TITLE PAGE

(a) Manuscript title: Identifying Predictors of Acquired Velopharyngeal Insufficiency in Cleft Lip and Palate following Maxillary Osteotomy using Multiple Regression Analyses

(b) Authors' full names, highest academic degrees, and affiliations

Valerie J Pereira, PhD, Senior Lecturer, Department of Otorhinolaryngology, Head & Neck Surgery, Faculty of Medicine, The Chinese University of Hong Kong

Tel: +852 3943 3966, Fax: 852 3942 0969, Email: valeriepereira@ent.cuhk.edu.hk

Jyrki Tuomainen, PhD, Senior Lecturer, Speech, Hearing and Phonetic Sciences, Division of Psychology & Language Sciences, University College London, London UK.

Tel: +4476794214, Fax: +442076794238, Email: j.tuomainen@ucl.ac.uk

Norman Hay, BDS MSc F Orth RCS, Clinical Director, North Thames Cleft Lip and Palate Team, Great Ormond Street Hospital for Children Foundation Trust, London UK

Tel: +447795567446, Fax: +442078298804, email: norman.hay@gosh.nhs.uk

Michael Mars, DSc PhD FDS RCS Eng, Formerly Lead Consultant Orthodontist, North Thames Cleft Lip and Palate Team, Great Ormond Street Hospital for Children Foundation Trust, London UK

Tel: N. A., Fax: N.A., Email: michael_mars@msn.com

Archna Suchak, MSc MOrth RCS FOrth RCS, Formerly Senior Registrar in Orthodontics, Royal London Hospital and Great Ormond Street Hospital for Children Foundation Trust, London UK

Tel: +44209524596, Fax: N.A., Email: archna20@hotmail.com

Debbie A Sell, LCSLT PhD FRCSLT, Senior Research Fellow, Centre for Outcomes and Experience Research in Children's Health, Illness and Disability (ORCHID), London UK

Tel: +442074059280, Fax: N.A., Email: debbie@debbiesell.com

(c) Name and address for correspondence

Valerie J Pereira, Department of Otorhinolaryngology, Head & Neck Surgery, Division of Speech Therapy, Room 303, Academic Building No.2, The Chinese University of Hong Kong, Shatin, New Territories, Hong Kong SAR, China

Valerie Jane Pereira, PhD https://orcid.org/0000-0001-7527-053X

Email: valeriepereira@ent.cuhk.edu.hk

Tel: (852) 3943 3966, Fax: (852) 3942 0969

(d) Address for reprints if different from that of corresponding author

Same as corresponding author

(e) Sources of support that require acknowledgement

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 Lee Yuet Sheung Kathy and Professor Tong Chi Fai Michael for supporting her continued work and interest in Cleft.

Conflicts of Interest and Source of Funding

The author(s) declare no potential conflicts of interest with respect to 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 Dr. D. Sell.

INTRODUCTION

It is well documented that maxillary osteotomy can impact negatively on velopharyngeal function in individuals with Cleft Lip and Palate (CLP) resulting in perceptual speech sequelae such as hypernasality, facial grimace, nasal airflow errors and/or loss of pressure consonants. Pereira et al[1] reported significant deterioration in hypernasality and nasal turbulence following maxillary osteotomy and more recently, Alaluusua et al[2] found that the proportion of patients with moderate or severe velopharyngeal insufficiency (VPI) increased from 3% to 14% post-osteotomy. Several factors have been proposed as possible predictors of acquired VPI following maxillary osteotomy: amount of forward advancement[3,4] need-adequate ratio based on lateral cephalographs[5]; and pre-surgery borderline velopharyngeal status including the "adaptation" or the ability of the soft palate to compensate for post-surgery changes in the nasopharynx[6].

In terms of amount of forward advancement, there is no consensus with some studies reporting an association with a cut-off of 10mm[7] and others reporting no one-to-one relationship[3,4,8,9]. The concept of need-adequate ratio was first discussed by Simpson and Austin[10] who defined need ratio as the relationship between pharyngeal depth and velar length. Schendel and colleagues[5] subsequently proposed that a pre-surgery need ratio of greater than one would be a risk factor for velopharyngeal deterioration post-surgery, however, other studies have reported the reverse[7]. With regards to pre-surgery borderline velopharyngeal function, the first advocate of this was Witzel[6] who defined "borderline" velopharyngeal or marginal velopharyngeal closure as "small pinhole gaps in the velopharyngeal valve through which bubbles or mucus [is] observed during videofluoroscopy, nasendoscopy or both" (p.200). Alaluusua et al [2] also confirmed velopharyngeal deterioration following maxillary osteotomy in patients with pre-surgery VPI including those with a "borderline" rating.

Statistical methods in the form of logistic regression was used by Phillips et al [3], McComb et al [11] and more recently, de Medeiros-Santana et al[12]. Phillips et al[3] reported that perceptual speech assessment was a good predictor of acquired VPI but not nasendoscopy or amount of forward advancement. McComb et al[11] reported that a short soft palate length was a good predictor too. However, the measurement of soft palate length was made on still lateral cephalographs with the velum at rest, thereby, giving no consideration to movement and function of the soft palate during speech. Both studies by Phillips et al[3] and McComb et al[11] were retrospective in nature, acknowledged as weaker level of evidence compared with prospective studies[13]. More recently, de Medeiros-Santana and colleagues[12] undertook a prospective study using cone-beam computed tomography imaging and identified levator veli palatini (LVP) mobility as a risk factor of worsening velopharyngeal function as evidenced by increased hypernasality. The mobility of the LVP was rated on a 3-point ordinal scale although no operational definitions were provided for each scalar point. Although ratings were undertaken by three experienced raters, no reliability data were reported for any of the parameters.

The aims of the current prospective study were to use multiple regression analyses based on a comprehensive set of perceptual and instrumental speech and velar outcomes to identify potentially valid predictors of acquired VPI following maxillary advancement by osteotomy methods in patients with CLP.

METHODS

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

20 participants with non-syndromic CLP were recruited from a single regional cleft service. Written consent was obtained from all participants. The cohort was a consecutive series of patients undergoing maxillary osteotomy with or without a mandibular setback by a single surgeon. All participants were native speakers of English. There were 16 males and four females. The mean age was 20;2 years (range = 18;1 – 30 years, SD = 2;6 years). None of the participants had any hearing and/or learning difficulties making them unable to participate in any of the tasks. Only one participant had had secondary velopharyngeal surgery prior to orthognathic surgery.

Measurement time points

Participants were seen at three time points for a battery of speech investigations: pre-surgery (T1), 3-months (T2) and 12-months post-surgery (T3). The first post-surgery time point (T2) was to capture early or immediate speech changes and a late time point (T3) was to capture permanent changes in speech. The maxilla is considered to be relatively stable by 12-months post-surgery (Eurocran, 2003). Two participants missed the 3-month post-surgery appointments.

Speech methodology

Assessment of speech was undertaken using a battery of perceptual and instrumental investigations (Table 1).

Speech data was collected using a standardized speech protocol based on the Cleft Audit Protocol for Speech-Augmented (CAPS-A), a validated and reliable speech outcome tool[14]. All speech data was recorded using a Panasonic digital video camera model NV-GS70 and a Rode NT-3 hypercardiod condenser microphone. In the CAPS-A, articulation errors are described as cleft speech characteristics (CSCs) e.g. glottal articulation or palatal/palatalization. Non-articulation parameters include hypernasal resonance, nasal emission, nasal turbulence, voice and intelligibility[14]. International standard recording recommendations were adhered to[15,16].

Two approaches to the reliability of these perceptual ratings were undertaken: CAPS-A ordinal ratings and for resonance, Visual Analog Scale (VAS) ratings. In both approaches, only the audio component was presented as including the visual component would bias raters as to whether participants were at the pre- or post-surgery stage. Composite scores from specified CAPS-A parameters were derived to reflect overall velopharyngeal function. This is described below.

CAPS-A ratings study. Articulation data was coded on an ordinal scale according to the number of consonants affected by the cleft speech characteristic. A score of 0 was assigned if no consonants were affected, 1 if two or less consonants were affected and 2 if three or more consonants were produced incorrectly. Non-articulation parameters were rated according to the CAPS-A guidelines (Table 2). For the purposes of this study, only passive CSCs (e.g. a pressure consonant such as /p/ is replaced with a nasal consonant [m]) hypernasality, nasal emission and nasal airflow, parameters associated with velopharyngeal function, were included in the analyses. Two experienced CAPS-A trained specialist speech and language therapists (SLTs) participated in this ratings study: rater 1 rated 104 samples and rater 2 served as the rater for the inter-rater reliability study and rated 48% of the total number of speech samples rated by rater 1. Both raters were blinded to the nature of the study, the participants and time points and all ratings were undertaken independently using studio professional headphones (Beyerdynamic DT100). Inter rater reliability was calculated using Spearman's correlation coefficient (r_s) and percent agreement. Two percent agreement statistics were used: Po (perfect agreement) and Po – 1 (whether raters agreed to a precision of -1 to +1 scores).

Visual analog scale ratings study. Two blocks of CAPS-A sentences were created. Block 1 (- Nasals) consisted of oral consonant sentences only and Block 2 (+ Nasals) consisted of sentences with oral and nasal consonants, considered to be more representative of normal conversational speech (Table 3). Block 1 consisted of sentences with plosive targets as fricatives are known to be at risk of being affected as a result of maxillary advancement surgery[17,18] and could be distracting for raters when rating resonance. The audio component of the six sentences of each participant at each time point was extracted using RER Audio Converter 3.7.5.0412 (A-S) and converted to wav format (uncompressed CD Audio Quality). The intensity (root mean square) of each sentence was equalized using PRAAT version 4.6.36[19].

Four specialist SLTs from two regional cleft services were convened as listeners. Each listener was assigned their own computer and high-quality stereo headphones. Special software presented the stimuli in random order and captured the ratings of individual raters. The order of presentation of the blocks was counterbalanced across raters. Raters used a sliding scale where the extreme left-hand side of the scale represented normal or oral resonance with a score of "1", and the extreme right-hand side of the scale represented severe hypernasality with a score of "100". The raters could listen to each sentence as many times as they needed by clicking on the "repeat" button on the screen. Each listener rated 127 sentence sets including repeats (15% for Block 1 and 10% for Block 2). Inter-rater reliability was calculated for each of the four listeners for Blocks 1 and 2 and was also averaged across listeners and blocks.

Velopharyngeal composite score-summary (VPC-SUM CAPS-A). A composite score reflecting overall velopharyngeal function was derived from four identified speech parameters rated on the CAPS-A: hypernasality, nasal airflow errors, non-oral and passive cleft speech characteristics. The method of obtaining an overall velopharyngeal composite score was based on that described by Lohmander et al[20] (Table 4). Individual ratings were recoded and summed to obtain a velopharyngeal composite score-summary (VPC-SUM CAPS-A) with "0" indicating adequate velopharyngeal function and a maximum of "4" indicating inadequate velopharyngeal function[21].

Nasalance. A standardized set of 16 sentences which included subsets of sentences with high pressure consonants, low pressure consonants or mixed consonants were collected[22,23]. Calibration was undertaken according to the manufacturer's guidelines. Test-retest reliability of the nasometer with head gear change was measured. The mean nasalance score in percentage form was obtained for each test situation and time point using the 'analyze' function. Boundaries to demarcate the start and end of the speech sample set were undertaken manually.

Lateral videofluoroscopy and nasendoscopy. Videofluoroscopy was undertaken according to routine clinical protocol. An external hypercardiod condenser microphone was used for all recordings and images were recorded onto Super (S) VHS tapes or onto DVD following an upgrade of the fluoroscopic system. For quantitative ratiometric measurements (QRMs), spatial calibration was undertaken. A calibration ring (circular test object) which was attached to a head alignment device[24] and located in the same midsagittal plane as the participant's head was screened at the same magnification. Nasendoscopic evaluation was undertaken using a flexible nasendoscope and attached to a DVD Recorder and electret condenser tie-clip microphone. Speech sampling was based on a standard clinical protocol[25] and recommendations by the International Working Group[26].

Datasets: Two sets of data were created for analyses: dataset 1 used the rest vowel /i/ positions and dataset 2, the full speech sample. Dataset 1 was required for quantitative ratiometric measurements (QRMs) [27,28] and dataset 2, for clinical visual perceptual judgments (VPRs). Dataset 2 contained a range of phonemes across the linguistic hierarchy which would be more valid as it is a better representative of normal conversational speech. Speech samples were randomized to prevent recognition of participant and time point.

Visual perceptual ratings. The rating proformas for both videofluoroscopy and nasendoscopy are shown in Table 5. Two raters were convened for each instrumental measure and ratings were undertaken independently on each dataset. Raters initially made independent judgements for each parameter followed by a consensus judgment. Where disagreement occurred, the raters watched the video samples again until a consensus was reached. An additional ten samples for videofluoroscopy and 12 samples for nasendoscopy were randomly identified and repeated for the calculation of intra-rater reliability.

Quantitative Ratiometric Measurements. All videofluoroscopic and nasendoscopic images were converted to Audio Video Interleave (AVI) container format and velar measurements on /i/ with the soft palate at maximum closure and at rest[27-29] were undertaken using Image Pro 6.3[30]. Formalized measurement proformas were created to standardize the complex procedures for gaining the videofluoroscopy and nasendoscopy measurements.

For lateral videofluoroscopic images, four velar parameters were measured (Figure 1). Two raters were trained to undertake these measurements independently by the author of the methodology (M. Birch). Ten additional video samples were randomly selected and included for the calculation of inter- and intra-rater reliability. For nasendoscopic images, only one QRM, *closure ratio*, was made using the formula Area 1 (shape of velopharyngeal gap if any at maximum closure) / Area 2 (shape of velopharyngeal gap at rest) where a closure ratio of 1 indicates incomplete closure and no movement of the velopharyngeal sphincter, and 0 indicates complete closure[31].

A summary of the inter- and intra-rater reliability for the range of speech investigations is provided in Appendix A.

Lateral cephalometric radiographs

Standard lateral cephalometric radiographs were undertaken by the team's orthodontists. These radiographs were taken as clinically indicated at four time points: pre-treatment (prior to the start of orthodontic treatment for maxillary osteotomy), pre-surgery (prior to maxillary osteotomy), post-surgery (following maxillary osteotomy) and end/near end treatment (following post-osteotomy orthodontic treatment). Films were traced onto acetate paper by a single examiner under standard conditions. Cephalometric measurements were recorded to the nearest 0.5mm or 0.5 degrees using a hand-held ruler or protractor. The following points were traced: centre of sella turcica (S); nasion, the intersection of the internasal and interfrontal suture (N); deepest point on the anterior contour of the maxillary arch (A); and the deepest point on the anterior contour of the mandibular arch (B). An approximate Frankfort plane was constructed 7 degrees from the line SN and a perpendicular was dropped through S. The vertical position of point A and point B was measured along this line and the horizontal position of point A (Hor A) and point B (Hor B) was measured at right angles to this line for each cephalogram (Figure 2). Hor A measurements were used to calculate the *amount of forward advancement* (Hor A post-surgery – Hor A pre-surgery) and *amount of relapse* (Hor A at end of treatment – Hor A immediately post-surgery). As two pre-surgery measurements were taken, independent t-tests were used to determine if the two measurements differed significantly. A non-significant difference was expected as facial growth is complete in this cohort, and as such, the Hor A just prior to surgery was taken as the pre-surgery measure. Where this data was missing, clinical measurements taken pre-treatment were used. The *amount of advancement* was used in the regression analysis as a potential predictor of acquired velopharyngeal deterioration following advancement surgery.

Eight samples (9% of the total available) were randomly identified for use in the inter- and intra-rater reliability studies. Two orthodontists acted as raters. Rater 1 undertook repeat ratings on the 8 samples for the intra-rater reliability study and rater 2 undertook ratings on the same 8 samples for the inter-rater reliability study. Only the Hor A measurement was made. Intra- and inter-rater reliability were $r = 0.989$ and $r = 0.989$ respectively, both statistically significant at $p < .001$.

RESULTS

Approach to regression analyses

Multiple linear regression analyses were undertaken to identify valid predictors of acquired VPI or deterioration, as a result of maxillary osteotomy in CLP. Potential predictor variables are outlined in Table 6. The specified dependent variables were hypernasality ratings from CAPS-A, VAS Block 1 (- Nasals), Block 2 (+ Nasals) and Nasalance. This resulted in eight regression models: four dependent variables and two post-surgery time points. As no a priori hypotheses or assumptions were made as to the importance or relevance of any one predictor, the Enter model was used in the regression analyses. Bivariate correlations between the dependent variable *hypernasality* and each possible predictor variable were undertaken for T2 and T3 separately. Variables with significant correlations with hypernasality were used in the regression analyses as possible predictor variables. Collinearity diagnostics were checked at this point where the variance inflation factor (VIF) should be ≤ 10 and the tolerance value should be ≥ 0.1[2].

Significant models

Only three of the eight models were a good fit and were statistically significant. All were T3 (12 months post-surgery) models (Table 7). The predictors accounted for 92% of the variance both when CAPS-A hypernasality ratings was the dependent variable (R^2 = .920, F(11,7) = 7.303, p = .007) and when VAS Block 2 (+ Nasals) was the dependent variable (R^2 = .920, F(11,7) = 7.551, p = .007). The accounted for variance was 89.3% when VAS Block 1 (non-nasal consonants) was the dependent variable (R^2 = .893, F(11,7) = 5.317, p = .018). Two variables appeared to be significant, or near significant, predictors in all three models:

proportion of palate contacting the posterior pharyngeal wall (PWW) (from VPRs of videofluoroscopy) and closure ratio (from QRMs of videofluoroscopy). A third variable, firmness of closure (from VPRs of videofluoroscopy) was a significant predictor for CAPS-A hypernasality ratings and almost reached significance for VAS Block 1. A fourth variable, palate lift angle (from QRMs of videofluoroscopy) significantly predicted CAPS-A hypernasality ratings (Table 7). Collinearity diagnostics indicated no problem with multicollinearity for any of these four variables.

Model 1: CAPS-A hypernasality ratings at 12-months as the dependent variable. Four velar parameters, *palatal lift angle, closure ratio, firmness of closure,* and *proportion of palate contacting the posterior pharyngeal wall*, were identified as significant or near significant predictors and accounted for 92% of the variance in this model. *Closure ratio* appeared to be the most significant predictor with the highest standardized beta value (β= -1.118).

Model 2: Hypernasality based on VAS Block 1 (- Nasals) at 12-months as the dependent variable. Three velar parameters, *closure ratio, proportion of palate contacting the posterior pharyngeal wall* and *firmness of closure* were identified as significant or near significant predictors in this model and together, accounted for 89.3% of the variance in this model. The first two were statistically significant predictors whilst *firmness of closure* almost reached statistical significance at p = .043. Both *proportion of palate contacting the posterior pharyngeal wall* and *firmness of closure* appear to have a comparable degree of importance as predictors in this model with fairly similar standardized beta values at β= -0.768 and β= 0.788 respectively.

Model 3: Hypernasality based on VAS Block 2 (+ Nasals) at 12-months as the dependent variable. Two velar parameters were identified as significant/almost significant predictors: *closure ratio* and *proportion of palate contacting the posterior pharyngeal wall* and in combination, accounted for 92% of the variance in this model. Both predictors appear to have a comparable degree of importance as predictors in this model with almost similar standardized beta values at β= -0.565 and β= -0.507 respectively.

DISCUSSION

Approximately 10 to 50% of individuals with CLP require orthognathic surgery to correct maxillary hypoplasia and the accompanying class III occlusion[33]. The risk of acquiring VPI following maxillary osteotomy has been reported by several authors[1,2]. Predictors of acquired VPI following surgery are currently ill-defined, adversely impacting on the patient informed consent process to surgical intervention.

The aim of this prospective study was to use multiple regression analyses to identify potentially valid predictors of acquired VPI following maxillary advancement by osteotomy methods in patients with CLP. Participants were seen at three time points: pre-surgery, 3-months and 12 -months post-surgery for a comprehensive battery of perceptual and instrumental speech investigations. Wherever possible, international guidelines to speech methodology were adhered to and inter- and intra- reliability studies undertaken.

Multiple regression models showed that only the T3 (12 months after surgery) regression models were statistically significant at 12-months post-surgery, when the dependent variable, or outcome, was the perceptual rating of hypernasality. Nasalance as a dependent variable, or outcome, did not result in any significant models. A likely explanation for this is that the presence of nasal airflow errors are known to be a potential source of error, thereby invalidating the procedure and results[34]. In this study, there was a significant increase in nasal turbulence post-surgery, which may explain the nasalance results.

Two velar parameters were found to be significant, or almost significant predictors in all three models (CAPS-A ratings; VAS Block 1(- Nasals); VAS Block 2 (+ Nasals). Both parameters are measurements made on lateral videofluoroscopic images either as a VPR (proportion of palate contacting the posterior pharyngeal wall) or a QRM (closure ratio).

The first predictor, *proportion of palate contacting the posterior pharyngeal wall*, was defined by Kummer[35] 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" and a small proportion of the soft palate contacting the posterior pharyngeal wall signifies that velopharyngeal closure is "tenuous" (p.456). This is often referred to as "touch" closure. Six of seven cases with normal resonance before surgery, but who acquired hypernasality or high VAS ratings 12-months post-surgery, were rated as having no or only a small proportion of their palate contacting the posterior pharyngeal wall pre-surgery. Kummer[35] also described *firmness of closure*, which in this study, was found to be a significant predictor in two of the three models. Although bivariate correlations showed a strong association between these two velar parameters ($r_s = .852$), multicollinearity was not an issue (variation inflation factor: *firmness of closure* = 9.143, *proportion of palate contacting the posterior pharyngeal wall* = 4.902; tolerance: 0.109 and 0.204, respectively) indicating that the two parameters are distinct entities.

The second velar parameter which was a significant predictor in all three models was *closure ratio*. Five of the seven cases who acquired hypernasality or increased VAS ratings post-surgery, had incomplete closure ratios ranging from 0.6 to 0.8 pre-surgery. Witzel[6] identified pre-surgery borderline velopharyngeal closure as a risk factor and defined "borderline" or marginal velopharyngeal closure as "small pinhole gaps in the velopharyngeal valve through which bubbles or barium or mucus [is] observed during videofluoroscopy, nasendoscopy or both" (p.200). Whilst some later studies showed support for this view[2,36], others have not[37]. The controversy arises probably due to the definition of "borderline" closure which differs across authors and studies, the instrumentation used to evaluate velopharyngeal function, and the psychometrics of the measurement methods used. This problem is exacerbated also by the nature of ordinal rating scales used in the VPRs of velopharyngeal function or gap size based on videofluoroscopic and/or nasendoscopic images. For example, ordinal scales may not have sufficient scalar points to capture small or subtle changes to velopharyngeal (dys)function, relevant to borderline velopharyngeal closure. Additionally, operational definitions for each scalar point may not be well-defined and certainly not comparable across scales and studies. These reasons may also explain why VPRs of velopharyngeal function and gap size were not identified as significant predictors of any of the dependent variables in the current study.

In terms of instrumentation, Gilleard et al[31] reported that absolute measurements of gap size from uncalibrated nasendoscopy edits are unreliable as "even relatively small differences in object-lens distance and variation in object position within the field of view cause significant changes in magnification and distortion" (p. 179). As such, the validity of judgments of "adequacy of velopharyngeal closure" which is correlated with judgments on "gap size" on both videofluoroscopic and nasendoscopic images, is questionable too. In contrast, QRM of *closure ratio* uses a standardized methodology and specialized software[27,28]. Measurements based on this methodology have been shown to be reliable and accurate with an acceptably small source of uncertainty. Furthermore, in the current study, the use of the fixed head alignment device in highly co-operative participants, also served to reduce errors due to alignment[28].

The parameter of *velar stretch* measured quantitatively on videofluoroscopic images requires some discussion. Although this increased post-surgery suggesting an increase in effective velum length, it was not identified as a valid predictor in any of the regression models. A plausible reason could be how the parameter was measured in this study: effective velar length during speech divided by the entire length of the soft palate (see Figure 1). This formula subsumes both the anterior and posterior portions of the palate. Tian and Redett[38], however, criticized using the posterior length of the palate (beyond the knee) arguing that it does not contribute to velopharyngeal closure, and proposed *effective velopharyngeal ratio* i.e. anterior portion of soft palate at rest divided by effective velar length during speech, which they found was a good predictor of velopharyngeal function.

None of the velar parameters based on nasendoscopy, whether as VPRs or QRMs, were identified as plausible or valid predictors in the current study. Similar findings were reported by Phillips et al[3] in their logistic regression study. One possible explanation relates to the lack of reliability and validity of VPRs based on nasendoscopic images[39]. With nasendoscopic images, the whole of the velopharyngeal portal needs to be viewed and recorded and this is often not possible. Additionally, as already discussed, ordinal type rating scales may not be sufficiently sensitive to small or subtle changes in velopharyngeal status or function. Inherent in the measurement of velar parameters are uncertainties related to magnification and barrel distortion[40]. For instance, absolute measurements of gap size on uncalibrated nasendosopic images are unreliable as “even relatively small differences in object-lens distance and variation in object position within the field of view cause significant changes in magnification and distortion” (p.179). This means that the validity of judgments of “adequacy of velopharyngeal closure” which is correlated with judgments on “gap size”, is questionable too.

In this study, reliability coefficients were large and statistically significant at $p < .001$ for VPRs with percent agreement ranging from 75% to 88% for perfect agreement and from 82% to 97% for precision of +1 to -1 agreement. For QRM data, reliability coefficients were also large (inter-rater: $r = .94$ and intra-rater: $r = .88$). Although no statistically significant results were obtained, the velar parameter *proportion of palate contacting the posterior pharyngeal wall* almost reached statistical significance with a medium effect size between the pre- and post-surgery, suggesting a possible true effect of maxillary osteotomy on the amount of palate contact with the posterior pharyngeal wall. This parameter is therefore worth considering, given that it was identified as a valid predictor of acquired VPI following maxillary osteotomy when based on videofluoroscopic images.

None of the predictors in isolation accounted for all the variance in the dependent variable or outcome. The four predictor variables acted in differing combinations, depending on which outcome (CAPS-A or VAS) was used to explain the proportion of variation seen in hypernasality at 12-months post-surgery. However, the identification of *closure ratio* and *proportion of palate contacting the posterior pharyngeal wall* as statistically significant or almost statistically significant predictors in all three models, suggest the combined usefulness of VPRs and QRMs of lateral videofluoroscopic images.

Speech and study methodology

In this prospective speech osteotomy study, a range of perceptual and instrumental measures were included with two post-surgery follow-up time points. Reliability studies were undertaken for all speech measures. For videofluoroscopy and nasendoscopy, both VPRs and QRMs using a published and validated methodology were utilized. Four dependent variables or outcomes were used: perceptual ratings of hypernasality on an ordinal scale, which is the most commonly reported method in the literature, visual analog scaling based on two types of speech samples and an acoustic and instrumental measure, nasalance. Multiple regression analyses were undertaken to identify valid predictors with standardized beta coefficients calculated to allow for direct comparisons across predictors. Collinearity diagnostics were checked to ensure that predictors did not correlate too highly with one another and that any one predictor should not have strong linear relationships with any other predictor.

With QRMs of velar parameters based on videofluorosopic images, inter-rater reliability was variable, ranging from r = .423 to r = .827, reflecting a medium sized correlation at a minimum. However, given that Birch et al. [28] reported high agreement between raters for the range of velar parameters (except for velar stretch which was not a validated parameter), the method requires perhaps more intensive training of raters, particularly if raters are not medically trained. An additional consideration is with regards to the parameter of *velar stretch*. A proposed alternative measurement is to omit use of the posterior portion of the soft palate in any measurements and a plausible alternative is L1 (P1-P3) at rest (anterior portion of soft palate at rest) / L2 (P1-P2) during speech (effective velar length during speech) (Figure 1). Obviously, this measure will require validation.

Clinical implications

We found that maxillary osteotomy can result in VPI in patients with CLP and that lateral videofluoroscopy is useful in the identification of valid predictors of this acquired VPI following the surgery. This instrumental measure is a well-established and a routine clinical tool in many cleft services around the world and can easily be integrated into the speech maxillary osteotomy protocol of assessment and review. Evidence from the study indicates that a combination of VPRs and QRMs using a specified methodology is indicated to be able to identify predictors of acquired VPI following maxillary osteotomy. Two predictors, *proportion of palate contacting the posterior pharyngeal wall* and *closure ratio*, were identified to be significant or almost significant predictors for resonance outcomes rated either on an ordinal scale or VAS. For hypernasality rated using the CAPS-A[14], regression analyses identified two additional predictors, *firmness of closure* and *palatal lift angle*, that act in combination to predict the resonance outcome following maxillary osteotomy. These findings have direct implications on the osteotomy speech assessment protocol and the informed consent process to surgical intervention.

CONCLUSIONS

The study provides further evidence of the potential adverse impact of maxillary osteotomy on velopharyngeal function in individuals with CLP. A range of reliable perceptual and instrumental speech outcome measures were included in the multiple regression analyses to identify valid predictors of this acquired VPI. Lateral videofluoroscopy was found to have both clinical and research utility with the undertaking of both VPRs as well as QRMs. The study has identified specific velar parameters as valid predictors of acquired VPI following maxillary osteotomy for both ordinal and VAS measurement methods of resonance as the outcome measure. Future maxillary osteotomy studies should include adapted measurements of effective velar stretch or velopharyngeal ratio based on videofluoroscopic images.

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 XXXX and Professor XXXX for supporting her continued work and interest in the field of speech and cleft lip and palate.

Research at Great Ormond Street Hospital NHS Foundation Trust is supported by the NIHR. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health.

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 Dr. XXXXX.

References

1. Pereira V, Sell D, Tuomainen J. Effect of maxillary osteotomy on speech in cleft lip and palate: perceptual outcomes of velopharyngeal function. *Int J of Lang and Comm Dis* 2013; 48: 640-650
2. Aluluusua S, Turunen L, Saarikko A, et al. The effects of Le Fort I osteotomy on velopharyngeal function in cleft patients. *Cleft Palate-Craniofac J* 2019: 47, 239-244
3. Phillips JH, Klaiman P, Delorey R, et al. Predictors of velopharyngeal insufficiency in cleft palate orthognathic surgery. *Plast and Reconstr Surg* 2005; 115: 681-686
4. Chua HD, Whitehill TL, Samman N, et al. Maxillary distraction versus orthognathic surgery in cleft lip and palate patients: effects on speech and velopharyngeal function. *Inter J of Oral and Maxillofac Surg* 2010: 39, 633-640
5. Schendel SA, Oeschlaeger M, Wolford LM, et al. Velopharyngeal anatomy and maxillary advancement, *Journal of Maxillofacial* Surgery 1979: 7, 116-124
6. Witzel MA. Commentary. *The Cleft Palate Journal* 1989: 26: 199-200
7. Okazaki K, Satoh K, Kato M, et al. Speech and velopharyngeal function following maxillary advancement in patients with cleft lip and palate. *Annals of Plast Surg* 1993; 30: 304-311
8. Schwarz C, Gruner E. Logopaedic findings following advancement of the maxilla. *J of Maxillofac Surg* 1976; 4: 40-55
9. McCarthy JG, Cocaaro PJ, Schwartz MD. Velopharyngeal function following maxillary advancement. *Plast and Reconst Surg* 1979; 64: 180-189
10. Simpson RK, Austin AA. A cephalometric investigation of velar stretch. *Cleft Palate J* 1972; 9: 341-351

11. McComb RW, Marrinan EM, Nuss RC, et al. Predictors of velopharyngeal insufficiency after Le Fort I maxillary advancement in patients with cleft palate. *J of Oral and Maxillofac* Surg 2011; 69 : 2226-2232
12. de Medeiros-Santana MNL, Jamie LP, Renato Yassutaka Faria Y et al. Predictors of velopharyngeal dysfunction in individuals with Cleft Palate following surgical maxillary advancement: clinical tomographic assessments 2019: 56: 1314-1321.
13. Vandenbroucke JP. Observational research, randomised trials, and two views of medical science. *PLoS Medicine* 2008; 11: e67
14. John A, Sell D, Sweeney T, et al. The cleft audit protocol for speech-augmented: A validated and reliable measure for auditing cleft speech. *Cleft Palate-Craniofac J* 2006; 43: 272-288
15. Sell D, John A, Harding-Bell A, et al. Cleft audit protocol for speech (CAPS-A): a comprehensive training package for speech analysis. *Int J of Lang and Comm Dis* 2009: 44: 529-548
16. Cleft Palate International Speech Issues (CLISPI). [online]. Available: https://clispi.com. [accessed 9 August 2019]
17. Vallino LD. Speech, velopharyngeal function, and hearing before and after orthognathic surgery. *J of Oral and Maxillofac Surg* 1990: 48: 1274-1281
18. Lee AS, Whitehill TL, Ciocca V, et al. Acoustic and perceptual analysis of the sibilant sound /s/ before and after orthognathic surgery. *J of Oral and Maxillofac Surg* 2002; 60: 364-372
19. PRAAT Doing Phonetics by Computer [computer program]. Version 4.6.36. Available at: http://www.fon.hum.uva.nl/praat/

20. Lohmander A, Willadsen E, Persson C, et al. 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:** 347-362
21. XXXX. The effect of maxillary advancement on speech and velopharyngeal function in Cleft Lip and Palate [PhD thesis] London UK: UCL Institute of Child Health; XXXX
22. Sweeney T. The perceptual and instrumental assessment of nasality and nasal airflow errors associated with velopharyngeal dysfunction [Phd thesis]. Dublin, Ireland: School of Clinical Speech and Language Studies, University of Dublin, Ireland; 2000
23. Sweeney T, Sell D. Relationship between perceptual ratings of nasality and nasometry in children/adolescents with cleft palate and/or velopharyngeal dysfunction. *Int J of Lang and Comm Dis* 2008; 43: 265-282
24. Sommerlad B, Rowland H, Harland K. Lateral videofluoroscopy: a modification to aid in velopharyngeal assessment and measurement. *Cleft Palate J* 1994; 31: 134-135
25. Sell D, Ma L. Great Ormond Street Hospital multiview videofluroscopic assessment of velopharyngeal function proforma [unpublished clinical protocol] 1990
26. Golding-Kushner KJ, Argamaso RV, Cotton RT, et al. Standardization for the reporting of nasopharyngoscopy and multiview videofluoroscopy: a report from an International Working Group. *Cleft Palate J* 1990; 27: 337-347
27. Birch MJ, Sommerlad BC, Bhatt A. Image analysis of lateral velopharyngeal closure in repaired cleft palates and normal palates. *Br J of Plast Surg*, 1994; 47, 400-405
28. Birch MJ, Sommerlad BC, Fenn C, et al. A study of the measurement errors associated with the analysis of velar movements assessed from lateral videofluoroscopic investigations. *Cleft Palate-Craniofac J* 1999: 36, 499-507

29. Gilleard O. The use of nasendoscopy in Submucous Cleft Palate [masters thesis]. London UK: University of London; 2008

30. Image Pro [computer program]. Version 6.3. Media Cybernetics Inc.; 2007

31. Gilleard O, Sommerlad B, Sell D, et al. Nasendoscopic image analysis: an analysis of measurement uncertainties. In: Raposo Do Amaral E, Albino De Amaral AB eds. *11th International Congress on Cleft Lip and Palate and related Craniofacial Anomalies InternationalPproceedings*. Bologna, Italy: Medimond s.r.l. Monduzzi Editore, 2009: 179-184

32. Field A. Discovering Statistics Using SPSS (3rd ed.). Dubai: Sage Publications Ltd 2009

33. 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* 2007; 44: 396-401

34. Karnell MP. Nasometric discrimination of hypernasality and turbulent nasal airflow. *The Cleft Palate-Craniofac J* 1995; 32: 145-148

35. Kummer AW. *Cleft Palate and Craniofacial Anomalies* (2nd ed.). NY: Delmar Cengage Learning 2008

36. Trindade IE, Yamashita RP, Suguimoto RM et al. 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: 54-64

37. Chanchareonsook N, Whitehill TL, Samman N. Speech outcome and velopharyngeal function in cleft palate: comparison of Le Fort I maxillary osteotomy and distraction osteogenesis--early results. *Cleft Palate-Craniofac J* 2007: 44, 23-32

38. Tian W, Redett RJ. New velopharyngeal measurements at rest and during speech: implications and applications. *J of Craniofac Surg* 2007; 20: 532-539

39. Sell D, Pereira V. Instrumentation in the analysis of the structure and function of the velopharyngeal mechanism. In:Lohmander-Agerskov A, Howard S. eds. *Cleft Palate Speech: Assessment and Intervention*. West Sussex, UK: John Wiley & Sons Ltd, 2011: 145-162

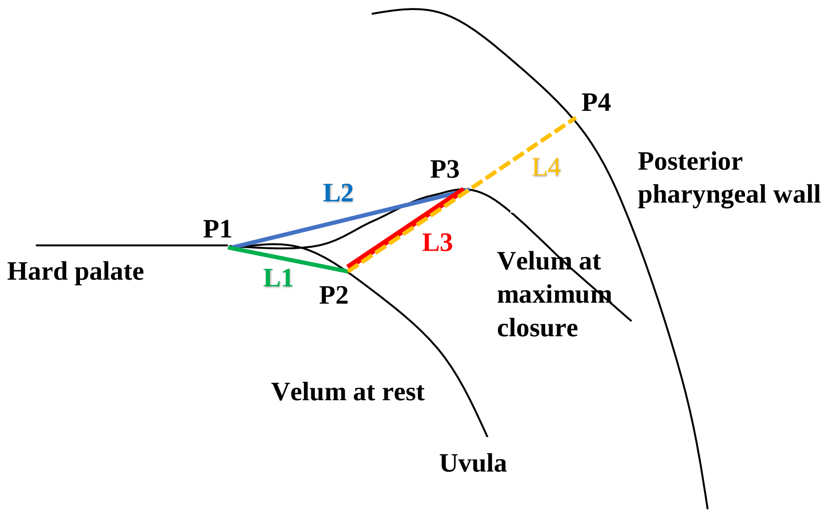
40. Gilleard O, Sommerlad BC, Sell D, et al. Nasendoscopic image analysis: an analysis of measurement uncertainties. *Cleft Palate-Craniofac J* 2013; 50, 351-317

41. Sell D, Sweeney T, Harding-Bell A et al. An update on cleft audit protocol for speech (CAPS-A). Paper presented at the Craniofacial Society of Great Britain and Ireland Annual Conference, Pre-Conference Special Interest Group Meeting; April, 2011; Ireland.

Figure Legends

Figure 1. Markings of anatomical landmarks and lines drawn for quantitative ratiometric velar measurements. Note: P1= point of transition between hard and soft palate; P2= point of velar knee at rest; P3= point of maximum (incomplete) velopharyngeal closure; P4= point of maximum (complete) closure with contact against the posterior pharyngeal wall.

Figure 2. Cephalometric points traced. S = sella turcica; N = nasion; A = deepest point on the anterior contour of the maxillary arch; B = deepest point on the anterior contour of the mandibular arch. Drawing by Suchak, A. (2011).



Extensibility = L2/L1

Palatal Lift Angle = Angle formed by P2-P1-P3

Closure Ratio = L3/L3 for complete closure and L3/L4 for incomplete closure

Velar Stretch = L2/P1 to tip of uvula

Figure 1. Markings of anatomical landmarks and lines drawn for quantitative ratiometric velar measurements. Note: P1= point of transition between hard and soft palate; P2= point of velar knee at rest; P3= point of maximum (incomplete) velopharyngeal closure; P4= point of maximum (complete) closure with contact against the posterior pharyngeal wall.

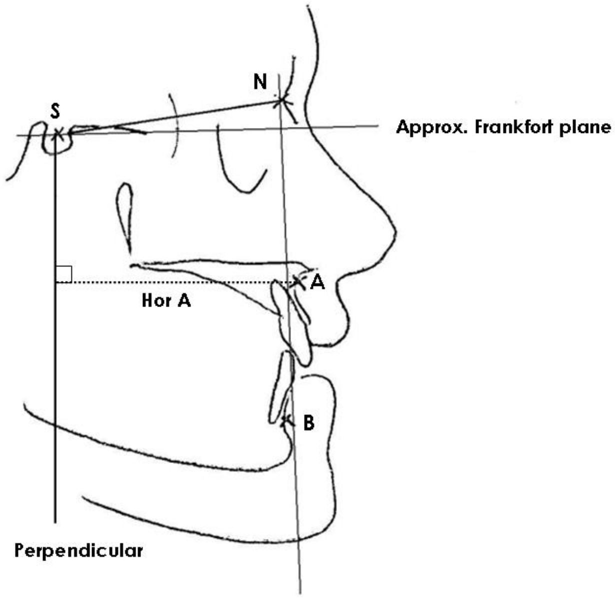


Figure 2. Cephalometric points traced. S = sella turcica; N = nasion; A = deepest point on the anterior contour of the maxillary arch; B = deepest point on the anterior contour of the mandibular arch. Drawing by Suchak, A. (2011)

Table 1. List of Perceptual and Instrumental Speech Assessments

Speech Parameter	Perceptual	Instrumental
Articulation	CAPS-A	-----
Resonance	CAPS-A	-----
Nasalance	-----	Nasometer[a]
Nasal Airflow	CAPS-A	-----
Velopharyngeal Function	VPC-SUM CAPS-A[b]	Lateral Videofluoroscopy
		Nasendoscopy

Notes: [a]Nasometer II 6400 (Kay Elemetrics, Pentax UK).

CAPS-A, Cleft Audit Protocol for Speech-Augmented[14]; VPC-SUM CAPS-A, Velopharyngeal Composite - Summary Score based on CAPS-A Ratings.

Table 2. Coding of Non-Articulation Speech Data from the CAPS-A Ratings Study[a]

Hypernasality
0 = absent, 1 = borderline-minimal, 2 = mild, 3 = moderate, 4 = severe
Nasal Emission and Turbulence
0 = absent on pressure consonants, 1 = heard on <10% of target phonemes, 2 = heard on >10% of target phonemes

Notes: [a]Only parameters from the CAPS-A relevant to the purpose of the study are indicated here. Scalar points and definitions for hypernasality are from the CAPS-A[14] and those for nasal emission and turbulence are from Sell et al[41].

Table 3. Sentences in the VAS Ratings Study

Block 1 (- Nasals)	Block 2 (+ Nasals)
Bob is a baby boy.	Daddy mended the door.
Gary's got a bag of Lego.	Karen is making a cake.
We were away all year.	Mary came home early.

Table 4

Table 4. Scoring of the Velopharyngeal Composite Score (VPC-SUM CAPSA)[a]

Speech Parameter	Description of Each Scalar Point	CAPS-A Ratings[b]	Scoring
Hypernasality	Absent	Rating of 0-1	0
	Borderline-minimal	Rating of 2-4	1
	Mild		
	Moderate		
	Severe		
Nasal Emission (NE) Nasal Turbulence (NT) (rated separately)	0: Absent	Both ratings of NE and NT = 0	0
	1: <10% of target phonemes	Rating of either NE or NT = 1	
		Both ratings of NE and NT = 1	1
	2: <10% of target phonemes	Ratings of NE/NT = 2	
Non-oral CSCs (e.g. pharyngeal or glottal articulation) and Passive CSCs (e.g. weak or nasalized consonants)		If absent or if only one or two consonants are affected	0
		If $\geq$ 3 consonants affected	1

Notes: [a]Adapted from Lohmander et al[20] and Pereira[21].

[b]CAPS-A parameters and scalar point descriptions from John et al[14] and Sell et al[41].

CSCs, Cleft speech characteristics[15]..

Table 5. Visual Perceptual Ratings of Lateral Videfluoroscopic and Nasendoscopic Images

Velar Parameter	Description of Each Scalar Point
Lateral Videofluoroscopy	
Velopharyngeal Closure	Definitely adequate/Probably adequate/Borderline/Probably inadequate/Definitely inadequate
Firmness of Closure	None/Touch type/Firm/Very firm
Proportion of Palate Contacting the PPW	None/Small/Moderate/Large
Nasendoscopy	
Velopharyngeal Closure	Definitely adequate/Probably adequate/Borderline/Probably inadequate/Definitely inadequate
Firmness of Closure	None/Touch type/Firm/Very firm

Table 6. Predictor variables with those used in the Refined Regression Models

Assessment/Tool	Predictor Variables
CAPS-A	Passive CSCs
	Nasal Emission
	Nasal Turbulence
	VPC-SUM
Lateral Videofluorosopy	Adequacy of VP Closure*
(Visual Perceptual Ratings)	Size of VP Gap
	Firmness of Closure*
	Proportion of Palate contacting the PPW*
Lateral Videofluorosopy	Extensibility
(Ratiometric Measurement)	Palatal Lift Angle*
	Closure Ratio*
	Best Closure Ratio*
	Velar Stretch
Nasendoscopy	Adequacy of VP Closure*
(Visual Perceptual Ratings)	Size of VP Gap*
	Firmness of Closure
Nasendoscopy	Closure Ratio*
(Ratiometric Measurement)	
Lateral Cephalometric Radiographs	Amount of Advancement

Notes: * Indicates significant correlation with at least one dependent variable.

VP, velopharyngeal; PPW, posterior pharyngeal wall.

Table 7

Table 7. Model Summary and Coefficients of Significant Predictors.

Dependent Variable	Time Point	Model Sig.	Model R^2	Significant/Almost Significant Predictors in Each Model	Std. Coeff. (β)	t	Sig.
CAPS-A	T2	.426	.734	None	---	---	---
Hypernasality	T3	.007**	.92	Firmness of Closure (LVF perceptual)	0.855	2.644	.033*
				Proportion of Palate (LVF perceptual)	-0.644	-2.718	.030*
				Palatal Lift Angle (LVF ratiometric)	-0.623	-3.655	.008**
				Closure Ratio (LVF ratiometric)	-1.118	-4.841	.002**
Nasalance	T2	.468	.717	None	---	---	---
	T3	.917	.384	None	---	---	---
VAS Block 1	T2	.569	.858	None	---	---	---
(Non-Nasals)	T3	.018*	.893	Proportion of Palate (LVF perceptual)	-0.768	-2.806	.026*
				Firmness of Closure (LVF perceptual)	0.788	2.109	.073
				Closure Ratio (LVF ratiometric)	-0.659	-2.471	.043*
VAS Block 2	T2	.251	.806	None	---	---	---
(Mixed Consonants)	T3	.007**	.922	Proportion of Palate (LVF perceptual)	-0.507	-2.171	.066
				Closure Ratio (LVF ratiometric)	-0.565	-2.487	.042*

Notes: VAS, visual analog scale; LVF, lateral videofluoroscopy.

*p < .05, **p < .01, ***p < .001.

Click here to access/download
Supplemental Data File (.doc, .tif, pdf, etc.)
Appendix A.docx