over time for *closure ratio* $\chi^2(2) = 2.179$, p = .336 and no significant differences across pairs of time points.

Nasalance, Lateral Videofluoroscopy and Nasendoscopy

In addition to the two participants (C12 and C18) who failed to attend the 3-month post-surgery appointments, Case 1 was excluded from the videofluorosopic and nasendoscopic analyses in view of previous velopharyngeal surgery (pharyngeal flap in-situ). Parametric and non-parametric correlations were run between Nasalance and Lateral Videofluoroscopy range of VPRs and QRMs, between Nasalance and Nasendoscopy VPRs and QRMs and between Nasendoscopy and Lateral Videofluoroscopy VPRs and QRMs across the three time points. Only the following correlations were found to be statistically significant.

Nasalance and Lateral Videofluoroscopy. Significant correlations were found between nasalance and *presence/absence of velopharyngeal defect* at T2, $r_s = .828$, p = .042 and *adequacy of velopharyngeal closure*, $r_s = .510$, p = .036, and at T3 for *proportion of palate contacting the posterior pharyngeal wall*, $r_s = -.705$, p =.005. For QRMs, the only significant correlation was at T1 for and *closure ratio*, r = -.581, p =.009.

Nasalance and Nasendoscopy. Significant correlations were found for *presence/absence of a velopharyngeal gap* at T3, $r_s = .604$, p = .005, *size of velopharyngeal gap* at T2, $r_s = .632$, p =.006 and at T3, $r_s = .635$, p = .003, *adequacy of velopharyngeal closure* at T2, $r_s = .627$, p = .007 and at T3, $r_s = .675$, p = .001, and *firmness of closure against posterior pharyngeal wall* at T2, $r_s = -.494$, p = .044 and at T3, $r_s = -.562$, p = .010. For QRMs, this was significant at T2, r = .667, p =.002.

Lateral Videofluoroscopy and Nasendoscopy. There were significant correlations for *size of velopharyngeal gap* at T2, $r_s = .880$, p =.021, *adequacy of velopharyngeal closure* at

T1, r_s = .496, p = .036 and for *firmness of closure against posterior pharyngeal wall* at T1, r_s = -.572, p = .021. For QRMs, *closure ratio* was significant at T1, r = -.535, p = .018 and T3, r = -.596, p = .009.

Skeletal Relapse

The formula used for measuring skeletal relapse was: Hor A_PS_2 – Hor A_PS_1. No significant difference was found between Hor A measurements at PO_1 and at PO_2, t(11) = 1.983, p = .073. indicating no skeletal relapse that could have impacted on speech changes.

Discussion

The risk of acquiring velopharyngeal insufficiency following maxillary osteotomy has been reported by several authors (e.g. Trindade et al., 2003; Pereira et al., 2013b). The evidence for the impact of the surgery on instrumental outcomes of velopharyngeal function has been mixed, due to speech and study methodological issues as well as the use of different types of instrumental methods (Pereira et al., 2013b). The current study attempted to address some of these methodological issues.

The findings of the study provide evidence that maxillary osteotomy can result in increased nasalance post-surgery. In this study, of the 16 participants who had normal nasalance scores before surgery, six participants acquired increases beyond one standard deviation above the normal mean right after surgery. Statistical analyses further showed that the increase in nasalance was maintained a year post-surgery. Calculation of effect sizes provided further evidence, with large or almost large effect sizes, between the pre- and post-surgery time points compared with a small effect size only between the two post-surgery time points. These findings contrast with those reported by Chua et al. (2010) and Smedberg et al. (2014) but are similar to those by Trindade et al. (2003) who also reported maintenance of nasalance findings between 45 days and 9-months post-surgery. The large effect sizes seen in this study between the pre- and post-surgery time points indicate a possible true effect of maxillary osteotomy on this speech parameter in CLP. The results also parallel the perceptual speech results reported in our earlier paper (XXXXXXX) where resonance rated perceptually and velopharyngeal composite scores deteriorated significantly post-surgery at a group level and similarly, were maintained between the two post-surgery time points.

In terms of direct assessment of velopharyngeal function, analyses of visual perceptual ratings based on lateral videofluoroscopy and nasendoscopy at a group level showed no significant differences across time points for any of the parameters. However, for one parameter, *proportion of palate contacting the posterior pharyngeal wall*, based on videofluoroscopic images, the difference between the pre-surgery and 3-months post-surgery time point was approaching statistical significance with a medium effect size suggesting a possible true effect of maxillary osteotomy on this parameter. This parameter is defined as "the extent of contact between the velar eminence (the high point on the top of the "knee") down through the vertical part of the velum" (Kummer, 2008: p.456). A small proportion of contact would signify tenuous velopharyngeal closure or what is known clinically as "touch" closure. For the six cases with elevated nasalance at twelve months post-surgery (+1 sd above the normal mean), the proportion of palate contacting the posterior pharyngeal wall, was rated as 'small' for five of the cases (C5, C12, C17, C19 and C20) at the pre-surgery data point, suggesting a possible relationship between the two parameters. Of further interest is that this rating of 'small' was in the context of normal nasalance scores. Although bivariate correlations were not significant between the two parameters, these findings suggest that *proportion of palate contacting the pharyngeal wall* may be a plausible risk factor or predictive factor in the acquisition of velopharyngeal insufficiency following maxillary advancement surgery, implicating the importance of direct visualization.

Ratiometric analyses showed significant changes post-surgery for *closure ratio*, *velar stretch* and *palatal lift angle*. For the same six cases who had increased nasalance post-surgery, four cases (C12, C16, C17 and C19), had a pre-surgery *closure ratio* measurement score of '1.0' indicating complete velopharyngeal closure, decreasing to 0.73 (C16), 0.50 (C17) and 0.73 (C19) a year after surgery. Bivariate correlation between nasalance and closure ratio was large and significant at pre-surgery but not significant at either post-surgery data point, implicating the role and contribution of other plausible factors in the increased nasalance seen post-surgery.

Another interesting phenomenon observed in the study was Passavant's Ridge. The measurement of closure ratio did not take into account Passavant's Ridge. For the six cases with increased nasalance post-surgery, Passavant's Ridge was not observed in any of the cases pre-surgery. Post-surgery, four (C16, C17, C19 and C20) of the six cases presented with Passavant's Ridge, which did not appear to aid in velopharyngeal closure, given the elevated nasalance scores. In fact, Skolnick (1989) reported seeing ridges on the posterior pharyngeal wall appear following maxillary osteotomy unrelated to any compensatory velopharyngeal phenomenon.

In terms of *velar stretch*, the mean value increased post-surgery from 0.582 to 0.663. Pruzansky and Mason (1969) were the first to describe increases in the intrinsic length of the soft palate during speech in individuals with velopharyngeal dysfunction based on lateral cephalometric x-rays, which they referred to as "stretch factor". This stretch factor, which is vital in achieving velopharyngeal closure for speech, appears to be adversely affected in individuals with CLP, due to the possible hypoplastic nature of the velum and associated tissues and presence of scarring as a result of the primary surgery. It is hypothesized this may have a tethering effect impeding muscular stretch and movement (Schendel et al., 1979; Witzel et al., 1989). In contrast, a decrease in *palatal lift angle* was found in the current study. The measured angle, represented by AF1 (angle formed by P2-P1-P3, Figure 1) decreases as the stretch phenomenon increases. This is because as velar stretch (represented by P2-P1/P1 to tip of uvula) increases, the angle AF1 becomes more acute, decreasing its measured value. For the same six cases with increased nasalance post-surgery, palatal lift angle measurements fell below the group mean of 33.34 (sd = 9.03) post-surgery at 29.3. Bivariate correlation between nasalance and palatal lift angle was not significant implying no direct clinical relationship between the two parameters.

In terms of nasendoscopy, there were no statistically significant changes seen post-surgery for any of the visual perceptually rated parameters, reflecting the results for lateral videofluoroscopy. The results were also non-significant for *closure ratio* which was measured quantitatively. Although bivariate correlations were significant between nasalance and nasendoscopic velar parameters, these statistical findings do not reflect the clinical picture, in comparison with findings from videofluoroscopy. For example, C17 had an elevated nasalance score of 36% (>+1sd above the normal mean) twelve-months after surgery, but visual perceptual ratings indicated 'no velopharyngeal defect', 'definitely adequate velopharyngeal closure', 'firm' firmness of closure and quantitative ratiometric measurement of closure ratio was approaching '0' at 0.07, indicating almost complete velopharyngeal closure.

The non-significant statistical results over time for nasendoscopy are in line with those reported by Chua et al. (2010) and Smedberg et al. (2014). With nasendoscopic images, the whole of the velopharyngeal portal needs to be recorded at rest and during speech for valid or more accurate measurements. Unfortunately, it may not always be possible to get the ideal distance of the scope from the velopharyngeal mechanism, compounded by a poor angle of the scope (Sell and Pereira, 2011:p.154). Additionally, with visual perceptual ratings ordinal type scales may not be sensitive enough to capture subtle changes. In this study, no operational definitions or descriptions were provided for each scalar point for both videofluoroscopy and nasendoscopy, in contrast to detailed descriptions provided for the perceptual rating of hypernasal resonance (e.g. John et al., 2006).

With regard to the inter-relationship between the three instrumental measures, the results were variable. A significant correlation was found between nasalance and 'closure ratio' based on pre-surgery lateral videofluoroscopic images. As already described above, a plausible reason is that other velar factors may play a contributory role in velopharyngeal closure, particularly post-surgery, where other velar parameters potentially play a significant contributory role to velopharyngeal closure e.g. palate extensibility. Additionally, the presence of Passavant's ridge was found to aid in velopharyngeal closure, albeit for some cases only. The lack of consistent relationship between nasalance and 'closure ratio' measured on nasendoscopic images is attributable to the difficulty in always visualizing the entire velopharyngeal portal, assessing the cephalocaudal position of maximum closure during quantitative measurement, and issues with lens or barrel distortion (Lam et al., 2006; Gilleard, 2008; 2009; Sell and Pereira, 2011). Significant relationships were found for 'closure ratio' between measurements made on nasendoscopic images and those made on videofluoroscopic images using the methodology based on Birch et al. (1994;1999), at both pre-surgery and post-surgery time points, suggesting clinical and research validity and utility of the measurement method.

Study Limitations

One of the study limitations is regarding the reliability of the ratiometric quantitative analyses of velar parameters based on videofluorosopy. Inter-rater reliability was variable, ranging from r = .423 to r = .827. Although a value of r = .423 still reflects a medium sized correlation, Birch et al. (1999) reported high agreement between raters for the range of velar parameters except for velar stretch which was not a published or validated parameter. The method requires accurate identification of anatomical landmarks and perhaps more intensive training is indicated.

It would have been ideal to stratify the group according to surgery type ie. maxillary or bimaxillary. However, in the field of cleft speech osteotomy studies where sample sizes are generally small, stratifying the group would result in smaller sample sizes per group. In the case of our cohort, this would also result in unequal sample sizes of N=7 (Bimaxillary) and N = 13 (Maxillary osteotomy only), where the risk, potentially, is in committing a Type I error (Glass et al., 1972).

A final limitation was the focus only on impairment-based outcomes. There is a need to consider and include functional speech outcomes such as intelligibility and acceptability (e.g. Henningson et al., 2008) but also patient reported outcomes such as CLEFT-Q© which measures *speech function* and *health-related quality of life scales* (e.g. speech distress) (Wong et al., 2013; Klassen et al., 2018). The meaning for the patient of these speech changes would be better understood.

Conclusions

The results of this study suggest that maxillary osteotomy can potentially have an adverse impact on velopharyngeal function as measured instrumentally either indirectly, nasalance, or directly, lateral videofluoroscopy, using both visual perceptual ratings of velopharyngeal function and quantitative ratiometric measurements of velar parameters. Reporting of effect sizes is important to understand the strength of the evidence. The larger effect sizes seen between the pre- and three-month and pre- and 12-month post-surgery time points as compared with the smaller effect sizes and non-significant results between both post-surgery time points suggest that speech changes seen early on at three-months post-surgery are stable and permanent. The results provide evidence that an earlier post-surgery speech review is valid, even as early on as 3-months post-surgery. The study findings also support the use of instrumentation in the assessment of velopharyngeal function in the osteotomy care pathway with the evidence pointing to the clinical and research utility of the Nasometer and lateral videofluoroscopy.

Acknowledgements

The authors would like to thank all the raters and participants in this study and Dr Malcolm Birch for his support and training with quantitative measurement using Image-Pro software. The first author would also like to thank Professor XXXXXXXXXX and Professor XXXXXXXX for supporting her continued work and interest in the field of speech and cleft lip and palate.

Declaration of Conflicting Interests

The author(s) declare no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by Sparks Children's Medical Research Charity [grant number 2DTK] awarded to XXXXXX.

ORCID iD

XXXXXXXXXXXXXX, PhD XXXXXXXXXXXXXXXX

References

Birch M, Sommerlad B C, Bhatt A. Image analysis of lateral velopharyngeal closure in repaired cleft palates and normal palates. *Br J Plast Surg*. 1994;47(6):400-405.

Birch M J, Sommerlad B C, Fenn C, Butterworth M. A study of the measurement errors associated with the analysis of velar movements assessed from lateral videofluoroscopic investigations. *Cleft Palate Craniofac J*. 1999;36(6):499-507.

Cheung L K, Chua H D. A meta-analysis of cleft maxillary osteotomy and distraction osteogenesis. *Int J Oral Maxillofac Surg.* 2006;35(1):14-24.

Cleft Palate International Speech Issues (CLISPI). Available at https://clispi.com. Accessed August 9, 2019.

Coe R. (2002). *It's the Effect Size, Stupid: What Effect Size is and Why it is Important*. Available at http://www.leeds.ac.uk/educol/documents/00002182.htm.

Chua H D, Whitehill T L, Samman N, Cheung L K. 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.

Cohen J. A power primer. *Psychol Bull.* 1992;112(1):155-159.

Dalston RM, March JL, Vig KW, Witzel MA, Bumsted RM. Minimal standards for reporting the results of surgery on patients with cleft lip, cleft palate, or both: a proposal. *The Cleft Palate J.* 1988; *25:* 3-7.

Egger M, Smith G D, Phillips A N. Meta-analysis: principles and procedures. *BMJ.* 1997;315(7121):1533-1537.

Eurocran (2003). *Timing of Minimal Records*. Available at http://www.eurocran.org/content.asp?contentID=779.

Faul F, Erdfelder E, Lang A G, Buchner A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods.* 2007;39(2):175-191.

Glass G, Peckham P, Sanders J. Consequences of failure to meet assumptions underlying the fixed effects of analyses of variance and covariance. *Review of Educational Research,* 1972; 42: 237-288.

Gilleard, O. The Use of Nasendoscopy in Submucous Cleft Palate. London UK: University of London; 2008. Masters Dissertation.

Gilleard O, Sommerlad B, Sell D, Ghanem A, Birch M. Nasendoscopic image analysis: an analysis of measurement uncertainties. In C.E. Raposo do Amaral and A. B. Albino de Amaral (Eds.), *11th International Congress on Cleft Lip and Palate and Related Craniofacial Anomalies International Proceedings*, Medimond s.r.l. Monduzzi Editore, Bologna, Italy, 2009:179-184.

Greenhalgh, T. *How to Read A Paper.* (2nd ed.). London: BMJ Books, 2001.

Golding-Kushner K J, Argamaso R V, Cotton R T, Grames L M, Henningsson G, Jones D L et al. Standardization for the reporting of nasopharyngoscopy and multiview videofluoroscopy: a report from an International Working Group. *Cleft Palate J.* 1990;27(4):337-347.

Good P M, Mulliken J B, Padwa B L. Frequency of Le Fort I osteotomy after repaired cleft lip and palate or cleft palate. *Cleft Palate-Craniofac J.* 2007;44(4):396-401.

Grunwell P, Sell D. Speech and cleft palate/velopharyngeal anomalies. In A C H Watson, D Sell, P Grunwell (Eds.), *Management of Cleft Lip and Palate*. London and Philadelphia: Whurr Publishers, 2005:68-86.

Harding A, Grunwell P. Active versus passive cleft type speech characteristics. *International Journal of Language and Communication Disorders*, 1988;33(3):329–352.

Henningsson G, Kuehn D P, Sell D, Sweeney T, Trost-Cardamone J E, Whitehill T L. Universal parameters for reporting speech outcomes in individuals with cleft palate. *Cleft Palate-Craniofac J.* 2008;45(4):1-17.

Impieri D, Tonseth KA, Hide O, Brinck EL, Hogevold HE, 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.

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.

John Wiley & Sons. (2020, Jan 14). *Dummies: how to interpret a correlation coefficient r.* https://www.dummies.com/education/math/statistics/how-to-interpret-a-correlation-coefficient-r/

Klassen AF, Riff KWW, Longmire NM, Albert A, Allen GC, Aydin MA, Baker SB, Cano SJ, Chan AJ, Courtemanche DJ, Dreise MM, Goldstein JA, Goodacre TEE, Harman KE, Munill M, Mahony AO, Aguilera MP, Peterson P, Pusic AL, Slator R, Stiernman M, Tsangaris E, Tholpady SS, Vargas F, Forrest CR. Psychometric findings and normative values for the CLEFT-Q based on 2434 children and young adult patients with cleft lip and/or palate from 12 countries. *CMAJ*. 2018;190(15): E455-E462.

Kuehn D, Moller K. Speech and Language Issues in the Cleft Population: The State of the Art. *Cleft Palate-Craniofac J.* 2000; 37(4):1-35.

Kummer A W. *Cleft Palate and Craniofacial Anomalies*. (2nd ed.). NY: Delmar Cengage Learning, 2008.

Kummer A W. Current practice in assessing and reporting speech outcomes of Cleft Palate and velopharyngeal surgery: a survey of Cleft Palate/Craniofacial professionals. *Cleft Palate-Craniofac. J.* 2012; 49(2):146-152.

Lam DJ, Starr JR, Perkins JA, Lewis CW, Eblen LE, Dunlap J, Sie KCY. A comparison of nasendoscopy and Multiview videofluoroscopy in assessing velopharyngeal insufficiency. *Otolaryngology-Head and Neck Surg*., 2006:134:394-402.

Lohmander A, Willadsen E, Persson C, Henningsson G, Bowden M, Hutters B. Methodology for speech assessment in the Scandcleft project--an international randomized clinical trial on palatal surgery: experiences from a pilot study. *Cleft Palate-Craniofac J.* 2009;46(4):347-362.

Lord C, Corsello C. Diagnostic instruments in autistic spectrum disorders. In F.R.Volkmar, R. Paul, A. Klin, & D. J. Cohen (Eds.), *Handbook of Autism and Pervasive Developmental Disorders*. New York: Wiley. 2005:730-771.

Mercer N S G, Pigott R W. Assessment and surgical management of velopharyngeal dysfunction. In A.C.H. Watson, D. Sell, & P. Grunwell (Eds.), *Management of Cleft Lip and Palate*. London and Philadelphia: Whurr Publishers, 2005:258-285.

Orticochea M. Construction of a dynamic muscle sphincter in cleft palates. *Plast Reconstr Surg*. 1968;41(4):323–327.

Pereira V. *The effect of maxillary advancement on speech, nasality and velopharyngeal function in cleft lip and palate*. [PhD Thesis], London; UCL Institute of Child Health; 2012.

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

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.

Pruzansky S, Mason R M. (1969). The "Stretch Factor" in soft palate function. *J Dent Res.* 1969;48(5):972.

Ross R B. Treatment variables affecting facial growth in complete unilateral cleft lip and palate. *The Cleft Palate J.* 1987;24(1):5-77.

Schendel S A, Oeschlaeger, M., Wolford, L. M., Epker, B. N. Velopharyngeal anatomy and maxillary advancement. *J Maxillofac Surg.* 1979;7(2):116-124.

Sell D, Ma L. Great Ormond Street Hospital Multiview Videofluroscopic Assessment of Velopharyngeal Function Proforma. Unpublished Clinical Protocol, 1990.

Sell D, Harding A, Grunwell P. A screening assessment of cleft palate speech (Great Ormond Street speech assessment). *Eur J of Dis of Commun.* 1994;29(1): 1-15.

Sell D, Harding A, Grunwell P. GOS.SP.ASS.'98: an assessment for speech disorders associated with cleft palate and/or velopharyngeal dysfunction (revised). *Int J Lang Commun Disord.* 1999;34(1): 17-33.

Sell D. Issues in perceptual speech analysis in cleft palate and related disorders: a review. *Int J Lang Commun Disord.* 2005;40(2): 103-121.

Sell D, Pereira V. Instrumentation in the analysis of velopharyngeal mechanism. In S. Howard and A. Lohmander (Eds.), *Cleft Palate Speech Assessment and Intervention.* UK: John Wile & Sons Ltd.

Semb G. A study of facial growth in patients with unilateral cleft lip and palate treated by the Oslo CLP Team. *Cleft Palate-Craniofac J.* 1991;28(1):1-21.

Skolnick L, Cohn E. *Videofluoroscopic Studies of Speech in Patients with Cleft Palate*. New York: Springer-Verlag, 1989.

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.

Sommerlad B C, Mehendele F V, Birch M J, Sell D, Hattee C, Harland K. Palate re-repair revisited. *Cleft Palate Craniofac J*. 2002;39(3):295-307.

Sweeney T. The Perceptual and Instrumental Assessment of Nasality and Nasal Airflow Errors Associated with Velopharyngeal Dysfunction. Dublin: School of Clinical Speech and Language Studies, University of Dublin; 2000. Dissertation.

Sweeney T, Sell D. Relationship between perceptual ratings of nasality and nasometry in children/adolescents with cleft palate and/or velopharyngeal dysfunction. *Int J Lang Commun Disord.* 2008;43(3):265-282.

Trindade I E, Yamashita R P, Suguimoto R M, Mazzottini R, Trindade A S Jr. Effects of orthognathic surgery on speech and breathing of subjects with cleft lip and palate: acoustic and aerodynamic assessment. *Cleft Palate-Craniofac J.* 2003;40(1):54-64.

Trost J E. Articulatory additions to the classical description of the speech of persons with cleft palate. *Cleft Palate J.* 1981;18(3):193-203.

Watterson T, Lewis K E. Test-retest nasalance score variability in hypernasal speakers. *Cleft Palate-Craniofac J.* 2006;43(4):415-419.

Witzel M. A. Commentary. *The Cleft Palate J.* 1989;26(3):199-200.

Wong K W, Forrest C R, Goodacre T E, Klassen A F. Measuring outcomes in craniofacial and pediatric plastic surgery. *Clin Plast Surg*. 2013;40(2):305-312.

Zarcone J R, Rodgers T A, Iwata B A, Rourke D A, Dorsey M F. Reliability analysis of the Motivation Assessment Scale: a failure to replicate. *Res Dev Disabil.* 1991;12(4):349-360.

Zeichmester J S, Zeichmester E B, Shaughnessy J J. *Essentials of Research Methods in Psychology*. New York: McGraw-Hill; 2001.

Figure Legends

Figure 1a. Marking of anatomical landmarks and lines drawn for quantitative ratiometric velar measurements.

Figure 1b. An example of the quantitative ratiometric velar measurements made in the context of complete velopharyngeal closure.

Figure 2. Manual tracing of velopharyngeal gap on production of /i/. The software calculates the area within the traced shape and provides a value in pixels.

Figure 1a

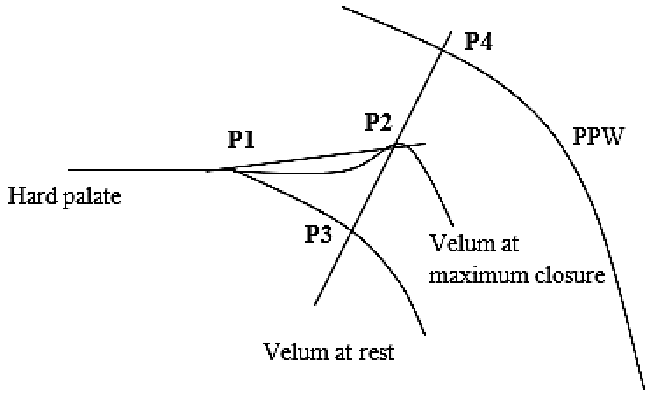
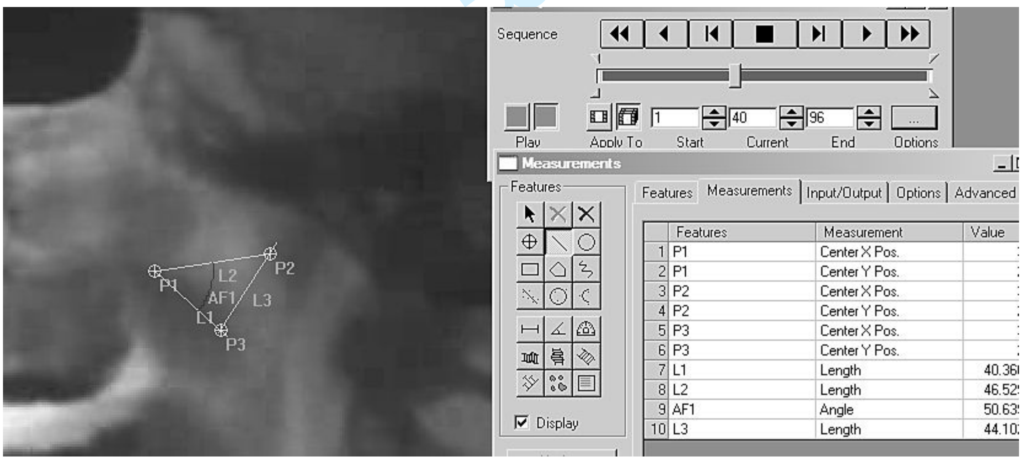


Figure 1b



Extensibility = L2 (P1-P2) / L1 (P1-P3)

Palatal Lift Angle = AF1 (angle formed by P2-P1-P3)

Closure Ratio = L3 (P2-P3) / L3 (P2-P3)

Velar Stretch = L2 (P2-P1)/ P1 to tip of uvula

Figure 2

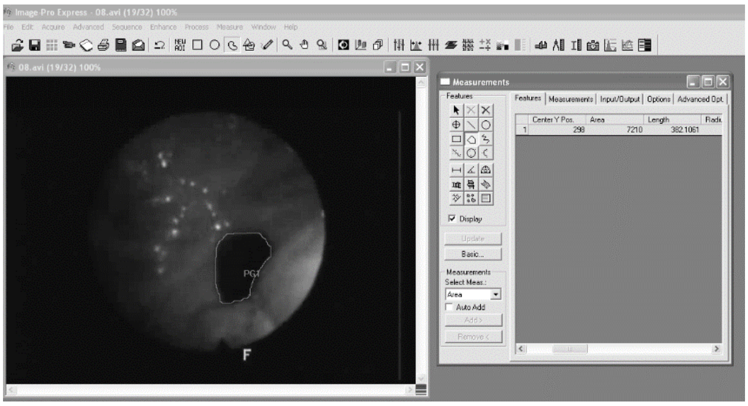


Figure 2

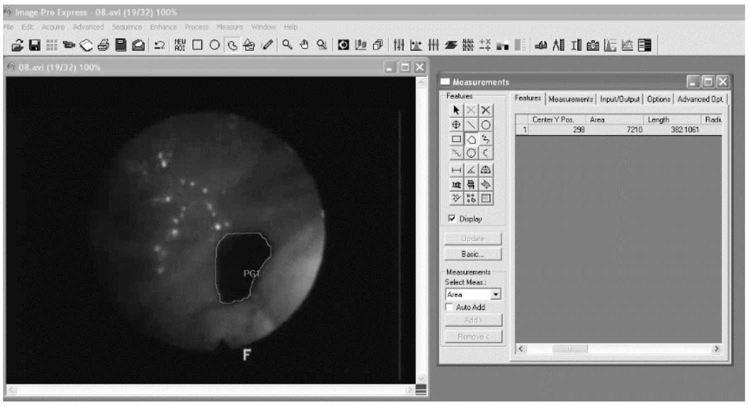


Image-Pro Express - 08.avi (19/32) 100%
File Edit Acquire Advanced Sequence Enhance Process Measure Window Help
08.avi (19/32) 100%
PG1
F
Measurements
Features
Features
Measurements
Input/Output
Options
Advanced Opt.
Center Y Pos.
Area
Length
1
298
7210
382.1061
Display
Update
Basic...
Measurements
Select Meas.:
Area
Auto Add
Add >
Remove <

Table 1. Participant and surgical details.

Case Number	Sex	Cleft Diagnosis	Orthognathic Surgery	Age at Orthognathic Surgery (years)
1	M	BCLP	Bimaxillary	21:1
2	M	RCLP	Le Fort I	20:3
3	F	RCLP	Le Fort I	19:3
4	M	LCLP	Le Fort I	18:3
5	M	BCLP	Le Fort I	19:3
6	F	LCLP	Le Fort I	19:0
7	F	RCLP	Le Fort I	18:4
8	M	LCLP	Le Fort I	20:6
9	M	RCLP	Bimaxillary	20:3
10	M	LCLP	Le Fort I	18:11
11	M	LCLP	Bimaxillary	21:9
12	M	LCLP	Bimaxillary	20:9
13	M	LCL+SPC	Bimaxillary	19:1
14	M	LCLP	Le Fort I	22:0
15	M	LCLP	Le Fort I	21:1
16	M	LCLP	Bimaxillary	19:5
17	F	BCLP	Le Fort I	18:5
18	M	LCLP	Le Fort I	18:1
19	M	BCLP	Le Fort I	20:1

20	M	RCLP	Bimaxillary	30:1

Abbreviations: BCLP, bilateral cleft lip and palate; RCLP, right-sided cleft lip and palate; LCLP, left-sided cleft lip and palate; LCL+SPC, left sided cleft lip and soft palate cleft.

Table 2. Nasalance Scores (%) for Each Participant Across Time.

Case No.	T1	T2	T3
1	44	45	63
2	39	55	34
3	7	38	31
4	21	29	26
5	28	23	41
6	27	30	20
7	19	24	31
8	33	29	29
9	16	17	29
10	20	28	29
11	21	20	20
12	18	---	33
13	42	44	46
14	17	11	20
15	13	20	23
16	14	43	32
17	19	35	36
18	24	---	21
19	12	39	53
20	27	34	34

Table 3. Bivariate Correlation Results for Visual Perceptual Ratings of Videofluoroscopic Images on /i/ and on the Full Speech Sample Set Across Time Points

Velar Parameter	Time Point		
	T1 (r / r_{pb})	T2 (r / r_{pb})	T3 (r / r_{pb})
Presence or absence of a VP defect	.215	.633**	.633**
Size of VP defect	.081	.748**	.662**
Adequacy of VP closure	.659**	.652**	.773***
Firmness of closure	.543*	.820***	.653**
Proportion of palate contacting PPW	.155	.650**	.703**
Presence or absence of PR	.322	.450	.783***
PR aiding in closure (yes/no)	.667	.730	not computed

Abbreviations: VP, velopharyngeal; PPW, posterior pharyngeal wall; PR, Passavant's Ridge.
*p<.05, **p<.01, ***p<.001

Table 4. Ratiometric Measurement of Lateral Videofluoroscopic Images: Planned Comparisons Across Pairs of Time Points for Each Velar Parameter.

Velar Parameter	Time Points	Mean (SD)	T	Sig.	Effect Size (*d*)
Extensibility	T1-T2	1.25(0.1188) – 1.30(0.1264)	-1.782	.094	0.407
	T1-T3	1.256(0.126) – 1.287(0.1608)	-0.699	.494	0.184
	T2-T3	1.30(0.1282) – 1.282(0.1705)	380.000	.709	0.117
Palatal Lift Angle	T1-T2	43.22(14.0890) – 33.92(7.8511)	3.269	.005**	0.761
	T1-T3	43.03(13.4083) – 33.37(9.0335)	2.752	.014*	0.816
	T2-T3	33.37(7.7624) – 32.53(8.6644)	0.405	.691	0.102
Closure Ratio	T1-T2	0.812(0.1958) – 0.68(0.1946)	2.754	.014*	0.676
	T1-T3	0.82(0.1945) – 0.68(0.2030)	3.265	.005**	0.704
	T2-T3	0.68(0.2007) – 0.65(0.1967)	0.762	.458	0.151
Velar Stretch	T1-T2	0.58(0.1382) – 0.66(0.1283)	-2.243	.039*	0.599
	T1-T3	0.58(0.1350) – 0.64(0.1154)	-2.082	.053	0.475
	T2-T3		0.484	.636	0.157
		0.66(0.1312) – 0.64(0.1228)			

*p < .05, **p < .001.