The primary goal of palatoplasty is anatomical reconstruction of an intact palate to allow for development of normal speech while minimizing the incidence of oronasal fistula and ensuring long-term harmonious facial growth. Kriens first drew attention to the muscular anatomy of the soft palate, describing the abnormal positioning of the levator veli palatini muscle and loss of integrity of the supporting structures as the primary causes for impairment of speech production following palate repair. Therefore, he advocated for release and reorientation of the levator muscles. This method, termed intravelar veloplasty, became an inherent component of the most commonly used cleft palate repair techniques, including von Langenbeck palatoplasty, Veau-Wardill-Kilner pushback palatoplasty, and the Bardach two-flap palatoplasty.

In 1976, in an attempt to improve the speech results following palate repair while allowing adequate maxillary growth, Furlow devised the double-opposing Z-plasty method for the repair of clefts involving the soft palate. All of these techniques produce a straight-line closure of the soft palate and are often grouped under the single designation of “straight-line” repairs.

A systematic review was conducted to compare the speech outcomes and fistula rates following repair of the cleft palate with Furlow double-opposing Z-plasty and straight-line intravelar veloplasty techniques. A systematic search of the English literature published in the MEDLINE, Ovid, and Embase electronic databases was performed using the following keywords: “cleft palate,” “intravelar veloplasty,” “velopharyngeal insufficiency,” and “speech outcome.” The exclusion criteria were as follows: syndromic patients, no description or poor description of the technique used, data not stratified by cleft palate type, two-stage cleft palate repair, average age at repair younger than 9 months or older than 18 months, and age at the last follow-up younger than 4 years. Statistical analysis was used to compare the rate of secondary operations and the incidence of velopharyngeal insufficiency.

Twelve studies satisfied the inclusion criteria. In the isolated cleft palate group, the mean failure rates were 9.7 and 16.5 percent for Furlow double-opposing Z-plasty and straight-line intravelar veloplasty closure, respectively. In the unilateral cleft lip–cleft palate group, the mean failure rates were 11.1 and 17.1 percent for Furlow and straight-line intravelar veloplasty closure, respectively. The difference in the odds of requiring secondary surgery in the straight-line intravelar veloplasty repair group versus the Furlow group was statistically significant ($p = 0.03$) in unilateral cleft lip–cleft palate.

This systematic review indicated an increased incidence of velopharyngeal insufficiency as revealed by higher odds of secondary operations in the straight-line intravelar veloplasty repair of unilateral cleft lip–cleft palate when compared with Furlow Z-plasty. (Plast. Reconstr. Surg. 134: 1014, 2014.)
incisions, optimizing the vascularity of both the mucosal and muscular layers, and reconstructing the functional levator veli palatini sling.

Despite the wealth of literature on palate repair techniques, there is lack of consensus as to which surgical technique yields optimal speech results (i.e., the lowest incidence of velopharyngeal insufficiency). Currently, numerous cleft palate teams report an average failure rate of approximately 5 to 38 percent in achieving a competent velopharyngeal mechanism, with an average range of 10 to 25 percent of the patients needing a second operation to correct velopharyngeal insufficiency.\(^4\) Given the emerging number of publications on improved outcomes following Furlow soft palate repair, the authors performed a systematic review to compare the speech outcomes between the Furlow and straight-line intravelar veloplasty methods of soft palate closure in isolated cleft palate and unilateral cleft lip–cleft palate.

**PATIENTS AND METHODS**

A comprehensive review of the English literature published in the MEDLINE, Ovid, and Embase electronic databases from January 1, 1978, to October 17, 2013, was performed. Pertinent references from articles found were reviewed and selected if they satisfied the inclusion criteria. The search strategy used the following terms as both subjects and keywords: “cleft palate,” “intravelar veloplasty,” “velopharyngeal insufficiency,” and “speech outcome.” The following exclusion criteria were applied to the search results: not in the English language, nonhuman studies, syndromic patients, no description or poor description of the surgical technique used, data not stratified by cleft palate type, two-stage cleft palate repair, average age at repair younger than 9 months or older than 18 months, and age at the last follow-up younger than 4 years. These exclusion criteria were sought to control for confounding variables and to capture a population of patients whose age at repair and protocol of palate repair were more consistent with the nationally accepted standards of care. The age of 4 years was chosen, as children with velopharyngeal insufficiency at this age are highly unlikely to develop subsequent velopharyngeal insufficiency.\(^7\)

Two authors (G.D.D. and M.R.T.) screened the articles based on the exclusion criteria, and independently reviewed all of the titles and abstracts. Any discrepancies between the authors were discussed thoroughly, and a decision was made as to whether a study should be included (Fig. 1). Because of the heterogeneity and subjectivity of methods establishing the presence of velopharyngeal insufficiency, the intention to undergo surgery for velopharyngeal insufficiency was used as the primary outcome measure for speech, assuming that only those patients who had prominent insufficiency interfering with daily functions would be advised to undergo secondary procedures. Reporting of moderate to severe hypernasality by a speech pathologist was used as

![Fig. 1. The stepwise method by which the final 12 articles were selected for the systematic review.](image-url)
a secondary outcome measure for the presence of insufficiency, because three studies fit all inclusion criteria except for reporting secondary operations. For studies that described both secondary operations and hypernasality, there was a direct correlation between patients who had moderate to severe hypernasality and those for whom secondary surgery was recommended.

Data were stratified by cleft type (isolated cleft palate or unilateral cleft lip–cleft palate) to avoid the bias inherent in the specific malformation. Because of the small number of patients affected by bilateral cleft lip–cleft palate, a decision was made to assess the outcomes only in patients with isolated clefts of the palate or unilateral cleft lip–cleft palate.

Furlow double-opposing Z-plasty repair of the soft palate was compared to straight-line methods of closure of the soft palate combined with intravelar veloplasty. Straight-line closures included the following specific types of palatoplasty techniques: von Langenbeck, V-Y pushback, and two-flap palatoplasty. The following data were extracted: year of publication, study design, sample size, mean age of patients, estimated age at the last speech assessment, type of cleft (Veau classification of cleft palate), method of palate repair, methods of speech assessment, rate of secondary operations for velopharyngeal insufficiency (performed or recommended), rate of moderate to severe hypernasality, fistula, and other complication rates.

**Statistical Analysis**

Event rates using a composite endpoint of secondary operations or hypernasality were calculated. Logistic regression models using generalized estimating equations that assume an independent correlation structure were used to calculate odds ratios for the event rates between groups. Sensitivity analyses were performed using only the primary endpoint (secondary surgery). Subsequently, the impact of each individual study was evaluated by calculating odds ratios, excluding one study at a time. The articles included in each group were compared with regard to the publication year (t test), to exclude for variability induced by the improvement of surgical instrumentation and technique over time. The analyses were performed using SAS Version 9.2 (SAS Institute, Inc., Cary, N.C.). A significance level of 0.05 was used for all tests.

**RESULTS**

The initial search of the databases yielded 5911 studies that included at least one of the aforementioned keywords. After enacting the general exclusion criteria, 3043 articles were selected based on the title and abstract and were read thoroughly. Applying the specific exclusion criteria strictly, 13 articles were further selected. One additional article was eventually excluded because it represented duplicate data from the same authors (Fig. 1).

The 12 studies that met the inclusion criteria are listed in Table 1. The studies included in the Furlow group were published between 1986 and 2013 (average publication year, 2001; median, 2001), and the studies included in the straight-line intravelar veloplasty group were published between 1984 and 2010 (average publication year, 2002; median, 2005). These differences were not statistically significant (p = 0.88). All studies were retrospective, with the exception of one prospective, randomized, clinical trial. Nine hundred twenty-seven patients underwent Furlow double-opposing Z-palatoplasty (five studies) (Table 1) and 1205 patients underwent straight-line intravelar veloplasty repair (seven studies) (Table 1). From this pool of patients, 697 of the patients in the Furlow and 1040 patients in the straight-line intravelar veloplasty repair groups had speech evaluation data and were used to compare the speech outcomes. The average age at palatoplasty was 13.5 and 13.9 months, respectively. Detailed data on age range at repair and speech evaluation are listed in Table 1.

**Oronasal Fistula**

The overall oronasal fistula rate was 7.87 percent in the Furlow repair group compared with 9.81 percent in the straight-line intravelar veloplasty repair group. This difference was not statistically significant (p = 0.14).

**Effect of Cleft Type on Fistula Rate**

In the Furlow group, the fistula rate was significantly lower (p < 0.0001) in Veau I/II (2.60 percent) compared with Veau III cleft palates (11.62 percent). In the straight-line intravelar veloplasty group, the mean oronasal fistula rate was 4 percent in Veau I/II and 13.71 percent in Veau III cleft palates (p = 0.0004). When the data were stratified by Veau classification, there was no statistical difference in fistula rate between the Furlow and straight-line intravelar veloplasty repair groups (Veau I/II, p = 0.33; Veau III, p = 0.38).

**Speech Outcomes**

All studies used perceptual speech evaluation performed by experienced speech therapists for
Table 1. Final 12 Articles Selected for Systematic Review

<table>
<thead>
<tr>
<th>Reference</th>
<th>Study Design</th>
<th>Total No. of Patients per Palate Repair Technique</th>
<th>No. of Patients by Cleft Type</th>
<th>Age at Palate Repair (mo)</th>
<th>Estimated Age at Speech Assessment (yr)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jackson et al., 2013</td>
<td>Retrospective</td>
<td>Furlow, 656</td>
<td>Veau I/II, 312; Veau III, 344</td>
<td>13.2</td>
<td>7</td>
</tr>
<tr>
<td>Furlow, 1986</td>
<td>Retrospective</td>
<td>Furlow, 14</td>
<td>Veau I/II, 6; Veau III, 8</td>
<td>10.8</td>
<td>4</td>
</tr>
<tr>
<td>Brothers et al., 1995</td>
<td>Retrospective</td>
<td>Furlow, 21; V-Y pushback, 10</td>
<td>Veau I, 31</td>
<td>15.6</td>
<td>4</td>
</tr>
<tr>
<td>Yu et al., 2014</td>
<td>Retrospective</td>
<td>Furlow, 46; von Langenbeck, 50</td>
<td>Veau I/II, 96</td>
<td>13.5</td>
<td>4</td>
</tr>
<tr>
<td>Williams et al., 2011</td>
<td>Prospective RCT</td>
<td>Furlow, 190; von Langenbeck, 269</td>
<td>Veau III: 459</td>
<td>13.5</td>
<td>4</td>
</tr>
<tr>
<td>Koh et al., 2009</td>
<td>Retrospective</td>
<td>Two-flap palatoplasty, 30</td>
<td>Veau III, 30</td>
<td>13.5</td>
<td>5</td>
</tr>
<tr>
<td>Inman et al., 2005</td>
<td>Retrospective</td>
<td>V-Y pushback/von Langenbeck, 148</td>
<td>Veau I/II, 148</td>
<td>12.0</td>
<td>&gt;4 (range, 3–17)</td>
</tr>
<tr>
<td>Ito, 2006</td>
<td>Retrospective</td>
<td>V-Y pushback, 109</td>
<td>Veau I/II, 52; Veau III, 57</td>
<td>16.0</td>
<td>&gt;4 (range, 3–12)</td>
</tr>
<tr>
<td>Salver et al., 2006</td>
<td>Retrospective</td>
<td>Two-flap palatoplasty, 381</td>
<td>Veau III, 381</td>
<td>9.4</td>
<td>&gt;4 (range, 4–8)</td>
</tr>
<tr>
<td>Dreyer and Trier, 1984</td>
<td>Retrospective</td>
<td>von Langenbeck, 22</td>
<td>Veau I/II, 22</td>
<td>18.9</td>
<td>4.6</td>
</tr>
<tr>
<td>Marrinan et al., 1998</td>
<td>Retrospective</td>
<td>von Langenbeck, 52; V-Y pushback, 116</td>
<td>Veau I/II, 77 (21 von Langenbeck; 56 V-Y pushback); Veau III, 100 (31 von Langenbeck; 69 V-Y pushback)</td>
<td>11.4</td>
<td>&gt;4</td>
</tr>
<tr>
<td>Polzer et al., 2006</td>
<td>Retrospective</td>
<td>V-Y pushback/von Langenbeck, 9</td>
<td>Veau I/II, 9</td>
<td>13.9</td>
<td>4.3</td>
</tr>
</tbody>
</table>

RCT, randomized controlled trial.

assessments of velopharyngeal function. When velopharyngeal insufficiency was suspected, nasoendoscopy or videofluoroscopy was performed in five studies. Velopharyngeal function was not scored uniformly across the studies with the same scale; therefore, the need for secondary surgery (performed or recommended) was selected as the primary outcome measure and hypernasality was selected as the secondary outcome measure. In all articles, “secondary surgery” was used to group the procedures to correct velopharyngeal insufficiency. Techniques used for secondary correction of velopharyngeal insufficiency included sphincter pharyngoplasty, pharyngeal flap, and Furlow Z-plasty. Other indications for surgery such as fistula repair were listed separately.

The need for secondary procedures to correct velopharyngeal insufficiency in the Furlow group ranged between 0% and 11.4 percent in isolated cleft palate, and between 0% and 6.7 percent in unilateral cleft lip–cleft palate. The prevalence of hypernasality ranged between 13% and 14.3 percent in isolated cleft palate, and between 8.9% and 18.5 percent of patients in unilateral cleft lip–cleft palate. In the straight-line intravelar veloplasty repair group, the need for secondary procedures ranged between 9.1% and 29.2 percent in isolated cleft palate patients, and between 6.7% and 19.4 percent in patients affected by unilateral cleft lip–cleft palate. The prevalence of hypernasality ranged between 11.1% and 20.0 percent in patients suffering from isolated cleft palate and between 29.1% and 33.3 percent of patients in unilateral cleft lip–cleft palate (Table 2).

The odds ratios for the composite endpoint and the primary outcome in Furlow and straight-line intravelar veloplasty groups are listed in Table 3. Although all odds ratios were above 1, indicating increased risk of an event for the latter group, only the analysis of the primary outcome among unilateral cleft lip–cleft palate patients was statistically significant ($p = 0.03$). Plots of all studies are shown in Figures 2 and 3. Although we could not test formally for heterogeneity in effect among the studies, the figures both indicate a wide variety of event rates within both surgery types. The sensitivity analysis excluding one study at a time for each endpoint demonstrated that the study by Jackson et al. was highly influential in all of the analyses.

**DISCUSSION**

Numerous techniques have been described for the repair of cleft palate. All these aim at anatomical closure of the palate, avoiding fistulas, providing a competent velum for adequate speech, and allowing harmonious facial growth. Refinements in the basic principles of repair and attention to anatomical restoration of the abnormally directed
velar muscles have improved the functional outcomes over the years. Currently, the techniques commonly adopted by cleft surgeons for repair of the soft palate in Veau II and Veau III clefts include straight-line closure combined with intravelar veloplasty and double-opposing Z-palatoplasty. These techniques seem to have contributed to better speech outcomes. It has been reported that the Furlow method has been adopted by 42 percent of the surgeon members of the American Cleft Palate–Craniofacial Association as their first choice in primary palatoplasty.

There is equipoise among cleft surgeons as to whether the double-opposing Z-palatoplasty improves speech through lengthening of the palate, redirection and overlapping of levator muscles, and increasing the vascularity of musculomucosal flaps, and causing an increased incidence of oronasal fistula because of increased tension on mucosal closures in wide cleft palate. This systematic review was undertaken to compare the outcomes of the Furlow double-opposing Z-plasty and the straight-line intravelar veloplasty techniques. Data were controlled for confounding variables such as severity of the cleft and age at repair, as previous studies have demonstrated that persistent velopharyngeal insufficiency was associated with Veau classification and the patient’s age at palatoplasty.

This systematic review revealed that in patients affected by unilateral cleft lip–cleft palate, straight-line repair combined with intravelar veloplasty was associated with an increased risk of a secondary operation (1.64 times) compared with the Furlow group (95 percent CI, 1.05 to 2.58; \( p = 0.030 \)). The increased odds of the primary outcomes (1.94 times) neared statistical significance in the isolated cleft palate group (95 percent CI, 0.96 to 3.91; \( p = 0.065 \)).

The importance of palatal length is often emphasized in performing palatoplasty. Randall et al. showed that patients with longer palates had statistically better speech outcomes compared with patients with shorter palates. However, palatal length is not the only variable necessary to achieve normal closure of the velopharyngeal port. Salyer et al. noted that complete tensionless closure of the entire palate and construction of an adequately functioning soft palate with a muscle...
sling, at an early age, were even more crucial than the length of the palate.\textsuperscript{18} Two techniques allow reorientation of the velar muscles: intravelar veloplasty and double-opposing Z-plasty. Compared with the original method described by Kriens,\textsuperscript{1} Furlow double-opposing Z-plasty is believed to create a tighter pharyngeal sphincter because the levator veli palatini muscles are fully mobilized each time, irrespective of the operator, and are substantially overlapped within the posterior soft palate. This technique is in direct opposition to that of the midline suture repair in intravelar veloplasty.\textsuperscript{11} Furthermore, preserving an intact muscularomucosal layer in each flap optimizes vascularity and thereby decreases postoperative scarring and subsequent stiffness of the soft palate. Jackson et al.\textsuperscript{11} published the largest series on speech outcomes following the Furlow palatoplasty technique. According to their findings, 72.5 percent of patients had a competent to borderline competent velopharyngeal mechanism, 85.3 percent demonstrated no or inaudible nasal emission, and 8.1 percent of all patients underwent secondary surgery to correct velopharyngeal dysfunction.

Intravelar veloplasty is a widely used technique, although there is much variability among surgeons in how the musculature is dissected and repositioned. The large variations in the outcomes of intravelar veloplasty could be secondary to improper identification, mishandling, or incomplete posterior repositioning of the levator veli palatini muscles.\textsuperscript{25} Furthermore, the physiologic length of the muscle fibers (tension at which the muscles optimally contract) could be lost when the muscles are sutured in the midline if done so at the incorrect tension. Thus, intravelar veloplasty is more operator dependent than the double-opposing Z-palatoplasty. Sommerlad\textsuperscript{26} reported on 442 palate repairs performed between 1978 and 1992, with at least 10 years of follow-up. He observed a decrease in the rate of secondary surgery for velopharyngeal
insufficiency from 10.2 percent to 4.6 percent and attributed the change to an evolution in technique with increasingly radical surgery and greater operator experience. Similar results have been reported by Cutting et al.,27 with rates of secondary surgical intervention of 6.3 percent using straight-line repair and levator muscle repositioning. Andrades et al.28 demonstrated a significantly lower rate of secondary palate surgery for velopharyngeal insufficiency (6.7 percent compared with 29 percent; \( p = 0.008 \)) after the introduction of radical intravelar veloplasty. These reports showed that the extent of retention of the levator muscles achieved with intravelar veloplasty affects velopharyngeal function, whereas incomplete mobilization of the muscle is associated with less favorable speech outcomes.29

The overall incidence of palatal fistulae following palatoplasty has been reported to range from 0 to 34 percent of patients.29,30 The occurrence of fistula not only can cause regurgitation of food but also can contribute to symptoms of audible nasal air escape during speech and the perception of hypernasality,3 compromising speech. Cleft width, experience of the surgeon, closure under tension, mucosal tearing, large tissue displacement and creation of dead spaces, and postoperative hemorrhage and tissue necrosis are all risk factors for postpalatoplasty fistulae.30

Criticisms of the Furlow method have included higher fistula rates when double-opposing Z-plasty is used in wider clefts and without relaxing incisions.31,32 This systematic review did not show evidence of a statistically significant difference in fistula rate between the Furlow and straight-line repair (7.87 and 9.81 percent, respectively). However, in each group, the fistula rate was significantly higher in complete clefts compared with incomplete clefts. Because of a lack of detail in describing the fistulae in most of the studies included in this review, it was not possible to compare the location of fistulae between the two groups in an

Fig. 3. Event rates for unilateral cleft lip–cleft palate. Event rate is measured by the number of secondary operations for patients with unilateral cleft lip–cleft palate. The percentage of secondary operations (failure rate = secondary operations/total primary operations) is shown on the x axis. On the y axis, the name of the first author for each of the articles selected for the systematic review is shown. Dark squares represent Furlow failure rates and circles represent straight-line intravelar veloplasty failure rates. The straight-line intravelar veloplasty and Furlow averages are calculated and depicted with a dashed vertical line. The mean failure rate for straight-line closure is calculated as 17.1 percent and that for the Furlow method is 11.1 percent.
attempt to delineate any correlation with the technique. However, in the Furlow group, most of the fistulae were located in the anterior hard palate, followed by junctional fistulae. This finding is similar to the distribution reported by Smith et al., in a retrospective review of 611 cleft palates repaired by different techniques. This could support our observation that the occurrence of fistulae correlated more with the width of the cleft (Veau classification) than with the repair technique.

There are limitations to the prior conclusions. The wide amount of variability between studies prevented use of traditional meta-analysis. A prior systematic review concluded that the variety in age of treatment, cleft types, short follow-up, and strategy of repair among studies posed difficulties for objective comparisons. Based on these findings, we established our exclusion criteria to limit the variability among study samples and select a population of patients whose ages and protocols would resemble the globally accepted standard of care for cleft palate repair. Although the enforced exclusion criteria attempted to limit large variability between studies, there are other inherent differences that might affect the outcomes but that are seldom reported (such as cleft width) or that would be difficult to measure or eliminate completely in any systematic review, including the surgeon’s experience, subtle differences in application of the same technique by different operators, and medical comorbidities. The search strategy used by the authors yielded a small amount of articles for review, and the sensitivity analysis demonstrated that the study by Jackson et al. appeared highly influential in determining outcomes, given the small number of patients in the remaining studies. We are aware that the inclusion of this study strongly influenced the significance of our findings; however, we believe that the inclusion of this study also provided a large cohort of patients who were uniformly treated and evaluated with more reliable outcomes, which increased the power of the final analysis. Furthermore, the outcomes of the latter study were in agreement with the other studies considered in the same repair group.

Finally, it could be argued that the different hard palate repair techniques included in the straight-line group could have had a variable impact on the outcomes. Among these, only the pushback technique was devised to lengthen the palate and could have had a theoretical advantage over others. However, in a study comparing the palatal lengthening following Furlow double-opposing Z-plasty and pushback palatoplasty, Bae et al. proved that the former produced a significantly greater lengthening of the palate (16.2 percent and 8.4 percent of incomplete clefts, respectively). Furthermore, Witzel et al. proved that V-Y pushback techniques do not improve speech over other hard palate repairs. Thus, all straight-line repairs were considered in one group to increase the power of the statistical analysis.

CONCLUSIONS

This systematic review demonstrated that there is a scientific basis for improved outcomes in Furlow double-opposing Z-plasty compared with the straight-line intravelar veloplasty technique; however, the clinical relevance of the moderately increased risk in the latter group has yet to be determined. The review uncovered once again a pressing need for standardizing the methods of reporting outcomes to allow sound comparisons between results in different centers.

**REFERENCES**


Gaby Doumit, M.D., M.Sc.
Cleveland Clinic
9500 Euclid Avenue
Cleveland, Ohio 44195
doumitg@ccf.org


