Pediatric Facial Plastic and Reconstructive Surgery

James D Smith, Robert M Bumsted

Chapter 13: Evaluation and Management of Velopharyngeal Inadequacy

Linda L D'Antonio and Dennis M Crockett

One of the major goals of cleft palate repair is to create a velopharyngeal mechanism that is capable of separating the oral and nasal cavities during speech. When symptoms of velopharyngeal inadequacy are observed, most surgeons view this as a surgical failure requiring secondary surgical management. This chapter presents a broader view of the problem and illustrates that not all symptoms of velopharyngeal inadequacy are the result of the absence of adequate tissue and length. The need for *differential diagnosis* is emphasized as the first and most critical step in management planning. Diagnostic evaluation methods are discussed and suggestions are presented for determining which intervention strategy is appropriate, ie, speech therapy, prosthetic management, or surgical intervention. Several surgical approaches are discussed along with the issues of patient selection, surgical technique, complications, and postoperative care.

Most otolaryngologists-head and neck surgeons traditionally have seen their involvement with the evaluation and management of the velopharynx in children as limited to those surgeons who were directly involved in the care of patients with cleft palate. On the contrary, head and neck surgeons who see children in routine clinical practice will be asked several times in their careers for an opinion concerning the competence of the velopharynx for speech. The opinion given (whether right or wrong) is likely to have profound ramifications for each patient.

Velopharyngeal symptoms may have a variety of etiologies including over cleft palate, submucous cleft, neuromotor impairment, interference from other vocal tract structures (particularly the palatine tonsils), and mislearning. It is particularly important for the clinician to know that velopharyngeal symptoms do not occur solely as the result of overt or submucous cleft palate. For example, velopharyngeal inadequacy that may appear to occur in isolation of any obvious orofacial anomaly, may in fact be the single presenting symptom that can point to a multiple malformation syndrome. For example, Jones studied 428 patients who presented to a large cleft program at Children's Hospital and Health Center, San Diego. She showed that of the 428 patients, 75% of those who presented with velopharyngeal inadequacy as the primary symptom were later shown to have multiple malformation syndromes. These data suggest that the head and neck surgeon should be particularly thoughtful and thorough in the evaluation of a patient who presents with velopharyngeal symptoms. Although the concept that velopharyngeal symptoms may point to a more global disorder may seem academic for the physician in routine clinical practice, it is not. For example, velocardio-facial syndrome (VCF) is a recently delineated congenital malformation syndrome first identified by Shprintzen et al. It can be associated with clefting, cardiac anomalies, characteristic facies, learning disabilities, and speech disorders. Frequently, these children first present in pediatric or otolaryngology offices with complaints of velopharyngeal symptoms. Positive identification of this syndrome is particularly crucial for any surgical considerations. Specifically, many of these children have medially displaced internal carotid arteries, which can be an obvious contraindication for many surgical procedures and therefore of great importance and interest for surgeons, especially head and neck surgeons who may be called upon to perform tonsillectomy, adenoidectomy, or velopharyngeal surgery to remediate the presenting symptoms.

Although the focus of this chapter is on the evaluation and management of the velopharynx, it must be understood from the outset that the velopharynx does not function in isolation and cannot be evaluated separately from the entire vocal tract and communication process. Rather, disturbances in anatomy and function that have the potential for affecting velopharyngeal function are far more complex and interrelated than has been acknowledged historically.

The Velopharynx Revisited

For clinical purposes, the speech production mechanism may be thought of as a large air-filled container always closed at the bottom with two openings to the atmosphere, the lips and the nares. Within the container, there are several valves that can be opened or closed to varying degrees, thus changing the shape of the container and the resistance to airflow. All of these valves must move in a highly coordinated manner creating a series of rapidly changing air pressures and airflows, which we ultimately perceive as the sounds of speech.

As viewed quite simplistically, the role of the velopharynx is to separate the oral and nasal cavities during speech and swallowing. For speech, the velopharynx directs air from the lungs and larynx through the mouth for oral sounds and through the nose for nasal sounds. When this valving is disturbed, speech can be affected in a number of ways including hypernasal resonance, nasal air emission, hyponasal resonance, nasal substitutions, compensatory articulation, sibilant distortion, and increased risk for disorders of phonation, ie, laryngeal voice disorders. Table 1 presents definitions of several speech symptoms associated with cleft palate and/or velopharyngeal dysfunction.

Table 1. *Definitions of resonance, articulation, and phonation disorders frequently associated with cleft palata and/or velopharyngeal dysfunction*

Hypernasality: The perception of inordinate nasal resonance during the production of *vowels.* This results from inappropriate coupling of the oral and nasal cavities. (The term *inordinate* is used to allude to the fact that low vowels and vowels in nasal consonant contexts are normally somewhat nasalized.)

Nasal emission: Nasal air escape associated with production of *consonants* requiring high oral pressure. It occurs when air is forced through an incompletely closed velopharyngeal port. Nasal emission may be audible or not. (Note: *Hypernasality* and *nasal emission* are not synonymous although they often occur together and are both symptoms of velopharyngeal dysfunction.)

Hyponasality: A reduction in normal nasal resonance usually resulting from blockage or partial blockage of the nasal airway by any number of causes including upper respiratory infection, hypertrophied turbinate or hypertrophied adenoids, a wide, obstructing pharyngeal flap.

Hyper-hyponasality: The simultaneous occurrence of hypernasality and hyponasality in the same speaker usually as the result of incomplete velopharyngeal closure in the presence of high nasal resistance, which is not sufficient to block nasal resonance completely.

Cul-de-sac resonance: A variation of hyponasality usually associated with tight anterior nasal constriction often resulting in a "muffled" quality.

Nasal substitution: The articulators are placed appropriately for an intended oral consonant. However, incomplete velopharyngeal closure causes the sound to be produced as a nasal consonant. For example, "b" becomes "m" and "d" becomes "n". Such substitutions frequently are called "homorganic nasals".

Compensatory articulation: The articulators are placed inappropriately so as to enable creation of the plosive or fricative characteristics of the sounds they replace. For example, if a patient cannot build up oral pressure for the fricatives (eg, "s") or plosives (eg, "p") because of velopharyngeal dysfunction, they may create those pressures below the level of the velopharyngeal port. Such substitutes include glottal stops, pharyngeal stops, and pharyngeal fricatives among others.

Sibilant distortion: Inappropriate tongue placement for the sounds "s" and "z".

Laryngeal/voice symptoms: A variety of phonation disorders may accompany velopharyngeal dysfunction including hoarseness, low speaking volume, strained or strangled voice quality, and unusual pitch alterations. The most recent theory for the co-occurrence of velopharyngeal and laryngeal symptoms is that speakers with velopharyngeal dysfunction may attempt to compensate for the inability to achieve complete closure and maintain adequate speech pressures by compensatory activity at the level of the larynx.

Historically, the velopharynx has been viewed as a simple binary valve with two possible positions, open or closed. However, research and clinical observation over time have shown that the velopharynx is a complex three-dimensional valve with a variety of shapes and patterns of activity that vary between speakers. Just as the lips and tongue are considered speech articulators, so too is the velopharynx, as it takes different positions or shapes for different sounds. Furthermore, it is not enough that the velopharynx be capable of achieving complete closure of the valve, but it must do so in a tightly controlled time domain in coordination with other articulators.

Terminology: More Than a "Mere" Semantic Difference

Traditionally, when symptoms have been present that suggested that the velopharynx was not functioning correctly, we referred to the problem as "velopharyngeal incompetence" (VPI). However, hearing hypernasality or nasal emission (ie, symptoms associated with VPI) does not necessarily indicate that the velopharynx cannot work. These symptoms merely indicate that the velopharynx is not functioning at that time. Wendel Johnson, one of the fathers of modern speech pathology and a great semanticist, explored the relationship between language and science. He taught us that *the way we talk about a topic influence the way we think about it.* Therefore, when a clinician hears hypernasality or nasal emission, ie, symptoms of VPI, and in turn makes a diagnosis of "VPI" the semantic label he or she has used can have far-reaching implications. For example, the diagnosis of VPI suggests that the velopharyngeal mechanism cannot achieve closure. For most clinicians this implies that only physical management (either surgical or prosthetic) will correct the problem. Additionally, such a label may bias the inexperienced speech pathologist to discontinue speech therapy on the grounds that additional therapy will be of no value until physical management is completed. Furthermore, if the diagnosis is incorrect, any surgical or prosthetic attempts to manage the problem will be of little value. In this instance the clinician has used imprecise diagnostic labels, which can unintentionally lead to erroneous assumptions and inappropriate actions.

Previous authors have made recommendations for standardizing the nomenclature. For example, Loney and Bloem reviewed the literature and found neither consensus nor precise definitions of the terms *velopharyngeal incompetence, velopharyngeal inadequacy,* or *velopharyngeal insufficiency.* They found that authors frequently used all three terms interchangeably or one term to describe all types of velopharyngeal malfunction.

In a response to Loney and Bloem, Trost-Cardamone suggested a taxonomy for velopharyngeal disorders based on etiology. She states:

> Impaired velopharyngeal closure for speech can result from a variety of etiologies. Moreover, there are perceptual speech characteristics that are pathognomonic of velopharyngeal impairment and that can distinguish among certain subtypes of velopharyngeal function problems. In both diagnosis and treatment, it is necessary for the clinician to have a taxonomic system for reference, which should serve to relate etiology to deviant velopharyngeal and speech production patterns. This is especially important because the nature of the velopharyngeal function disorder allows for certain treatment alternatives and excludes others.
>
> Trost-Cardamone suggests the following taxonomy:
>
> *Velopharyngeal inadequacy:* the generic term used to denote any type of abnormal velopharyngeal function ... Within the broad group of inadequacies, there are subgroups of structural, neurogenic, and mislearning or functional origins.

Under the broad classification of velopharyngeal inadequacy she delineates three etiologic categories: velopharyngeal insufficiency, verlopharyngeal incompetence, and velopharyngeal mislearning. Each of these is defined as follows:

> *Velopharyngeal insufficiency:* Includes any structural defect of the velum or pharyngeal wall at the level of the nasopharynx; there is not enough tissue to

accomplish closure, or there is some type of mechanical interference to closure. Most often these problems are congenital.

> *Velopharyngeal incompetence:* Includes neurogenic etiologies that result in impaired motor control or impaired motor programming of the velopharynx ... Motor control disorders can cause partial or total paresis of the soft palate and pharyngeal walls. Depending upon the nature, level, and locus of the nervous system lesion, velopharyngeal incompetence often disturbs velopharyngeal closure for protective and reflexive acts of gagging and swallowing, as well as for speech.
>
> *Velopharyngeal mislearning:* Includes etiologies that are not caused by structural defects or by neuromotor pathologies of the velopharyngeal complex.

In this latter category of velopharyngeal mislearning, Trost-Cardamone includes sound-specific nasal emission and velopharyngeal symptoms associated with deafness or hearing impairment.

The use of standardized, precisely defined terms is one of the prerequisites to scientific investigation. However, the use of clearly defined nomenclature is also a critical factor in clinical practice where unclear terminology may lead to inappropriate assumptions and inappropriate management. Loney and Bloem and Trost-Cardamone should be congratulated for their discussion of the lack of standardized terminology and for their attempt to define terms. However, in clinical practice the delineation they propose often may be too cumbersome for daily use and may not accomplish their goal of "improved interprofessional communication". Trost-Cardamone's system offers an important, meaningful classification of velopharyngeal inadequacies based on etiology, which is missing in the literature. It is sound and useful, especially once differential diagnosis is known. However, accurate use of the levels of definition frequently requires diagnostic information that may not be available at the first presentation of velopharyngeal symptoms.

Therefore, a simpler and more direct approach for routine, clinical practice is to use the generic and all-encompassing term *velopharyngeal dysfunction.* This term does not assume, or rule out, any possible cause of the perceived speech symptoms or any management approach. As described by Netsell, velopharyngeal dysfunction during speech may be the result of structural deficits, neurological disorders, faulty learning, or a combination of sources. Dalston uses the term *velopharyngeal dysfunction* and defines it as follows:

> ... any impairment of the velopharyngeal complex. It may result from a lack of sufficient tissue to enable contact to be effected between the soft palate and the posterior pharyngeal wall ("velopharyngeal insufficiency"), a lack of neuromuscular competency in moving velopharyngeal structures into contact with one another ("velopharyngeal incompetency") or both. Finally, it may be due to maladaptive articulatory habits that do not reflect physical or neuromuscular impairment (eg, phoneme-specific nasal emission).

By not implying the source of the symptoms, the term *dysfunction* acknowledges that some features of the velopharyngeal valving mechanisms are not functioning appropriately, but no cause or treatment approach is implied or suggested until appropriate diagnostic testing can be conducted.

One more note on terminology seems appropriate in the discussion of velopharyngeal function. In the book *Cleft Palate Speech* McWilliams et al point out that individuals with velopharyngeal incompetence are typically described as displaying "hypernasal voice quality". However, the word *voice* more accurately refers to problems associated with phonation. Therefore, they suggest the use of the term *phonation disorders* to refer to problems that occur at the level of the larynx and the term *resonance disorders* to refer to hypernasality and other disturbances that occur supraglottally. Once again, the use of clearly defined, physiologically based terminology leads to a more accurate and precise nomenclature for both clinical and research applications.

Differential Diagnosis

In the evaluation and treatment of velopharyngeal dysfunction, the clinician must be aware that speech symptoms often attributed to velopharyngeal incompetence may come from a variety of sources or combination of sources. For example, hypernasality and/or nasal emission may be the result of the lack of sufficient tissue to allow closure of the velopharyngeal port; neuromotor impairment involving innervation of muscles of the velopharyngeal port as in many congenital anomalies, neurologic diseases, or head injury; mislearning or other behavioral factors as in instances of phoneme-specific nasal emission or lack of oral/nasal discrimination; or from other structural involvement, such as a palatal fistula, that may allow nasal air escape or enlarged tonsils, which have been shown to result in impaired velopharyngeal function in some instances.

The process of differential diagnosis can be quite difficult since speech symptoms that may appear to be quite similar and indistinct to the casual listener or inexperienced observer or in some instances even the experienced listener may in fact be quite diverse in both etiology and, therefore, appropriate management. For example, in the cleft population, it is not uncommon to observe hypernasality and nasal emission in a patient with a repaired cleft and a residual oronasal fistula. Casual perceptual observation cannot determine the source of these symptoms. Therefore, it would be inappropriate to label the phenomenon as velopharyngeal incompetence without investigating the symptoms further. In some instances, the symptoms may be solely attributable to air escape through the fistula. In this case the diagnosis would be hypernasality and nasal emission due to a patent oronasal fistula and may be unrelated to the velopharyngeal mechanism. In this diagnostic situation, the appropriate management would be repair or obturation of the fistula. In other cases, the symptoms may appear to be attributable to a combination of air escape through the fistula and a lack of proper velopharyngeal function. In these cases, diagnosis and management planning become far more complicated.

Another illustration of the need for differential diagnosis is the phenomena of "phoneme-specific nasal emission". This is the perception of nasal emission isolated to specific pressure consonants such as /s/ and /z/, as opposed to the consistent nasal emission generally seen when the velopharyngeal valve is incapable of closure. The inexperienced clinician (and until recent years, even some experienced clinicians) have mistaken this articulation-based error for true velopharyngeal incompetence. Many patients have been referred for or have received inappropriate physical management for what is basically an articulation error that requires behavioral therapy for effective remediation.

As these examples illustrate quite clearly, the evaluation of velopharyngeal function for speech must be carried out by a speech pathologist with expertise in this area. Although some cases may be straightforward, most are not. It is a detailed description of symptoms, their frequency and severity, and their response to behavioral probes that defines the problem and leads to appropriate management suggestions. This description or differential diagnosis is made by addressing a number of diagnostic questions using a variety of evaluation methods. For example, the speech pathologist might ask:

1. Can the patient achieve complete velopharyngeal closure?

2. If the patient can achieve closure, what are the speech symptoms that are resulting in the perception of a speech disorder?

3. If the patient cannot achieve closure, what are the possible reasons why not?

4. If closure is intermittent, which speech environments facilitate closure and which contexts are detrimental?

5. How much do symptoms of velopharyngeal dysfunction contribute to the overall communication handicap?

6. Has previous behavioral management for the symptoms in question been attempted? If so, what was the outcome?

7. Is management indicated? If so, is behavioral management, ie, speech therapy, indicated in isolation or in conjunction with physical management.

Ultimately, the speech pathologist must interpret the findings of the comprehensive evaluation of velopharyngeal function along with the patient's history and physical findings and arrive at a provisional differential diagnosis. This information should then be discussed with members of a multidisciplinary team and interpreted along with other pertinent physical and social information regarding the patient. Only then can an appropriate, individualized management and follow-up plan be established.

The Rationale for Multilevel Multimethod Evaluation

Speech production is one of the most intricate of all human behaviors. A block diagram of the speech production process as conceptualized by Netsell: CNS organization --> Nerve impulses --> Muscular events and structural movements --> Air pressures and air flows --> Acoustic waveform. The process originates in the central nervous system governed by various linguistic processes. Nerve impulses travel to various muscles and the muscular events result in structural movements. As many as 100 muscles may be involved in this process. These muscle and structural movements then result in air pressure and airflow events, which result in the acoustic output that ultimately is perceived as speech. Although this process occurs over a number of levels, generally only the latter two or three levels are available for observation and measurement. For many years we were limited to observations at one level of this complex speech production process, ie, the auditory perception of the acoustic output, or listener judgments of speech. With the increased application of instrumentation in clinical practice we have moved more routinely into observation of aerodynamic events and structural movements.

This view of the speech production process reminds us that in clinical practice, the events we most frequently observe as "speech" represent the most peripheral level of observation in the speech production process. This fact provides the rationale for the use of instrumental methods that allow us to make observations of the speech production process at earlier levels. Information from instrumental assessments at various levels of this complex process not only provide relevant clinical information about speech production but also provide insight into speech motor control processes.

Which Method is Best?

Many clinicians and researchers who are experienced in the use of instrumental methods, frequently are asked which method is best. In this era of shrinking budgets and ever-reducing reimbursement, this question might at first appear quite reasonable. However, the answer is not simple. Instrumental methods can differ from one another on a number of features including the level of the speech production process accessed, whether the method provides direct or indirect observation, clinical practicality, invasiveness, cost-to-benefit ratio, validity, and reliability.

There is increasing interest in the use of instrumental methods for assessing velopharyngeal function for speech. It has been suggested that instrumental methods are more reliable and informative than listener judgments alone, and therefore lead to a more accurate diagnosis and improved treatment planning. Critics of the trend toward instrumental assessment often cite a study by McWilliams et al to show the poor correlation among instrumental methods, particularly aerodynamic methods. In this study, McWilliams and her coauthors compared listener judgments of velopharyngeal function (specifically hypernasality) with nasal manometry, pressure-flow, and multiview videofluoroscopy. They showed complete agreement among methods for only three of the 35 speakers studied.

Although these data have been interpreted to point out the poor validity of some instrumental methods, an alternative interpretation has been proposed. D'Antonio and her colleagues have suggested that the important issue is not how often the methods agree completely or even which method is best. Rather, the user must ask:

1. What unique and what overlapping information does each method provide?

2. When is it appropriate to use a given method or combination of methods? That is, what information is being sought? And is this the most appropriate method and observers to provide that information?

3. When the information from multiple methods does not appear to agree, what information does this provide and how might this alter our diagnostic and management decisions?

Results from two studies reported by D'Antonio and colleagues suggest that the reliability and validity of evaluation methods may be affected to a greater degree than we have acknowledged by variables such as an individual user's training, background experience, and visual and auditory perceptual strengths. The authors suggest also that in clinical practice, it is likely that the effects of these and other variables that may result in instances of "apparent disagreement" may be diminished by operation factors such as the combined use of multiple evaluation methods and the use of an interdisciplinary team approach for the collection and interpretation of evaluation data.

Another important question that must be addressed in clinical practice is whether a methodology is being employed correctly. As with all areas of medicine, there is great potential for misuse and damage when instrumentation is used incorrectly. With the increased utilization of instrumental methods for the evaluation of velopharyngeal function comes the distressing reality that in several instances the advantages and limitations of a method and the recommendations for its appropriate use have been ignored. For example, D'Antonio et al conducted a national survey concerning methods used for the evaluation of velopharyngeal function and showed 90% of cleft palate teams now have nasendoscopic evaluation capabilities. This is a significant increase in availability compared with studies conducted in 1980 and 1984, which showed only 8% availability. Unfortunately, the results also showed that increased availability did not necessarily assure optimal use.

For example, the principal advantage of direct visualization techniques such as nasendoscopy is the assessment of velopharyngeal function during dynamic speech activity. However, three cleft palate teams reported the use of general anesthesia or heavy sedation during endoscopic examination. Utilization such as this clearly demonstrates the potential for significant misuse when the clinicians involved have not clearly identified the questions being asked and the most appropriate method to answer those questions. Additionally, responses to the same survey were disappointing with regard to the interpretation of results. In the majority of centers, endoscopic evaluations of velopharyngeal function were performed by a physician alone. This finding is of particular concern since most physicians are not likely to be familiar with many of the behavioral variables that can affect velopharyngeal function or with how to interpret physiologic findings in conjunction with the overall speech production processes and communication skills present in a given patient. Results such as these indicate that the advantages instrumental methods can and often do have can be overshadowed by inappropriate use. Once again, clinicians must insure that the most suitable method is being utilized by the most appropriate user in order to answer the question under investigation. In this way, we can help to insure that the method being employed is "the best method".

Evaluation Methods

Numerous methods exist for evaluating velopharyngeal function at different levels of observation within the speech production process. Frequently reported methods include:

- Listener judgments
- Intraoral examination
- Acoustic analysis
  - Spectrograph
  - TONAR

- Accelerometer
- Aerodynamic measures
- Observation of structural movement
    - Multiview video fluoroscopy
    - High-speed x-ray
    - Ultrasound
    - Photodetector
    - Nasendoscopy
    - Electromyography.

It can be noted to some extent, that the least invasive procedures allow observation at the most peripheral stages of the speech production process and the relative invasiveness of the methods increase as we move up to higher, or earlier, stages.

Obviously, not all, or even most available evaluation methods can or should be employed in a given patient or patient population. As discussed previously, each method has inherent strengths and weaknesses and can provide both unique and overlapping information pertaining to velopharyngeal function.

The rationale for selection of assessment methods appropriate to a given population or question has already been discussed in the section Which Method is Best? However, one topic that was not covered is the strategy for selecting a combination of methods for clinical use that complement each other's unique contributions and limitations. In this way, the outcome can be one where "the whole is greater than the sum of its parts". For example, we have seen that nasendscopy provides direct visualization of the velopharyngeal port but relies on subjective, visual interpretation, usually by a single observer. On the other hand, aerodynamic assessment is an example of a method that may provide quantitative data regarding the effects of velopharyngeal closure, but in this case the data about velopharyngeal closure are indirect. As illustrated below, these two methods can compliment one another and if utilized together, can provide direct and quantitative information about velopharyngeal function.

Nasendoscopy	Aerodynamic Measures
direct	indirect
subjective	quantitative.

This examples illustrates once again that the selection and combination of instrumental evaluation methods should be given great consideration prior to their inclusion in the diagnostic armamentarium.

In current practice, the most commonly employed methods for clinical evaluation of velopharyngeal function are perceptual assessment by a trained speech pathologist, aerodynamic measures, nasendoscopy, and multiview videofluoroscopy.

Peroral and Head and Neck Examination

A thorough peroral and head and neck examination is essential for any child who presents with known or suspected velopharyngeal dysfunction. The oropharynx is visualized and palpated, and the length of the soft palate gauged with respect to the depth of the nasopharynx, both at rest and during phonation. Symmetry of palatal excursion during phonation is noted, keeping in mind that normal palatal motion with swallowing and gag often is present in children even with severe velopharyngeal dysfunction and does not relate to, or predict, velopharyngeal activity during speech. Any signs of palatal clefting are noted: bifurcation or notching of the uvula, notched or U-shaped posterior border of the hard palate, and any midline mucosal color change (zona pellucida) or muscle bulge indicating muscle diastasis.

The general head and neck examination should place special emphasis on three areas: the upper airway; the ears, including audiological evaluation; and any structural anomalies that may indicate the presence of a recognized pattern of human malformation or syndrome of which cleft palate is a known feature. The upper airway (including the septum, turbinates, adenoid and palatine tonsils, and tongue base) is examined for obstruction that may affect speech quality or speech articulation and for any potential obstruction that might further compromise the airway if any surgery is performed to improve velopharyngeal function. The relationship between the nasal airway and breathing and speech is far more complicated in patients with cleft palate than has been understood historically. An excellent review of current clinical data of importance for the head and neck surgeon involved in the care of patients with cleft palate can be found in the work of Warren et al.

The ears are examined for middle ear effusions, and audiologic testing is completed, if needed. General craniofacial configuration is examined; any anomalies such as synostoses, telecanthus, maxillary or malar hypoplasia, abnormal pinnae, abnormal mandibular shape and excursion, or malocclusion may be reason for referral to a pediatric geneticist.

Perceptual Speech Evaluation

Evaluation by an experienced speech pathologist is essential and should include a comprehensive evaluation of communication skills, including symptoms of velopharyngeal dysfunction. Particular attention should be paid to assessment or resonance, nasal emission, articulation, the presence of compensatory articulations (such as glottal stops and pharyngeal fricatives), the presence of facial grimacing, and overall speech quality and intelligibility. The head and neck surgeon should be capable of performing simple and brief listener judgments of these variables so that he or she is aware of when to refer for a speech language evaluation or to an expert in the evaluation of velopharyngeal function. However, under no circumstances should the head and neck surgeon attempt to perform a speech evaluation for the purposes of differential diagnosis of velopharyngeal dysfunction, just as a speech pathologist would never perform surgery.

The advantage of the listener evaluation is that it has high "face validity", that is, the human auditory perceptual system is excellent at determining when there is a speech disorder, particularly impaired velopharyngeal function. It is the way that speech sounds that is the "gold standard" for whether treatment is indicated. In most instances, regardless of the results of more sophisticated instrumental measures, it is the contribution that an impairment in velopharyngeal function causes to the overall communication competence of an individual that will determine the severity of the disorder and whether management of any kind is indicated. Unfortunately, in spite of the fact that it is "how the speech sounds" that is the critical variable, there are numerous criticisms that indicate that subjective, listener evaluations, by a single clinician (without known inter- and intrarater reliability) are frequently unreliable. Listeners vary in tolerance for various speech parameters and differ greatly in personal biases of what may constitute a communication impairment. Therefore, although listener judgments of speech are essential, they should always be accompanied by instrumental assessments that may help to validate the listener's perceptions. Shprintzen and Golding-Kushner summarize this point quite well:

> The correct approach to treatment of VPI is dependent upon adequate diagnostic information, and potentially dangerous or fatal complications may occur when VPI is treated inappropriately. Therefore, it is imperative that clinicians apply state-of-the art diagnostic procedures to the patient with VPI. This is of particular importance when one considers that hypernasality is one of the few speech disorders that is often treated with some form of physical management, including a large number of surgical approaches.

Aerodynamic Measures

As described previously, the human vocal tract is an air-filled container closed at the bottom by the lungs. The lungs generate air pressure that is valved by the vocal tract at a number of levels including the larynx, velopharynx, nasal cavity, and lips. A number of aerodynamic methods have been employed such as active posterior rhinomanometry to asses the function of the velopharynx and nasal cavities during speech. Warren and DuBois introduced the most commonly applied methodology for the measurement of oral and nasal pressure differential and nasal flow, which can be placed into a hydrokinetic formula to estimate the area of velopharyngeal opening at the moment of peak pressure of a stop consonant. Other methods exist that are an adaptation. Although this methodology requires sophisticated equipment and significant operator competence, its major advantage is that it is a noninvasive method that can be used with small children to provide quantitative data concerning velopharyngeal function. Studies have shown these methods (when used and interpreted properly) provide valid and reliable information regarding velopharyngeal function. These methods are particularly valuable because they can subtract out the contribution of the nasal passages to the impedance of airflow from the velopharynx. This is especially useful in the common situation of the patient with velopharyngeal incompetence in the presence of hypertrophied turbinates or a structurally occluded nasal airway. In these instances, aerodynamic measures can provide critical information that the human ear simply cannot sort out. Aerodynamic data also provide important information about the timing of closure of the velopharyngeal port that is of great value in patients with neurologic contributions to their velopharyngeal symptoms.

Although aerodynamic methods have numerous advantages and important applications for the cleft and noncleft populations, they have not gained widespread clinical popularity because of the expense involved in instrumentation and the amount of training necessary for accurate interpretation of test data. Additionally, although aerodynamic measures are quantitative, valid, and reliable, they tell the surgeon nothing about the anatomy and structure of the mechanism he or she may be about to perform surgery on.

Nasendoscopy

Movement of the velopharyngeal mechanism during speech is complex and includes musculature contractions of not only the soft palate, but also of the lateral and posterior pharyngeal walls. This complexity is reflected in the fact that the patterns of closure of the velopharyngeal port are highly variable among both normal speakers and patients with velopharyngeal dysfunction. Every child with known or suspected velopharyngeal dysfunction should undergo nasendoscopy, especially if a surgical procedure is anticipated. The examination should be performed with both the speech pathologist and surgeon in attendance, with a permanent record of the examination recorded on videotape. Permanent documentation on videotape is important for reviewing results of the evaluation after the stress of the actual procedure is over. It allows for multidisciplinary review of the results and provides a means of monitoring magnitude and direction of change. This last advantage cannot be emphasized enough in today's medical-legal climate. Additionally, if surgical intervention is recommended, a videotape record of the evaluation allows the surgeon to view the tape just prior to the operation to recall the specifics of the pattern of closure and the function of the mechanism to be operated on.

Most children 3 years of age and older can be examined successfully with the nasopharyngoscope. D'Antonio et al found that 87% of children between 3 and 6 years of age can be examined. The diameter of the scope can be a limiting factor. The 3.6-mm scope gives a superior video image, and can be used in younger children. The 3.0- to 3.2-mm scope results in a video image of slightly less quality, but can be used in younger children. Important considerations for successful nasendoscopy in children include a comfortable surrounding, a slow and easy approach to the examination, and appropriate rewards for a successful examination. Nasal mucosal anesthesia is accomplished by means of a hand-held atomizer, which is used to gently spray 2% cocaine solution into the anterior portion of the nasal cavity. If this does not provide adequate anesthetization, a tightly rolled cocaine-soaked cotton pledget may be gently introduced into the nasal cavity and passed posteriorly to further anesthetize the most sensitive, posterior areas of the nasal cavity.

Following removal of the pledget, the nasopharyngoscope is passed above the inferior turbinates, opposite the middle meatus, into the nasopharynx. The scope is positioned so that a complete circumferential view of the nasopharynx is obtained. The child is then asked to produce a sample of words and short sentences that comprise an adequate speech sample. The speech pathologist should be available to conduct the speech evaluation during the endoscopic examination. Behavioral, diagnostic probes should be used to ascertain the child's performance limits. If possible, attempts toward velopharyngeal closure should be assessed during correctly articulated speech sound productions only. Finally, function should be observed in single sounds, single syllable words, short phrases and sentences, and in increasing levels of complexity. The speech sample and evaluation should be designed to place minimum and maximum demands on the velopharyngeal mechanism during speech production. Consistency of closure and patterns of variability are *critical* findings for the speech pathologist to help establish a differential diagnosis.

Direct visualization of velopharyngeal movement in this manner allows a mapping of the defect that is recorded on videotape. The degree and symmetry of soft palate motion, lateral pharyngeal wall motion, and the pattern of velopharyngeal closure, or attempted closure, in the coronal and sagittal planes are noted. In addition to the permanent videotaped record, movement of the soft palate, lateral pharyngeal walls, and posterior pharyngeal walls may be rated on a numerical scale and recorded on a diagram. The vertical level of most prominent lateral wall motion is estimated with respect to its distance from the rostrum of the torus tubarius.

It is important to pass the scope into the inferior nasopharynx, as the vertical level of most prominent lateral wall motion may be obscured by movement of the soft palate superiorly. If large palatine tonsils have been noted previously during the peroral examination, special attention is afforded them during videonasendoscopy. Some children with mild velopharyngeal symptoms do not exhibit adequate lateral seal between the velum and the lateral pharyngeal walls because the superior pole of one or both palatine tonsils projects into the nasopharynx.

Following evaluation of the velopharyngeal mechanism, the head and neck surgeon should evaluate laryngeal and vocal fold function during rest and phonation. Over 40% of patients with VPI will exhibit abnormal voice characteristics, as well as pathologic vocal fold changes such as nodules, thickening, edema, and various patterns of incomplete glottic closure. This information can be important for directing behavioral speech therapy in selected cases.

Videonasendoscopy should be repeated approximately 3 months to 6 months postoperatively. This is necessary to confirm the success of the surgical procedure in those children who have achieved postoperative velopharyngeal competency. In patients who demonstrate persistent velopharyngeal dysfunction postoperatively, videonasendoscopy allows the surgeon and the speech pathologist to visualize the mechanism during speech and to determine whether the residual symptoms are physical or behavioral. This allows proper planning of a revision procedure if necessary or for additional speech therapy.

The advantages of nasendoscopy are obvious. It allows for direct observation of the velopharyngeal mechanism during dynamic speech activity. It is relatively noninvasive (under appropriate circumstances) and gives clinicians important information about the physical structure to be modified if physical management is necessary. Another advantage that has not been discussed is the potential for nasendoscopy to be used for diagnostic therapy or for biofeedback therapy.

In spite of the significant advantages of the use of nasendoscopy in the evaluation of velopharyngeal dysfunction, the method does have some limitations that should be remembered. At the present time it is still a subjective, visual-perceptual evaluation. Quantification methods do not exist for this procedure. There are numerous problems with image distortion, scope placement, repeatability etc. Additionally, as discussed previously, since most otolaryngologists are familiar with the procedure and scopes are readily available in almost all otolaryngology offices, there is the danger that many head and neck surgeons feel qualified to perform the complex differential diagnosis of velopharyngeal dysfunction without the assistance of a qualified speech pathologist.

Multiview Videofluoroscopy

This method is a radiologic technique that facilitates evaluation of the velopharyngeal mechanism in several planes during dynamic speech activity. In most settings, high-density barium is placed into one or both nostrils with a syringe and the patient is instructed to "sniff" the contrast material into the nasopharynx. The patient is positioned and as many views as necessary (and practical) are conducted in order to obtain an adequate representation of velopharyngeal function. The most common views employed are the lateral view, frontal view, Towne's view, and base view. (In some institutions, the Towne view will be used as a substitute for the base view and the Waters projection will be used as an alternative for the frontal view.)

The lateral view usually is taken first, and is the most common. It provides information about the relative length of the soft palate, its thickness, and the depth of the pharynx. Tonsillar and adenoid tissue are also noted. The frontal view provides important information about the vertical location of most active movements of the lateral pharyngeal walls. The combination of the lateral and frontal views supplement one another to give a more accurate impression of the velopharyngeal mechanism. However, in some cases important information is still missing and the base (or Towne or Waters view) is necessary to clarify the three-dimensional nature of velopharyngeal movements.

Many of the advantages and cautions associated with videofluoroscopy are similar to those discussed for nasendoscopy. For instance it should always be performed by the radiologist in cooperation with the speech pathologist. Interpretation of results should never be attempted without sound accompanying the fluoroscopic images or in isolation of a more comprehensive speech and velopharyngeal evaluation. Additionally, lateral, static, two-dimensional x-rays are not a substitute for providing information regarding the dynamic activity of the velopharyngeal mechanism.

Earlier multiview cineradiography allowed frame-by-frame analysis of articulatory movements over time (not only of the velopharynx but of the tongue, lips, and jaw as well). Videofluoroscopy continues to allow the assessment of the velopharynx in relation to other articulators but does not lend itself well to accurate quantitative analysis over time. However, this limitation is overridden by the reduction in radiation necessary for cineradiographic studies compared with videofluoroscopic examinations. Nevertheless, even in the latter case the amount of radiation exposure is the principal disadvantage of this method, especially in young children for whom repeat studies may be necessary.

Issues Related to Management Decisions

Once a diagnosis has been established, attention turns toward the question of whether management of any kind is indicated. McWilliams et al address the interdisciplinary roles in diagnosis and management:

> We want to emphasize our conviction that it is the role of the speech pathologist to judge the adequacy of velopharyngeal function for speech production. This is done in collaboration with other professionals who also provide valuable information about the mechanism and its function and suggest

> alternatives for management. If physical management is chosen as the preferred method of treatment, the surgeon or the dentist is legally and ethically accountable to the patient and family for proper treatment and must always have the last word about whether or not to perform the surgery or construct the prosthesis. However, it is the speech pathologist who decides whether or not the proposed physical management is indicated for speech improvement.

No form of management is without its own costs. Although surgical management may be the most permanent and carry the most physical risks, prosthetic management and speech therapy also require significant compliance, time commitment, and often emotional commitment. Therefore, no management alternative should be undertaken casually.

It is incumbent upon the speech pathologist to interpret the available evaluation data and to determine whether management of any form will improve speech intelligibility, speech quality, or quality of life. Unfortunately, there is no formula for making such decisions easily. For example, in one instance, a patient may present with a significant communication impairment of which velopharyngeal dysfunction is only a small source of the overall decreased communication competence. In such a case a cost-benefit analysis might lead to the decision not to offer physical management, especially surgical management. Another patient might present with mild velopharyngeal dysfunction with no other communication impairment. Although the symptoms may be slight the patient and/or the patient's family may feel that even a small observable stigmata will result in detrimental effects and reduced life options. Again a cost-benefit analysis may suggest that physical management will be likely to result in only a minimal improvement in speech quality. However, in this case even small improvement may be enough to translate into significant improvements in quality of life. In this situation, it may be appropriate to offer physical management, even surgical management.

This discussion illustrates the point that in spite of our ability to collect detailed scientific data concerning the source and magnitude of velopharyngeal symptoms, in the final analysis decisions regarding management are as much a part of the art of cleft care as the surgical act itself.

As discussed previously, there are three broad categories of management for velopharyngeal dysfunction - speech therapy, prosthetic management, and surgical management - which will be discussed in the following sections. All three of these management categories are familiar to clinicians involved in cleft care. Most readers will have personal biases about the particular effectiveness and value of each method. However, in reality, the reader should realize there has been little prospective research on treatment outcome for any of these methods. Much of what we, as clinicians, believe to be true about each method's effectiveness usually is based on personal experience.

McWilliams et al quite accurately point to the need for prospective research concerning treatment effectiveness. They remind us that the same methods available for the evaluation of velopharyngeal function

> ... are also appropriate for assessing changes in speech and velopharyngeal function in response to therapy and to surgical and dental treatments. Prospective research predictive of treatment outcome is needed. It is difficult

> to accomplish because many variables in addition to treatment influence the results. For example, patients with different characteristics respond differently to a given treatment, and measurement error is not unknown. Two clinicians may differ in the skill with which they use a particular technique. These confounding variables are difficult to control experimentally, but the development of well-founded treatments depends on sound treatment research.

Speech Therapy

The role of the speech-language pathologist and the field of speech pathology in general often appears to be an enigma to most medical specialists. There are a variety of potential speech-language pathology interventions for patients with craniofacial anomalies that may come under the broad label of "speech therapy". These include parent counseling, speech and language stimulation, joint parent-child communication therapy (individual or group), articulation therapy, phonological processes therapy, resonance therapy, voice therapy, velopharyngeal function therapy, and language therapy.

In discussing the role of behavioral therapy in the management of velopharyngeal dysfunction Van Demark and Hardin provide an excellent definition of the role of speech intervention:

> The goal of speech therapy for any child is to establish age-appropriate speech production patterns through behavioral modification. In addition to addressing any developmental articulation deficits that the child may demonstrate, therapy for the child with cleft lip and palate is typically directed toward the objectives of eliminating or reducing inappropriate patterns of nasalization, oral distortions, and "compensatory" articulation gestures.

In discussing the role of speech therapy in the management of velopharyngeal dysfunction, it must be understood that *in most cases the velopharyngeal symptoms cannot easily be separated out from other speech production disorders.* Therefore, therapy usually must focus on more than one source of speech errors and clinicians may take diverse approaches to staging therapy goals and/or managing different aspects of the global disorder. Just as there are numerous approaches to the surgical correction of cleft lip and palate, there are a variety of opinions and approaches to the treatment of speech impairments associated with clefting.

Diagnostic Speech Therapy

A common surprise for most nonspeech pathologists is that a child with velopharyngeal dysfunction who currently is receiving speech therapy or has received many years of therapy, may never have engaged in training focused specifically on improved velopharyngeal function. Unfortunately, without direct knowledge of a patient's therapy program or appropriate documentation of programming, the mere fact that a child has received speech therapy is no insurance that velopharyngeal function itself has ever been a target of therapy.

Therefore, when velopharyngeal function is borderline or variable, it is appropriate to determine whether it can be improved. A careful review of previous therapy may suggest that suitable efforts have been made to stimulate improved closure. In such cases, further diagnostic attempts may not be necessary. However, frequently such a history cannot be documented and a period of "diagnostic therapy" may be warranted.

When velopharyngeal function is variable, as it often is, particularly in young children with cleft palate, an important source of diagnostic information is "stimulability testing". A cornerstone of modern speech therapy is that a child's ability to be stimulated for improved speech production through auditory and visual (and in some instances tactile) models and cues is a good prognostic indicator of potential for long-term improvement. Such stimulability testing is particularly useful in the child with variable velopharyngeal function. It can provide valuable information about whether behavioral management is likely to remediate the velopharyngeal symptoms or if such amelioration of symptoms appears unlikely.

Morris suggests there are two major subgroups of patients with marginal velopharyngeal dysfunction and proposes that the two groups can be distinguished by the response to short-term therapeutic intervention.

The first group is the *almost-but-not-quite* (ABNQ) subgroup. This group tends to present with mild and consistent nasalization of speech, which is highly consistent among and within tasks. Morris suggests:

> speech training is not successful with the ABNQ group for the purpose of improving velopharyngeal function, that is, because the patient apparently has already extended the mechanism to the physiologic limits and does so consistently. If trial training for that purpose is provided to confirm the diagnosis of ABNQ marginal incompetence, it should be discontinued after 6 hours of treatment if no improvement in velopharyngeal function is observed.

In this case, the lack of response to highly focused therapy suggests that physical management is the treatment of choice if the impairment is significant enough to warrant intervention.

The second diagnostic group of marginal velopharyngeal function described by Morris is the *sometimes-but-not-always* (SBNA) subgroup. Patients in this group generally show marked inconsistency in velopharyngeal function. Morris indicates that some patients in this category will show improvement with training and some will not. He believes that the major diagnostic determinant in this group is the lack of positive response to directed speech therapy focused on improving velopharyngeal function. Morris suggests that patients in this group usually demonstrate that they are capable of increasing oral productions within their speech repertoire at the single sound or single word level. However, they often are unable to generalize this pattern into connected, conversational speech. Morris explains the importance of diagnostic therapy in this group:

> Diagnosis of these patients is frequently controversial. Inexperienced speech pathologists fail to interpret correctly the lack of improvement from training. Parents are falsely optimistic about the outcome because they observe the

> variance in speech production that is typical of the group. Surgeons and dentists who work often with cleft patients overinterpret the observation that the SBNA patient can perform well on single word tasks or highly specific speech activities. As a consequence, it is vital that observations about response to speech training (or rather, lack of it) be included as part of the diagnostic findings for these patients.

Morris suggests that a child in the SBNA subgroup who will be capable of achieving complete, consistent velopharyngeal closure should do so after approximately 10 hr of intensive, focused training. On the other hand, he believes that if no improvement is observed in connected speech within 10 to 20 hr of training, the diagnosis of SBNA should be made.

Unfortunately, it is these patients who present the greatest dilemma for the speech pathologist and other team members in attempting to establish an appropriate management plan. Oftentimes these patients' inability to achieve consistent closure is a complex, interrelated problem with several contributing variables. For instance it is likely that many of these patients demonstrate poor timing of velopharyngeal movements and poor coordination with other articulators and vocal tract components, which may result in the observed variability. Since the velopharyngeal symptoms in this group are quite complex in origin and manifestation, it is this group of patients that often shows limited improvement following surgical intervention. Therefore, the decision to manage these patients surgically should be undertaken with caution. Detailed counseling of the patient and the patient's family concerning realistic expectations from surgical intervention should be a high priority.

Articulation Therapy

All speech sounds can be characterized by the *place of articulation* (ie, the position of the lips and tongue for a given sound), the *manner of articulation* (ie, the way the vocal tract is altered to constrict the flow of air), and whether the sound is accompanied by laryngeal vibration (ie, voiced or voiceless).

Articulation therapy is the process of training an individual to produce a sound in the correct place of articulation and to direct the airflow in the correct manner with appropriate laryngeal voicing. Articulation therapy for individuals with cleft palate uses similar principles and techniques as articulation therapy for other clinical populations. However, the types of errors, the severity of the disorder, and the staging of therapy may be complicated by the presence or possible presence of physical impairments. Physical contributors that may impair articulation development or articulation skills include past or present velopharyngeal dysfunction, palatal fistulae, abnormalities in the skeletal relationship between the maxilla and mandible, dental abnormalities and nasal airway deviations.

Not all children with cleft lip and palate necessarily develop misarticulations. However, when articulation errors occur, they traditionally have been categorized as omissions, distortions, and substitutions (including compensatory articulations). In general, for most children with cleft palate, tongue placement is shifted posteriorly in an attempt to valve the airstream before it escapes or is diminished through an open velopharyngeal port or through an open cleft. Most commonly, the correct place of articulation is sacrificed while the correct manner of articulation is preserved.

There are several opinions in the literature regarding the relationship between velopharyngeal dysfunction and articulation disorders. Peterson-Falzone sums up the controversy:

> The relationship between velopharyngeal closure and speech is rarely as clear-cut as both the clinician and the research worker would like. Studies relating specific speech problems (and the severity of those problems) to measurements of velopharyngeal closure have often yielded contradictory and confusing results because of problems in measuring both speech output and the function of the velopharyngeal system.

In spite of the research data necessary to confirm a causal relationship between velopharyngeal insufficiency and resulting articulation errors, most clinicians would agree that many of the misarticulation patterns found in patients with cleft palate are the result of past or present velopharyngeal insufficiency. There is no disagreement that these articulation errors should be addressed in therapy. However, there are several opinions in the literature regarding the timing of articulation therapy in general and, in particular, its value (if any) in managing velopharyngeal dysfunction.

No one would argue that articulation therapy cannot eliminate a large velopharyngeal gap due to the absence of adequate tissue. However, Van Demark and Hardin cite several studies that show decreased perception of hypernasality and/or audible nasal emission (ie, symptoms of velopharyngeal dysfunction) in some children following intensive articulation therapy. They suggest, "This improvement, along with concomitant improvement in articulation, may facilitate speech intelligibility for select children and minimize the need for surgical management".

Hoch et al point out that traditionally it was believed that articulation therapy for children with documented velopharyngeal insufficiency should be deferred until physical management of the velopharyngeal mechanism was accomplished. One argument for this approach is that inappropriate therapy (especially for young children) may lead to posterior tongue posturing to facilitate velopharyngeal closure, which can then result in undesirable compensatory articulation patterns that may be difficult to remediate after surgery. Hoch et al and Shprintzen et al cite evidence to suggest that articulation therapy can improve velopharyngeal function prior to physical management. These authors assert that the changes in velopharyngeal function associated with the improved articulation may alter subsequent surgical planning. Based on their experience with articulation therapy prior to surgical intervention. Hoch et al recommend the following:

> To summarize, when VPI is observed in association with compensatory articulation errors (most specifically glottal stop substitutions), we believe that surgical correction of the VPI should be deferred. This approach is contrary to the widely accepted practice of first physically managing the VPI. We suggest that it is more appropriate to eliminate the compensatory articulation errors first with a well-planned speech-therapy program. The rationale is that improvement of oral articulation may have the benefit of improving the valving of the velopharyngeal sphincter. Furthermore, in at least a small percentage of cases, correction of the oral articulation disorder will eliminate

> the need for surgery altogether. In addition, when oral articulation has been normalized the effects of surgery can be immediately appreciated and assessed.

The controversy between these two approaches continues to exist. However, with the use of direct visualization techniques, there is increasing clinical evidence that marked improvements in velopharyngeal function can be observed over time with speech therapy. It is unclear whether these changes are truly the result of "articulation therapy" alone or other behavioral therapeutic techniques included during the process of articulation therapy. However, these numerous and frequent clinical observations support the approach espoused by Shprintzen and Hoch et al and suggest that at a minimum a trial of intensive articulation therapy is warranted for some patients prior to secondary surgical management for velopharyngeal insufficiency. Certainly, well-controlled experimental studies are needed to confirm these observations and to establish a databased approach to treatment planning.

Behavioral Therapy for Velopharyngeal Dysfunction

As discussed in the section on speech therapy as a diagnostic tool behavioral therapy is often appropriate for a limited duration for patients with inconsistent velopharyngeal closure. Van Demark and Hardin suggest that the goals of behavioral management in these cases are:

> 1. Maximize the range of velar or pharyngeal wall movement when small portal openings are evident.
>
> 2. Generalize oral responses obtained during simple speech tasks (single-word production) to more complex tasks (connected speech) when inconsistent closure is demonstrated.
>
> Proponents of behavioral therapy have argued that a patient who evidences marginal closure should be able to enforce purposeful movement of the velopharyngeal structures. Opponents have pointed out, however, that despite clinical case reports of success, limited data are available to support the notion that velopharyngeal competence can be taught.

Speech pathologists who may read this chapter and wish more information regarding behavioral management techniques for velopharyngeal dysfunction should refer to the writings of McWilliams et al, Van Demark and Hardin, and Hoch et al.

In addition to traditional behavioral techniques, a variety of biofeedback methods have been described. However, little experimental data are available concerning the effectiveness of instrumental methods in managing velopharyngeal dysfunction. Until such data are available, biofeedback therapy should be considered experimental. Nevertheless, individual reports from several centers and investigators suggest that instrumental biofeedback techniques may have significant promise as treatment modalities and should be investigated further.

The goal of all of these methods - articulation therapy, behavioral therapy, and biofeedback alike - is to facilitate consistent velopharyngeal closure first at the single-sound level, then expanded to all sounds in all sound contexts, single words, short phrases, modeled connected speech, and, ultimately, spontaneous conversational speech. Therapy should begin at the lowest level at which the child can achieve success and progress at a rate that challenges the child but allows continued success. If progress is consistently observed it is reasonable to continue with behavioral management. However, although behavioral management for children with marginal, variable velopharyngeal dysfunction can be quite successful and rewarding in some cases, *therapy should be continued only as long as progress can be documented.* It is just as unethical to relegate a child to prolonged speech therapy with no success as it is to refer a child with behavioral velopharyngeal symptoms to surgery. The critical variable in engaging in behavioral therapy for these patients with variable, marginal velopharyngeal dysfunction is to know when "enough is enough". In many cases, surgical intervention is far more humane than the cruelty of continued, prolonged behavioral therapy that is doomed to failure. In these cases, the potential for reduced self-esteem and a personal sense of failure outweighs any commitment to the ideology of behavioral management strategies.

Postsurgical Speech Therapy

Many patients undergo surgical intervention for velopharyngeal dysfunction with or without the benefit of a comprehensive presurgical evaluation. Some of these patients will present postoperatively with continued symptoms of velopharyngeal dysfunction. In these cases, comprehensive evaluation is even more critical than in the preoperative situation. Depending on the approach to presurgical speech therapy, many, perhaps even most, cleft patients will require continued therapy for remediation of developmental errors, placement and manner errors, compensatory articulation errors, or phonation disorders that could not be remediated presurgically. However, in many cases with persistent velopharyngeal dysfunction, where comprehensive presurgical evaluation was not conducted, it is possible that the etiology of the velopharyngeal symptoms was not properly symptoms may be diminished but residual articulation errors and learned habits persist and must be addressed through behavioral therapy. In other cases surgery may have been the appropriate management choice but failed to provide an adequate closure mechanism. And, in some cases, surgical intervention may not have been the appropriate course of management and therefore, no noticeable change in velopharyngeal function will be evident.

It is often necessary and appropriate to conduct a period of postsurgical evaluation and possibly more diagnostic therapy to understand the "apparent surgical failure". However, it should be understood that such cases will be less likely to occur if adequate differential diagnostic information was obtained and utilized by the surgeon prior to surgery. Said another way, appropriate differential diagnosis and presurgical diagnostic speech therapy should result in a higher percentage of surgical success. This leads to happier patients, satisfied families, and much happier surgeons.

However, if postsurgical therapy appears indicated, it should be based on the same comprehensive evaluation process suggested for the new, unoperated patient. *It should never be assumed "a priori" that the patient has been given an adequate velopharyngeal mechanism that he/she now simply needs to learn to use effectively.* Without data to support such a claim, this assumption is simplistic and dangerous. And, in many cases, postsurgical speech therapy will be of little value. Once again, it is unfair to the patient, the family, and the speech pathologist to relegate the patient to a prolonged period of therapy that has little chance of effecting improvement or diminishing symptoms. This is a far too common occurrence after an "apparent surgical failure". Too often, the surgeon seems to act as if he/she has run out of options. So, the patient is turned over to the speech pathologist to "fix the remaining problem". This set of circumstances is a setup for failure for both the patient and the therapist.

Prosthetic Management

Prosthetic treatment of velopharyngeal dysfunction has been available for many decades. LaVelle and Hardy define satisfactory prosthetic management as follows:

> 1. An *optimum result* occurs when the prosthesis results in palatopharyngeal port closure during speech production except in association with production of nasal consonants; that is, the resulting pattern of closure would be essentially normal.
>
> 2. The result is considered *successful* when there is palatopharyngeal closure throughout speech production.
>
> 3. The result is considered only *desirable* when the palatopharyngeal port area is reduced so that incompetence is a relatively minor speech physiology problem.
>
> Improvement in speech is expected, but the multiple speech physiology deficits of many of the patients for whom this management procedure is appropriate frequently negate the goal of normal speech production.

There are two basic types of speech prostheses: speech bulbs (or obturators) and palatal lifts. The first option is used when there is inadequate tissue to achieve velopharyngeal closure, resulting in a gap that can be filled with the speech bulb. A palatal lift, on the other hand, generally is reserved for cases where there is adequate tissue, but the control, coordination, or timing of movements is impaired.

Traditionally and frequently, a speech bulb or obturator is prescribed for a patient when there is clear velopharyngeal insufficiency but there is some contraindication for surgical intervention. Although historically this has been the most common use of a speech prosthesis, prosthodontic appliances can and are being employed in at least two diagnostic situations.

In the first circumstance the patient has been evaluated and shown to have incomplete velopharyngeal closure, either mild or severe. However, for any number of reasons it is unclear to the speech pathologist and other team members whether management of the velopharynx alone will provide a noticeable or significant improvement in speech intelligibility or speech quality. In these cases a trial period of prosthetic management usually will provide the diagnostic information needed to establish a management plan. Essentially then, this use of a prosthesis is a "reversible" test of the effect of surgical management of the velopharynx.

A second role for prosthetic management in the diagnostic process is to employ a prosthesis to determine whether dynamic velopharyngeal activity can be stimulated or improved. Some reports have shown that prosthetic management may improve or facilitate motion of the velum, posterior pharyngeal wall, or lateral walls in some patients. Although these findings are not universal and remain controversial, a trial of prosthetic management may be appropriate in some patients for whom little or no velopharyngeal movement is noted, and surgical intervention, therefore, would require near or complete obstruction of the nasopharyngeal airway. If improved muscle or structural function can be demonstrated, this information might be useful in altering an existing diagnosis or management plan.

Additionally, in cases of patients with dysarthria (either congenital or acquired), it is possible in some cases for management of the velopharynx with a palatal lift to free other vocal tract components to begin to function more normally.

As has been discussed earlier in this chapter, the velopharynx is only one valve in a series of interrelated valves that make up the vocal tract. For some patients with complex, multifactorial speech production disorders, as in cerebral palsy or other conditions that include paralysis or paresis of the vocal tract, management of the velopharynx may stimulate activity of other vocal tract components as well as activity of the actual velopharyngeal mechanism itself. In selected cases such as these, a trial period of prosthetic management will provide critical diagnostic information about the continued course of management.

Although there have been advances in materials and techniques for manufacturing the actual prosthetic appliances, the dental specialist historically has been dependent primarily on perceptual judgments of speech to determine the function and fit of a prosthesis in a given patient. Oftentimes this was an arduous process, requiring multiple visits to both the prosthodontist and the speech pathologist. With the advent of more objective, instrumental methods, several reports emerged describing the benefits of instrumental methods for facilitating prosthesis manufacture and assessing prosthesis fit.

Perhaps the greatest advance, which apparently has increased both interest in and utilization of speech prostheses, is the increasing availability of fiberoptic nasendoscopy in the management of patients with velopharyngeal dysfunction. Direct visualization of the velopharynx in conjunction with auditory perceptual ratings allows more precise evaluation, fabrication, and tailoring of prostheses for palatal management. Flexible fiberoptic nasendoscopy is the method of choice for direct observation of the velopharynx during dynamic speech activity. Endoscopic evaluation prior to fabrication of the initial prosthesis defines asymmetric or unusual closure patterns. This information is used by the prosthodontist in the fabrication of the prosthesis to provide the best possible fit. Video documentation of the endoscopic evaluations also allows review by any specialists involved in the patient's care prior to decisions regarding management. Additionally, if the endoscopic evaluations are videotaped, there is permanent documentation, and successive evaluations allow for accurate assessment of change over time. This approach allows the prosthodontist and speech pathologist to work together closely to provide a highly individualized and tailored prosthesis for the patient.

D'Antonio et al analyzed retrospective data comparing prosthesis fitting with and without the benefit of nasendoscopy used in a team approach with a speech pathologist and prosthodontist for fitting palatal prostheses. The authors interpreted their data to show that the number of patient visits and the time required to achieve optimal fit were reduced when using direct visualization in a team approach compared with the traditional trial and error approach by clinicians working in isolation. The authors suggested that the reduction of time between initial evaluation and adequate fit benefited the patient and the clinician in time, cost, and particularly, in patient compliance.

The procedure of speech bulb reduction has gained popularity recently. It was first described by Blakely in the 1960s when he suggested that for a small number of patients gradual reduction of the bulb portion of a pharyngeal bulb obturator can result in increased velopharyngeal movements sufficient enough over time that the prosthesis may become unnecessary in some cases. McGrath and Anderson have become the main proponents of this method. Investigators from several centers have reported cases where speech bulb reduction has been effective in improving velopharyngeal function prior to surgical intervention or completely eliminating all symptoms of velopharyngeal dysfunction so that the speech appliance could be discontinued. This method appears to have great clinical potential. Unfortunately, there is still a paucity of experimental and/or prospective data to support many of the contentions made by the proponents of speech bulb reduction as a routine clinical practice. Necessary information is needed regarding the patients for whom this procedure is most likely to be of benefit, objective documentation of change in structural movements, and the long-term stability of improvements in function.

Of most appeal is the claim that prostheses can be used in patients with little or no lateral wall motion to prepare them for pharyngeal flap surgery such that the prescribed flap will not necessarily need to be as wide and potentially obstructive if lateral wall activity is not present. Although this concept has great intuitive appeal and may in the future prove to be correct, there are no prospective data to support or negate this hypothesis at the present time. In spite of the controversy surrounding some of the claims of prosthetic management of velopharyngeal dysfunction, this continues to be an underutilized form of management and is an active and exciting area for future research.

Surgical Management

The obvious first goal of any surgical procedure designed to correct velopharyngeal dysfunction is obtaining velopharyngeal competency. The not-so-obvious second goal should be the creation of velopharyngeal competency with minimal obstruction of the airway. The pharyngeal flap and pharyngoplasty operations are antiphysiologic in terms of their effect on nasal airway physiology; that is, some degree of nasal airway obstruction will result. Depending on individual patient differences in etiology of the velopharyngeal inadequacy, anatomy, type of operation performed, and postoperative healing, the degree of nasal airway obstruction will vary from imperceptible to significant.

The maximum resultant airway effect may be obstructive sleep apnea. The head and neck surgeon should correct any mechanical airway obstruction evident on physical examination prior to performance of a pharyngeal flap or pharyngoplasty, so as to lessen the detrimental airway effects of the procedure as much as possible. This may include conservative nasal septal reconstruction, turbinoplasty, and nasal valve reconstruction, especially in the cleft child. Children with large "kissing" palatine tonsils should undergo tonsillectomy prior to performance of a pharyngeal flap or pharyngoplasty. Prior removal of tonsils improves the technical performance of the pharyngeal flap or pharyngoplasty procedure, as well as avoids obstruction of the ports by large tonsils following healing. Reassessment of speech should follow the tonsillectomy once the oropharynx has completely healed (3 to 6 months), as some children may show significant improvement of hypernasality and nasal emission and may not require further operative therapy. If further surgical therapy is indicated to improve speech following complete healing of the oropharynx after tonsillectomy, instrumental methods may be used to reassess the velopharyngeal defect, allowing "improved tailoring" of the pharyngeal flap or pharyngoplasty.

A rare child with clinical evidence of adenoiditis and recurrent upper respiratory infection may undergo adenoidectomy approximately 3 months prior to the pharyngeal flap or pharyngoplasty (performance of the adenoidectomy at the same time as a pharyngeal flap would not allow access to the nasopharynx in the case of postoperative bleeding). Rarely a child will have hypertrophied lingual tonsils and require laser lingual tonsillectomy.

Following a pharyngoplasty or pharyngeal flap, children who have significant nasal obstruction or obstructive sleep apnea should undergo videonasendoscopy to assess the size of the port (pharyngoplasty) or ports (pharyngeal flap), as well as be reexamined for obstruction at another site that was not appreciated preoperatively.

Numerous methods of surgical therapy have been utilized for the correction of velopharyngeal inadequacy. Five basic categories of operations may be utilized: palatine tonsillectomy, veloplasty with redirection of velar musculature, the double-opposing Z-plasty of Furlow, the superiorly based pharyngeal flap, and pharyngoplasty. The choice of operation is largely based on direct observation of attempted velopharyngeal closure by videonasendoscopy for each individual patient.

Can surgeons consistently obtain adequate velopharyngeal function and maximal nasal airway function by tailoring the surgical procedure to the pattern of velopharyngeal closure for each patient? The surgical techniques described in this chapter attempt to create a tailor-made speech operation to fit each patient's individual velopharyngeal defect. Individual patient differences in wound healing, tissue contraction, and ultimate scar formation are not subject to control by the surgeon and may therefore make the techniques described less precise than one might wish. The final answer to the question will come only from prospective studies involving preoperative and postoperative measurements of speech and velopharyngeal function employing perceptual and instrumental evaluation methods.

Palatine Tonsillectomy

If videonasendoscopy shows inadequate seal between the velum and the lateral pharyngeal walls because of projecting superior poles of the palatine tonsils, tonsillectomy should be considered. This should be a "class act" tonsillectomy, with fine dissection of the tonsil, including the capsule, and without damage to the underlying palatoglossus, palatopharyngeus, and constrictor musculature. Dissection by unipolar cautery is avoided. Bleeding points are controlled with gentle packing and bipolar cautery. Failure of tonsillectomy to correct velopharyngeal dysfunction does not preclude the use of another operation at a later date.

Veloplasty with Redirection of Velar Musculature

The generally accepted definition of submucous cleft palate consists of the triad of bifid uvula, notching of the posterior border of the bony palate with loss of the posterior nasal spine, and diastasis of the soft palate musculature. These findings are usually obvious on physical examination. Congenital palatopharyngeal disproportion has been used to explain the presence of velopharyngeal dysfunction in a child without obvious evidence of clefting, either over or submucous. In this condition, palatal function is inadequate either because the soft palate is too short or the anteroposterior dimension of the nasopharynx is too deep. However, substantial evidence exists to demonstrate that the majority of so-called short palates (congenital palatopharyngeal disproportion) represent unrecognized submucous cleft palate, or at least a microform of over cleft palate: (a) muscle diastasis is visible on the nasal surface of the soft palate during videonasendoscopy, (b) muscle diastasis or abnormal insertion of velar musculature is directly visible at the time of the pharyngeal flap operation when the soft palate is transected and dissected, (c) the frequency of middle ear disease approaches that for overt cleft, and (d) the occurrence of short palate in family lines where overt cleft is present.

If a diastasis of velar musculature is present (such as for the case of a classic submucous cleft palate), videonasendoscopy will show either a midline furrow on the nasal surface of the soft palate, or a lack of musculus uvulae bulge, or both. This anatomic defect, coupled with good palatal motion in general and good lateral wall motion, will result in a small midline defect in velopharyngeal closure. These patients can be corrected with transection of the soft palate and redirection of the velar musculature. The operation begins with a midline through-and-through incision of the soft palate, bisecting the bifid uvula if present. The velar musculature is dissected from its attachments to the posterior edge of the hard palate, as well as the overlying nasal and oral mucosa. The dissection should be accomplished with fine forceps and scissors, utilizing adequate illumination and loupe magnification. The velar musculature is thus redirected from an anteroposterior direction to a horizontal direction, and sutured in the midline. The nasal mucosa is closed prior to the suturing of the muscle. Final closure involves uvuloplasty if needed, and closure of the oral mucosa. This technique is best utilized for correction of velopharyngeal dysfunction secondary to the unoperated submucous cleft palate. For the case of a previously repaired cleft palate, another option should be chosen.

Furlow Double-Opposing Z-plasty

In 1978, Furlow introduced a technique of cleft palate repair by double-opposing Z-plasty, which he described in detail. As noted by other authors who have utilized this method, speech results have improved when compared with other palatoplasty techniques.

The surgical technique is not described in detail here, as the surgical steps are thoroughly explained in Furlow's classic paper. Although probably not the best method for correction of velopharyngeal inadequacy following cleft palate repair, Furlow's method may be ideal for surgical repair of the unoperated submucous cleft palate, due to several theoretical advantages. When compared with veloplasty with redirection of the velar musculature, the double-opposing Z-plasty method reorients the musculature and simultaneously lengthens the soft palate. Although palatal length is an inconsistent determinate of velopharyngeal adequacy, healing via Z-plasty may be superior to healing via a straight-line closure that may tether the soft palate and result in less mobility. Another possible advantage involves the fact that the bulk of velar musculature is relatively small in many children with a submucous cleft palate. Furlow's method accomplished redirection of the velar muscles via separation of the overlying mucosa on only one versus both sides, which may result in less surgical trauma, less postoperative scarring with healing, and thus improve palatal mobility.

Superiorly Based Pharyngeal Flap

If videonasendoscopy reveals moderate (approximately one-third the distance to the midsagittal plane) to excellent (two-thirds the distance to the midsagittal plane) lateral pharyngeal wall motion, construction of a pharyngeal flap is probably the best option for surgical management. Patients with minimal lateral wall motion may be candidates, but are at risk for hyponasality and airway obstruction. The degree of competency obtained from a pharyngeal flap will depend on the size of the constructed lateral port, tailored to the degree of lateral pharyngeal wall motion. The goal is to perform an operation that will create openings between the nasopharynx and the oropharynx just small enough to allow competency, yet large enough to minimize the amount of nasal obstruction. Three elements are essential to the creation of a flap that, it is hoped, will achieve this: flap width, height or level of the flap base, and control of the lateral port size.

Based on the knowledge of the defect in velopharyngeal closure (and the degree of lateral wall motion) observed by videonasendoscopy for the particular patient, the width of the flap is constructed relatively thin (for excellent or better lateral pharyngeal wall motion), of moderate width (for moderate pharyngeal wall motion), or relatively wide (for less than one-third the distance to the midsagittal plane lateral pharyngeal wall motion). Based on the preoperative knowledge of the vertical level of maximal lateral wall closure, the length of the flap is determined. The base of the flap is estimated such that the flap will leave the posterior pharyngeal wall at a right angle to join the palate without tension, and at the same time be opposite the position of maximum lateral wall movement. The markings (flap width and length) are completed prior to the incisions.

The soft palate is sharply incised in the midline, from the posterior edge of the hard palate through to and bisecting the uvula. The oral mucosa is incised to a point approximately 2 mm from the posterior edge of the bony palate, whereas the nasal mucosa is incised approximately 0.5 cm from the edge of the posterior hard palate, in order to facilitate later placement of the pharyngeal flap. The tissue contents of the soft palate are carefully inspected at this time, and the presence of muscle fibers are noted, as well as their direction. Children with cleft palates who have undergone previous primary palatoplasty with intravelar veloplasty, will be noted to have muscle fibers that run in a transverse direction as a result of the previously created levator sling. This levator sling will be carefully reapproximated at the conclusion of the procedure. Children with submucous cleft palate will be noted to have the classic misdirection of the levator muscle fibers, curving cephalad into the posterior edge of the hard palate. These muscles are dissected at a later time in the procedure and an intravelar veloplasty is performed.

The previously marked lateral borders of the pharyngeal flap are incised through mucosa. The divided palate allows easy access into the nasopharynx for extension of the incisions in a cephalad direction to define the level of the flap base. By scissor dissection, the incisions are extended deep through the underlying superior constrictor muscle to the plane of the alar layer of prevertebral fascia. Blunt dissection is completed in the plane between the superior constrictor muscles and the alar layer of prevertebral fascia that connects the two lateral incisions. Good hemostasis is obtained, and the nasopharynx and soft palate are carefully inspected a third time. The lateral incisions are carried to a point cephalad in the nasopharynx, such that the base of the flap will be opposite the position of maximum lateral pharyngeal wall motion, as noted previously by nasendoscopy. A useful landmark is the prominence of the first cervical vertebra. If lateral pharyngeal wall movement is relatively high in the nasopharynx, the base of the flap must also be high in the nasopharynx, opposite the level of maximum lateral pharyngeal wall motion. Once the final base of the pharyngeal flap is selected, the length of the flap is estimated and the lower end of the flap is then transected.

Turnover flaps of soft palate nasal mucosa are outlined and dissected. The plane of dissection is between the nasal mucosa and submucosa and the musculature of the soft palate. The borders of the turnover flaps are created beginning several millimeters from the peak of the nasal mucosa incision and carried approximately 1.5 cm directly lateral. The incision is then turned 90° inferiorly and laterally into the area of the upper portion of the posterior tonsillar pillar. The incision may be extended into the lateral pharyngeal wall as necessary, depending upon the intended size of the lateral pharyngeal port (as described below). If the patient has a submucous cleft palate and the levator muscles are abnormally inserted, they are dissected free at this time. Scissor dissection is used to create a plane between the levator muscles and the overlying oral mucosa. The abnormal insertion of the muscle fibers into the posterior hard palate is transected and the flap of fibers is moved posteriorly until they line a transverse plane.

The flap is attached to the palate. Different sizes of standard endotracheal tubes are used to gauge the size of the lateral port. Small pediatric endotracheal tubes (sizes 3 to 3.5) are used in patients who demonstrate minimal to moderate lateral wall motion. Larger endotracheal tubes (sizes 4 to 4.5) are used in patients who demonstrate moderate lateral movement, and sizes 5 to 5.5 are used in patients who demonstrate excellent lateral wall motion. The flap is attached to the nasal layer of the palate, beginning at the anterior peak and progressing laterally to the level of the lateral port. The final "lateral port control stitch" attaches the appropriate point of the lateral edge of the flap to the appropriate point of the incision in the superior aspect of the posterior tonsillar pillar or lateral pharyngeal wall, so that the tissues are snug around the endotracheal tube.

To assist adaptation of the flap to the palate, a single absorbable suture is placed, grasping the flap muscle layer in the midline, and brought out through the peak of the palatal incision and tied over the oral mucosa. The turnover flaps are then sutured into place so that they line the raw inferior surface of the pharyngeal flap. The flaps may be turned in an arc - anywhere from 90° to 180°, and trimmed appropriately for the final fit. The nasal mucosa on the posterior surface of the uvula is then approximated. Mattress sutures are used to close the divided levator sling, or to join the muscle fibers if an intervelar veloplasty has been performed. Several of the sutures may incorporate the underlying flap muscle layer to aid in adaptation of the flap to the palate.

The anterior oral mucosa is closed. The posterior pharyngeal wall defect is then closed. The defect is sutured loosely toward the base of the pharyngeal flap, to avoid tubing of the flap base and unnecessary strain on the lateral port.

Good postoperative care is paramount for the success of pharyngeal flap surgery and includes room mist, adequate hydration, intravenous antibiotics, and close monitoring of the airways. The endotracheal tubes used to control the lateral port size are removed the morning after the surgical procedure.

Pharyngoplasty

If videonasendoscopy shows relatively good velar movement and poor to absent lateral wall motion, pharyngoplasty may be the best surgical option for rehabilitation of velopharyngeal inadequacy. The technique described and illustrated is similar to that described by Jackson. A horizontal incision is accomplished in the posterior pharyngeal wall, opposite the level of maximum attempted closure of the velopharyngeal port, as previously viewed via videonasendoscopy. Bilateral superiorly based mucosal-palatopharyngeus muscle flaps are elevated, with the medial incision connected to the horizontal posterior pharyngeal wall incision. The mucosal-palatopharyngeus muscle flaps are turned 90° and sutured end to end in the midline posterior pharyngeal wall. The technique narrows the nasopharyngeal dimension laterally and posteriorly so that the velum can make contact. It is important that the mucosal muscle flaps are placed opposite the level of the maximum velopharyngeal port closure.

Conclusion

In summary, the goal of this chapter was to explain the need for *differential diagnosis* in the evaluation and management of velopharyngeal dysfunction. Through comprehensive, multimethod evaluation, appropriate diagnosis can lead to effective individualized management. It is the role of the speech pathologist to collect and interpret as much data as necessary to establish a provisional differential diagnosis and to interact with other medical specialists, prosthodontists, and surgeons, to develop an appropriate intervention strategy. In brief, the message of this chapter is *differential diagnosis is the first step toward differential management.*