Sainsbury, D. C. G., Williams, C. C., Butterworth, S., de Blacam, C., Fell, M., Mullen, J., Breakey, W., Murphy, C., Hodgkinson, P., & Wren, Y. E. (2023). Patient Factors Influencing Speech Outcomes in Velopharyngeal Function Following Initial Cleft Palate Repair: A Systematic Review and Meta-Analysis. *Cleft Palate-Craniofacial Journal*. https://doi.org/10.1177/10556656231191384

Patient Factors Influencing Speech Outcomes in Velopharyngeal Function Following Initial Cleft Palate Repair: a Systematic Review and Meta-Analysis

ABSTRACT

Objective: Identification of patient factors influencing velopharyngeal function for speech following initial cleft palate repair.

Design: A literature search of relevant databases from inception until 2018 was performed using medical subject headings and keywords related to cleft palate, palatoplasty and speech assessment. Following three stage screening data extraction was performed.

Setting: Systematic review and meta-analysis of relevant literature.

Patients / Participants: Three hundred and eighty-three studies met the inclusion criteria, comprising data on 47658 participants.

Interventions: Individuals undergoing initial palatoplasty.

Main outcome measures: Studies including participants undergoing initial cleft palate repair where the frequency of secondary speech surgery and/or velopharyngeal function for speech was recorded.

Results: Patient factors reported included cleft phenotype (95% studies), biological sex (64%), syndrome diagnosis (44%), hearing loss (28%), developmental delay (16%), Robin Sequence (16%) and 22q11.2 microdeletion syndrome (11%). Meta-analysis provided strong evidence that rates of secondary surgery and velopharyngeal dysfunction varied according to

cleft phenotype (Veau I best outcomes, Veau IV worst outcomes), Robin Sequence and syndrome diagnosis. There was no evidence that biological sex was associated with worse outcomes. Many studies were poor quality with minimal follow-up.

Conclusions: Meta-analysis demonstrated the association of certain patient factors with speech outcome, however the quality of the evidence was low. Uniform, prospective, multi-centre documentation of preoperative characteristics and speech outcomes is required to characterise risk factors for post-palatoplasty velopharyngeal insufficiency for speech.

Word Count: 226

Systematic review registration: Registered with PROSPERO CRD42017051624

Introduction

Published rates of velopharyngeal insufficiency for speech (henceforth referred to as VPI) following initial cleft palate surgery range between 0-66%.[1-5] Possible reasons for this disparity include the broad phenotype of orofacial clefts, the range of classification systems used when assessing cleft severity and VPI, the inclusion of multiple surgeons, techniques, or institutions within the same studies, and the inherent limitations of retrospective data collection in many studies.[6] However, whilst certain factors are implicated, the reasons for such variation remains unclear. Sitzman *et al.* noted that whilst secondary surgery was associated with cleft phenotype (hazard ratios for unilateral cleft lip and palate (ULCP) and bilateral cleft lip and palate (BCLP) were 1.69 and 2.06, respectively, when compared to isolated cleft palate) and age at initial palatoplasty, the largest variation in secondary surgery was attributed to undetermined differences among surgeons and hospitals performing the initial palatoplasty.[7] Furthermore, the proportion of children undergoing secondary surgery varied five-fold across surgeons and hospitals.

It is likely that a diverse range of patient factors influence the success of a cleft palate repair with respect to speech.[8] There is evidence that cleft phenoype, cleft extent in an isolated cleft palate, cleft width and length, biological sex, socio-economic status, age at initial surgery, Robin Sequence (RS), previous general anaesthesia, presence of a syndrome and additional diagnosis may influence outcomes.[2-4, 7, 9-22]

It would be highly beneficial if it were possible to pre-operatively determine which children were likely to develop VPI following cleft palate repair. This information would facilitate

pre-operative counselling, informing parents/caregivers pre-operatively about the risk factors and likelihood of success or the potential need for secondary surgery or further intervention.

The aim of this systematic review was to collate the available evidence about the influence of patient factors on velopharyngeal function for speech following initial cleft palate repair. Non-patient factors and interventions, including surgical procedures, medications, non-surgical devices such as naso-alveolar moulding, arm splints, or speech and language therapy input were not the focus of this systematic review. The hypothesis for the review was: *"There are patient factors which influence velopharyngeal function for speech following cleft palate repair."*

METHODS

PRISMA 2020 guidelines for the conduct of a systematic review were followed.[23] The protocol registered with PROSPERO (CRD42017051624) and published with open access.[24] Deviations from the original protocol are recorded in the results section.

Ethical Approval / Patient Consent Statement

This was deemed not applicable as this was a systematic review using previously published material.

Eligibility Criteria

The objectives for this systematic review were summarised using a PECO framework (Supplementary Table 1).[25] Eligible studies were defined as full text primary data publications reporting: (a) the proportion of patients born with cleft palate, who were

recommended or underwent re-operation or secondary speech surgery for VPI and/or (b) perceptual speech outcomes following initial cleft palate reconstruction. Randomised controlled trials (RCTs), cluster RCTs, non-randomised controlled clinical trials, longitudinal studies, cross-sectional investigations, retrospective studies with prospective data collection and case series with 10 or more participants were included. Literature reviews, case reports, case series with less than 10 participants, review articles, commentaries, letters, editorials, dissertation abstracts or conference proceedings were excluded.

Included studies were grouped with respect to variables of interest such as cleft phenotype, biological sex, Robin sequence, and secondary surgery rates and speech outcomes calculated for the variable. Meta-analysis was performed on studies which reported outcomes for each variable sub-type. For example, if a study reported secondary surgery or speech outcomes for males and females then meta-analysis was performed on the biological sex variable. For variables with binary sub-domains (i.e. the presence or absence of a syndrome, presence or absence of Robin Sequence or presence or absence of 22q11.2 microdeletion syndrome) studies were included for meta-analysis if the study included outcomes for both the presence and absence of a particular variable.

Information Sources

The following databases were included in the search: MEDLINE, EMBASE, AMED (Allied and Complimentary Medicine Database), PsycINFO, CINAHL (Cumulative Index to Nursing & Allied Health Literature), Health Technology Assessment Database, SpeechBITE, CENTRAL (Cochrane Central Register of Controlled Trials), Cochrane Database of Systematic Reviews and Scopus from inception until 2018.

Search Strategy

Literature search strategies were developed using medical subject headings (MeSH) and text words related to cleft palate, cleft palate surgery and cleft speech assessment (Supplementary Table 2). The literature search was limited to articles written in English and human subjects. Publications of potential relevance to the review were identified by using both exploded MeSH headings and text words. Reference lists from included studies and additionally systematic reviews identified from the Health Technology Assessment Database and Cochrane Database of Systematic Reviews were searched for any relevant articles. The search strategy was peer-reviewed by two independent senior medical librarians and the final search was performed by one of these librarians (JM) in conjunction with the first author. EndNote X7 (Thompson Reuters, Philadelphia, PA) bibliographic software program was used to manage citations.

Selection Process

The search results were uploaded to Distiller SR systematic review software (Evidence Partners, Ottawa, ON) and de-duplicated. Thereafter, a three-stage screening process was performed in duplicate by two independent reviewers (DS: Cleft Surgeon; CW: Speech and Language Therapist). Study titles during the first stage, abstracts during the second stage and full texts in the third stage were screened against the inclusion criteria. Reasons for excluding papers were recorded. Discrepancies in study selection were dealt with by discussion between the two reviewers. If there was more than one paper reporting on a patient population then the paper with the most comprehensive data was included. The reviewers were not blinded to the journal titles, the study authors or institutions.

Data Collection Process

Following three stage screening, data extraction was performed by eight independent reviewers with cleft clinical experience (two cleft surgeons, a specialist cleft speech and language therapist, a cleft fellow, two plastic surgery trainees with an interest in cleft and two oral and maxillofacial trainees with an interest in cleft) using systematic review software (Distiller SR). Data extraction was performed in duplicate in 10% of studies to assess inter-rater reliability; overall agreement was rated as "*almost perfect*" (kappa = 0.95).

Data Items

The primary outcome was the proportion of patients who had, or who were recommended, further revisional surgery or speech surgery for VPI. The secondary outcome was the proportion of patients with normal velopharyngeal function for speech, as determined by perceptual assessment of resonance and nasal airflow during speech, following initial cleft palate repair. Any perceptual speech assessment was included within the review to capture all relevant papers. However, studies were excluded if the reported speech outcomes did not relate to velopharyngeal insufficiency. To facilitate comparison between studies, resonance outcomes for each study were recorded under five categories: normal, mild hypernasality, moderate hypernasality, severe hypernasality or any grade of hyponasality. If it was not possible to determine the degree of hypernasality this was recorded as "hypernasality of unknown severity". Nasal airflow errors (emission and turbulence) were recorded as normal, mild, moderate or severe. If the authors had not stratified this then this was recorded as 'presence of nasal emission', 'presence of nasal turbulence', 'no airflow errors' or 'not reported'. Supplementary Table 3 shows the data points extracted from each study.

Study Risk of Bias Assessment

Study quality was assessed using the Oxford Centre for Evidence based Medicine 2011 level of evidence guidelines (OCEBM Levels of Evidence Working Group, 2011). Risk of bias was appraised using modified criteria based on the Cochrane Collaboration's tool for assessing risk of bias in therapeutic studies.[26] The following domains were assessed: selection bias, detection bias (blinded assessment of outcomes), outcome assessment (recording of key outcomes—in this case, secondary surgery and normalised resonance), attrition bias (inclusion of all patients undergoing cleft palate surgery).

Effect Measures

Heterogeneity of the included studies was analysed by exploring the study characteristics and the I^2 statistic. The quantitative impact of exposures on outcomes was investigated using meta-analysis techniques where data from a minimum number of two studies were sufficiently homogeneous. Proportions of binary outcomes were principally extracted. The Mantel-Haenszel method was used to calculate pooled effect sizes. A fixed effects model was used where levels of statistical heterogeneity were low (I^2 <50%); otherwise a random effects model was used. Funnel plots were used to visually assess the likelihood of small study publication bias if more than 10 studies were included. Egger's test was calculated to quantify funnel plot asymmetry. Subgroup meta-analysis of cleft phenotype categories was performed using a fixed-effects (pleural) model using the random effects model for each subgroup. All meta-analysis was performed using the "meta" package via the R Project for Statistical Computing (www.R-project.org/).[27]

Synthesis Methods

The proportion of studies and participants with respect to decade of publication, exclusion rates and number of participants with available data were grouped (Table 1).

The number and proportion of studies reporting follow-up duration and / or age at final review were noted (Supplementary Table 4).

The number of studies reporting each variable (patient factor) and its influence on rate of secondary speech surgery (Supplementary Table 5) or perceptual speech outcomes (Table 2), was recorded. The number of participants in all studies assessing each variable were combined (Supplementary Table 6).

Reporting bias assessment

The proportion of studies reporting a blinded speech assessment, a speech outcome for velopharyngeal function (as previously defined) and/or a secondary surgery rate for all subjects was recorded. Reasons for exclusion of participants was explored.

RESULTS

Deviations from Original Protocol

In the original protocol we planned to report outcomes for participants aged five years or older.[24] However, as the included studies used a heterogeneous mix of descriptive statistics to describe follow-up duration (i.e. youngest age at follow-up, oldest age at follow-up, mean or median age at follow-up) all ages of participants with reported speech outcomes (with respect to velopharyngeal function) and secondary surgery rates were included. Risk of bias and level of evidence analysis was performed as documented below.

Study Selection

The results of the searches and study selection is summarised in Figure 1. Following database searches, 1882 papers were identified; 19 papers were found through other sources, including back searches. After three levels of screening by two reviewers, 383 papers were included for data extraction.

Study Characteristics

A complete list of references is provided in Supplementary Table 7 and details of included studies are presented in Supplementary Table 8. No papers meeting the inclusion criteria were published before 1960, with the majority of relevant studies being published between 2010-2018 (Table 1).

The regions of the world where studies were conducted were as follows: Europe (n=145), North America (n=131), Asia (n=73), Africa (n=14), South America (n=13), Australasia (n=8) and Middle East(n=6). The top 10 countries where included studies were conducted were as follows: USA (n=120), Sweden (n=33), United Kingdom (n=31), Finland (n=22), China (n=20), Japan (n=14), Germany (n=13), Egypt (n=10), India (n=10) and the Netherlands (n=10). Three hundred and forty one studies were reports from a single centre, 18 were multi-centre studies (2 centres n=7, 3 centres n=1, 4 centres n=3, 6 centres n=2, 9 centres n=1, 10 centres n=2, 12 centres n=1, 13 centres n=1) and in 24 studies, the number of centres involved was unclear. The mean number of participants per study was 173.3 (range 10-2616).

A total of 102 studies (26.6%) recorded mean follow-up duration. Fifteen studies (3.9%) recorded median follow-up duration. Five studies reported both mean and median follow-up

duration;271 studies (70.8%) recorded neither a mean nor a median follow-up duration. Of the 86 studies recording a minimum follow-up duration, eight studies had a minimum follow-up duration of 5 years or greater. While 74 studies recorded a maximum follow-up duration, this was less than 5 years in 29 studies. The youngest age at final review was recorded in 151 studies; in 60 studies this was 5 years of age or older. Of the 125 studies which recorded the oldest age at final review most (104) reported 5 years of age or older. Further details regarding follow-up duration are presented in Supplementary Table 4.

Risk of Bias in Studies

The following study types were recorded: case series (n=289; 75.4%), cohort studies (n=29; 7.6%), case-control studies (n=28; 7.3%), longitudinal studies (n=18; 4.7%), randomised controlled trials (n=11; 2.9%) and cross-sectional studies (n=6; 1.6%). For two studies the study type was unclear. Consequently, the studies included for data analysis had the following levels of evidence (https://www.cebm.ox.ac.uk/resources/levels-of-evidence/oxford-centre-for-evidence-based-medicine-levels-of-evidence-march-2009): level 1 n=11 (2.9%), level 2 n=29 (7.6%), level 3 n=28 (7.3%), level 4 n=313 (81.7%) level 5 n=0. It was not possible to determine the level of evidence in two studies (0.5%). For outcome reporting the majority of studies reporting perceptual speech assessment used an in-house scale (35%) or published scale (30%). The remainder of studies did not report the assessment tool (26%), used an undefined scale (2%), performed an informal assessment (4%) or used a method of speech assessment that was unclear to the reviewer (2%) (Supplementary Table 9).

Risk of bias is summarised in Table 3. Participant exclusions were evident in 174 studies. It is possible that some studies pre-selected their cohort and consequently exclusion criteria may not have been mentioned (86 studies did not report whether participants were excluded).

The commonest reasons reported for excluding participants were: presence of a syndrome (other than 22q11.2 microdeletion syndrome, n=90), lost to follow-up (n=63) and presence of hearing loss (n=61). The number of participants excluded also varied by decade from 17.9% in 2010-2018 to 45.2% in 1990-1999. Overall, 25% of participants were excluded from speech assessment or reporting of secondary surgery rate. With respect to attrition bias, most authors reported on a subset of patients operated on at their centre, with only 15% of studies reporting speech outcomes for all patients and only 12% reporting secondary surgery rates for all patients.

Results of Individual Studies

Cleft phenotype was the most-frequently reported variable (95% of studies). Further information regarding the recording of specific cleft phenotypes and the total number of participants in these studies when combined is shown in Supplementary Table 6. Sixty-four percent of studies reported the biological sex of participants and 44% documented syndromic status. The proportion of studies reporting the presence or absence of the following variables was as follows: hearing loss 28%, developmental delay 16%, Robin Sequence 16% and 22q11.2 microdeletion syndrome 11%. Parental smoking, birth order and head circumference were not reported in any study. The remaining variables, including ethnicity, adoption, family history of cleft and previous general anaesthesia, were reported in less than 10% studies (Supplementary Table 6). It was therefore not possible to carry out meta-analysis on these variables.

For each variable, Supplementary Tables 2 and 5 summarise the number of studies that reported secondary speech surgery and speech outcomes and highlight those studies which report a variable to have no influence on outcome or a negative influence on outcome.

Results of Syntheses

The number of participants included in the studies undergoing meta-analysis with respect to secondary surgery for particular variables was as follows: cleft phenotype (Veau I n=1151, Veau II n=1410, Veau III n=2450, Veau IV n=1069, submucous cleft palate (SMCP) n=475), biological sex (male n=1589, female n=1482), Robin Sequence (present n=600, absent n=1358), syndrome diagnosis (present n=213, absent n=3385) and 22q11.2 microdeletion syndrome (present n=117, absent n=406). The number of participants included in the meta-analysis with respect to normal resonance for particular variables was as follows: cleft phenotype (Veau I n=401, Veau II n=550, Veau III n=1859, Veau IV n=759, SMCP n=570), biological sex (males n=330, females n=389) and syndrome diagnosis (present n=150, absent n=818).

Cleft phenotype

The overall combined secondary surgery rate for all cleft phenotypes was 19.6% (2028/10321 patients). The combined secondary surgery rate for all included studies reporting this outcome with respect to cleft phenotype showed a statistically significant difference with (Chi-Squared test; $p<0.00001$): Veau I (12.5%), Veau II (23.8%), Veau III (18.9%), Veau IV (24.8%) and SMCP (21.1%). Following meta-analysis (Supplementary Figures 1-4) there was strong evidence to suggest that Veau I clefts were associated with less secondary surgery than other cleft phenotypes (RR=0.62, 95% CI 0.54-0.70; $P<0.0001$) and that Veau IV clefts were

associated with more secondary surgery than other cleft phenotypes (RR=1.22, 95% CI 1.08-1.37; P<0.0001). The data for submucous clefts was highly heterogeneous and therefore was excluded from analysis.

Overall, normal resonance was achieved by a higher proportion of participants with Veau I clefts (70.8%), when compared to Veau II (57.5%), Veau III (58.8%) and Veau IV (42.6%) clefts and those with submucous cleft palate (66.1%). Meta-analysis demonstrated strong evidence that Veau I clefts were associated with improved resonance scores (RR1.16, 95% CI 1.07-1.26; P<0.0001) and Veau IV clefts were associated with poorer resonance outcomes when compared to the other phenotypes (RR 0.86, 95% CI 0.81-0.92; P<0.001) (Supplementary Figures 5-8).

Biological Sex

Meta-analysis of 17 studies found no evidence that biological sex influenced secondary surgery rates (Supplementary Figure 9). There was no statistically significant difference between the 19.2% (284/1482) overall combined secondary surgery rate in female participants compared to 20.1% (320/1589) in male participants (RR = 1.07, 95% CI 0.93-1.24; P=0.32). There was no evidence of publication bias in these studies.

Six studies recorded the presence of normal resonance with respect to biological sex. Combining the results from these studies normal resonance was achieved in 63.3% (145/263) females and 56.3% (219/389) males. Meta-analysis of these studies found no evidence that males had increased hypernasality when compared to females (RR 0.91, 95% CI 0.83-1.01; P=0.07) (Supplementary Figure 10). When studies reported '*acceptable resonance'* rather

than '*normal resonance'* this increased to 74.5% (207/278) for females and 71.8% (216/301) for males. This again was not statistically significant (P>0.05; Chi-squared test).

Robin Sequence

Eleven studies had data reporting the secondary surgery rates in participants with RS and eight studies reported secondary surgery rates in participants without RS. Combined secondary surgery rates were 21.7% (130/600) and 13.7% (186/1358), respectively. This was statistically significant (P=0.00001; Chi-Squared test). Meta-analysis of seven studies, which included participants with and without RS, provided strong evidence RS was associated with more secondary surgery (RR 1.54, 95% CI 1.21-1.96; P<0.001) (Supplementary Figure 11). However, this calculation was based on seven studies with moderate heterogeneity (I^2=49.7%).

The combined data for three studies in those with RS gave a normal speech outcome of 53.8% (98/183) and normal nasal airflow of 42.7% (90/211). Meta-analysis of studies was not possible.

Syndrome Diagnosis

Fourteen studies recording syndrome diagnosis had a combined secondary surgery rate of 25.4% (54/213) compared to 26 studies with a combined secondary surgery rate of 14.7% (499/3385) in participants without a syndrome. This was statistically significant (P<0.00001; Chi-Squared test). Meta-analysis of 12 studies suggested strong evidence that syndromes were associated with an increase in secondary surgery (RR 1.69, 95% CI 1.25-2.28; P<0.001) (Supplementary Figure 12). There was no evidence of publication bias in included studies.

Normal resonance was 60.6% (91/150; 4 studies) in participants with a syndrome and 77.9% (638/818; 6 studies) in participants without a syndrome. The maximum number of studies in the other speech parameters (acceptable resonance, normal nasal airflow, passive cleft speech characteristics) was three and with small total participant numbers (See Table 2), making further analysis not possible.

22q11.2 Microdeletion Syndrome

Three studies recorded 22q11.2DS syndrome presence giving a combined mean secondary surgery rate of 39.5% (15/38) compared to two studies with a combined mean secondary surgery rate of 14.0% (50/357) in participants without 22q11.2DS. This was statistically significant ($P<0.00006$; Chi-Squared test). Meta analysis of four studies, where each individual study included participants with and without a diagnosis of 22q11.2DS, suggested moderate evidence for 22q11.2DS increasing the risk for secondary surgery (RR 3.18, 95% CI 1.85-5.44; $P<0.001$). However, caution is required as only four studies were included (Supplementary Figure 13).

Three studies reported the proportion of participants with 22q11.2DS who had normal resonance. The combined mean across studies was 25.6% (11/43). No studies reported speech outcomes for participants who were specifically identified not to have 22q11.2DS.

Adoption

Five studies reported secondary surgery rates in adopted individuals, giving a combined mean secondary surgery rate of 35.7% (127/356). In comparison, the combined mean secondary surgery rate in two studies including non-adopted participants was 15.9% (44/276). This was statistically significant ($P<0.00001$; Chi-Squared test). There was minimal reporting of

studies with respect to speech outcomes (zero or one study for each speech parameter) in adopted participants so meta-analysis was not possible (Table 2).

DISCUSSION

There is strong evidence to suggest that cleft phenotype impacts on speech outcomes (normal resonance and secondary speech surgery rate) in children born with a cleft palate. Veau I clefts were associated with less secondary surgery and Veau IV clefts were associated with more secondary surgery when compared to other cleft phenotypes. Interestingly, the combined secondary surgery rates for all included studies for Veau II clefts was higher (23.8%, 22 studies) than for Veau III clefts (18.9%, 34 studies). Under reporting of Robin Sequence, more commonly associated with Veau II clefts (as opposed to Veau III clefts), and the association with wider palatal clefts and poorer speech outcomes, might partially explain this. Wider palatal clefts may be more technically challenging to reconstruct, are associated with increasing cleft severity (higher Veau rank) and decreased palatal length.[6] Consequently, wider palatal clefts may result in higher rates of oro-nasal fistula and velopharyngeal disproportion and therefore potentially poorer speech outcomes with respect to velopharyngeal function.

In contrast to the findings of this meta-analysis Marrinan *et al.* and Hardin-Jones *et al.* reported higher rates of secondary speech surgery in patients with a cleft of the hard and soft palate (without a cleft lip) compared to bilateral cleft lip and palate or unilateral cleft lip and palate.[28,29] However, Sitzman *et al.* conducted a retrospective cohort study of 49 paediatric hospitals and reported in comparison to children with cleft palate only (Veau I and Veau II clefts) those with UCLP (Veau III) had 1.69-fold increase of secondary surgery (95% CI:

1.54-1.85) and children with a BCLP (Veau IV) had 2.06-fold increase in secondary surgery (95% CI: 1.85-2.28).[7] In our study the overall secondary surgery rate for all cleft phenotypes (including SMCP) was 19.6% (157 studies, 2028/10321 having secondary surgery). Although there is wide variation in reported secondary surgery rates, a figure of approximately 20% is often quoted and our findings reinforce this.[30] Sitzman *et al.* noted the predicted proportion of children undergoing secondary surgery by 5 years was 18.7% for the median surgeon.[7]

Relatively high numbers of studies reported normal resonance with respect to cleft phenotype. When all studies reporting normal resonance outcomes were combined, patients with Veau I clefts had the best outcomes, followed by Veau III and Veau II clefts. Veau IV clefts had the worst outcomes. This does not entirely fit with other reported evidence, namely that more extensive clefts were directly associated with worse speech outcomes.[6, 16,31-34] Studies investigating children with cleft palate only, found that those with a cleft of soft palate (Veau I), had significantly better outcomes, with respect to the need for extra speech therapy, articulation and secondary surgery, compared to those children with clefts of the soft and hard palate (Veau II).[28,31,33,35-40] Baillie and Sell reported that in 391 eligible participants hypernasality was relatively evenly distributed across cleft types (Veau I/II: 5%, Veau III: 8%, Veau IV: 6%) at five years.[8] However, they reported a statistically significant relationship between cleft type and articulation severity score (normal articulation was reported in 39% Veau IV clefts, 69% Veau III clefts and 85% Veau I/II clefts). More extensive clefts in the isolated cleft palate group were associated with perceptual features of VPI at five years.[8] Our findings support the articles that found poorer resonance outcomes for more extensive clefts. However, we found that nasal airflow results for Veau II clefts appeared to be worse than those for Veau III and IV clefts (Veau I clefts achieved the best

rate (73.7%; five studies) followed by Veau III clefts (66.0%; 14 studies), Veau IV clefts (65.1%; 10 studies) and Veau II clefts (57.3%; five studies)).

It is well documented that cleft phenotype varies with respect to biological sex. For example, the prevalence of isolated cleft palate is typically reported as 1.5 times more common in females when compared to males, whereas the prevalence of cleft lip with or without a cleft palate is twice as common in males.[41-46] Literature exploring the impact of biological sex on speech outcome is divided, with no statistically significant difference being found between biological sex and speech outcomes including hypernasality or velopharyngeal competence (VPC) rate in some compared to significant differences in hypernasality and articulation in other studies.[37, 47-53] Biological sex was not associated with VPI surgery or fistula repair rate in the following studies but did equate to a greater number of secondary surgery procedures for males in one study.[37, 39, 52, 54] Sitzman *et al.* documented biological sex was not associated with time to secondary surgery.[7] Our meta-analysis showed no association between biological sex and velopharyngeal outcomes although a large proportion of studies did not include this important patient characteristic. This emphasizes the need for prospective studies that record speech outcomes and secondary surgery to report on all baseline characteristics.

Our review found strong evidence that RS was associated with higher rates of secondary surgery. However, it is challenging to draw conclusions on the influence of RS on speech outcomes as there were very few included studies specifically reporting the speech variables of interest in participants recorded as having the presence or absence of RS. At least two recent studies demonstrated VPI and secondary surgery rates that were higher in patients with RS compared to those with isolated cleft palate.[55,56] Additionally, we found moderate evidence that 22q11.2DS is associated with increased secondary surgery rates, although the

3.3 overall effect of 22q11.2DS should be interpreted with caution, as this was based on only four case series (across which 31 participants had a diagnosis of 22q11.2DS and 406 did not have a diagnosis of 22q11.2DS); three of which were retrospective studies Hanes *et al.* 2015).[15,57-59] It was challenging to report speech results in those with 22q11.2DS due to the low number of studies reporting this. Poorer speech outcomes with higher revision rates are widely reported in 22q11.2DS patients when compared to non-syndromic cohorts.[60-63] Given the predisposition for multifactorial velopharyngeal dysfunction in children with 22q11.2DS (i.e. submucous or overt cleft palate, hypotonia, platybasia, adenoid hypoplasia), it is unsurprising that those with an overt cleft have worse outcomes overall. There remains work to be done in determining the relative contribution of each component to the overall presentation of VPD in these patients. Patients with 22q11.2DS are a heterogenous group of patients worthy of additional speech and language therapy input and support. Beyond RS and 22q11.DS the review found strong evidence that any syndrome diagnosis was associated with increased need for secondary surgery.

Although the number study of studies was low, secondary surgery rates for those who were adopted was higher than non-adopted participants. This an area which warrants further investigation. For example, were there differences in the timing of cleft palate repair and this was associated with an adverse effect on speech outcomes, was palatoplasty performed before or after adoption, were children adopted in their country of birth or elsewhere?

There is evidence to suggest that certain variables may influence outcomes following cleft palate repair including hearing loss, socioeconomic status, developmental delay and presence of cardiac co-morbidities. However, there was insufficient data within our search to conduct a

meta-analysis for these variables of interest. Sitzman *et al.* noted that median family income and the presence of additional congenital anomalies were not associated with time to secondary surgery.[7] It is likely that most databases in the US will in some way be connected to billing information thereby increasing the accuracy of this domain. However, this may exclude those with no health insurance and if this group were included, there might be a difference in the rate of secondary surgery. The reality of insufficient data to conduct meta-analysis for many of the variables of interest highlights that this is an area that would hugely benefit from further research. Together with poor levels of evidence (i.e. mostly retrospective case series), the limited number of studies reporting patient factors diminishes the applicability of a significant body of previous research. To be able to stratify patients from the outset would be invaluable in all healthcare environments.

Our results demonstrate heterogeneity in documentation of patient factors within studies which may influence speech outcomes. Additionally, we found that a minority of studies reported speech outcomes beyond five years of age and that just over a quarter of studies reported an average follow-up duration for their cohort. Given the evidence that speech outcomes change over time following cleft palate repair and secondary surgery rate increases with age it is important that follow-up duration is documented.[2,3,21] Progress should be made towards the inclusion of vulnerable patient groups to allow smaller subsets of patients to be analysed in a more robust manner. In an enlightened and compassionate society the outcomes for all individuals, regardless of co-morbidity, syndrome, or developmental delay should be reported.

Limitations

The major limitation of this review is that the results are based on the analysis and synthesis

of studies that are considered to be low level evidence. Within the studies differing methodologies for assessing and interpreting speech make meta-analysis challenging. The rate of secondary speech surgery was intentionally chosen as the primary outcome measure to try and avoid these challenges.[22, 64] However, secondary surgery rates must be interpreted with caution as there will be some children with poor speech outcomes who do not proceed to speech surgery. There is potential variability amongst raters when describing VPI severity and consequently, the threshold for secondary speech surgery is likely to vary between centres, clinicians, families and patients.[65] Additionally, as Sitzman *et al.* noted, the rates of secondary surgery may be underestimated due to lost to follow-up or secondary surgery being performed at another institution.[7]

To pool the results of the various speech assessments we used the authors' categorisation of normal, mild, moderate or severe hypernasality and nasal emission. If the authors did not use this categorisation then their results were not included in the analysis. Where speech was described as 'normal' or 'acceptable' in Table 2, this description was also taken directly from the text of the respective articles. We appreciate that these definitions are subjective and, therefore, that the objective speech outcome could vary significantly between these articles.

Many of the studies did not have reliability measures for the hypernasality, or indeed other perceptual outcome measures such as documentation of inter- or intra-rater reliability, blinded or consensus listening.

As regards study selection, it is acknowledged that only studies written in English were included, so some valid literature may have been omitted. By using back searching and validating the search criteria with an experienced medical librarian, it is felt that within the

limits of language, the majority of papers were captured. Additionally, one must always consider the possibility of a publication bias (i.e. a bias towards teams only reporting favourable outcomes). We attempted to quantify this by noting the number of studies which excluded patients. We would recommend that the 'gold standard' should be the reporting of outcomes for consecutive patients with no exclusions.

Future Work and Developing a Core Outcome Set and Clinical Prediction Tool

Debate abounds as to the optimal treatment protocol for initial cleft palate repair. Cleft management varies between surgeons, teams and countries and non-adherence to locally mandated management recommendations and guidelines is frequent.[66] Mossey *et al.* reported that for unilateral complete cleft lip and palate, 194 different protocols were used amongst 201 teams undertaking initial surgical repair internationally.[67] To provide the optimal treatment for individuals born with cleft, there needs to be a mechanism allowing meaningful comparison of outcomes, to allow recognition of the best protocol. In the absence of large-scale randomised controlled trials, this will only be possible if a minimum reporting data set is followed. Such a process is likely to require a Delphi process which is gaining interest in cleft care.[68, 69] On the basis of the current work, the authors suggest that establishing a core set of baseline characteristics and standardised speech outcome reporting. Hypernasality is considered to be one of the primary indictors of VPI.[70] This, along with posterior non-oral articulation, nasal emission and weak pressure consonants form the Velopharyngeal Composite Score Summary which may facilitate post-palatoplasty speech outcome assessment and allow international comparison given that different languages typically have dissimilar phonetic transcription and sound placement.

CONCLUSIONS

This review and meta-analysis have successfully pooled data from 47658 participants and three hundred and eighty-three studies to consider the evidence for associations between a range of exposure variables and outcomes for VPI. While the robustness of some of the included studies could have been stronger, combining the results across studies provides a robust analysis of the importance of the candidate exposure variables with a high level of evidence. The results showed that Veau classification (Veau I type clefts best outcomes, Veau IV type clefts worst outcomes), RS or a diagnosed syndrome are important predictors of outcome for VPI. There was no evidence that one biological sex was associated with higher rates of VPI.

The evidence could be strengthened further by defining the threshold for secondary speech surgery, establishing minimum post-palatoplasty follow-up duration (for example, 5 years of age), establishing standardised speech outcome reporting and developing a core set of baseline characteristics which might include: cleft phenotype, biological sex, presence of syndrome / genetic diagnosis, presence of Robin Sequence and hearing loss. Including teams from low middle income countries in such work is crucial.

Word Count: 5592 (7000)

REFERENCES

1. Mackay D, Mazahari M, Graham WP, Jeffords K, Leber D, Gorman P, Lieser JD, Wrye S, Kutz R, Saggers G. Incidence of operative procedures on cleft lip and palate patients. Ann Plast Surg. 1999;42(4):445-448
2. Sommerlad BC. A technique for cleft palate repair. Plast Reconstr Surg. 2003;112(6):1542-1548
3. Sommerlad BC. The use of the operating microscope for cleft palate repair and pharyngoplasty. Plast Reconstr Surg. 2003;112(6):1540-1541
4. Lithovius RH, Ylikontiola LP, Sandor GK. Frequency of pharyngoplasty after primary repair of cleft palate in northern Finland. Oral Surg Oral Med Oral Pathol Oral Radiol. 2014;117(4):430-434
5. Smith DM, Losee JE. Cleft palate repair. Clin Plast Surg 2014;41:189-210
6. Mahoney M-H, Swan MC, Fisher DMF. Prospective analysis of presurgical risk factors for outcomes in primary palatoplasty. Plastic Reconstr Surg. 2013;132(1):165-171
7. Sitzman TJ, Hossain M, Carle AC, Heaton PC, Britto MT. Five-Fold Variation Among Surgeons and Hospitals in the Use of Secondary Palate Surgery. Cleft Palate-Craniofacial J. 2019;56(5):586-594
8. Baillie L, Sell D. Benchmarking Speech, Velopharyngeal Function Outcomes and Surgical Characteristics Following the Sommerlad Protocol and Palate Repair Technique. Cleft Palate Craniofac J. 2020 Oct;57(10):1197-1215
9. Williams AC, Sandy J, Thomas S, Sell D, Sterne JA. Influence of surgeon's

experience on speech outcome in cleft lip and palate. Lancet. 1999;354(9191):1697-1698

10. Persson C, Lohmander A, Elander A. Speech in children with an isolated cleft palate: a longitudinal perspective. Cleft Palate-Craniofacial J. 2006;43(3):295-309
11. Brandao GR, de Souza Freitas JA, Genaro KF, Yamashita RP, Fukushiro AP, Lauris JR. Speech outcomes and velopharyngeal function after surgical treatment of velopharyngeal insufficiency in individuals with signs of velocardiofacial syndrome. J Craniofac Surg. 2011;22(5):1736-1742.
12. Bezuhly M, Fischbach S, Klaiman P, Fisher DM. Impact of 22q deletion syndrome on speech outcomes following primary surgery for submucous cleft palate. Plast Reconstr Surg. 2012;129(3):502e-510e
13. Patel KB, Sullivan SR, Murthy AS, Marrinan E, Mulliken JB. Speech outcome after palatal repair in nonsyndromic versus syndromic Robin sequence. Plast Reconstr Surg. 2012;130(4):577e-584e.
14. Moar KK, Sweet C, Beale V. Fistula rate after primary palatal repair with intravelarveloplasty: a retrospective three-year audit of six units (NorCleft) in the UK. British J Oral Maxillofac Surg. 2016;54(6):634-637
15. Basta MN, Silvestre J, Stransky C, Solot C, Cohen M, McDonald-McGinn D, Zackai E, Kirschner R, Low DW, Randall P, et al. A 35-year experience with syndromic cleft palate repair: operative outcomes and long-term speech function. Ann Plast Surg. 2014;73(suppl 2):S130-S135
16. Choa RM, Slator R, Jeremy A, Robinson S, Franklin D, Roberts A, Winter R, Extence H, Collard M, Clark V. Identifying the effect of cleft type, deprivation and ethnicity on speech and dental outcomes in UK cleft patients: a multi-centred study. J Plast Reconstr Aesthet Surg. 2014 Dec;67(12):1637-43

17. Feragen KB, Stock NM. When there is more than a cleft: psychological adjustment when a cleft is associated with an additional condition. Cleft Palate-Craniofac J. 2014;51(1):5-14

18. Smillie I, Yong K, Harris K, Wynne DM, Russell CJ. Socioeconomic influence on orofacial cleft patient care. Scott Med J. 2015;60(2):70-4

19. Rautio J, Andersen M, Bolund S, Hukki J, Vindenes H, Davenport P, Arctander K, Larson O, Berggren A, Abyholm F, et al. Scand cleft randomised trials of primary surgery for unilateral cleft lip and palate: 2. Surgical results. J Plastic Surg Hand Surg. 2017;51(1):14-20

20. Shaw W, Semb G. The Scandcleft randomised trials of primary surgery for unilateral cleft lip and palate: 11. What next? J Plastic Surg Hand Surg. 2017;51(1):88-93

21. Sitzman TJ, Hossain M, Carle AC, Heaton PC, Britto MT. Variation among cleft centres in the use of secondary surgery for children with cleft palate: a retrospective cohort study. BMJ Paediatr Open. 2017;1(1):e000063

22. Stein MJ, Zhang Z, Fell M, Mercer N, Malic C. Determining postoperative outcomes after cleft palate repair: A systematic review and meta-analysis. J Plast Reconstr Aesthet Surg. 2019;72(1):85-91

23. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, Shamseer L, Tetzlaff JM, Akl EA, Brennan SE, Chou R, Glanville J, Grimshaw JM, Hróbjartsson A, Lalu MM, Li T, Loder EW, Mayo-Wilson E, McDonald S, McGuinness LA, Stewart LA, Thomas J, Tricco AC, Welch VA, Whiting P, Moher D. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021; 372 :n71

24. Sainsbury D, Williams C, Blacam C, Mullen J, Chadha A, Wren Y, et al. Non-Interventional Factors Influencing Velopharyngeal Function For Speech In Initial

Cleft Palate Repair: A Systematic Review Protocol. Systematic Reviews 2019;8(261):1–8

25. Morgan RL, Whaley P, Thayer KA, Schünemann HJ. Identifying the PECO: A framework for formulating good questions to explore the association of environmental and other exposures with health outcomes. Environ Int. 2018;121(Pt 1):1027-1031
26. Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from www.handbook.cochrane.org.
27. Harrer M, Cuijpers P, Kanzawa T, Ebert DD. Doing meta-analysis with R. A hands-on guide. CRC Press. 2021.
28. Marrinan E, LaBrie RA, Mulliken JB. Velopharyngeal function in nonsyndromic cleft palate: relevance of surgical technique, age at repair, and cleft type. Cleft Palate-Craniofac J. 1998; 35(2):95-100
29. Hardin-Jones MA, Brown CK, Van Demark DR, Morris HL. Long-term speech results of cleft palate patients with primary palatoplasty. Cleft Palate Craniofac J. 1993 Jan;30(1):55-63
30. Kuehn D, Moller K. Speech and Language Issues in the Cleft Palate Population: The State of the Art. Cleft Palate-Craniofacial J. 2000;37(4):348-348
31. Hardin-Jones MA, Jones DL. Speech production of preschoolers with cleft palate. Cleft Palate-Craniofac J. 2005;42(1):7-13
32. Sullivan SR, Marrinan EM, La Brie RA, Rogers GF, Mulliken JB. Palatoplasty outcomes in nonsyndromic patients with cleft palate: a 29-year assessment of one surgeon’s experience. J Craniofac Surg. 2009;20(3):612-616
33. Klintö K, Falk E, Wilhelmsson S, Schönmeyr B, Becker M. Speech in 5-year-olds

with cleft palate with or without cleft lip treated with primary palatal surgery with muscle reconstruction according to Sommerlad. Cleft Palate-Craniofacial J 2018;55(10):1399-1408

34. Nyberg J, Westberg LR, Neovius E, Larson O, Henningsson G. Speech results after one-stage palatoplasty with or without muscle reconstruction for isolated cleft palate. Cleft Palate-Craniofac J. 2010;47(1):92-103
35. Jakobsson OP, Pontén B. Closure of the cleft palate in one or two stages: the surgical methods. Scand J Plast Reconstr Surg Hand Surg. 1990;24(2):149-55
36. Lohmander-Agerskov A, Havstam C, Söderpalm E, Elander A, Lilja J, Friede H, Persson EC. Assessment of speech in children after repair of isolated cleft palate. Scand J Plast Reconstr Surg Hand Surg. 1993;27(4):307-10
37. Bicknell S, McFadden LR, Curran JB. Frequency of pharyngoplasty after primary repair of cleft palate. J Can Dent Assoc. 2002 Dec;68(11):688-92
38. Persson C, Elander A, Lohmander-Agerskov A, Soderpalm E. Speech outcomes in isolated cleft palate: impact of cleft extent and additional malformations. Cleft Palate-Craniofacial J. 2002;39(4):397-408
39. Gustafsson C, Heliovaara A, Leikola J, Rautio J. Incidence of speech correcting surgery in children with isolated cleft palate. Cleft Palate-Craniofac J. 2018;55(8):1115-1121
40. Andersson EM, Sandvik L, Tørdal IB, Abyholm F. Pharyngoplasty after primary repair of clefts of the secondary palate. Scand J Plast Reconstr Surg Hand Surg. 2010;44(1):26-30
41. Shaw GM, Croen LA, Curry CJ. Isolated oral cleft malformations: associations maternal and infant characteristics in a Calfornia populations. Tertaology. 1991;43(3):225-228

42. Freitas JA, Dalben Gda S, Santamaria M, Freitas PZ. Current data on the characterization of oral clefts in Brazil. Braz Oral Res. 2004;18(2):128–133

43. Hashmi SS, Waller DK, Langlois P, Canfield M, Hecht JT. Prevalence of nonsyndromic oral clefts in Texas: 1995-1999. Am J Med Genet A. 2005;134(10):368–372

44. Dixon MJ, Marazita ML, Beaty TH, Murray JC. Cleft lip and palate: understanding genetic and environmental influences. Nat Rev Genet. 2011;12(3):167–178

45. Mossey PA, Modell B. Epidemiology of oral clefts 2012: an international perspective. Front Oral Biol. 2012;16(4):1–18

46. Pool SMW, der Lek LM van, de Jong K, Vermeij-Keers C, Mouës-Vink CM. Embryologically Based Classification Specifies Gender Differences in the Prevalence of Orofacial Cleft Subphenotypes. The Cleft Palate-Craniofacial Journal. 2021;58(1):54-60.

47. David DJ, Anderson PJ, Schnitt DE, Nugent MA, Sells R. From birth to maturity: a group of patients who have completed their protocol management. Part II. Isolated cleft palate. Plast Reconstr Surg. 2006 Feb;117(2):515-26

48. Joos U, Wermker K, Kruse-Löesler B, Kleinheinz J. Influence of treatment concept, velopharyngoplasty, gender and age on hypernasality in patients with cleft lip, alveolus and palate. J Craniomaxillofac Surg. 2006 Dec;34(8):472-7

49. Chen Q, Zheng Q, Shi B, Yin H, Meng T, Zheng GN. Study of relationship between clinical factors and velopharyngeal closure in cleft palate patients. J Res Med Sci. 2011 Jul;16(7):945-50

50. Yang Y, Li Y, Wu Y, Gu Y, Yin H, Long H, Shi B, Zheng Q. Velopharyngeal function of patients with cleft palate after primary palatoplasty: relevance of sex, age, and cleft type. J Craniofac Surg. 2013 May;24(3):923-8

51. Follmar KE, Yuan N, Pendleton CS, Dorafshar AH, Kolk CV, Redett RJ 3rd. Velopharyngeal Insufficiency Rates After Delayed Cleft Palate Repair: Lessons Learned From Internationally Adopted Patients. Ann Plast Surg. 2015 Sep;75(3):302-5.

52. Hosseinabad HH, Derakhshandeh F, Mostaajeran F, Abdali H, Davari HA, Hassanzadeh A, Kummer AW. Incidence of velopharyngeal insufficiency and oronasal fistulae after cleft palate repair: A retrospective study of children referred to Isfahan Cleft Care Team between 2005 and 2009. Int J Pediatr Otorhinolaryngol. 2015 Oct;79(10):1722-6

53. Willadsen E, Lohmander A, Persson C, Lundeborg I, Alaluusua S, Aukner R, Bau A, Boers M, Bowden M, Davies J, Emborg B, Havstam C, Hayden C, Henningsson G, Holmefjord A, Hölttä E, Kisling-Møller M, Kjøll L, Lundberg M, McAleer E, Nyberg J, Paaso M, Pedersen NH, Rasmussen T, Reisæter S, Andersen HS, Schöps A, Tørdal IB, Semb G. Scandcleft randomised trials of primary surgery for unilateral cleft lip and palate: 5. Speech outcomes in 5-year-olds - consonant proficiency and errors. J Plast Surg Hand Surg. 2017 Feb;51(1):38-51

54. Andersson EM, Sandvik L, Semb G, Abyholm F. Palatal fistulas after primary repair of clefts of the secondary palate. Scand J Plast Reconstr Surg Hand Surg. 2008;42(6):296-9.

55. Logjes RJH, Upton S, Mendelsohn BA, Badiee RK, Breugem CC, Hoffman WY, Pomerantz JH. Long-term Speech Outcomes of Cleft Palate Repair in Robin Sequence versus Isolated Cleft Palate. Plast Reconstr Surg Glob Open. 2021 Jan 21;9(1):e3351. doi: 10.1097/GOX.0000000000003351. Erratum in: Plast Reconstr Surg Glob Open. 2021 Apr 23;9(4):e3583. PMID: 33564582; PMCID: PMC7859383.

56. Schwaiger M, Cook H, Jordan Z, Edmondson SJ, Mischak I, Wallner J, Atherton D,

Haers P, Timoney N, Ferguson L. Robin Sequence: 5-Year Speech Outcomes-A Case-Control Study. Plast Reconstr Surg. 2021;147(3):676-686

57. Sommerlad BC, Fenn C, Harland K, Sell D, Birch MJ, Dave R, Lees M, Barnett A. Submucous cleft palate: a grading system and review of 40 consecutive submucous cleft palate repairs. Cleft Palate Craniofac J. 2004;41(2):114-23
58. Oji T, Sakamoto Y, Ogata H, Tamada I, Kishi K. A 25-year review of cases with submucous cleft palate. Int J Pediatr Otorhinolaryngol. 2013;77(7):1183-5
59. Hanes LA, Murphy A, Hatchette JE, Delorey R, Wilson KL, Hong P, Bezuhly M. Chronic Otitis Media with Effusion Is Associated with Increased Risk of Secondary Speech Surgery. Plast Reconstr Surg. 2015;136(2):343-349
60. Losken A, Williams JK, Burnstein FD, Malick DN, Riski JE. Surgical correction of velopharyngeal insufficiency in chil-dren with velocardiofacial syndrome. Plast Reconstr Surg 2006;117:1493–1498
61. Widdershoven JCC, Stubenitsky BM, Breugem CC, Mink van der Molen AB. Outcome of velopharyngoplasty in patients with velocardiofacial syndrome. Arch Otolaryngol Head Neck Surg 2008;134:1159–1164
62. Nugent N, McGillivary A, Earley MJ. 22q11 chromosome abnormalities and the cleft service. J Plast Reconstr Aesthet Surg 2010;63:598–602
63. Sainsbury DCG, Filson S, Butterworth S. et al. Velopharyngoplasty in patients with 22q11.2 microdeletion syndrome: outcomes following the Newcastle protocol. Eur J Plast Surg 2013;36:607–618
64. de Blacam C, Smith S, Orr D. Surgery for Velopharyngeal Dysfunction: A Systematic Review of Interventions and Outcomes. Cleft Palate Craniofac J. 2018 Mar;55(3):405-422
65. Ahl R, Harding-Bell A. Comparing methodologies in a series of speech outcome

studies: challenges and lessons learned. Cleft Palate Craniofac J 2018;55(1):35-44

66. Shaw WC, Semb G, Nelson P, Brattström V, Mølsted K, Prahl-Andersen B, Gundlach KK. The Eurocleft project 1996-2000: overview. J Craniomaxillofac Surg. 2001 Jun;29(3):131-40; discussion 141-2
67. Mossey PA, Little J, Munger RG, Dixon MJ, Shaw WC. Cleft lip and palate. Lancet. 2009 Nov 21;374(9703):1773-85
68. Harman NL, Bruce IA, Callery P, Tierney S, Sharif MO, O'Brien K, Williamson PR. MOMENT--Management of Otitis Media with Effusion in Cleft Palate: protocol for a systematic review of the literature and identification of a core outcome set using a Delphi survey. Trials. 2013 Mar 12;14:70.
69. de Blacam C, Baylis AL, Kirschner RE, et al Protocol for the development of a core outcome set for reporting outcomes of management of velopharyngeal dysfunction. BMJ Open 2020;10:e036824
70. 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):347-62

Figure Legends

Figure 1. Study selection PRISMA-P flow diagram.

Supplementary Figure 1. Meta-analysis of studies reporting secondary surgery as an outcome and cleft phenotype as an exposure (Veau I compared against Veau II (upper

figure), Veau I compared against Veau III (middle figure) and Veau I compared against Veau IV (lower figure).

Supplementary Figure 2. Meta-analysis of studies reporting secondary surgery as an outcome and cleft phenotype as an exposure (Veau II compared against Veau I (upper figure), Veau II compared against Veau III (middle figure) and Veau II compared against Veau IV (lower figure).

Supplementary Figure 3. Meta-analysis of studies reporting secondary surgery as an outcome and cleft phenotype as an exposure (Veau III compared against Veau I (upper figure), Veau III compared against Veau II (middle figure) and Veau III compared against Veau IV (lower figure).

Supplementary Figure 4. Meta-analysis of studies reporting secondary surgery as an outcome and cleft phenotype as an exposure (Veau IV compared against Veau I (upper figure), Veau IV compared against Veau II (middle figure) and Veau IV compared against Veau III (lower figure).

Supplementary Figure 5. Meta-analysis of studies reporting normal resonance as an outcome and cleft phenotype as an exposure (Veau I compared against Veau II (upper figure), Veau I compared against Veau III (middle figure) and Veau I compared against Veau IV) (lower figure)

Supplementary Figure 6. Meta-analysis of studies reporting normal resonance as an outcome and cleft phenotype as an exposure (Veau II compared against Veau I (upper

figure), Veau II compared against Veau III (middle figure) and Veau II compared against Veau IV (lower figure).

Supplementary Figure 7. Meta-analysis of studies reporting normal resonance as an outcome and cleft phenotype as an exposure (Veau III compared against Veau I (upper figure), Veau III compared against Veau II (middle figure) and Veau III compared against Veau IV (lower figure).

Supplementary Figure 8. Meta-analysis of studies reporting normal resonance as an outcome and cleft phenotype as an exposure (Veau IV compared against Veau I (upper figure), Veau IV compared against Veau II (middle figure) and Veau IV compared against Veau III (lower figure).

Supplementary Figure 9. Meta-analysis of studies reporting secondary surgery as an outcome and biological sex as an exposure.

Supplementary Figure 10. Meta-analysis of studies reporting normal resonance as an outcome and biological sex as an exposure.

Supplementary Figure 11. Meta-analysis of studies reporting secondary surgery as an outcome and Robin Sequence as an exposure.

Supplementary Figure 12. Meta-analysis of studies reporting secondary surgery as an outcome and syndrome diagnosis as an exposure.

Supplementary Figure 13. Meta-analysis of studies reporting secondary surgery as an outcome and 22q11.2 microdeletion syndrome as an exposure.