Analysis and Classification of the Upper Lip Aesthetic Unit

Peter Raphael, M.D.
Ryan Harris, M.S.
Scott W. Harris, M.D.
Plano and Dallas, Texas

Background: Disharmonies of the upper lip aesthetic unit generally stem from tall ergotrids and/or thin lips. Comprehension and correction of such defects has been stifled by a lack of metrics and organized systems of diagnosis and treatment.

Methods: The philtral-labial score was devised to better analyze the upper lip region. Measurements were made with Adobe Photoshop CS6, and computations were performed on a standard calculator. A retrospective medical records review identified 908 patients of the senior authors’ (P.R. and S.W.H.) practices who underwent perioral rejuvenation between January 1, 2001, and July 31, 2012. Two hundred patients were randomly selected and assessed for disharmonies in three surveys that sequentially built on data points provided.

Results: When preoperative anteroposterior and lateral photographs, dental show measurements, and philtral-labial scores were available, diagnostic concordance between the authors approached 100 percent. Pattern analysis resulted in a classification system (labral classification system), designating patients as either type 0 (no defects), type 1 (thin upper lip), type 2 (long philtrum), or type 3 (both) defects. Characteristic dental show values, philtral-labial scores, and suggested treatments were paired with each type.

Conclusions: The labral classification system and its associated analytical tools serve as useful references in consultation, simplify discussion of patients with upper lip defects, furnish a practical alternative to complex algorithms, enable documentation of changes, and facilitate analysis of large sample sizes. When implemented judiciously, the tools described in this article will help surgeons confidently address upper lip problems by streamlining accurate diagnosis and guiding proper treatment. (Plast. Reconstr. Surg. 132: 543, 2013.)

CLINICAL QUESTION/LEVEL OF EVIDENCE: Diagnostic, III.

To keep pace with trends toward greater emphasis on the perioral region, fresh modes of analysis are needed. We therefore introduce a metric based on relative philtral and labial dimensions; merge it into an original classification system; and determine what information permits effective, universally consistent diagnosis.

To simplify discourse, we term the upper lip aesthetic unit “labrum,” derived from Latin for lip. This structure comprises the upper lip and ergotrid.1 Upper lip is a context-dependent label that can denote just the dry vermilion (pink lip) or further encompass the ergotrid and labial mucosa. Because this article focuses on external surface anatomy, we define it as the pigmented skin spanning the vermilion border to the labial fissure. Above the upper lip lies the ergotrid, also called the mustache area, white lip, cutaneous upper lip, or labial ledge.2–4 This trapezoidal region is delimited by the vermilion border caudally, the nasal base cephalically, and the nasolabial folds bilaterally (Fig. 1).

Dissatisfaction with labral appearance usually stems from tall ergotrids or thin lips, for which we propose the term “leptocheilos,” borrowed from Greek roots lepto and cheilos connoting slender and lip, respectively. Their causes can be congenital, although acquired causes have been described extensively.5–10 Namely, philtral elongation is a byproduct of gravity, bone retraction, elastosis,
and weakening fascial attachments. Other changes in the senescent ergotrid include diminution of pout and philtral flattening. Concurrently, the lips undergo soft-tissue atrophy, a process accelerated by sun exposure and smoking.

Identifying labral disproportions hinges on an understanding of normal dimensions. Several authors have attempted to define “normal” through structural comparisons, but not without criticism. One theory suggests that ergotrid height should not exceed the distance between the lower lash line and the supratarsal crease. However, periorbital and perioral aging may not occur in tandem. According to another, upper lip volume is ideally 75 to 80 percent that of the lower lip. However, the latter may be irregularly sized, and volumetric software is not readily available. We prefer an alternate approach that avoids comparing disparate anatomy, operates in linear dimensions, and explores “abnormal” territory more fully.

PATIENTS AND METHODS

Before presenting our metric, several parameters must be established. Provided the lips are touching at rest, three soft-tissue loci are of relevance: base of columella (subnasale); midpoint of superior vermilion border (labiale superius); and center of labial fissure (stomion). These define important vertical measurements: philtral height and labial height (Fig. 1). We believe the ratio of these components, philtral height divided by labial height, is a reliable indicator of normalcy and accordingly label it the philtral-labial score.

After securing institutional review board approval, we identified 908 patients of the senior authors (P.R. and S.W.H.) who underwent lip augmentation and/or lip lift surgery between January 1, 2001, and July 31, 2012. The study population consisted predominately of 25- to 75-year-old Caucasian women. Two hundred medical records containing preoperative anteroposterior and lateral photographs were selected randomly, provided that the patients possessed normal maxillary height. Using Adobe Photoshop CS6 (Adobe Systems, Inc., San Jose, Calif.), we superimposed brackets onto all anteroposterior photographs, manipulated them to span each philtrum and upper lip, and recorded their heights. Philtral-labial scores were then computed on a nonscientific calculator and rounded to the nearest tenth.

The senior authors independently evaluated said patients for labral defects in three surveys that included (1) photographs only; (2) photographs and philtral-labial scores; and (3) photographs, philtral-labial scores, and dental show values. The goal was to determine how much information was necessary to reach a significant degree of unanimity. Results were then analyzed for meaningful, consistent trends.

RESULTS

In surveys 1 to 3, the authors disagreed on 85 (42.5 percent), 23 (11.5 percent), and six (3.0 percent) subjects, respectively. The six discrepancies from the latter survey were reevaluated to obtain full consensus. This resulted in the defect-free, leptocheilos, philtral excess, and double-defect cohorts containing 66, 15, 71, and 48 members with average ages of 40, 41, 52, and 56 years and philtral-labial score ranges of 1.1 to 3.3 (mean, 2.3), 3.1 to 4.9 (mean, 4.0), 2.7 to 4.8 (mean, 3.7), and 4.7 to 42.3 (mean, 7.5), respectively (Fig. 2).

Age likely accounts for the increasing number of defects across the four cohorts. Interestingly, leptocheilos patients presented much earlier than those with philtral excess. This is not necessarily

Fig. 1. (Left) Relevant perioral surface anatomy (blue) with defined borders (green). (Right) Midline reference points (black) and defined heights (red). sn, subnasale; ls, labiale superius; st, stomion.
because of faster aging of the lips; rather, we suspect that innate philtral excess is relatively less common and recognition of this defect often delayed. Comparing the same groups, the slightly higher philtral-labial scores of leptocheilios patients is perhaps attributable to the formula itself; that is, incremental differences in small denominators carry more weight than equivalent changes to larger numerators.

Table 1 summarizes our findings with the labral classification system, a diagnostic guide that ascribes patients to one of four pairs of labral traits, inclusive of dental show and philtral-labial score correlates (ignoring outliers responsible for overlap). Philtral-labial scores add objectivity to the labral classification system. As scores escalate, so does the extent of imbalance. Dental show represents another vital quantitative element. Although 2 mm is normal in younger patients,11 we consider 1 mm acceptable in those older than 40 years. Assuming normal maxillary height, leptocheilios patients exhibit normal to excessive incisal show (type 1), whereas philtral redundancy veils dentition (types 2 and 3). Thus, lip volumizing and lifting procedures have antagonistic effects.

Generic treatments are also outlined in Table 1 and elaborated below. Although specifics are saved for future articles, suffice it to say that all our augmentations used alloplastic implants, and all lifts were a permutation of the subnasal resection. In practice, we use “bivertical” as a vertical-axis counterpart to “bilateral” when describing simultaneous augmentation of both lips. We accordingly document such cases as bivertical lip augmentations.

Type 0

These patients present with normal philtral and labial heights, a combination signifying a youthful labrum (Fig. 3). Incisor show is 1 to 2 mm, and philtral-labial scores remain below 3. Scores between 1.2 and 2.3 appear to embody “ideal” aesthetics according to calculations on young patients and professional models. Although not necessarily indicated, lip augmentation is sometimes desirable.

Type 1

These patients possess either an isolated thin upper lip or thin lips in general. Dental show

<table>
<thead>
<tr>
<th>Type</th>
<th>Philtral Height</th>
<th>Labial Height</th>
<th>Dental Show (mm)†</th>
<th>Philtral-Labial Score</th>
<th>Treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Normal</td>
<td>Normal</td>
<td>1–2</td>
<td>&lt;3</td>
<td>None</td>
</tr>
<tr>
<td>1</td>
<td>Normal</td>
<td>Short</td>
<td>≥1</td>
<td>3–5</td>
<td>Lip augmentation</td>
</tr>
<tr>
<td>2</td>
<td>Tall</td>
<td>Normal</td>
<td>0</td>
<td>3–5</td>
<td>Lip lift</td>
</tr>
<tr>
<td>3</td>
<td>Tall</td>
<td>Short</td>
<td>0</td>
<td>&gt;5</td>
<td>Combination‡</td>
</tr>
</tbody>
</table>

*Assumes normal maxillary height.
†In repose with lips slightly parted; age- and sex-dependent.
‡Lip augmentation and lip lift.
values are normal to high, and the philtral-labial score falls between 3 and 5. Clues that these patients desire fuller lips include perivermilion lipstick, lip liner, or cosmetic tattoo. Two subsets exist. Congenital patients (Fig. 4) have always had hypoplastic lips and generally lack perioral wrinkles. They are usually more dissatisfied with the upper lip, but may complain of both. In contrast, patients with the acquired type (Fig. 5) once had fuller lips but now note involution. They often exhibit perioral rhytides and diminished pout as well.

Recommended surgical management is lip augmentation, which achieves balance relative to the ergotrid, but lower lip enhancement may also be warranted to maintain labial harmony. Assessing lip skin availability will determine whether augmentation without dry mucosa recruitment is prudent, especially in those with “razor-thin” lips. Despite its benefits, many patients are reluctant to undergo mucosal advancement, in which case conservative augmentation using fillers is acceptable.

Type 2

Philtral excess characterizes this patient type. Incisor display is negligible, and philtral-labial scores lie between 3 and 5. Such patients are often borderline and difficult to classify, but clinical judgment and incisor display may be helpful. In its purest form, type 2 is least commonly seen in consultation, as many patients do not even recognize their defect. These patients sometimes complain of labial deficiency, which is often a reflection of redundant ergotrid causing the lip to invert. Once again, congenital (Fig. 6) and acquired (Fig. 7) subsets exist.

The applicable treatment is a lip lift. A resection of 3 to 12 mm will restore 1 to 2 mm of incisor
display, depending on severity. However, when augmentation is performed in place of a lift, the result may be large, procumbent lips relative to an uncorrected ergotrid (i.e., “duckbill”). Conversely, it may impart slight improvement, particularly in cases of excessive labial rhytides. Following a thorough discussion of the ergotrid’s role in facial harmony and a preview of the expected results using advanced imaging software or before-and-after photograph books, many will agree to a lip lift.

Type 3

This category exhibits leptocheilos, of either the upper lip exclusively or more commonly both, coupled with philtral excess. Several etiologic permutations exist, although consultation is nearly identical for each. Dental show is absent, and philtral-labial scores exceed 5 (Fig. 8). Frequent accompaniments include poor philtral definition, negligible pout, and radial wrinkles (Fig. 9). It is imperative that surgeons recognize this category to design judicious treatment plans and explain the anticipated results. Lengthy discussions may be warranted to offset unrealistic expectations.

Given their anatomical complexity, these patients present the greatest challenge. Standard treatment entails a lift and augmentation, which may be staged concomitantly. Resection of 5 to 12 mm may sufficiently unfurl the upper lip to safely perform an augmentation without mucosal advancement, even in extreme cases of leptocheilos. These patients often invoke our entire range of surgical techniques, including mucosal advancement, laser skin resurfacing, filler injections, and an aggressive skin care program.

---

Fig. 5. (Left) This woman, in her fifties, fell in the classic 30- to 60-year-old age range for type 1 patients. She presented for lip enhancement, relating that she had lost her lip fullness of youth. (Right) The same patient shown 1 year after bivertical lip augmentation using 4- and 5-mm implants in the upper and lower lips, respectively. PLS, philtral-labial score.

Fig. 6. (Left) This woman, in her late fifties, complained of always having a long lip. She had no dental show preoperatively. (Right) The same patient shown 1 year after having undergone a 6-mm lip lift. She declined any other perioral surgery (specifically, a lower lip augmentation). PLS, philtral-labial score.
DISCUSSION

As set forth in the Introduction, we sought to enhance diagnosis of labral disharmony. This is best explored by addressing each section of Table 1, beginning with the labral classification system. Review of the perioral literature revealed classification systems, rating scales, and surgical algorithms for filler placement, musculature, rhytides, reconstruction, and cosmetic modification of the lips and ergotrid separately.3,6,11,16–22 However, the labral classification system is the first to address the entire labrum.

Among its strengths, the labral classification system allows surgeons to circumvent convoluted decision trees, communicate with peers concisely and on common ground (akin to the Baker scale of capsular contracture or Regnault system for breast ptosis), relate treatment goals to patients, follow them prospectively as they change types, and study large cohorts of given categories according to many variables (e.g., age, race, sex, philtral-labial score, dental show, treatment modality, or complication and satisfaction rates thereof).

Moreover, the labral classification system is a useful reference in consultation, especially with patients who have misguided beliefs or vague complaints. Imagine a type 2 patient who desires augmentation or a woman who wants perioral rejuvenation but cannot pinpoint the source of her dissatisfaction. In both common scenarios, the labral classification system promotes mutual understanding of the disharmony at hand, thus guiding sound therapy.

Fig. 7. (Left) This woman, in her sixties, had nonspecific complaints of labral aging. Her elongated ergotrid is not immediately obvious, a testament to the fact that philtral excess is harder to discern than leptocheilosis. (Right) The same patient shown 1 year after a 6-mm lip lift and bivertical lip augmentation using 4-mm implants. Note her shortened ergotrid and improvement in vermilion lines and wrinkles. PLS, philtral-labial score.

Fig. 8. (Left) This 62-year-old woman presented for perioral rejuvenation. (Right) The same patient shown 1 year after an 8-mm lip lift and bivertical lip augmentation using 4-mm implants. Significant improvement was obtained without establishing a normal philtral-labial score (PLS).
Insofar as shortfalls are concerned, subjectivity is the theme. Typing patients is usually clear-cut, yet labeling a lip as too thin or an ergotrid as too long can sometimes be challenging. Disagreement will invariably occur, and to account for all cultural, training, and aesthetic biases in a single system is impractical. Nevertheless, our results suggest that subjectivity can be minimized if preoperative anteroposterior and lateral photographs (or meticulous live appraisal) are weighed against dental show and philtral-labial scores. When such information is available, we expect high concordance among surgeons worldwide. Rates may not be on par with that of our study, but more accurate classification will surely ensue. Other caveats are worth mentioning: (1) labral structures must be assessed on a case-by-case basis in relation to the entire face, together with underlying skeletal and dental support; (2) it has been implied that the goal is to convert all patients to type 0, but sometimes it is sufficient to improve a type 3 to a type 1 or 2 (Fig. 8); and (3) the labral classification system ignores large lips, short ergotrids, and the lower lip complex altogether because, in our experience, such complaints are rare.

Finally, the labral classification system presupposes normal maxillary height because patients with vertical maxillary excess or deficiency are highly complex. Whereas soft-tissue surgical intervention may be required, it does not always produce optimal results alone. Maxillary abnormalities can have a profound impact on facial appearance, and our philosophy is that bony anomalies require bony solutions. Once skeletal abnormalities are addressed, these patients become more easily classifiable.

Proceeding to the two columns of diagnostic tools, we will first inspect the philtral-labial score. In creating this metric, we invoked anatomical landmarks and heights established centuries ago, and photogrammetric analysis, a more recent methodology that has been used for performing such measurements. The use of indices to ascertain populational averages also has historical precedent. For example, Farkas et al. examined the ratio of labial to labral height in his study of Caucasian adults (note the slight modification of the philtral-labial score formula).

The dental show column is examined next. Preoperatively, this is a crucial factor in patient typing and may represent the final arbiter as to whether a labral anomaly exists. Postoperatively, it is a gauge of improvement; indeed, the surgical plan sometimes hinges on its restoration. Nevertheless, given its dependency on incisal, maxillary, labial, and philtral heights, this multifaceted variable remains an adjunct rather than cornerstone to the labral classification system.

Farkas et al. deduced that normal lips span over 25 percent of labral height, which equates to a philtral-labial score below 3 and corroborates our defect-free range. By logical extension, scores above 3 must be abnormal. Furthermore, average philtral height is said to fall between 18 and 20 mm, whereas ideal upper lip height has been placed at 7 to 8 mm. These numbers give rise to philtral-labial scores between 2.5 and 2.9, also paralleling the defect-free range. Thus, the literature validates our metric and empirically supports...
our ranges, especially our characterization of low scores.

That said, this study transcends others on several levels. Whereas Farkas et al. determined normal values, we extract meaning from high scores, which are clinically more relevant. Preoperatively, the implications become manifest in ostensibly borderline cases, where scores surpassing conveniently memorable cutoffs, 3 and 5, confirm suspicion as to the presence and quantity of labral flaws. The philtral-labial score also reflects the severity of disharmony, a useful feature in consultations warranting patient education regarding phenotypic and surgical complexity. Lastly, its calculation directly incorporates the upper lip, underscoring that labral disharmony can entail a labial component.

However, the philtral-labial score is not infallible. In light of our sample population, we concede that our metric’s purview is limited and that philtral-labial score breakouts are needed across various demographics, such as African Americans, Hispanics, male patients, and transgender patients, who represent a growing percentage of perioral rejuvenation candidates. Second, that the philtral-labial score analyzes the upper lip relative to its superior neighbor should not obviate a careful examination of the lower lip. Third, the philtral-labial score formula only factors in midline heights; however, a hanging columella may belie a lengthy ergotrid, and a sharp, central indentation of the vermilion border can have the opposite effect. Nevertheless, we assume that our chosen parameters are valid proxies for vertical measurements along the full span of labral structures.

Let us also consider exceptions at each score range. For example, low scores do not always denote labral harmony. Duckbill patients’ large lips, for instance, can numerically offset their ptotic ergotrids. Moreover, scores above 5 do not always imply both defects, as razor-thin lips may overshadow normal ergotrid height. Between these extremes, the philtral-labial score is unable to differentiate philtral excess from leptocheilos, and isolating the anomaly may necessitate dental show information, palpation, or observation of the lips in excursion.

Regarding the role of the philtral-labial score in surgical management, some cautionary advice is fitting. For example, deferring to the philtral-labial score as a criterion standard may mislead surgeons into robotically treating scores rather than patients. As for how much correction is warranted, the metric offers only rough indications.

Bringing the philtral-labial scores down below 3 generally results in more pleasing appearances, yet such aims are not mandatory—particularly in severe outliers or the aged population, where less drastic changes can still produce excellent outcomes (Fig. 8). Likewise, recall that a score below 3 does not obviate the need for surgery (e.g., duckbill patients).

Despite its diagnostic shortcomings, the philtral-labial score plays an important role in the postoperative setting, where it provides a means of tracking patients over time and documenting/authenticating surgical improvement. Moreover, given that the philtral-labial score is a unitless ratio independent of face size and shape, images need not be life size when performing height measurements, and interpopulation analysis is possible.

When interpreting or influencing philtral-labial scores in practice, familiarity with its drawbacks is crucial. However, to extract the metric’s full potential in areas where it shines, sound practices must be observed. *Photographic Standards in Plastic Surgery* is an excellent resource for obtaining a complete perioperative set of photographs, although we recommend an additional one with slightly parted lips for dental show. Rather than measure directly on the patient, we prefer anteroposterior digital images because they hide pout and avoid contour distortion. Finally, software capable of extreme sensitivity is preferred.

The final section of Table 1 is treatment, in which we recommend augmentation for leptocheilos and lifting for philtral excess. As with any classification system, however, gray zones exist. Surgeons may (1) opt not to perform a lip lift on a type 2 patient lest it increase an already adequate degree of dental show; (2) perform a lip augmentation on a type 0 or 2 patient with severe wrinkles or a significantly thinner upper lip relative to lower; or (3) perform a lip lift on a type 0 or 1 patient because of an absence of visible dentition. As such, our suggestions are intended as guidelines tailored to patient goals rather than mandates.

**CONCLUSIONS**

To help objectify the subjective task of labral evaluation, we submitted a new metric that does not require complex measurements, high-tech equipment, or sophisticated analysis. The philtral-labial score is also capable of pretreatment and posttreatment comparison. Furthermore, we ascertained the most beneficial information for patient typing and presented it within the
context of a classification system. As a testament to their validity and reliability, the labral classification system and its accompanying analytic tools have been applied successfully for 10 years in over 2000 consultations. We hope that their adoption will improve evaluation, treatment, and follow-up.

REFERENCES