Anatomic Distribution of the Morphologic Variation of the Upper Lip Frenulum Among Healthy Newborns

Shagnik Ray, BA; William Christopher Golden, MD; Jonathan Walsh, MD

IMPORTANCE The maxillary labial frenulum and its potential contribution to breastfeeding difficulty may substantially affect public health. However, objective studies of the frenulum are limited.

OBJECTIVE To measure the variations in length, thickness, and attachments of the maxillary labial frenulum in healthy newborns and to identify which anatomic measurements could be used in further research investigating the maxillary labial frenulum.

DESIGN, SETTING, AND PARTICIPANTS This prospective cross-sectional study conducted measurements on images of maxillary labial frenula captured by digital photography from 150 healthy newborns admitted to the newborn nursery at a tertiary care children's hospital in Maryland between September 1, 2017, and April 1, 2018.

MAIN OUTCOMES AND MEASURES The primary outcome was the measurement of numerous frenulum morphologic components.

RESULTS Of 150 newborns enrolled, 77 were female, the mean (SD) gestational age was 38.60 (1.72) weeks, and the mean (SD) birth weight was 3180 (570) g. The means and SDs of the morphologic components of the frenulum with the broadest distributions, which were most helpful in differentiating degrees of lip tethering, included the following: alveolar edge to frenulum gingival attachment, 1.53 (0.85) mm; frenulum length on stretch, 5.19 (1.68) mm; frenulum gingival attachment thickness, 0.84 (0.36) mm; frenulum labial attachment thickness, 2.83 (1.33) mm; and the percentage of free lip to total lip length, 87.38% (7.67%). Gingival attachment mean (SD) thickness differed between late-preterm (0.69 [0.24] mm) and term (0.88 [0.37] mm) infants (Cohen $d$, −0.52; 95% CI, −0.94 to −0.10).

CONCLUSIONS AND RELEVANCE To our knowledge, this cross-sectional study was the first to objectively measure the numerous morphologic components of the upper lip anatomy in healthy newborns. Variations in maxillary labial frenulum morphology were identified, and some combination of the stated measurements may be used to create a more robust classification system to advance quality research in the association of lip-tie with breastfeeding difficulty.
As more mothers are encouraged to breastfeed their infants, tethering and anatomic positioning of the labial frenulum and lingual frenulum have surfaced as issues at the center of the breastfeeding difficulty discourse. Upper lip frenulum tethering (termed lip-tie) has been implicated in addition to ankyloglossia (termed tongue-tie) in various childhood conditions, including breastfeeding difficulty. Tethering of the upper lip by the maxillary labial frenulum has been postulated to cause improper latching of the newborn to the mother’s breast, preventing seal formation around the maternal areolar tissue and hypothetically causing a concomitant increase in infant reflux and aerophagia. With many infants being diagnosed as having upper lip-tie, the role of the labial frenulum in impaired nursing necessitates further study.

Originating as a posteruptive remnant of embryonic ectodermal bands, the maxillary labial frenulum is a small, somewhat triangular fold of nonmuscular connective tissue extending from the midline maxillary gingiva into the vestibule and central upper lip. At present, the typical anatomic variation of the maxillary labial frenulum has been described in 2 limited studies. In a study of 1021 Swedish newborns, Flinck et al noted 76.7% of maxillary labial frenula inserted into the crest of the alveolar ridge, 6.7% inserted into the buccal mucosa of the alveolar ridge, and 16.7% inserted into the palatal mucosa of the alveolar ridge. More recently, in a study of 100 newborns, Santa Maria et al found that 83% of newborn maxillary labial frenula attach at the gingival margin, whereas 6% attach near the mucogingival junction and 11% attach along the inferior margin at the alveolar papilla and beyond to the posterior surface. However, both of those studies described frenulum insertion points into the gingiva with nonspecific anatomic locations and without measurements of frenulum insertions relative to gingival and alveolar edge landmarks. Furthermore, neither study assessed frenulum length and thickness, which equally may play a role in frenulum tethering. Finally, those studies did not assess the potential association of the maxillary labial frenulum with ankyloglossia.

Some health care professionals, in an attempt to improve breastfeeding in neonates, have proposed and performed surgical modification, release, or removal of the maxillary labial frenulum in procedures known as labial frenotomy or frenectomy. In May of 2015, the Agency for Healthcare Research and Quality surveyed the literature regarding the labial frenotomy, finding the strength of evidence generally low to insufficient based on small, short-term studies with insufficient randomized controlled trials. Ghaheri et al recently found significant improvement in breastfeeding outcomes in a prospective cohort study after combined tongue-tie and lip-tie release; however, that study lacked a control cohort, and only 1 participant had isolated lip-tie release as opposed to combined lip-tie and tongue-tie release. Limited conclusions can be drawn on the effect of lip-tie release based on studies with such confounding data. Ultimately, studies on possible negative effects of the maxillary labial frenulum and surgical amelioration of these effects require an objective, consistent, and thorough classification system to enable proper clinical decision-making and consistency among future studies.

Currently, 2 classification systems for the maxillary labial frenulum exist. The more commonly used Kotlow classification system denotes 4 frenulum types and focuses on the insertion point of the gingival attachment of the maxillary labial frenulum. However, this system does not correlate the epidemiologic variation of frenula with poor breastfeeding outcomes. Santa Maria et al subsequently proposed a classification system with 3 frenulum types, again with a focus on the gingival insertion point of the maxillary labial frenulum and attempting to simplify the Kotlow classification system and improve interrater reliability. However, this tool also did not analyze other factors beyond the insertion point, and the absence of objective measurements resulted in an interrater reliability of only 38%. In the present article, we report the anatomic distribution of different morphologic components of the maxillary labial frenulum that may be associated with lip mobility. We used objective measurements to help assess the components of the maxillary labial frenulum that could potentially be used in a classification system and in further research. We also analyzed these measurements by race/ethnicity, sex, gestational age, birth weight, presence or absence of ankyloglossia, and LATCH score (a commonly used measure of breastfeeding success).

Methods
We performed a prospective cross-sectional study of 150 healthy newborns admitted to the Johns Hopkins Hospital Newborn Nursery (Baltimore, Maryland) between September 1, 2017, and April 1, 2018. The Johns Hopkins institutional review board approved this study. Verbal informed consent was obtained from the parents or guardians. Verbal consent was used because of the minimal risk of the study and to reduce the burden on mothers in the perinatal period. No one received compensation or was offered any incentive for participating in this study.
Newborn nursery pediatric nurse practitioners assessed all infants twice weekly for study enrollment, including late-preterm newborns (born at 34 0/7 to 36 6/7 weeks’ gestation). Infants with possible or probable craniofacial anomalies or those admitted to the neonatal intensive care unit were excluded. A member of our team (S.R.) imaged the frenulum of each infant using a standardized protocol, beginning with elevation of each infant’s upper lip and retraction to the level of the alveolar sulcus. A standardized ruler was then placed along the alveolus, and high-definition digital photographs of the participants’ upper lips and gums were obtained with a Canon PowerShot A4000 IS HD camera (Canon Inc). Numerous images were obtained, with the highest-quality image selected for measurements. LATCH scores, newborn birth weight, gestational age, presence of ankyloglossia, and demographic data were collected from each infant’s medical record.

ImageJ software, version 1.51j8 (National Institutes of Health) was used to calibrate the measurement scale for each photograph and to adjust for slight variations in focal distance. Digital measurements were obtained for the following components: distance from alveolar edge to frenulum attachment, length of frenulum, distance from frenulum lip attachment to vermilion border, frenulum gingival attachment thickness, frenulum labial attachment thickness, ratio of free gingiva to total gingival length, and ratio of free lip to total lip length. In addition, the frenula were scored based on the Kotlow and Stanford (Santa Maria et al) classification systems.

All statistical analyses were performed with STATA/SE, version 15.1 (StataCorp LLC) for Windows. The Shapiro-Wilk test was used to assess whether measurements were normally distributed. Cohen d, 95% confidence intervals (CIs), and η² were used to analyze differences between measured frenulum components and participant characteristics. A Cohen d value of 0.2 is considered a small effect size, 0.5 a medium effect size, and 0.8 a large effect size. Eta-squared reflects the proportion of variation in the dependent variable that is accounted for by the groups defined by the independent variable. A value of 0.01 is considered small, 0.06 is medium, and 0.14 or above is large. In addition, η² was used to determine the association between LATCH scores and measurements for the combined and ankyloglossia subgroups. We determined means, percentiles, and SDs for the measurements. Cohen d was used to determine the difference between LATCH scores for patients with or without ankyloglossia. In addition, potential differences in LATCH scores between different Kotlow and Stanford scale scores were determined with η². A validated breastfeeding assessment scale, LATCH has 5 items with a total score range of 0 to 10. Scores of 6 or higher indicate successful breastfeeding.

Results

We enrolled 150 newborns in the study based on a population estimate of 5% incidence of lip-tie. Of the study patients, 77 participants (51.33%) were female. The race/ethnicity of the patients included 7 Asian (4.67%), 65 black (43.33%), 76 white (50.67%), and 2 (1.33%) other newborns. The mean (SD) gestational age of the study participants was 38.60 (1.72) weeks (range, 34.0-41.6 weeks’ gestation). The mean (SD) birth weight of the participants was 3180 (570) g (range, 1850-4480 g). Thirty-one newborns (20.67%) were diagnosed by newborn nursery clinician assessment as having ankyloglossia (based on a modified Coryllos system along with functional assessment). The LATCH scores were routinely obtained only for newborns whose mothers elected to attempt to breastfeed; 129 of 150 newborns had documented LATCH scores. There was no statistical difference between frenulum morphologic components for newborns with or without documented LATCH scores. The mean (SD) LATCH score of the participants was 6.78 (1.62). The mean (SD) LATCH score for black infants (7.84 [1.64]) was slightly lower than that for white infants (8.13 [1.45]), and late preterm infants had slightly lower mean (SD) LATCH scores (7.19 [1.81]) than term infants (8.28 [1.81]). All statistics involving LATCH scores excluded newborns without a documented LATCH score.

The 7 measurements of the maxillary labial frenulum components obtained from the captured images are shown in Figure 1 along with several examples of photographed maxillary labial frenula. The ratio of free gingiva (ie, gingiva length not covered by the maxillary labial frenulum) to total gingiva was calculated as 100 ÷ [(alveolar edge to gingival attachment) (alveolus to sulcus)]. The mean (SD) ratio of free lip (ie, lip length not covered by the maxillary labial frenulum) to total lip, calculated as 100 ÷ [(lip attachment to vermilion border) (sulcus to vermilion border)], was 87.38% (7.67%). The mean and SD for each component measured can be found in the Table. For example, the mean (SD) of the distance from the alveolar edge to the frenulum gingival attachment was 1.53 (0.85) mm, the frenulum length on stretch was 5.19 (1.68) mm, the frenulum gingival attachment thickness was 0.84 (0.36) mm, and the frenulum labial attachment thickness was 2.83 (1.33) mm. Most frenula attached less than 2 mm from the alveolar edge and had a relatively small mean (SD) value. Visual frequency distributions for each measurement can be found in Figure 2 and Figure 3.

Statistical analysis showed that the length from alveolar edge to frenulum gingival attachment, gingival attachment thickness, and lip attachment thickness were log normally distributed. Frenulum length on stretch, length from lip attachment to vermilion border, distance from alveolus to sulcus, and distance from sulcus to vermilion border were normally distributed. The frenula were graded with the Kotlow and Stanford classification systems, and the results are shown in Figure 4. Using the Kotlow scale, most neonates (101 of 150) scored 3 of 4; when graded with the Stanford scale, 140 neonates scored 2 of 3, with only 2 newborns scoring 3.

There was a medium size effect difference for gingival attachment thickness between mean (SD) late-preterm (0.69 [0.24] mm) and term (0.88 [0.37] mm) infants (Cohen d, −0.52; 95% CI, −0.94 to −0.10), with term defined as gestational age greater than or equal to 37 weeks. Otherwise, there were no meaningful differences for variable measurements between late-preterm and term newborns. Of note, when comparing differences between black infants and white infants, there were small to medium effect size differences in alveolar attachment (mean, 1.66 [95% CI, 1.46-1.86] vs 1.36 [95% CI, 1.19-
1.53 mm; Cohen’s d = −0.39 [95% CI, −0.72 to −0.05]), frenulum length (4.79 [95% CI, 4.37-5.20] vs 5.54 [95% CI, 5.17-5.92] mm; Cohen’s d = 0.46 [95% CI, 0.12-0.79]), and the ratio of lip attachment to vermilion border (0.89 [95% CI, 0.87-0.91] vs 0.86 [95% CI, 0.84-0.87] mm; Cohen’s d = −0.39 [95% CI, −0.73 to −0.06]), respectively. Furthermore, we found no clinically meaningful difference between newborns diagnosed as having or as not having ankyloglossia for each variable measurement, including LATCH score, except lip attachment thickness (mean, 3.3 [95% CI, 2.8-3.9] vs 2.7 [95% CI, 2.5-2.9] mm; Cohen’s d = 0.46 [95% CI, 0.06-0.86]). The very small η² effect size estimates (eTable in the Supplement) described no clinically meaningful association between measurements and LATCH scores. These estimates include the alveolar edge to frenulum gingival attachment, which is the functional correlate to the Kotlow and Stanford scales. In addition, η² showed no meaningful differences in LATCH scores between different Kotlow classification system scores and different Stanford classification system scores.

### Discussion

Evidence regarding an association between the maxillary labial frenulum and lip-tie is currently sparse. To improve diagnosis and more rigorous study, the present study provided reference values of the various morphologic components of the frenulum from a sample population. Considering that tethering necessitates 2 attachment points, numerous aspects of frenulum anatomy are needed rather than simply the gingival attachment site. Both the Kotlow classification system and the Stanford system proposed by Santa Maria et al involve the distance from the alveolar edge to the frenulum gingival attachment. In our cohort, as shown in Figure 2, the distance from the alveolar edge to

### Figure 1. How Frenulum Morphologic Components Were Measured and Examples of Maxillary Labial Frenula

#### Table. Data on Measured Frenulum Morphologic Components

<table>
<thead>
<tr>
<th>Component</th>
<th>Mean (SD)</th>
<th>Combined (n = 20)</th>
<th>Ankyloglossia (n = 10)</th>
<th>Nonankyloglossia (n = 10)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distance from alveolar edge to frenulum gingival attachment, mm</td>
<td>1.53 (0.85)</td>
<td>1.57 (0.91)</td>
<td>1.52 (0.83)</td>
<td></td>
</tr>
<tr>
<td>Length of frenulum on stretch, mm</td>
<td>5.19 (1.68)</td>
<td>5.56 (1.51)</td>
<td>5.09 (1.72)</td>
<td></td>
</tr>
<tr>
<td>Distance from frenulum labial attachment to vermilion border, mm</td>
<td>6.02 (2.01)</td>
<td>5.67 (2.03)</td>
<td>6.11 (2.01)</td>
<td></td>
</tr>
<tr>
<td>Gingival attachment thickness, mm</td>
<td>0.84 (0.36)</td>
<td>0.86 (0.33)</td>
<td>0.84 (0.37)</td>
<td></td>
</tr>
<tr>
<td>Labial attachment thickness, mm</td>
<td>2.83 (1.33)</td>
<td>3.30 (1.50)</td>
<td>2.70 (1.26)</td>
<td></td>
</tr>
<tr>
<td>Free gingiva to total gingival length ratio, %</td>
<td>25.49 (13.65)</td>
<td>24.1 (0.12)</td>
<td>25.8 (0.14)</td>
<td></td>
</tr>
<tr>
<td>Free lip to total lip length ratio, %</td>
<td>87.38 (7.67)</td>
<td>85.6 (0.10)</td>
<td>87.8 (0.07)</td>
<td></td>
</tr>
</tbody>
</table>
to the frenulum gingival attachment had a narrow range of attachment variability, with a mean value of 1.53 mm and an SD of 0.85 mm. Given that most frenula attached less than 2 mm from the alveolar edge and had a relatively small mean and SD, the attachment site for the maxillary labial frenulum alone was not a sufficient population discriminator for lip tethering. Furthermore, the small magnitude of the mean and SD could explain the low interrater reliability found in the Kotlow and Stanford scales because such a small value can be difficult to accurately assess on a crying neonate. When this cohort was graded with the Kotlow scale as shown in Figure 4, the vast majority (101 of 150) of individuals scored 3 of 4, which could suggest a more severe lip-tie by the Kotlow scale. When graded with the Stanford scale as shown in Figure 4, 140 patients scored 2 of 3, with only 2 patients scoring 3, overall providing little ability to discriminate within our population. For both scales, statistical analyses showed no difference in LATCH scores within each scale in the present study. Thus, the current maxillary labial frenulum classification systems may be inadequate to properly assess variation in a general population of newborns.
An ideal maxillary labial frenulum classification system should involve numerous appropriately selected frenulum component measurements based on epidemiologic distributions to achieve accurate description and population discrimination of lip tethering. Establishing a population-based frenulum classification system is critical before any further research can be undertaken regarding the role of the maxillary frenulum and breastfeeding difficulty.

On the basis of findings in the present study, certain measurements appeared well suited to be used in such a classification system or for further study in general. Measurements with broader distributions appeared most helpful in differentiating between degrees of lip tethering in our cohort. Although not a sufficient population discriminator alone given its narrow distribution, the length from alveolar edge to gingival attachment may be useful because it serves as an important determinant of restriction in 3-dimensional space, describing one of the necessary points of attachment. The remaining frenulum measurements were not accounted for in the Kotlow and Stanford scales, but Figure 1 shows their variability within the population. Lip and gingival attachment thickness may affect lip tethering because thicker frenula are less likely to allow for lip mobility. As given in the Table, lip attachment thickness had a relatively large SD relative to the mean. Thus, this value may help differentiate between degrees of lip tethering within the population. Gingival attachment thickness had a smaller mean and SD but may still provide useful data regarding frenulum morphology. The length of the frenulum on stretch, with a large mean and SD, may help further distinguish degrees of lip tethering. Biomechanically, a frenulum that stretches easily has less risk of labial tethering, whereas a short frenulum that does not stretch much has a higher risk of labial tethering in a 3-dimensional space. Finally, the percentage of the lip that is “free” (ie, not covered by the frenulum) helps indicate how mobile a lip can be. Overall, some combination of these particular component measurements could be used to create a more comprehensive grading scale for both research and clinical purposes to characterize the maxillary labial frenulum.

Our study also analyzed the presence of ankyloglossia based on the frenulum measurements. Because both tongue and upper lip tethering models are based on improper midline attachments in the mouth and functional restriction, one could hypothesize that the 2 conditions could be interrelated. The presence or absence of ankyloglossia was determined by the newborn nursery clinicians, and they used a modified Coryllos system along with functional assessment. However, our study did not show any association between the measured frenulum morphologic components or the Kotlow and Stanford scales with the presence or lack of ankyloglossia.
These data and our identified anatomic criteria may be useful in understanding the role of the frenulum throughout childhood. Recent work has suggested that abnormal tethering of the upper lip attributable to the maxillary labial frenulum may lead to formation of dental caries in childhood.\(^1,2,9,14\) The labial frenulum also has been implicated in the formation of a midline diastema (a space between the maxillary central incisors), which may then in turn result in dental caries secondary to food trapping.\(^1,15\) However, there have been few studies regarding the natural history of the maxillary labial frenulum. In a study of children 1 to 8 years of age, Boutsi and Tataki\(^16\) found that the attachment of the frenulum differed across ages, with older children showing mucosal or gingival frenula rather than papillary penetrating frenula, suggesting that the maxillary labial frenulum may shift the insertion point as a child ages and the maxilla develops. Use of more comprehensive anatomic data may enable researchers to investigate concepts such as potential association with oral health feeding, longitudinal assessment of frenulum development, and characterization of changes that may affect oral health and feeding.

**Limitations**

We noted several limitations of our investigation. This prospective cross-sectional study used digital photography with a standardized examination technique. Minor variations in protocol and image quality could artificially increase variability in the measurements obtained, but the variability was smaller than expected for direct measurements without digital assistance. Given the low SDs noted in all measurements, this association was likely minimal. In addition, LATCH scores for the patients in our study were obtained at different times after birth and by numerous lactation consultants and postpartum nurses. LATCH scores did vary slightly between black infants and white infants as well as between term and late-preterm infants. These factors may have obscured any correlation between degree of frenulum and LATCH score. The present study was not designed to determine a correlation with lip-tie and breastfeeding difficulty but to describe the population’s anatomic variation, which limited the use of other breastfeeding assessment tools. It is possible that the lip anatomy may have limited association with breastfeeding given that no prospective study has rigorously demonstrated such an association. Future investigations should include a more comprehensive maxillary labial frenulum grading scale than the Kotlow and Stanford scales and should follow up with prospective case-control studies to assess for correlations between the more ideal scale and feeding assessments, such as the LATCH score, the Infant Breastfeeding Assessment Tool, or the short-form McGill Pain Questionnaire.\(^17\) The LATCH score alone does not fully encapsulate and quantify feeding difficulty among infants, and infant feeding is best assessed with a battery of tests and qualified clinical expertise. Further studies also are necessary to understand the role of the maxillary labial frenulum in breastfeeding and to assess the association of labial frenectomy with breastfeeding.

**Conclusions**

To our knowledge, the present study is the first to describe detailed, specific measurements of maxillary labial frenulum morphology in newborns. We found little variability in the alveolar attachment location across the study population, with most frenula attaching less than 2 mm from the alveolar edge. These findings may have substantial implications for treatment of the rapidly growing population of infants with suspected lip-tie. The currently available maxillary labial frenulum classification systems provided poor discrimination within the present study population. We identified additional anatomic criteria that may improve future classification.


