8-2009

Velopharyngeal Insufficiency: Diagnosis and Management

Robert J. Shprintzen
Sacred Heart University, shprintzenr@sacredheart.edu

Eileen Marrinan

Follow this and additional works at: http://digitalcommons.sacredheart.edu/speech_fac

Part of the Speech Pathology and Audiology Commons

Recommended Citation
Velopharyngeal insufficiency: diagnosis and management

Robert J. Shprintzen, Ph.D. and
Director, Communication Disorder Unit, Director, Velo-Cardio-Facial Syndrome International Center, Director, Center for Genetic Communicative Impairment, Professor of Otolaryngology and Communication Science, Professor of Pediatrics

Eileen Marrinan, M.S., M.P.H.
Director, Central New York Cleft and Craniofacial Center, Department of Otolaryngology and Communication Science

Abstract

Purpose of Review—Journal articles relevant to the diagnosis and treatment of velopharyngeal insufficiency (VPI) were reviewed. All articles ascertained by PubMed search were included.

Recent Findings—Articles reported on the application of magnetic resonance scanning, reliability tests of the International Working Group diagnostic protocol, the use of nasometry, and techniques designed to assess the function of the velopharyngeal mechanism. Treatment papers focused on outcomes in small samples of cases and complication rates from pharyngeal flap. One paper discussed ineffective speech therapy procedures.

Summary—There were relatively few papers this past year. Those that were published were hindered by small and heterogeneous sample sizes, and occasionally by inappropriate methods for assessing outcomes. None of the findings will have a major impact on the current state-of-the-art for diagnosis of VPI. The speech therapy paper has a very important message that should be taken to heart by all clinicians involved in the management of children with clefts and craniofacial disorders.

Keywords

Velopharyngeal insufficiency; palate; pharynx; speech therapy; hypernasality

INTRODUCTION

The history of the recognition of velopharyngeal insufficiency (VPI) as a disorder of speech production that can be diagnosed and treated is a long one, and considerable literature has been devoted to it. Last year in this journal, Rudnick and Sie1 did an excellent job of reviewing the literature, covering many topics of relevance. Rather than repeat a job already well done, this article will take a critical look at the state of the art in 2009 and how it relates to current practices. We will cover issues that relate to VPI as a specific condition rather than as an issue related to primary palatoplasty and its outcomes. It is interesting to note that the overall body of relevant literature this past year has been sparse. A PubMed search found fewer than 25 journal articles of direct relevance. There has been an increase in the number of articles citing VPI as a symptom of velo-cardio-facial syndrome (VCFS), but in most cases these descriptions are incidental to the primary purpose of the paper.

Contact Information: Robert J. Shprintzen, Ph.D., VCFS International Center, Upstate Medical University, 725 Irving Avenue, Suite 504, Syracuse, NY 13210 USA; Telephone: +315-464-6596, Fax: 315-464-6598, shprintr@upstate.edu.
Probably the first article to report on the mechanism of velopharyngeal closure was that of Hilton in 1836\textsuperscript{2}. Following the spontaneous separation of a maxillary tumor in an adult, Hilton illuminated the pharynx and described the movements of the soft palate and lateral pharyngeal walls. Although a single subject report, the observations were consistent with what scientists today recognize as the physiology of velopharyngeal closure. However, these observations were largely ignored until late in the 20\textsuperscript{th} century when advanced diagnostic techniques rediscovered the physiology of velopharyngeal closure. In part, this is because of the advent of ionizing radiation used to diagnose disorders of the head and neck. Probably the first published report of the use of a lateral neck radiograph to assess velopharyngeal function came from Scheier in 1909\textsuperscript{3}. The technique provided a view of the velopharyngeal mechanism that made sense to observers and was therefore adopted as being the gold standard for assessing hypernasal speech caused by VPI. Although the movements of the velum and posterior pharynx could be seen, it was a two-dimensional representation of a three-dimensional tube. The three-dimensional nature of velopharyngeal closure was not recognized until the publications of Skolnick who developed the technique for multi-view videofluoroscopy\textsuperscript{4–6}. Contemporaneous to Skolnick’s publications were those of Pigott who introduced the use of nasal endoscopy to assess the velopharyngeal mechanism from above\textsuperscript{7,8}. These landmark publications led to the widespread application of these procedures. The introduction of improved thinner fiberoptic endoscopes that could be applied without discomfort to pediatric cases made nasopharyngoscopy more commonplace; they could be performed in an office setting without a radiologist and fluoroscopic equipment. Nasopharyngoscopy became the gold standard for direct assessment of VPI. Although it has been found that nasopharyngoscopy and multi-view videofluoroscopy are not in perfect agreement when applied to the same patient\textsuperscript{9}, relatively few centers employ both techniques, preferring to rely on nasopharyngoscopy alone.

The use of indirect measures of VPI have gained popularity in speech clinics and universities because they measure the acoustic properties of hypernasality. The most commonly used technique is nasometry\textsuperscript{10}. Nasometry measures an acoustic phenomenon known as nasalance, a ratio of oral and nasal resonance measured from normally non-nasal speech samples. Nasometry is often used as an outcome measure for assessing treatment results, but is it truly a more valid outcome than listening and looking?

Nearly all of the treatment issues discussed in the literature relate to surgery. In years past, new operative approaches to the management of VPI were introduced frequently. Over this past year, there were really no distinctly different techniques reported. Most of the surgical articles were reports of outcomes, complications, or variations on previously reported procedures.

**Diagnostic Approaches to VPI**

In 2008, journal articles discussing direct observation of the velopharyngeal valve were limited to two reports on nasopharyngoscopy\textsuperscript{11,12} and two utilizing magnetic resonance imaging\textsuperscript{13,14}.

**Nasopharyngoscopy**—The two articles focused on the use of nasopharyngoscopy both addressed reliability and validity of the procedure. Pegoraro-Krook et al.\textsuperscript{11} described “diagnostic therapy.” A sample of 10 children with repaired unilateral cleft lip-palate who had VPI and normal articulation were assessed with nasopharyngoscopy. A speech sample was obtained and the video recorded. The subject was then given instructions to perform a maneuver referred to as “diagnostic therapy.” The procedure involved teaching the child to impound air in the oral cavity, positioning the tongue between the lips and puffing the cheeks out while maintaining pressure orally, then maneuvering the production of a /p/ sound. This technique is actually a minor modification of the “tongue anchor technique”\textsuperscript{15}. The movements in the velopharyngeal valve during increased intraoral breath pressure (during “diagnostic therapy”)
were compared to the movements during the speech productions. The authors reported that they followed the protocol of the International Working Group recommendations for assessing VPI. They found increased movement in the velopharyngeal mechanism during the therapy condition, particularly in the lateral pharyngeal walls, including two subjects who eliminated VPI. This observation is consistent with the hypothesis that increasing oral pressure demands can improve velopharyngeal during speech. The authors concluded that the diagnostic value of nasopharyngoscopy can be enhanced by performing the task described and that treatment outcomes would be affected if it was omitted. The question is if the technique is the only way of obtaining this information, or if it can be obtained by less complicated means. According to the authors, training during the endoscopic sessions required an additional five minutes. Although many children will be able to comply, many young children may not. If compliance is poor, will this type of information be lost? The answer may be found in the failure to follow the actual recommendations of the International Working Group that they said was implemented. The authors assessed the differences in movement using the syllable /pa/ applying the ratio rating scale described by the International Working Group. Their rationale was that this syllable requires air to be impounded in the oral cavity if it is to be produced correctly. However, that does not mean that the production of this syllable will produce maximum movement in the velopharyngeal valve. The International Working Group recommended using a standardized speech sample that included every sound in the speaker’s native language, including at least two sustained fricative sounds, such as /ssssss/ or /ffffff/. The reason for this recommendation is that experienced clinicians recognize that movement in the velopharyngeal valve often varies for different sounds, and maximum movement can be seen on sustained fricatives because the pressure demands are high. However, this study did not determine if the movement seen on sustained /ssssss/ or /ffffff/ was the same as that seen for the “diagnostic therapy.” It is far easier to obtain these sounds during the examination than performing the maneuver suggested by the authors. The International Working Group also recommended that assessment of the velopharyngeal mechanism should include both nasopharyngoscopy and multi-view videofluoroscopy in at least frontal and lateral views. Therefore, although the conclusions drawn by Pegoraro-Krook et al. are probably accurate, the methodology implemented to reach that decision cannot actually support it completely.

The other article reporting nasopharyngoscopic data was that of Sie et al who addressed the inter- and intrarater reliability of the International Working Group protocol. Nasopharyngoscopic studies of VPI from a single institution were distributed to a total of eight institutions with each having two otolaryngologists who were to assess the outcomes of the study according to the International Working Group rating system. The 16 reviewers rated each study twice, one month apart, to assess intra-rater reliability. Inter-rater reliability was also assessed for all reviewers. Correlation coefficients were calculated from the ratings. The authors reported that intra-rater reliability was better than inter-rater reliability and that the “fair” correlation coefficient obtained for inter-rater reliability implied that clinicians from different centers would not be able to accurately communicate the outcome of endoscopic studies to colleagues at other institutions. The entire purpose of the International Working Group was to develop a system that would allow the communication of examination results in a standardized manner. The International Working Group did not perform validation or reliability studies, so studies like that of Sie et al. are needed to determine if the system should be retained or a new one sought. Sie et al. detailed a number of potential weaknesses in their study including optical distortion of the endoscopes, the inclusion of the audio portion of the exam that the authors felt might influence the reviewers, and recall effect for the intra-rater study. Only the last point is actually a liability and may have been responsible for making intra-rater reliability better than inter-rater reliability. All reviewers watched the same video, so even if distortion were present, it would the same for all. The inclusion of audio is essential to the International Working Group protocol because movement must be rated for to the sound produced. Other criticisms could have been discussed but were not. Within the 10 centers
chosen, only otolaryngologists were chosen to participate. No other specialists were involved, and it was not revealed if those institutions relied on the otolaryngologists for those assessments. In order to account for experience, the authors rank-ordered the number of VPI assessments they had done in the previous year. Of the 20 otolaryngologists, four had done fewer than 4 VPI evaluations in the past year, eight had done fewer than 10 and only three had done more than 20 in one year. Although the authors claimed that experience did not affect the, it could be successfully argued that performing less than 20 evaluations in a year, or even doing less than 40 or 50 in a year, cannot be construed as “experienced.” In other words, the investigators may simply have been comparing degrees of inexperience for at least 17 out of the 20 reviewers. The members of the International Working Group who devised the scale were all highly experienced clinicians and the outcome of their work may have been assessed by a group of inexperienced clinicians. It would have been interesting to add a training period for all of the participants to see if that would improve both inter- and intra-rater reliability.

**Magnetic Resonance Imaging (MRI)**—Two articles used MRI to assess VPI\(^{13,14}\). Both described “dynamic” assessment of the velopharyngeal valve, a misuse of the word. Atik et al.\(^{13}\) performed MR scanning during two conditions, the production of a sustained /mmm/ sound and the sustained production of the /ssss/ sound. Acquisition of the MR data using a 1.5T scanner and a 0.8 second acquisition time yielded images in the lateral, frontal, and axial views. The procedure was performed on 32 cases, 16 with cleft palate and 16 controls. The authors stress the quality of the imaging of the soft tissue and the ability to define the anatomical landmarks. This same claim was made by Kao et al.\(^{14}\) who used a 3.0T scanner and surface coil to obtain images. They reported qualitative assessment of the levator veli palatini muscle and other soft tissues, and the ability to acquire images during “limited speech”\(^{14}\), two speech samples, a sustained e vowel (/i/) and sustained /nnnn/. Six subjects were scanned including a pre- and postoperative assessment of a case of submucous cleft palate who had Furlow palatoplasty. Kao et al.\(^{14}\) claimed that MRI was advantageous for assessing VPI because it was noninvasive (they arguably claimed that nasopharyngoscopy was an invasive procedure) and did not involve ionizing radiation. They also claimed that the major obstacle to the widespread use of MRI to assess VPI was poor image quality, a problem overcome by the use of a 3.0T scanner and proper technique.

Both articles make a number of significant errors. First, these are not dynamic studies. They are still images acquired nearly a second apart during two different tasks. It is not possible to know what the structures were doing at the precise moment of the acquisition of the image, nor is it possible to know if this represents typical, maximal, or minimal movement. MR cannot provide true dynamic information. Because the issue being studied is speech, for the diagnostic test to be valuable, it has to acquire images during normal speech production. There is no logical reason for Kao’s choice of speech sample. It has long been known that sustained vowel production does not require velopharyngeal closure\(^{17,18}\) and the consonant /n/ is a nasal consonant. The true reason why MR will not and should not be a front-line diagnostic procedure for assessing VPI is that connected speech and multiple speech contexts cannot be observed and the procedure is prohibitively expensive and complicated; nasopharyngoscopy and multi-view videofluoroscopy as advocated by the International Working Group are not. It is also true that endoscopy and fluoroscopy provide surgeons the information they require to plan operations for the correction of VPI. MR data may have a place in assessing the palate muscle size, distribution, and position in cases of submucous or occult submucous cleft, but it is not known if this has any value as a clinical tool.

**Acoustic Assesments**—Nasometry is used in many speech clinics to assess nasalance, a measure of resonance characteristics during speech. Three papers focus on research findings related to detectable nasal resonance during vowel production\(^{19,20}\). Lee et al.\(^{19}\) studied hypernasal resonance on vowels in speakers with cleft palate and VPI. Mandulak and

---

*Curr Opin Otolaryngol Head Neck Surg*. Author manuscript; available in PMC 2010 January 22.
Zajac\textsuperscript{20} studied nasalance in normal speakers for /i/ and /a/ vowels and the differences between male and female speakers. The implications for the perception of hypernasality were discussed but not assessed directly. Sweeney and Sell\textsuperscript{10} correlated nasalance with perceptual assessments of hypernasality. They reported a positive correlation coefficient ranging from 0.82 to 0.92. The question remains if the addition of nasalance scores in any way adds to perceptual judgments of nasality during normal speech production.

**Treatments for VPI**

In 2008, there were eight articles describing surgery for VPI and three articles describing complications. There was also one important article discussing speech therapy techniques commonly applied to the management of children with VPI.

**Surgical Management**—As is often the case with articles reporting surgical outcomes, sample sizes are small. There were 5 reports analyzing 10 or fewer cases\textsuperscript{21–25}. It is unclear how a recommendation or endorsement for a particular approach or a report of success can be determined with so small a sample. One would suppose that there is a learning curve for procedures so that early reports may not even be advantageous to those endorsing a procedure if better outcomes depend on experience. Additionally, the assessment of outcomes used techniques that were not necessarily adequate. For example, MR scans were utilized\textsuperscript{24} to visualize the pharynx postoperatively, and as discussed above, this is not a true assessment of speech. In fact, the phonation used for the MR scan was a vowel, /eeee/, a completely invalid assessment. A voice evaluation was included that measured vocal cord parameters (voice), an irrelevant assessment. No information was given about postoperative nasopharyngoscopy although it had been done preoperatively.

Three articles had larger sample sizes, including 17\textsuperscript{26}, 25\textsuperscript{27}, and 38 cases\textsuperscript{28}. Abdel-Azziz\textsuperscript{26} described a new procedure involving surgical dissection of the posterior faucial pillars and vertical segments of superior constrictor muscle bilaterally that were then inserted into incisions laterally in the soft palate in an attempt to increase velar motion based on hypothetical assumptions of synergistic and antagonistic contributions of the superior constrictor and palatopharyngeus. The results were reported in two ways, one utilizing an accurate description of the outcome, and a second describing “improvement.” Of the 17 cases, only 9 (53\%) had complete elimination of VPI, but the article cited an 88\% “improvement” rate because several patients went from moderately hypernasal to mildly hypernasal or severely hypernasal to moderately hypernasal. It has been suggested that defining “success” or a positive outcome really needs to be done within the context of what an operation is designed to do\textsuperscript{29}. If a patient “improves” from severely hypernasal to “moderately” or “mildly” hypernasal, their speech is still abnormal, and the goal of the operation is normal resonance. If the improvement is not from a pathologic state to normality, the operation has not succeeded. Therefore, the reported success rate is approximately 50\% and the reader needs to judge the value of the operation based on this statistic, not 88\%\textsuperscript{26}.

The article of de Buys Roessingh\textsuperscript{28} assessed speech outcomes in Robin sequence cases following primary palatoplasty and subsequent pharyngeal flap if required. They report data on 25 nonsyndromic Robin sequence cases and 13 syndromic cases. Of the 13 syndromic cases, the most common diagnosis was velo-cardio-facial syndrome (VCFS) in three, two cases of Stickler syndrome, and 7 other conditions in the remaining cases. This represents a very heterogeneous sample for comparison thereby invalidating the outcome of the study. Children in the syndromic group were not grouped because they had similarities; they were grouped because of what they were not—isolated Robin sequence. This type of grouping by exclusionary criteria is an inappropriate way of selecting a sample.
Widdershoven et al.\textsuperscript{27} reported on speech outcomes in VCFS following palatal lengthening surgery. Twenty-five children with VCFS had palatal lengthening procedures and according to the authors, only 16\% (4 cases) required additional surgery. However, according to the authors, “…15 of 25 of patients with VCFS did not show a decrease in hypernasality (58\%) \textsuperscript{27}. The remaining 10 showed an audible improvement (42\%), but none of them achieved normal speech after the operation.” It is unclear why only 4 cases went on to additional reconstructive surgery.

**Complications and Side Effects**—An article with a sample size large enough to produce meaningful statistics\textsuperscript{30} reported complications and outcomes from 222 consecutive pharyngeal flap operations. A 3.3\% frequency of obstructive sleep apnea based on all-night polysomnograms was reported, although nearly 10\% were referred for PSGs based on clinical findings. Other complications were similarly rare. The authors did not comment on performing tonsillectomy prior to pharyngeal flap, a factor cited in earlier research as being a major risk factor for postoperative obstructive sleep apnea and increased upper airway resistance\textsuperscript{31,32}. They did report that pharyngeal flap was a safe and effective operation for the treatment of VPI and that there was no reason to avoid applying it to children.

Yamashita and Trindade\textsuperscript{33} also studied postoperative airway complications following pharyngeal flap in 58 cases. Five surgeons had done the operations; there was no mention about those who had prior tonsillectomy or adenoidectomy. Assessment of airway patency was performed by rhinometry and pressure-flow studies. Although nasopharyngoscopy was performed preoperatively, there was no endoscopic assessment of airway patency or flap width postoperatively preventing the correlation of airflow characteristics to flap width, position, or the presence of tonsils. It is therefore not possible to know if the complications experienced postoperatively could have been avoided by tonsillectomy\textsuperscript{32} or if the flap was too low\textsuperscript{31}.

Pereira et al.\textsuperscript{34} reported on changes in articulation, nasal resonance, and velopharyngeal function following maxillary advancement. There were 15 subjects grouped together as “craniofacial dysostosis” and data was compared as if they represented a homogenous group. However, the individual diagnoses show that four subjects had Apert syndrome, seven subjects had Crouzon syndrome, three had Pfeiffer syndrome (type unknown), and one had Saethre-Chotzen syndrome. Although lumped together, the sample is sufficiently heterogeneous to present variables that affect the outcome. Individuals with Saethre-Chotzen syndrome do not have craniofacial findings consistent with Apert syndrome. Although there are some obvious similarities between individuals with Apert, Crouzon, and Pfeiffer syndromes, there are substantial differences between their cognitive and developmental profiles and speech phenotypes with or without surgery. Outcomes could potentially have as much to do with the syndromic diagnosis as with surgery. Of course, it must be stated that all of these conditions are rare so that obtaining large sample sizes is not possible in many cases. Should important data be withheld simply because the sample is small? It is possible that the outcome data is very important for a study of this type, and therefore should be reported, but the small number of each type of case suggests that rather than grouping the subjects together to perform data analysis on the aggregate sample, the data should be reported anecdotally as a series of case reports. This does not diminish their importance, but does prevent overinterpretation of the data.

**Speech Therapy**—Ruscello\textsuperscript{35} addressed the issue of speech therapy for improving velopharyngeal closure. This review article described a variety of techniques including experimental approaches with CPAP and biofeedback, but the primary emphasis was on the use of “oral-motor therapy,” labeled as NSOME (nonspeech oral motor exercises), such as blowing, sucking massage, etc. Ruscello appropriately discounts the usefulness of these approaches that are, unfortunately, applied frequently by speech pathologists in schools and
early intervention programs, but have never been demonstrated to be of any value. Ruscello questions why these techniques are still being applied and considers it “distressing” that given our current state of knowledge about speech in people with VPI that speech pathologists would still be applying a discredited procedure.

Conclusion

Review of the literature has yielded two major conclusions. The first is that many papers have been published that rely on small samples sizes. The conclusions of these studies often overstep the scope of the data collected. Moreover, study design is often flawed in terms of the ascertainment of research subjects and the techniques used to assess them. Although some of the conditions studied are rare, it is difficult to know what motivates researchers to publish small samples while applying statistical analyses that cannot stand up to close scrutiny.

Another issue is the relative scarcity of publications devoted to the assessment and treatment of VPI. It may be that clinicians are comfortable with tried and true diagnostic procedures and treatments so that new explorations are not necessary. It is also possible that fewer investigators are being produced from the fields that have typically been interested in these aspects of patient care. Either way, both the scientific community and our customers would benefit from a healthy curiosity and even skepticism amongst researchers in order to keep us from treading water for too long. The potential for innovative diagnostic procedures and surgical approaches for VPI remains, but cannot be realized without a systematic melding of perceptual and instrumental analyses of the only outcome that matters: normal speech.

References


35 **. Ruscello DM. An examination of nonspeech oral motor exercises for children with velopharyngeal inadequacy. Semin Speech Lang 2008;29:294–303. This paper explicitly discusses speech therapy approaches that do not work in the treatment of VPI and its related disorders. Every clinician, surgeon and non-surgeon alike, should take this information to heart and make sure that their patients are not receiving ineffective speech therapy. [PubMed: 19058116]