Pediatric Facial Plastic and Reconstructive Surgery

James D Smith, Robert M Bumsted

Chapter 12: Cleft Palate

Craig W Senders and Jonathan M Sykes

Historical Perspectives

A French dentist named Le Monnier Rouen was the first to attempt closure of a congenital cleft palate in the mid-18th century. The technique involved cauterization of the edges of the cleft and holding the edges together with suture. Presumably this was carried out on cleft of the soft palate only.

Attempts at closing the hard palate were not successful until the 19th century when the importance of reducing tension was discovered. In 1928, Dieffenbach recommended separating the oral mucosa from the bone of the hard palate as a method of closing the palate. Additionally, he used lateral osteotomies to aid in the closure. The importance of utilizing relaxing mucosal incisions to reduce tension, became popularized in the mid-19th century by Pancoast (1843), Warren (1828, 1843) and Fergusson (1845).

The 20th century has brought about refinements in techniques, with attention to improving function. In 1937, Wardill and Kilner independently described a technique to lengthen the soft palate involving V-to-Y advancement of palatal mucosa. This procedure is limited by the difficulty of providing a similar advancement of the nasal layer. Additionally, this technique requires leaving large areas of exposed palatal bone, which may inhibit future palatal growth. Attempts at increasing the length of the soft palate without inducing scarring in the hard palate resulted in numerous modifications, from using buccal or nasal flaps to skin grafts. A definitive improvement in velopharyngeal function has not been found.

In the 1930s, Veau was the first to focus upon the anatomy in clefts. Later, Kriens developed the intravelar veloplasty. This technique releases the attachments of the malpositioned tensor veli palatini and levator veli palatini muscles from the posterior portions of the hard palate and realigns these muscles. Retrospective studies based upon historical controls have shown improved velopharyngeal function. The only prospective randomized study found no improvement in function. Although realigning the malpositioned palatal muscles has a strong theoretical advantage, the scarring induced by isolation of the muscle may limit palatal function.

In a further attempt to improve function of the soft palate after repair, Furlow developed the double reversing Z-plasty palatoplasty. This technique utilizes lengthening of the palate by a Z-plasty with realignment of the palatal musculature. There is less muscle dissection than in the intravelar veloplasty. This may result in less scarring and improved function. Unfortunately, the only published studies are retrospective and based upon historical controls.

Another approach popularized by Schweckendiek is a two-stage repair of the palatal defect. In the first stage, only the soft palate is closed. The hard palate cleft narrows after soft palate repair. Typically, the hard palate is then closed at 5 to 6 years of age. This allows the benefit of early soft palate movement and function and minimizes the detrimental effects on facial growth of the hard palate repair. It is paramount to this technique to use a palatal obturator in the interval prior to hard palate repair. This requires a talented, interested, and dedicated prosthodontist. At most centers this is not practical.

At the present time, the most popular repairs are single-stage repairs, although there has been a resurgence of interest in the Schweckendiek-type procedure in a number of centers across the country. The Wardill-Kilner V-to-Y advancement is used at a number of centers, although studies have not shown this to improve velopharyngeal function. Efforts at reestablishing the muscular continuity via the intravelar veloplasty or the Furlow palatoplasty have gained significant popularity, although improvement of velopharyngeal function has not been absolutely determined.

As we approach the 21st century there are four basic principles of palatoplasty: (a) multiple layer closure, (b) elimination of tension at the suture line, (c) reconstruction to allow improved velopharyngeal function, and (d) minimization of scarring of the hard palate to limit inhibition of midfacial growth.

Anatomy

The hard palate is divided embryologically into the primary and secondary palate. The primary palate forms at the time of lip formation between the 4th and 6th week of gestation. The secondary palate forms between 7th and 12th week. The medial and lateral incisor teeth develop from the primary palate, the remainder of the teeth develop from the maxilla.

A cleft of the lip can result in a cleft alveolus and a cleft between the primary and secondary palate. Additionally, this cleft between the primary and secondary palate can inhibit the fusion of the palatal shelves, from anterior to posterior. Thereby a cleft of the lip can result in a complete cleft of the lip and palate.

In a complete bilateral cleft lip, the primary palate (often referred to as the premaxilla) is not fused with the secondary palate and is attached to the vomer. In severe cases the premaxilla is displaced anteriorly and the lateral alveolar segments are collapsed. In this situation the vomer is not attached to either palatal shelf. However, variations can occur where the vomer is attached to one or the other palatal shelf.

A cleft of the secondary palate occurs between the 7th and 12th week of gestation. This is a different embryological malformation than in a cleft lip or a cleft lip with a cleft palate. The expression is variable. A submucous cleft palate is a mild expression. In other cases the cleft will involve the entire secondary palate. In addition, this cleft can be bilateral, with the vomer unattached to either palatal shelf. Typically, the vomer is attached to one or the other palatal shelf.

The major blood supply of the hard palate is from the greater palatine artery. This artery passes through the greater palatine foramen, which is located next to the second molar

tooth. Additional sources of the blood supply of the hard palate come from the incisive artery as well as the lesser palatine artery and the ascending pharyngeal artery.

The fost palate is composed of five paired muscles: tensor veli palatini muscle (TVPM), levator veli palatini muscle (LVPM), palatoglossus muscle, palatopharyngeus muscle, and musculus uvulae. Both the TVPM and LVPM arise from the base of the skull and a portion from the auditory tube. A portion of the TVPM arises from the lateral portion of the auditory tube. It is thought that the TVPM opens the eustachian tube. The TVPM passes anterior to the hamulus of the pterygoid plate where it makes a 90° turn to become the palatal aponeurosis. The LVPM arises from the medial aspect of the base of skull and attaches to the medial aspect of the auditory tube. It travels anteriorly into the soft palate and interdigitates with the LVPM of the opposite side. This muscle is important in elevating the soft palate and providing velopharyngeal closure. It is thought that the LVPM and the salpingopharyngeus muscles have no effect on auditory tube opening, because of their anatomic origin.

The palatoglossus forms the anterior tonsillar pillar and inserts into the posterior portion of the soft palate. The palatopharyngeus is a large muscle that forms from the lateral portion of the pharynx and makes up the posterior tonsillar pillar. In the soft palate it forms two fasciculi, divided by the LVPM. The anterior fasciculus inserts into the palatal aponeurosis. The posterior fasciculus interdigitates with the opposite posterior fasciculus, posterior to the LVPM. A few fibers of the posterior fasciculus pass nasally over the LVPM to join the anterior fasciculus. The musculus uvulae is a small muscle that attaches to the posterior nasal spine and inserts into the uvula of the soft palate. It can play an important role in velopharyngeal closure.

A cleft palate distorts the musculature of the soft palate. Instead of the LVPM and the palatopharyngeus muscle interdigitating with the opposite muscle, these fibers tend to assume a more anterior projection and insert along the margin of the cleft and into the posterior portion of the hard palate. In a palate repair that does not realign the muscles, it may be difficult for these muscles to function normally.

Indications for Surgery

The most important reason to repair a cleft palate is to allow the patient to develop normal speech. Without repair of this large oronasal fistula, the patient is condemned to a life of poor communication. Unfortunately, hypernasal speech is less well tolerated than hyponasal speech.

Surgical repair of a cleft palate also serves to improve swallowing. However, by the time most patients have their palatal surgery, adaptation to the cleft has occurred and daily nourishment is no longer an issue. In carefully selected cases, a palatoplasty is performed primarily to improve swallowing in a child with developmental disabilities.

Auditory tube dysfunction with resultant chronic otitis media is highly prevalent in patients with cleft palate. To date, studies have not demonstrated that repair of the cleft palate results in improved auditory tube function. It is unknown whether this is due to an abnormal origin of the muscle in relation to the auditory tube or an inability to improve the function

of the muscle.

Timing of the Surgical Repair

Palatoplasty is performed on patients as young as 3 months of age and as old as 2 years of age. At the majority of centers the palatal repair occurs between 9 and 18 months of age. Those who perform the repair on 3-month-old patients generally utilize a two-stage repair. The soft palate is repaired first, and the hard palate is repaired at a second stage.

The theoretical benefit of repairing the palate early is improved muscular function and ultimately improved speech. Many believe that having an intact soft palate during swallowing, prior to the development of speech, will improve function during speech. However, palatal surgery is more difficult on younger patients and the risks of a blood transfusion increase.

There are no prospective randomized studies to compare early and late palatal closure. However, there are several retrospective studies that suggest there may be a benefit to closing the palate prior to 18 months of age. Overall there is a trend toward earlier palatal closure.

In over 150 palatoplasties that have been performed at the University of California (UC)-Davis, no patient has required a transfusion. An added benefit is that the older child does not require feeding by bottle, so the postoperative repair will not suffer additional trauma.

Preoperative Evaluation and Preparation

The greatest concern at the preoperative evaluation is whether the mandible has grown to an adequate size to minimize the risk for postoperative airway obstruction. In the UC Davis experience, approximately 1% to 2% of the patients have required reintubation for postoperative airway obstruction. In a patient with Pierre Robin sequence, if the mandible does not demonstrate catch-up growth, delaying surgery until 24 months of age is appropriate.

Many centers allow bottle feeding immediately following palatoplasty. The UC Davis craniofacial anomalies team has been slow to follow this trend as randomized studies have not been completed.

A child that is not above the 5th percentile for weight and height can be anticipated to have a complicated postoperative course. Typically these patients are neurologically delayed and a delay in the surgery will not further compromise their speech development. However, chronic airway obstruction and cardiac anomalies must be ruled out in these failure-to-thrive patients.

Surgical Technique

Von Langenbeck Procedure

The von Langenbeck procedure is a relatively popular procedure that was introduced in 1859. The procedure is optimally suited for clefts of the secondary palate only, but can be performed on complete clefts if the anterior portion of the anterior palate is not closed. The

technique involves a releasing incision laterally, leaving a bipedicle flap to provide blood supply to the palatal segments. This procedure does not add any lenght to the soft palate.

The edges of the cleft are incised down to bone over the hard palate and extend into the soft palate all the way to the tip of the uvula. Relaxing incisions are made just medial to the alveolus and extend around the maxillary tuberosity just lateral to the hamulus. Large mucopeiosteal flaps are elevated on the hard palate. Care is taken not to damage the greater palatine artery as it exits the greater palatine foramen. The soft palate muscular attachments to the hard palate are separated and blunt dissection is utilized along the posterior portion of the relaxing incision. The mucoperiosteum is elevated off of the nasal layer. The incision is closed in three layers with the nasal layer being closed from anterior to posterior using 4-0 chromic suture. The palatal musculature is closed with 3-0 chromic suture or 4-0 synthetic absorbable suture. The hard palate is closed with 3-0 chromic suture or a 4-0 synthetic absorbable suture in a vertical mattress fashion. To eliminate the dead space over the hard palate, the nasal layer can be included in two of the oral mucosal sutures.

Three-Flap Palatoplasty

This is the preferred technique of the UC Davis craniofacial anomalies team for closure of clefts of the secondary palate. The three-flap palatoplasty can be utilized to lengthen the soft palate using a V-to-Y technique (Wardill-Kilner operation). However, the V-to-Y technique leaves more bone exposed. Subsequent scarring of the hard palate will cause more problems with midfacial growth and is not worth the unproven benefit of lengthening the soft palate.

This procedure involves an incision along the cleft, extending down to the bone over the hard palate and dividing the border of the cleft of the soft palate up to the base of the uvula, the sides of which are excised. Both incision must join each other anteriorly. The lateral incision is very similar to the von Langenbeck lateral releasing incision. An incision is made just lateral to the hamulus around the maxillary tuberosity, medial to the alveolus, to approximately the level of the canine tooth. An incision is then made from the apex of the cleft to the canine tooth. Incision and complete elevation is performed on one side prior to the opposite to facilitate better hemostasis.

Using a periosteal elevator the mucoperiosteum is elevated, mobilizing the large palatal mucoperiosteal flap. Care is taken not to damage the greater palatine artery as it exists its foramen. Kitners are utilized to provide blunt dissection on both sides of the greater palatine artery and improve the mobilization. This often better defines the posterior palatal shelf. Using Padgett elevators, the muscle attachments to the posterior palatal shelf are dissected free and the nasal mucosa is elevated. Blunt dissection around the hamulus with a Kitner as well as complete dissection around the vascular pedicle usually will allow the flaps to meet easily in the midline. However, if there is tension, the greater palatine vessels can be sharply dissected from the mucoperiosteal flap using a right-angle Beaver blade for several millimeters. Rarely, it is necessary to dissect the tendon of the tensor veli palatini muscle from around the hamulus or fracture the hamulus to provide greater flap mobility. Closure of the nasal layer is performed in an anterior-to-posterior direction with 4-0 chromic suture.

The muscular layer is then closed with 3-0 chromic or 4-0 synthetic absorbable suture with two or three large bites. The oral mucosa is then closed using intrerrupted vertical mattress sutures of 3-0 chromic or 4-0 synthetic absorbable suture from posterior to anterior. The dead space over the hard palate can be eliminated by including the nasal layer in two of the oral mucosal sutures. Typically the lateral incisions can also be closed unless the cleft is quite large. The saloon door effect as described by Bardach and Nosal explains this surgical mystery. Because the palatal vault typically has a significant vertical height, this adds increaed width to the palatal flaps. At the time of closure the flaps assume a more horizontal orientation and effectively leave very little exposed bone. If bone is left exposed laterally it should be covered with an absorbable hemostatic agent.

Two-Flap Palatoplasty

A two-flap palatoplasty is utilized for unilateral complete clefts of the lip and palate. This technique is very similar to the three-flap palatoplasty technique with modified palatal incisions. Along the edge of the cleft, an incision to bone is made over the hard palate and extending into the soft palate at the midpoint between the oral and nasal layers. This incision is carried to the base of the uvula, the sides of which are excised. A lateral incision is made just lateral to the hamulus, passing around the maxillary tuberosity. The incision is made from posterior to anterior just medial to the alveolus until it joins the previously made cleft incision. On the noncleft side, the lateral incision is carried past the canine tooth, to the opposite canine where it joins the cleft incision. Mucoperiosteal flaps are elevated as in the three-flap palatoplasty. Again, attention is paid to eliminating tension on the closure line and allowing a three-layer closure over the soft palate and a two-layer closure over the hard palate. Care is taken to eliminate the dead space over the hard palate by including the nasal layer with the deep portion of the vertical mattress suture. Any large areas of exposed bone are covered with an absorbable hemostatic agent.

Intravelar Veloplasty

The technique of intravelar veloplasty was described by Kriens in 1969 and 1970. This technique gained popularity in the 1970s and 1980s. Dreyer and Trier in a retrospective study using historical controls, showed the need for a pharyngeal flap dropping from 35% to 9% when using the intravelar veloplasty. However, Marsh et al in 1986 reported in a randomized prospective study that there was no statistical difference between the two groups. This suggests that the increased scarring caused by dissection and plication of the levator veli palatini muscle inhibits the function, despite its new anatomical relationship.

The technique of intravelar veloplasty can be combined with the three-flap palatoplasty, two-flap palatoplasty, or von Langenbeck procedure. All of the mucosal incisions and elevations are the same, except that in the intravelar veloplasty attention is taken to completely mobilize the levator veli palatini muscle as well as the aponeurosis of the tensor veli palatini muscle. This involves sharp dissection for approximately 1 cm from the cleft margin. Care is then taken to approximate or plicate the muscles of the soft palate to provide better function.

Double Reversing Z-Plasty

The double reversing Z-plasty was described by Millard in 1978 at the annual meeting of the Southeastern Society of Plastic and Reconstructive Surgeons. However, it was not until 1986 that studes utilizing this technique were published. This technique utilizes a Z-plasty to lengthen the soft palate, and it realigns the levator veli palatini muscle. Because the muscle is dissected on one surface only, there should be less fibrosis than in the intravelar veloplasty.

As originally described, the procedure combines the palatal lengthening of a Z-plasty and realignment of the palatal muscles with a closure of the hard palate under tension. Thus, this procedure has taken two steps forward and one step backward. The procedure is excellent for clefts of the soft palate.

The procedure involves making a Z-plasty on both the oral and nasal surface in a reversed pattern from one another. Because of the difficulty in elevating the levator veli palatini muscle from the nasal mucosa, right- and left-handed surgeons would perform the operation in mirror image. The illustrations describe the procedure for a right-handed surgeon.

An incision is made along the margin of the cleft to the base of the uvula, the sides of which are excised. A lateral incision is then made from the apex of the cleft to the hamulus on the left side. This incision is carried down to the nasal mucosa. Care is taken to try and include all of the soft palate musculature and most of the palatal aponeurosis in this flap. Using a right-angled Beaver (R) blade, the nasal mucosa is separated from the palatal musculature. Laterally, blunt dissection using scissors is most optimal. Using right-angled scissors an incision is made in the nasal mucosa from the base of the uvula to within 1 cm of the hamulus.

On the right side an incision is made through the oral mucosa from the base of the uvula toward the hamulus. This incision is carried through the epithelium and the fibrous glandular tissue overlying the palatal musculature. This flap is relatively thick as the musculature is intimately associated with the nasal mucosa. This flap is elevated sharply to the hard palate, although care is taken not to damage the greater palatine vessels. Using a Padgett elevator, the palatal musculature is dissected from the posterior border of the palatal shelves. Using right-angled scissors or a #15 blade, an incision is made from just posterior to the hamulus to the apex of the cleft. Care is taken to save 1 or 2 mm of nasal mucosa anteriorly to allow suturing. This incision separates the right palatal musculature from the hard palate. The Z-plasty limbs are transposed and the nasal layer is closed with a 4-0 chromic suture. If there is significant tension, the lateral portion of these incisions can be closed upon itself rather than stretching the Z-plasty flaps to the apex. The left palatal musculature on the left is sutured to the right lateral soft tissue (palatopharyngeus muscle) with a 3-0 chromic suture. The Z-plasty flaps on the oral layer are then transposed. Again, the palatal musculature is sutured to the palatopharyngeus of the opposite side with a deep suture. The oral layer is closed with a 3-0 chromic or 4-0 synthetic absorbable suture in an interrupted fashion. If there is tension laterally, the incision is closed upon itself rather than pulling the tip of the flap into the apex. Care must be taken at the original apex of the cleft to insure that there is a good two-layer tension-free closure.

Postoperative Care

Typical postoperative care involves a pureed diet, per cup only, for 3 weeks. All patients are placed in arm restraints. A mist tent rather than a mist hood is utilized to markedly increase oral hydration. This is particularly important during the first day as clots in the pharynx can become dry and result in upper airway obstruction. The mist tent is discontinued when oral intake becomes regular. Oxygen saturation is routinely monitored. The patients are discharged from the hospital when they are taking 90% of the calculated maintenance fluid for their size.

Complications

Bleeding

Blood loss can be minimized by the careful use of a vasoconstrictive agent and allowing time for the vasoconstriction agent to take effect. If significant bleeding is encountered during elevation of flaps, pressure is applied for several minutes. After the flaps have been elevated, it is important to search carefully for bleeding points, as an uncontrolled bleeder resulting in only 2 cc of blood loss per minute would result in 240 cc of blood loss over a 2-hour period. For most children this would require a blood transfusion.

Postoperative bleeding of any significance requires a return to the operating room to determine the source.

Airway Obstruction

Some degree of airway obstruction following palatoplasty is relatively common. Significant airway compromise may be due to tongue swelling, laryngeal edema, pharyngeal secretions, or prolonged action of anesthetic agents. Tongue swelling has been thought to be due to ischemia induced by the mouth gag. Therefore, it is prudent to release the mouth gag periodeically during the procedure. Additionally, continuous suspension of the mouth gag from the Mayo stand is not recommended.

Significant airway obstruction requires treatment. In past years a suture was placed in the tongue and a resident or nurse held the tongue out to improve the airway in troublesome cases. However, in this day and age patients with significant airway obstruction should be reintubated.

Palatal Dehiscence

A palatal dehiscence due to poor blood supply infection, poor blood supply, and/or poor suturing technique is probably best treated by observation and reoperation when scarring becomes mature. In fact, looking at the palatal incision in the postoperative period, has been termed the "evil eye". This is because there is little benefit in diagnosing a dehiscence. Looking with a tongue blade in an uncooperative patient may damage the repair. A small dehiscence often will heal on its own without sequelae. A slight separation of the uvula is seen relatively frequently and has no functional consequences. If a dehiscence is caused by a trauma, such as falling onto a hard object, then immediate repair may be beneficial.

Oronasal Fistulae

Fortunately, with proper care, oronasal fistulae are quite unusual. Additionally, many fistulae are quite small and are of no functional significance. Often, on a complete cleft of the lip and palate, there is a small oronasal fistula just posterior to the alveolar cleft. In essence, this is a continuation of the alveolar cleft. This should be repaired atr the time of alveolar bone grafting, which is typically between 5 and 11 years of age.

A soft palate fistula typically can be excised and closed in three layers. However, a fistula of the hard palate requires a larger surgery. A fistula here is best treated with relatively large mucoperiosteal flaps and closed in two layers. Often it is necessary to almost repeat the two- or three-flap palatoplasty to allow a tension-free closure. Closures that involve lesser procedures often fail in the long run.