

Critical Review: Effectiveness of Sphincter Pharyngoplasty Compared to Pharyngeal Flap Surgery in Reducing Hypernasality in Individuals with Velopharyngeal Insufficiency

Bianchi, L.C.

M.Cl.Sc. (SLP) Candidate

School of Communication Sciences and Disorders, U.W.O.

This critical review examines several studies to determine the effectiveness of sphincter pharyngoplasty (SP) compared to pharyngeal flap surgery (PF) in reducing hypernasality in individuals with velopharyngeal insufficiency (VPI). Overall, research indicates preliminary evidence that both surgeries produce similar results in reducing hypernasality. However, some questions arose as to the strength of the research designs. Additionally, further research is recommended in a variety of areas.

Introduction

Velopharyngeal insufficiency (VPI) occurs when there is a deficit in the closure of the velopharyngeal port (i.e. the junction of the velum and the lateral and posterior pharyngeal walls). This may occur from inadequate movement or reduced length of the velum, resulting in a gap between the velum and the pharyngeal walls (Conley, Gosain, Marks, & Larson, 1997; Willging, 1999). Adequate closure is necessary for normal speech intelligibility, resonance and deglutition (Conley et al., 1997). VPI can result in articulatory compensations (e.g. pharyngeal fricatives, glottal stops), reduced speech intelligibility, hypernasality, nasal air emission and nasal regurgitation (Conley et al., 1997; Willging, 1999).

VPI is typically associated with various genetic conditions and syndromes that result in structural abnormalities (Willging, 2003). This includes velocardiofacial syndrome, autosomal dominant genetic inheritance, submucous cleft palate, and cleft lip and palate (Willging, 2003). VPI can also result from a short palate or a deep pharynx (Willging, 1999). Removal of the adenoids can cause transient VPI. Occasionally, symptoms persist and intervention is required. Adenoidectomies are a particular concern for patients with palatal repairs or evidence of submucous clefting (Willging, 2003). VPI may also result from innervation damage due to congenital or acquired neurological deficit, or muscular degeneration from progressive neuromuscular disorders (Willging, 1999). Furthermore, VPI may occur secondary to phoneme specific articulation errors, developmental delay and hearing loss (Willging, 1999).

Multiple treatment options are available for the correction of VPI, which include prosthetic devices (e.g. palatal lifts, obturators), pharyngeal

augmentation and surgical modification (Willging, 1999; Willging, 2003).

Pharyngeal flap (PF) surgery has been considered the standard surgical treatment for VPI. According to Muntz (2001) a “myomucosal flap is elevated from the posterior pharyngeal wall and then inserted into the palate. The resultant obturation of the velopharyngeal space is at the midline with two resultant lateral ports” (p. 208). Some complications of PF include persistent hypernasality, hyponasality, obstructive sleep symptoms and obstructive sleep apnea. Short term complications include nausea, vomiting, bleeding, and difficulty managing secretions (Willging, 2003).

The Jackson-Silverton sphincter pharyngoplasty (SP), a second surgical option for treating VPI, is an operation by which “bilateral superiorly based flaps are elevated from the posterior tonsillar pillars, including the palatopharyngeus muscles. These two flaps are then sutured together in the midline and attached to the undersurface of a superiorly based posterior pharyngeal flap” (Sloan, 2000; p. 118).

When considering outcome success of surgical approaches in the area of velopharyngeal insufficiency, two factors must be kept in mind. There is no standardized operational definition of outcome success. Many evaluators consider post-treatment hyponasality and mild hypernasality to be acceptable whereas other evaluators look for complete resolution (de Serres, L., Deleyiannis, F.W.B., Eblen, L.E., Gruss, J.S., Richardson, M.A., & Sie, K.C.Y., 1999). Secondly, speech-language pathologists use subjective measures when assessing VPI. Perceptual speech evaluations are used to determine resonance, nasal air escape and overall speech intelligibility. Nasometers, which measure airflow from the nasal and oral passages, also require interpretation (Willging, 1999; Willging, 2003).

While the aforementioned measures are subjective, they are considered to be the standard measures of evaluating VPI and therefore they are used across all VPI assessments (Willging, 1999;

Willing, 2003). Additionally, comparing papers that use different definitions of outcome success may assist researchers in developing a standardized definition for VPI correction. Therefore, despite these limitations, it is valuable to use these measures to continue researching and comparing various surgical techniques to ensure continued use of best practice for resolving velopharyngeal insufficiency.

Objectives

One of the major complications of VPI is hypernasality, which is “nasal resonance perceived during normally nonnasal speech” (Conley et al., 1997; p. 42). Altered resonance can affect speech intelligibility and volume causing mild to severe speech distortion, which can have social and emotional consequences (de Buys Roessingh, A.S., Cherpillod, J., Trichet-Zbinden, C., & Hohlfeld, J., 2006). Over the years, pharyngeal flap (PF) surgery has been viewed as the “gold standard” in treating hypernasality when speech therapy has been unsuccessful (Husein, M., Chang, E., Cable, B., Karnell, M., Hynds Karnell, L., & Canady, J.W., 2004). More recently sphincter pharyngoplasty (SP) has been established as a surgical option following concerns of postoperative complications with PF; however, considerable uncertainty exists as to which is the best approach (Husein et al., 2004; Sloan, 2000; VPI Surgical Trial Group, 2005). Therefore, the primary objective of this review is to critically evaluate the literature to compare the effectiveness of pharyngeal flap surgery and sphincter pharyngoplasty in reducing hypernasality in patients with velopharyngeal insufficiency.

Methods

Search Strategy

The following computerized data bases were searched: CINAHL, Medline, Cochrane, PsychInfo, Scholars Portal, ComDisDome, OVID, and SCOPUS.

Keywords used include:

(pharyngeal flap surgery) AND (hypernasality), (pharyngeal flap surgery), (velopharyngeal insufficiency), (surgical flaps), (hypernasality)

The search was limited to articles written in English between 1995 and present.

Selection Criteria

Studies were included in the critical review if they compared the effectiveness of PF and SP in

reducing hypernasality. This included studies that only reported results relating to the overall reduction of VPI, as adequate nasality was designated in all studies as a measure of VPI correction.

Data Collection

Results of the literature search yielded the following: retrospective mixed-design studies (2), and randomized control trials (3). Two of the RCT studies will be analyzed jointly as one is an expanded version of the original study. Five informational articles were used for background information.

Results

Retrospective Mixed-Design Studies

de Serres et al. (1999) sought to examine outcome results of SP as an alternative surgical approach to treating VPI. A retrospective chart review was conducted on fifty patients who underwent either PF or SP between January 1990 and December 1995. Thirty-four patients were identified as having both pre- and post-operative measures (PF=18; SP=16). Several outcome measures were analyzed including a perceptual speech analysis of nasal resonance. Post-operative results for change in nasal resonance were not directly reported in the study; however, overall VPI resolution encompassed normal nasal resonance as an outcome measure. The results indicated VPI resolution for SP was 50% compared to a 22.2% resolution for PF; the differences between the results were not statistically significant. Resulting hyponasality was also deemed to be not significantly different between the two surgical procedures (18.8% for SP versus 22.2% for PF). In terms of complications, obstructive sleep symptoms were observed equally in patients from both treatment groups, however, cases of obstructive sleep apnea only developed in patients who had undergone PF surgery. Despite the trends in the data, no specific conclusions were reported by the researchers. However, they suggested that decisions should be made based on the surgery with the most reliable results and lowest risk of complications and their own practice has ceased to use PF surgery.

Seagle, M.B., Mazaheri, M.K., Dixon-Wood, V.L., & Williams, W.N. (2002) used a retrospective chart review to examine the results of four surgical approaches (PF, SP, Furlow palatoplasty, and palatal pushback with pharyngeal flap) that had been individualized to the patients according to perceptual and videofluoroscopic assessments. Ninety-eight patients were identified as having undergone one of the four procedures between 1988 and 2000; however

only seventy-five had both pre- and post-operative assessment results. The measures used were the Bzoch base-10 index of Nasal Air Emission (NAE), Bzoch Cul-de-Sac Test, Bzoch Error Pattern Articulation Test, multiview videofluoroscopy and nasendoscopy. Outcome criteria for success was considered to be a score of 2/10 or less for NAE and hypernasality, and a score of 4/10 or less for hyponasality. The results indicated a 95.8% success rate for the 24 patients (21 were primary surgeries, 3 were secondary surgeries) who had undergone SP, 8% had resulting hyponasality. A 90.9% success rate was reported for the 11 patients who had undergone PF, 18.2% had resulting hyponasality. The statistical significance of the results was not reported. Chart reviews indicated that neither surgery resulted in obstructive sleep apnea. The researchers concluded that the success rates among the surgeries were similar. They regarded this as tentative support for individualizing a particular surgical approach based on thorough pre-operative assessment.

Randomized Control Trial Studies

The VPI Surgical Trial Group (2005) undertook a multi-centre randomized trial to investigate effectiveness and risk of PF and SP. Based on detailed inclusion criteria related to general development and diagnosis, 103 patients from a referred group of 144 were randomized into two treatment groups. A sample size of 97 was available for pre- and post-operative analysis. The surgical procedures were standardized between the surgeons from five treatment centres. Several outcome measures were examined including resonance. Resonance was evaluated along a four-point scale (normal, mild, moderate, severe) based on a speech sample collected using standard protocol between the centres. Results were compared using a Chi-Square test. At three months post-operative SP was 42% successful while PF was 82% successful at reducing hypernasality. This difference was statistically significant ($p < 0.01$). The patients were assessed again 12 months post-operative. The researchers analyzed these results along two outcome measures. When outcome measures designated resulting hyponasality as acceptable, SP was 78% successful and PF was 83% successful in reducing hypernasality. When outcome measures designated resulting hyponasality as an indication of unsuccessful treatment, SP was 76% successful and PF was 81% successful in reducing hypernasality. Therefore, by 12 months post-operative, neither analysis demonstrated a significant difference between the two surgeries (p -values = 0.45 and 0.81 respectively). The data of a subgroup of patients

($n=19$: SP=9, PF=10), who were classified as having severe pre-operative hypernasality, were also analyzed. Post-operative elimination of hypernasality was 33% successful for SP and 50% successful for PF. No statistical difference was noted between the two surgical approaches (p value = 0.72). The researchers concluded that both SP and PF are equally successful at eliminating hypernasality, however, the results from this study suggest PF may produce faster results.

Ysunza, A., Carmen Pamplona, M., Ramirez, E., Molina, F., Mendoza, M., & Silva, A. (2002) and Ysunza, A., Carmen Pamplona, M., Molina, I., Drucker, M., Felemovicius, J., Ramirez, E., & Patina, C. (2004) will be reported together. The 2004 study is an expanded version of the 2002 study employing the same research design and patient population. In the 2002 study 359 patients were evaluated between January 1995 and December 2000 using a perceptual speech evaluation, videonasopharyngoscopy (VNP) and multiview videofluoroscopy (MVVF). Fifty patients met criteria and were randomized into one of the two surgical treatments. The 2004 study evaluated an additional 200 patients from December 2000 to December 2002, which resulted in an increased sample size of 70 patients. In both studies a percentage of the participants received speech therapy prior to surgical intervention to reduce compensatory articulation. All patients were reassessed four months post-operative with a perceptual speech evaluation and VNP and MVVF. Specific results were not reported for hypernasality; however, criteria for VPI correction included normal resonance. Therefore, results indicating the persistence or elimination of VPI, which were analyzed using a Fischer's exact t-test, were examined for the purpose of this critical review. Results of the 2002 paper indicated residual VPI in 3/25 patients for PF and 4/25 patients for SP. The difference between the two surgical procedures was not statistically significant ($p > 0.05$). Similarly, results from the expanded 2004 study indicated VPI correction in 89% of the PF treatment group and 85% of the SP treatment group, the results were also not statistically significant ($p > 0.05$). Data on resulting hyponasality was not provided. Complications were not observed for any of the patients in either study. The researchers suggested that PF and SP demonstrate similar effective results in treating VPI.

Critical Appraisal of Research

Subject selection and sample size

Among the studies reviewed there were several concerns related to subject selection and sample size.

The sample sizes were small, which affected the ability of each study to detect a significant difference between the two surgical treatments and limited the researcher's capacity to make definite conclusions. The VPI Surgical Trial Group (2005) calculated a sample size required to achieve 80% power; however, they were limited by the number of available patients and were unable to achieve a sufficient sample. None of the other studies reported the level of power.

The VPI Surgical Trial Group (2005) selected patients who had undergone primary palatal repair from each of the five participating clinics. Ysunza et al. (2002 & 2004) evaluated patients at a cleft palate clinic who had a repaired unilateral cleft lip and palate. In both studies inclusion/exclusion criteria were used to help eliminate a priori differences between the groups, strengthening the comparison of outcome results between the two surgeries. However, generalization of evidence to the entire VPI population may be limited by the homogeneity of the sample groups. The retrospective studies were less able to control for extraneous variables. In terms of patient selection, patients' charts were reviewed if they had had one of the two surgical treatment options during the specified period and both pre-and post-operative assessment results were available. Seagle et al. (2002) additionally excluded patients with uncorrected submucous clefts and de Serres et al. (1999) excluded patients who had undergone concomitant speech procedures as both of these factors could inherently alter outcome results.

Method

Seagle et al. (2002) and de Serres et al. (1999) were both retrospective studies which inherently weakens the evidence reported. Retrospective studies limit the variables that a researcher can control and may lack important information such as patient characteristics, treatment protocol, and outcome measures. Furthermore, randomization of patients was not possible. Many of the patients had been matched to a treatment approach based on certain anatomical characteristics, which may have introduced confounding variables that could have affected the outcome of success for each surgical treatment. Post-operative evaluation may have been susceptible to bias as information regarding blinding of the evaluators was not reported.

Ysunza et al. (2002, 2004) and the VPI Surgical Trial Group (2005) randomized their patients to a surgical treatment using a block randomization design. Ysunza et al. (2002, 2004) balanced the group sizes every twenty patients, while the VPI Surgical Trial Group balanced group sizes following every six

patients allocated. This allowed for equal participants in each surgical treatment group if a study was stopped earlier than the proposed timeline. Evaluators for the three RCT studies were blinded for the perceptual speech analysis and all three studies attempted to blind the surgeons in some manner.

The reliability of the surgical procedures is an issue of contention in all of the studies except for the VPI Surgical Trial Group (2005). In this study the surgical procedures were standardized across all surgeons involved which would strengthen any comparisons the researchers made. The lack of standardized procedures for the remaining studies, as well as the lengthy timeline of the study (Seagle et al., 2002; de Serres et al., 1999; Ysunza et al., 2002 & 2004) weakens the evidence. Modifications to the surgical procedure between patients in an individual study reduces the ability to properly compare outcome results. In addition, a lengthy time span (Seagle et al., 2002, Ysunza et al., 2004) allows for surgical advancement, which may lessen the ability of the researchers to compare outcome results from the beginning and the end of the study. Further procedural confounds were present in the Ysunza et al. studies. Approximately half of the patients were required to undergo speech therapy prior to surgical treatment in order to reduce compensatory articulations. The rationale for this approach, to allow for greater speech success following surgery, was valid. However, data indicated that reductions in gap size prior to surgery were noted in those patients who received speech therapy. Therefore, the additional treatment may affect outcome results if these patients were not equally distributed between the two surgical treatment groups as milder pre-operative VPI generally results in greater success for restored speech functioning (de Serres et al., 1999).

Most of the studies reported on a sample size that was representative from the beginning of the study to the end. The VPI Surgical Trial Group (2005) was the only study that had patients drop out or excluded following the initial analysis, however, they delineated the numbers and reasons for attrition. The retrospective study by Seagle et al. (2002) reported on four surgical treatments with unequal group sizes. In addition, a number of the patients had undergone secondary procedures that differed from their initial procedure. The analysis of results included both the primary procedures and secondary procedures, which would affect outcome measures as some patients were reported on twice. While these factors were not controllable as it was a retrospective design, they weaken the evidence presented by this study.

Measurement tools and outcomes

All of the studies included an operational definition of successful VPI correction. de Serres et al. (1999), and Ysunza et al. (2002, 2004) concluded that residual hypernasality and resulting hyponasality were indications of unsuccessful treatment. VPI Surgical Trial Group recognized that various studies define successful VPI correction differently and analyzed results along two definitions of VPI correction, one did not accept residual hypernasality or resulting hyponasality while the other classified patients with resulting hyponasality as successful treatments. Seagle et al. (2002) used a more lenient operational definition of successful VPI correction accepting mild hypernasality and mild to moderate hyponasality in their success rates. Unfortunately in the area of VPI there is no agreed upon definition of successful resolution of VPI. This limits the ability to gain a true understanding of the success rates of these surgical procedures as well as the comparability between research studies.

In all of the studies, perceptual speech evaluations to assess nasal resonance were conducted by speech-language pathologists who specialized in the area of cleft palate. Inter-rater reliability for perceptual speech analysis was reported as adequate for all of the RCT studies. This strengthens the reliability of these studies. While the use of subjective measures reduces the reliability of evidence, perceptual speech evaluations are used to determine the initial diagnosis of VPI and therefore, consistency in measurement tools can be considered a strength. Furthermore, the nature of resonance and the lack of available objective measures are a limitation to this area of speech language pathology. As such the researchers have used what is considered the current gold standard in assessment. The use of highly experienced evaluators and the reporting of high inter-rater reliability further increase confidence that accurate evaluations were conducted. Inter-rater reliability was not noted for either of the retrospective studies, which weakens the reliability of these studies and the applicability of the evidence.

Statistical analysis

Several limitations were noted in terms of statistical analysis among all of the studies. Despite the measurement of pre- and post-operative resonance in all of the papers analyzed, statistical significance for within-group changes was not reported. In addition, de Serres et al., 1999, Ysunza et al., 2002, and Ysunza et al., 2004 failed to report post-operative resonance results. Therefore, readers were left to draw conclusions on the reduction of hypernasality based on overall results of successful

VPI correction. However, all of the studies except for Ysunza et al. (2002 & 2004) reported resulting incidents of hyponasality.

Ysunza et al., (2002 & 2004) and VPI Surgical Trial Group (2005) used appropriate statistical measures, Fisher's exact test and Chi-Square test respectively, to analyze outcome results between the two surgical procedures. In addition, Ysunza et al. (2002, 2004) and VPI Surgical Trial Group (2005) reported p-values. This allows for a more complete analysis of results, which lends to a stronger basis of evidence. Seagle et al. (2002) and de Serres et al. (1999) did not indicate what statistical measures were used to analyze results and what p-values were obtained. This makes it difficult to ascertain any confidence in these results or draw conclusions on differences between the surgical treatment approaches for successfully eliminating hypernasality. None of the studies reported effect size or confidence intervals which limits the ability to critically appraise the evidence.

Recommendations

Pharyngeal flap surgery has typically been considered the gold standard in correcting hypernasality related to VPI. From the above critical analysis, research evidence demonstrates preliminary evidence that sphincter pharyngoplasty is an equally effective surgical option in reducing hypernasality in individuals with velopharyngeal insufficiency. However, further research employing more solid research designs (e.g. greater sample size, comparison between pre- and post-operation measures) is necessary.

Furthermore, there are many variables that contribute to the overall correction of velopharyngeal insufficiency that are overlooked when considering only hypernasality. Surface evaluation of resulting hyponasality, resulting speech intelligibility, reduction of gap size, and incidence of complications suggest that there are similar results between the two surgeries for these additional variables. However, preliminary research by de Serres et al. (1999) has demonstrated some evidence that individuals who initially present with obstructive sleep symptoms have a higher rate of developing obstructive sleep apnea following surgical intervention if they have undergone pharyngeal flap surgery. Therefore, a more in depth, holistic, critical analysis of VPI correction would perhaps develop a clearer picture as to whether one approach is more advantageous.

In addition, further research is necessary to examine the possible benefits of individualizing surgical procedures. Technological advancements have allowed for the implementation of

videofluoroscopic (VF) and nasoendoscopic (NE) evaluations to identify the functioning of the velopharyngeal (VP) sphincter, including movement of the lateral and posterior pharyngeal walls, and the size and location of the VP closure deficit (Willging, 1999; Willging, 2003). There appears to be a divide among the research papers analyzed in this review on whether individual surgical planning based on patient characteristics visualized using NE and VF improves outcome success. Certain trends evident in de Serres et al., 1999, suggest that earlier determined anatomic criteria for selecting one surgical procedure over another may not be as relevant as previously thought for correcting VPI. In addition, studies that did not employ any individualization (VPI Surgical Trial Group, 2005) achieved comparable results. Alternately, those studies that supported the use of VF and NE (Seagle et al., 2002; Ysunza et al., 2002 & Ysunza et al., 2004) did not directly analyze these factors; instead they concluded that because the overall VPI correction results were positive that the use of these individualization techniques was beneficial. Therefore, the controversy around this topic implies that further research is necessary to verify the applicability of using technological advancements to individualize surgery according to specific anatomical and physiologic characteristics.

Additional research goals in the area of velopharyngeal insufficiency should include the development of a standardized definition of outcome success. Many of the studies use varying degrees of acceptance of resulting hyponasality and residual hypernasality that make comparison of results difficult and set different standards for surgical outcome success. Furthermore, the field of speech-language pathology would benefit from the development of a more objective measure of evaluation for hypernasality.

Conclusion

While the research papers examined in this review contain some limiting factors as discussed in the critical analysis, the evidence is suggestive that both pharyngeal flap surgery and sphincter pharyngoplasty are equitable surgical approaches to reducing hypernasality in individuals with velopharyngeal insufficiency. However, additional research in the areas discussed would further the field of speech language pathology as well as continue to advance the ability of craniofacial surgeons to improve the outcomes of surgical intervention for treating velopharyngeal insufficiency.

References

- Conley, S.F., Gosain, A.K., Marks, S.M., & Larson, D.L. (1997). Identification and assessment of velopharyngeal insufficiency. *American Journal of Otolaryngology*, 18(1): 38-46.
- de Buys Roessingh, A.S., Cherpillod, J., Trichet-Zbinden, C., & Hohlfeld, J. (2006). Speech outcome after cranial-based pharyngeal flap in children born with total cleft, cleft palate, or primary velopharyngeal insufficiency. *Journal of Oral Maxillofacial Surgery*, 64:1736-1742.
- de Serres, L., Deleyiannis, F.W.B., Eblen, L.E., Gruss, J.S., Richardson, M.A., & Sie, K.C.Y. (1999). Results with sphincter pharyngoplasty and pharyngeal flap. *International Journal of Pediatric Otorhinolaryngology*, 48: 17-25.
- Husein, M., Chang, E., Cable, B., Karnell, M., Hynds Karnell, L., & Canady, J.W. (2004). Outcomes for children with submucous cleft palate and velopharyngeal insufficiency. *The Journal of Otolaryngology*, 33(4): 222-226.
- Muntz, H.R. (2001). Current techniques for treatment of velopharyngeal insufficiency. *Operative Techniques in Otolaryngology-Head and Neck Surgery*, 12(4): 204-209.
- Seagle, M.B., Mazaheri, M.K., Dixon-Wood, V.L., & Williams, W.N. (2002). Evaluation and treatment of velopharyngeal insufficiency: The University of Florida experience. *Annals of Plastic Surgery*, 48(5): 464-470.
- Sloan, G.M. (2000). Posterior pharyngeal flap and sphincter pharyngoplasty: The state of the art. *Cleft Palate-Craniofacial Journal*, 37 (2): 112-122.
- VPI Surgical Trial Group. (2005). Pharyngeal flap and sphincterplasty for velopharyngeal insufficiency have equal outcome at 1 year postoperatively: Results of a randomized trial. *Cleft Palate-Craniofacial Journal*, 42(5): 501-571.
- Willging, J.P. (2003). Velopharyngeal insufficiency. *Current Opinion in Otolaryngology and Head and Neck Surgery*, 11: 452-455.

Willging, J.P. (1999). Velopharyngeal insufficiency. *International Journal of Pediatric Otorhinolaryngology*, 49(1): S307-S309.

Ysunza, A., Carmen Pamplona, M., Ramirez, E., Molina, F., Mendoza, M., & Silva, A. (2002). Velopharyngeal surgery: A prospective randomized study of pharyngeal flaps and sphincter pharyngoplasties. *Plastics and Reconstructive Surgery*, 110 (6): 1401-1407.

Ysunza, A., Carmen Pamplona, M., Molina, I., Drucker, M., Felemovicius, J., Ramirez, E., & Patina, C. (2004). Surgery for speech in cleft palate patients. *International Journal of Pediatric Otorhinolaryngology*, 68: 1499-1505.