

# Development and Validation of a Photographic Scale for Assessment of Lip Fullness

W. Philip Werschler, MD; Steven Fagien, MD; Jane Thomas, AAS; Deepali Paradkar-Mitragotri, PhD; Adam Rotunda, MD; and Frederick C. Beddingfield III, MD, PhD

Aesthetic Surgery Journal  
2015, Vol 35(3) 294–307  
© 2015 The American Society for  
Aesthetic Plastic Surgery, Inc.  
This is an Open Access article  
distributed under the terms of the  
Creative Commons Attribution  
Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com  
DOI: 10.1093/asj/sju025  
[www.aestheticsurgeryjournal.com](http://www.aestheticsurgeryjournal.com)

**OXFORD**  
UNIVERSITY PRESS

## Abstract

**Background:** As lip augmentation becomes more popular, validated measures of lip fullness for quantification of outcomes are needed.

**Objective:** Develop a scale for rating lip fullness and establish its reliability and sensitivity for assessing clinically meaningful differences.

**Methods:** The initial Allergan Lip Fullness Scale (iLFS; a four-point photographic scale with verbal descriptions) was validated by eight physicians rating 55 live subjects during two rounds, conducted on one day. In addition, subjects performed self-evaluations. The revised Allergan Lip Fullness Scale (LFS), a five-point scale with a broader range of lip presentations, was validated by 21 clinicians in two online image rating sessions,  $\geq 14$  days apart, in which they used the LFS to rate overall, upper, and lower lip fullness of 144 3-dimensional (3D) images. Physician inter- and intra-rater agreement, subject intra-rater agreement (iLFS), and subject-physician agreement (iLFS) were evaluated. Additionally, during online rating session 1, raters ranked 38 pairs of 3D images, taken before and after lip augmentation, as “clinically different” or “not clinically different.” The median LFS score difference for clinically different pairs was calculated to determine the clinically meaningful difference.

**Results:** Clinician inter- and intra-rater agreement for the iLFS and LFS was substantial to almost perfect. Subject self-assessments (iLFS) had substantial intra-rater reliability and a high level of agreement with physician assessments. Median LFS score differences for overall, upper, and lower lip fullness were 1 (mean: 0.63-0.69) for “clinically different” and 0 (mean: 0.28-0.36) for “not clinically different” image pairs; thus, clinical significance of a 1-point difference in LFS score was established.

**Conclusions:** The LFS is a reliable instrument for physician classification of lip fullness. A 1-point score difference can detect clinically meaningful differences in lip fullness.

Accepted for publication July 16, 2014.

The lips play a central role in the human perception of facial beauty and attractiveness. The definition of “perfect” or “ideal” lips varies, depending on factors such as age, ethnicity, culture, fashions, and current trends.<sup>1-5</sup> It is generally accepted, however, that shape, size, fullness, and symmetry are important aspects of lip attractiveness and that color and contrast may also play a role.<sup>5-7</sup>

With the availability of minimally invasive products for dermal filling and volume creation, lip enhancement and augmentation procedures have become increasingly popular.<sup>5,7</sup> Candidates for lip enhancement include patients whose lips have flattened or lost volume, show vermilion, and/or have lost definition as a result of age, photodamage, smoking, or a combination thereof; patients with congenitally thin and/or asymmetrical lips; and patients who desire enhanced lip shape or fullness.<sup>7,8</sup>

Dr Werschler is Assistant Clinical Professor of Dermatology at the University of Washington School of Medicine, Seattle, Washington. Dr Fagien is a plastic surgeon in private practice in Boca Raton, Florida. Ms. Thomas is an employee of Kythera Biopharmaceuticals, Inc., Calabasas, California. Dr Paradkar-Mitragotri is an Employee of Allergan, Inc., Santa Barbara, California. Dr Rotunda is an Assistant Clinical Professor, Division of Dermatology, David Geffen School of Medicine, UCLA, Los Angeles, California, and Assistant Clinical Professor, University of California, Irvine, California. Dr Beddingfield is an Associate Clinical Professor, Division of Dermatology, David Geffen School of Medicine (UCLA), Los Angeles, California, and is an employee of Kythera Biopharmaceuticals, Inc., Calabasas, California.

## Corresponding Author:

W. Philip Werschler, MD, Spokane Dermatology Clinic, 104 W 5th Ave., #330, Spokane, WA 99204, USA.  
E-mail: [pkwerschle@aol.com](mailto:pkwerschle@aol.com)

Validated measures of lip fullness, for quantification of lip augmentation outcomes in clinical trials and clinical practice, are needed. A handful of assessment scales have been developed to aid in the objective evaluation of lip enhancement treatments.<sup>9-11</sup> This report describes the development and validation of the Allergan Lip Fullness Scale (LFS), a tool for evaluating the effectiveness of lip augmentation in clinical trials of hyaluronic acid gel injection for lip augmentation.<sup>12,13</sup> Following validation studies, the initial version of the LFS (iLFS; a four-point photometric scale) was revised to the five-point LFS, in response to recommendations from the United States Food and Drug Administration (FDA). The objective of these validation studies was to establish inter- and intra-rater reliability of the scales and the appropriateness of LFS ratings for detecting changes in lip fullness that experienced physicians considered “clinically different”. These investigations were registered at [www.clinicaltrials.gov](http://www.clinicaltrials.gov) (ID: NCT01197495).

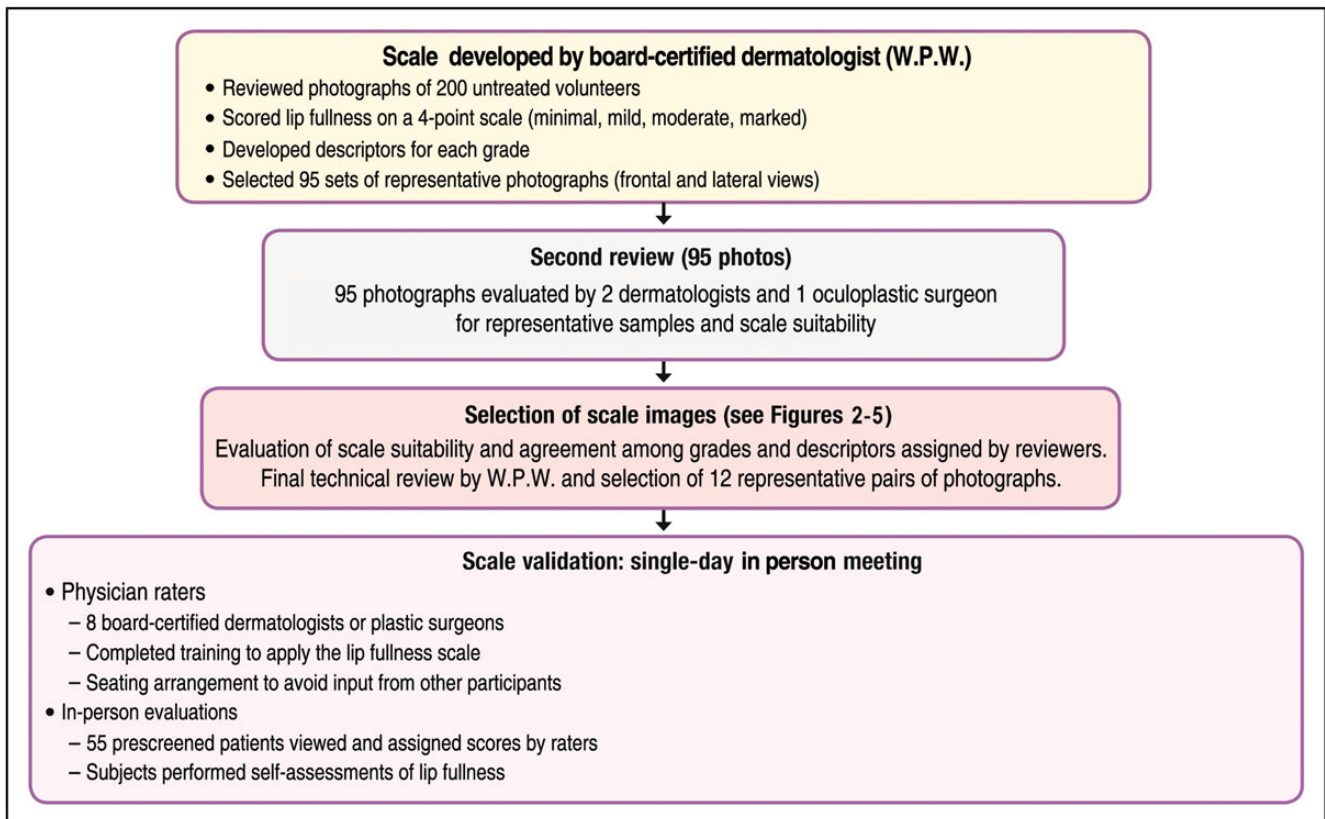
## METHODS

### Development and Validation of the Initial LFS

Photographs from 200 female volunteers who had not received lip augmentation treatment were taken with the

subjects’ mouths relaxed and with their lips gently closed. The images were scored by a clinically experienced, board-certified dermatologist, according to a four-point grading system that included verbal descriptions (Figure 1), and, based on the comments of the scoring dermatologist and technical review of photograph quality for standardized positioning, expression, skin tone evenness, and other features (W.P.W.), 95 photographs were selected to represent four grades of lip fullness. These photographs were then evaluated for inclusion in the iLFS by a panel of clinically experienced physicians that included two board-certified dermatologists and one oculoplastic surgeon (Figures 2-5). The photographs ultimately included in the initial scale were selected based on agreement between scores from the dermatologists and from the oculoplastic surgeon, their comments, and final review by the initial dermatologist. Grades of “minimal,” “mild,” “moderate,” and “marked” lip fullness are illustrated with three examples per grade in 12 frontal images paired with 12 lateral images.

Intra- and inter-rater reliability of the iLFS was evaluated in live subjects in two rounds of evaluations (morning and afternoon), conducted in a single day. The study enrolled subjects aged 18 years or older. All subjects signed an IRB-approved informed consent and authorization for use or release of health and research study information form and the

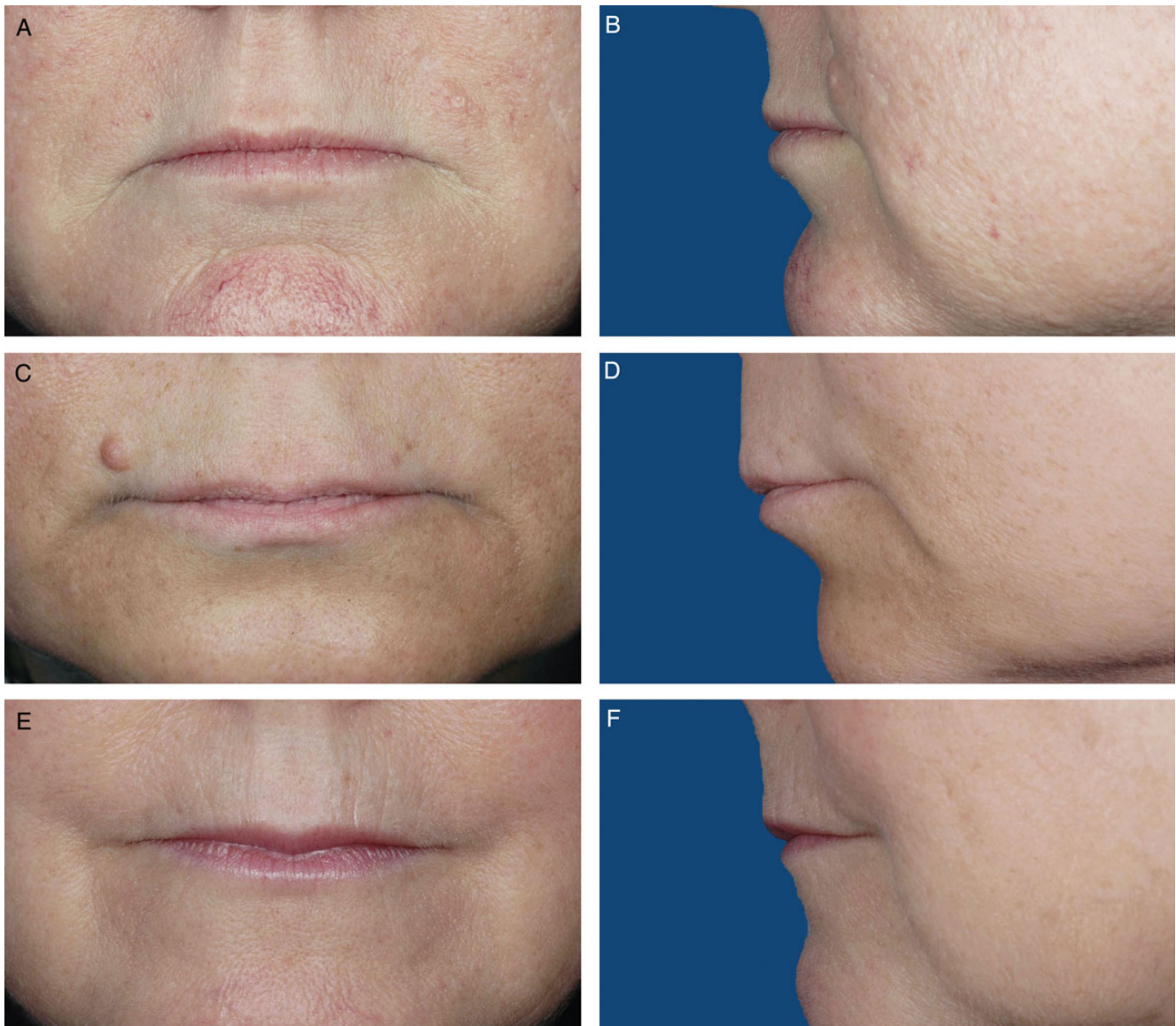


**Figure 1.** The initial Allergan Lip Fullness Scale development and validation process.

California Experimental Subject's Bill of Rights form. Subjects were excluded if they had facial hair covering the area around the mouth; permanent lipstick; lip tattoos; lip or tongue piercings; lip tint or dye application; permanent or semi-permanent lip implants; asymmetry of the lips or perioral rhytides; lip augmentation with a temporary dermal filler within 6 months prior to study entry; an infection, disorder, or scar in the lip or mouth area that would prevent adequate study assessments; or a planned dental or cosmetic procedure in the lower face between screening and the day of the iLFS evaluation. Before physician evaluation of lip fullness, subjects removed all make-up below the eyes and all jewelry from the face, ears, and neck and, if needed, wore a headband to hold their hair off their face. During screening, a physician evaluated each subject's lips and mouth area, in repose

and in animation, and assigned one of four possible iLFS scores (minimal, mild, moderate, marked).

The validation panel comprised eight board-certified aesthetic dermatologists or plastic surgeons who were instructed on scale application and the validation protocol before round 1. While evaluating subjects, panel members were seated in a large circle, facing away from each other (to prevent influence from other participants). Each rater received four stacks of cards, corresponding to each of the four scores on the scale. Each card included the rater number and a simple verbal description of the score. Subjects formed a queue based on a predetermined random order and moved from rater to rater in sequence. Each rater selected a card to score each subject, applied the subject's identifier label to the card, and dropped the card



**Figure 2.** (A-F) The initial Allergan Lip Fullness Scale. Minimal: Flat or nearly flat contour; minimal red lip shows.

into an envelope. This sequence was performed by all raters for each subject in the morning and was repeated for round 2 in the afternoon. In order to provide subject reliability ratings and evaluate concordance between physician and subject ratings, all subjects also provided self-assessments using the iLFS after rounds 1 and 2 of the panel assessments. All study investigators and subjects were blinded to the ratings assigned by the screening physician and other study investigators.

### Development and Validation of the Revised LFS

The iLFS was revised to accommodate a broader spectrum of facial presentations, such as the fuller lips observed in

the African American population, and to incorporate separate assessments of overall, upper, and lower lips (Figures 6-10). Two aesthetically experienced, board-certified dermatologists were consulted regarding revisions to the iLFS and recommended adding a lip fullness grade of “very marked” to the scale. The images that make up the LFS guide were selected from subjects who participated in the live training and validation exercise for the iLFS, whose images were collected for a different facial aesthetic scale, or who had participated in a clinical study of a temporary dermal filler for lip augmentation.<sup>12</sup> Two dermatologists selected subjects from photographs of African-American volunteers to represent “very marked” upper and lower lip fullness and “marked” upper and lower lip fullness. Grades



**Figure 3.** (A-F) The initial Allergan Lip Fullness Scale. Mild: Some red lip shows; no lower lip pout.

of minimal, mild, moderate, marked, and very marked are illustrated with two examples per grade in 10 frontal images paired with 10 lateral images.

The LFS was validated by a panel of 16 dermatologists and 5 plastic surgeons (Figure 11). The panel rated images, rather than live subjects, in the training and validation portions of the process. To simulate live assessment, all 3-dimensional (3D) images included multiple renderings from the original 3D images and could be rotated up to 180 degrees, to allow viewing the lips from different perspectives. Image capture was performed using the VECTRA system (Canfield Scientific, Fairfield, NJ). All subjects whose photographs were included in the LFS validation

signed an IRB-approved informed consent and authorization for use or release of health and research study information form and the California Experimental Subject's Bill of Rights form.

### Training

All panel members attended web-based training, which provided background information and instructions for employing the LFS. During training, raters evaluated 7 image pairs (14 unique 3D digital images) showing each subject before and after lip augmentation with temporary dermal filler. Images were selected such that the time period between before- and after-augmentation images ranged



**Figure 4.** (A-F) The initial Allergan Lip Fullness Scale. Moderate: Moderate red lip shows, with slight lower lip pout; may have curves.

from 2 weeks to 9 months. For each pair, raters assessed whether there was clinical difference (Yes/No) in overall lip fullness between images in that pair. The same assessment was done for upper and lower lip fullness, for each image pair. Raters then ranked the lip fullness (overall, upper, and lower) of each depicted subject in the same set of 14 images, viewed in a prespecified random order, with the LFS.

### Testing

Following training, each rater completed two independent online image rating sessions separated by an interval of at least 14 days. Raters accessed a different set of 3D images via the Canfield Scientific website and employed a validated electronic data capture system to record their ratings.

During each session, raters ranked overall, upper, and lower lip fullness of 144 3D images with the LFS. Additionally, during session 1, raters examined 38 pairs of before-and-after images. For each pair, they assessed whether the overall, upper, and lower lip fullness were clinically different (Yes/No) between images in a pair. These 38 image pairs were obtained from 29 subjects, 9 of whom provided 2 image pairs each (18 pairs) corresponding to different time points after treatment and 20 of whom provided 1 image pair each (20 pairs). All subjects, including seven whose pretreatment images were also utilized in the training session, had previously participated in a lip augmentation study.<sup>12</sup> The set of 144 3D images for rating with the LFS included 76 images from these 38 pairs.



**Figure 5.** (A-F) The initial Allergan Lip Fullness Scale. Marked: Significant red lip shows and lower lip pout; may be very curved.

## Statistics

Inter-rater agreement (ie, agreement among raters on the assessment of overall, upper, and lower lip fullness) using the iLFS and LFS was measured by Shrout-Fleiss intra-class correlation (ICC).<sup>14</sup> For the LFS, ICC(2,1) (Shrout-Fleiss random) was calculated for inter-rater agreement. Intra-rater agreement (ie, agreement between sessions 1 and 2 assessments of overall, upper, and lower lip fullness) was measured by weighted kappa co-efficients using Fleiss-Cohen weights.<sup>15</sup> Kappa scores in the range of 0.20 to 0.39 indicate “fair” agreement, 0.40 to 0.59 indicate “moderate” agreement, 0.60 to 0.79 indicate “substantial” agreement, and 0.80 to 1.00 indicate “almost perfect” agreement.<sup>14,16</sup>

To determine the value of a clinically meaningful difference in overall lip fullness using the LFS, the absolute

difference in LFS scores was calculated for each pair of pre- and post-augmentation images rated as “clinically different” for overall lip fullness. For example, if the score for pre-augmentation image was mild and post-augmentation was marked, the absolute difference in LFS score was 2 points. Absolute differences in LFS scores were summarized (mean, median, standard deviation, range) for all physicians. A similar summary of absolute differences was provided for pairs that were not deemed clinically different. Upon examination of these summaries, the clinically meaningful difference for overall lip fullness on the LFS was established. Similar analyses were performed separately for upper and lower lip fullness.

## Sample Size Consideration

Power calculations for the iLFS validation study in live subjects estimated that at least 40 subjects would be required to



**Figure 6.** (A-D) The Allergan Lip Fullness Scale. Minimal: Flat or nearly flat contour; minimal red lip shows.

complete the study in order to provide adequate power to demonstrate intra- and inter-rater reliability of the iLFS (lower boundary set at 0.70). Allowing a drop-out rate of 5% after screening, up to 100 subjects were to be enrolled and screened until at least 48 subjects (12 subjects for each of the four grades) were selected and had confirmed their availability to attend the scale evaluations. In the iLFS validation study, the average intra-rater agreement was observed to be 0.80; thus, a sample size of 144 images with two observations per image (corresponding to two rounds) would have provided >99% power to demonstrate that intra-rater agreement (weighted kappa) is significantly greater than 0.60 (ie, substantial agreement) using an F-test at a 0.05 level. The inter-rater agreement was observed to be 0.79; thus, a sample size of 144 images with 21 observations per image (corresponding

to 21 raters) would have provided >99% power to demonstrate that inter-rater agreement (ICC) is significantly greater than 0.60 (ie, substantial agreement) using an F-test at a 0.05 level. Power calculations were performed using PASS software (2008, Version 8.0.13, NCSS, LLC, Kaysville, Utah).

## RESULTS

### iLFS Validation

Based on iLFS evaluations of 55 live subjects, physician intra-rater agreement between scores from the morning and afternoon rating rounds was “substantial” according to prespecified criteria, with a mean weighted kappa value of 0.799 (standard deviation [SD]: 0.044; 95% confidence



**Figure 7.** (A-D) The Allergan Lip Fullness Scale. Mild: Some red lip shows; no lower lip pout.



interval [CI]: 0.762-0.836). Physician inter-rater agreement was “substantial” to “almost perfect,” with ICC values of 0.814 in the morning round and 0.787 in the afternoon round. Agreement between subject and physician iLFS scores was “substantial,” with Pearson correlation coefficients of 0.800 and 0.755 for rounds 1 and 2, respectively. Agreement between subject ratings for round 1 and 2 was also “substantial” (kappa: 0.733, 95% CI: 0.586-0.879; weighted kappa: 0.790, 95% CI: 0.667-0.912).

### LFS Validation

Twenty-one raters completed web-based training and one testing session; 20 completed two sessions, and these raters comprised the analysis dataset. Demographic characteristics

of the subjects in the 144 images that were part of the LFS validation study are summarized in Table 1.

Inter-rater agreement was  $\geq 0.69$  (substantial agreement) for all lip regions and intra-rater agreement between sessions 1 and 2 was almost perfect (Table 2). The median differences in LFS ratings for overall, upper, and lower lip fullness assessments were 1, 1, and 1, respectively, for subject pairs identified as “clinically different” and 0, 0, and 0, respectively, for subject pairs identified as “not clinically different” (Table 3). The mean LFS rating differences for overall, upper, and lower lip fullness ranged from 0.63 to 0.69 for “clinically different” pairs and from 0.28 to 0.36 for “not clinically different” pairs. Therefore, it can be interpreted that, for images deemed clinically different, the majority had an approximate 1-point difference in LFS score, and the remainder



**Figure 8.** (A-D) The Allergan Lip Fullness Scale. Moderate: Moderate red lip shows with slight lower lip pout.

were scored identically. This may have resulted from raters identifying a marginally meaningful difference between the image pairs but not identifying enough gradation on the scale to provide different LFS grades. Based on these observations, a 1-point change in overall, upper or lower lip fullness on the LFS can be established as a clinically meaningful difference.

## DISCUSSION

These validation studies demonstrated substantial to almost perfect inter- and intra-rater agreement for the iLFS and LFS, suggesting that multiple lip fullness assessments for the same subject or patient over time and across different raters should be reliable. Subject self-assessments using the iLFS also had substantial intra-

rater reliability and a high level of agreement with physician assessments. A 1-point change in the LFS was shown to reflect “clinically different” degrees in lip fullness, indicating that the LFS has adequate sensitivity for measuring physician-assessed, clinically meaningful differences in lip fullness.

The clinically meaningful difference in LFS scores was defined as observable differences in scale grades from the perspective of an experienced physician. This definition is somewhat subjective in nature. Although objective measurements, such as 3D imaging by VECTRA, could have been used, we believe that a graded photographic scale is easier to use and more cost-effective than 3D measurements. A recent clinical study of lip augmentation demonstrated that lip volume measured by 3D imaging correlated well with iLFS scores.<sup>12</sup>



**Figure 9.** (A-D) The Allergan Lip Fullness Scale. Marked: Significant red lip shows and lower lip pout; upper lip with moderate pout.

The iLFS was originally developed to assess lip fullness for the primary clinical study population indicated for lip augmentation treatment, which was expected to be female Caucasians. Therefore, the initial scale validation population included only female Caucasians. The grade of “very marked” and an example photograph of an African-American female were added to the LFS at the FDA’s request. The LFS scale validation study further included images of Asian and Hispanic subjects, indicating that assessing lip fullness of individuals of different races and ethnicities with the LFS is not likely to be a challenge. A clinical study of lip augmentation in males and subjects of various ethnicities is currently ongoing ([www.clinicaltrials.gov](http://www.clinicaltrials.gov) ID: NCT01197495).

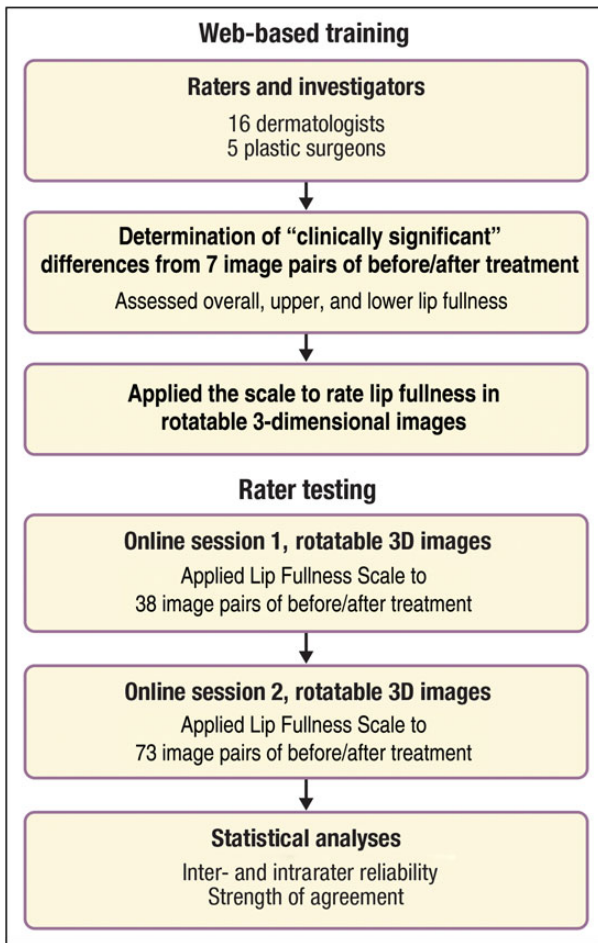
Before development of the LFS, there was a paucity of objective instruments for rating lip fullness. One such

pre-LFS scale is the Lip Fullness Grading Scale, which is a five-point scale developed by digitally morphing different lip images onto a single subject’s image.<sup>9</sup> This scale did not include African-American lip examples. Another scale, the Overall Lip Index, measures lip protrusion and height with a special ruler and may be useful in clinical practice, to follow treatment outcomes.<sup>17</sup> However, the need to perform calculations and the need for a special ruler could prove cumbersome in some situations. Unlike the development process for these other two scales, validation of the iLFS also incorporated subject assessments, which are valuable, given the importance of patient-reported outcomes in aesthetic medicine.

Two other scales have been published since the development of the LFS. The Medicis Lip Fullness Scale consists of



**Figure 10.** (A-D) The Allergan Lip Fullness Scale. Very Marked: Very significant red lip shows, lower lip pout, and lip pout.



**Figure 11.** Rater training and testing of the Allergan Lip Fullness Scale (LFS).

separate five-point scales for the upper and lower lips, with three photographs for each grade.<sup>11</sup> This scale was validated in both live subjects and photographs and was shown to have inter- and intra-rater consistency comparable to the LFS. Another lip fullness scale has a nine-level grading system (grades 1-5, with half points) with separate scales for the upper and lower lips.<sup>10</sup> It was shown to have good inter- and intra-rater consistency and correlated well with clinical evaluation and 3D measurements. Neither of these scales included examples of African-American individuals, making the LFS the only currently published scale that includes a rating grade for the marked and very marked lip fullness seen among African Americans.

### Study Limitations

The clinical significance of LFS scores was solely determined by clinicians and not by subjects. Although a 1-point change on the LFS is meaningful for a physician, it may or may not be meaningful for the subject. For some subjects, even a less than 1-point change may be meaningful if they

**Table 1.** Demographics of Subjects in 144 Images Assessed in the Allergan Lip Fullness Scale Validation Study

Characteristic	Number (%) of Images (N = 144)
Gender	
Female	131 (91.0)
Male	13 (9.0)
Age range	
20-29	24 (16.7)
30-39	26 (18.1)
40-49	36 (25.0)
50-59	38 (26.4)
60+	20 (13.9)
Race	
Caucasian	95 (66.0)
African American	27 (18.8)
Hispanic	14 (9.7)
Asian	8 (5.6)
Fitzpatrick skin phototype group	
I/II	46 (31.9)
III/IV	68 (47.2)
V/VI	30 (20.8)

**Table 2.** Physician Inter- and Intra-Rater Agreement for the Allergan Lip Fullness Scale

Lip Region	Inter-Rater Agreement ICC (2,1)		Intra-Rater Agreement Weighted Kappa
	Round 1 (N = 21)	Round 2 (N = 20)	Mean (SD) 95% CI
Overall	0.71	0.69	0.84 (0.075)
			0.809-0.880
Upper	0.74	0.71	0.86 (0.066)
			0.830-0.892
Lower	0.73	0.69	0.85 (0.054)
			0.826-0.877

CI, confidence interval; ICC (2,1), intraclass correlation (2,1) Shrout-Fleiss random; SD, standard deviation.

are looking for a subtle change, whereas some subjects may desire drastic changes and perceive those as meaningful. Hence, the LFS is not recommended for patient

**Table 3.** Differences in Allergan Lip Fullness Scale Scores Before and After Lip Augmentation – Image Pairs Rated “Clinically Different” or “Not Clinically Different” by 21 Experienced Physicians

Lip Region	“Clinically Different” Pairs			“Not Clinically Different” Pairs		
	n <sup>a</sup>	Mean (SD), 95% CI	Median	n <sup>a</sup>	Mean (SD), 95% CI	Median
Overall	577	0.63 (0.633)	1	221	0.35 (0.532)	0
		0.577-0.681			0.278-0.419	
Upper	562	0.63 (0.605)	1	236	0.36 (0.497)	0
		0.576-0.676			0.292-0.420	
Lower	545	0.69 (0.620)	1	253	0.28 (0.467)	0
		0.636-0.740			0.223-0.339	

CI, confidence interval; SD, standard deviation. <sup>a</sup>Number of pairs evaluated = 798 (38 pairs × 21 raters).

self-assessment of meaningful improvement in lip fullness, because patients will neither be trained on the scale nor have consistent agreement on meaningful results. Patient satisfaction with treatment outcomes may be assessed using existing questionnaires or scales for subject satisfaction with aesthetic outcomes.<sup>12</sup>

The verbal descriptions for each grade on the LFS are inherently subjective in nature. However, the scale is comprised of the description for each grade along with two examples of subject photographs that depict the grade, removing some of the subjectivity inherent in the descriptions, as illustrated by the substantial inter-rater agreement in the validation study.

In the study validating the first-generation iLFS, intra-rater reliability ratings were based on two rounds of evaluations of live subjects occurring on the same day. Because of FDA concerns that a greater interval between rating sessions might affect intra-rater reliability, LFS validation sessions were separated by an interval of at least 14 days. Training and testing for LFS validation studies used 3D photographic images rather than live subjects. However, a clinical trial is currently ongoing to determine the agreement between live assessment of subjects on the LFS and corresponding assessment based on subject images ([www.clinicaltrials.gov](http://www.clinicaltrials.gov) ID: NCT01197495).

Finally, information regarding the gender and ethnicity of the subjects appearing in the LFS is not available. At the time the LFS validation study was done, the approved protocol did not stipulate collection of demographic information for subjects, because fullness of the subject’s lips, and not gender or ethnicity, was the primary consideration, in terms of the utility of the scale.

## CONCLUSIONS

Intra- and inter-rater agreement for the iLFS and LFS was substantial to almost perfect among physicians. During

the iLFS validation process, substantial reproducibility and reliability for physician classification and subject self-evaluation of lip fullness were demonstrated in live subjects, thus making it a suitable measure of effectiveness in clinical studies. The LFS is also the only currently published scale that includes rating grades for the lip fullness of African Americans. The clinical significance of a 1-point difference in LFS score was established by experienced clinicians. The LFS may be a helpful tool for discussing lip treatments, setting lip treatment goals, and providing chart documentation in clinical practice settings.

## Disclosures

Dr Werschler has served as a clinical investigator, consultant, speaker, and advisory board member for, and/or received research support from, Allergan, Inc., Medcis, Merz, Suneva, and Sanofi-Aventis. Dr Fagien has served as a consultant and investigator for Allergan, Inc., has served as a clinical investigator and consultant for Medcis, and is a consultant for Merz. Ms. Thomas receives compensation in salary, as well as stock or stock options (or both), from Kythera Biopharmaceuticals. She was an employee of Allergan, Inc., and received compensation in salary, as well as stock or stock options (or both), at the time of the study. Dr Paradkar-Mitragotri is an employee of Allergan, Inc., and receives compensation in salary, as well as stock or stock options (or both). Dr Rotunda is a consultant for Kythera Biopharmaceuticals. Dr Beddingfield receives compensation in salary, as well as stock or stock options (or both), from Kythera Biopharmaceuticals. He was an employee of Allergan, Inc., and received compensation in salary, as well as stock or stock options (or both), at the time of the study.

## Funding

This study was funded by Allergan, Inc. (Irvine, CA). Funding for writing and editorial support was provided by Allergan, Inc. (Irvine, CA). Writing and editorial assistance was provided by Paula G. Davis, PhD, Ramana Yalamanchili, PhD, MBA, and Antoinette Campo of SCI Scientific Communications and

Information (Parsippany, NJ) and Linda Romagnano, PhD, and Lela Creutz, PhD of Peloton Advantage, LLC (Parsippany, NJ). Image processing assistance was provided by Canfield Imaging Systems (Fairfield, NJ).

## REFERENCES

- Romm S. The changing face of beauty. *Aesthetic Plast Surg*. 1989;13(2):91-98.
- Ioi H, Shimomura T, Nakata S, Nakasima A, Counts AL. Comparison of anteroposterior lip positions of the most-favored facial profiles of Korean and Japanese people. *Am J Orthod Dentofacial Orthop*. 2008;134(4):490-495.
- Wong WW, Davis DG, Camp MC, Gupta SC. Contribution of lip proportions to facial aesthetics in different ethnicities: a three-dimensional analysis. *J Plast Reconstr Aesthet Surg*. 2010;63(12):2032-2039.
- Holland E. Marquardt's Phi mask: pitfalls of relying on fashion models and the golden ratio to describe a beautiful face. *Aesthetic Plast Surg*. 2008;32(2):200-208.
- Klein AW. In search of the perfect lip: 2005. *Dermatol Surg*. 2005;31(11 Pt 2):1599-1603.
- Sforza C, Laino A, D'Alessio R, Grandi G, Binelli M, Ferrario VF. Soft-tissue facial characteristics of attractive Italian women as compared to normal women. *Angle Orthod*. 2009;79(1):17-23.
- Clymer MA. Evolution in techniques: lip augmentation. *Facial Plast Surg*. 2007;23(1):21-26.
- Beer KR. Rejuvenation of the lip with injectables. *Skin Therapy Lett*. 2007;12(3):5-7.
- Carruthers A, Carruthers J, Hardas B, et al. A validated lip fullness grading scale. *Dermatol Surg*. 2008;34(Suppl 2):S161-S166.
- Rossi AB, Nkengne A, Stamatas G, Bertin C. Development and validation of a photonumeric grading scale for assessing lip volume and thickness. *J Eur Acad Dermatol Venereol*. 2011;25(5):523-531.
- Kane MA, Lorenc ZP, Lin X, Smith SR. Validation of a lip fullness scale for assessment of lip augmentation. *Plast Reconstr Surg*. 2012;129(5):822e-828e.
- Fagien S, Maas C, Murphy DK, Thomas JA, Beddingfield FC III. Juvederm Ultra for lip enhancement: an open-label, multicenter study. *Aesthet Surg J*. 2013;33(3):414-420.
- Eccleston D, Murphy DK. Juvederm® Volbella in the perioral area: a 12-month prospective, multicenter, open-label study. *Clin Cosmet Invest Dermatol*. 2012;5:167-172.
- Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull*. 1979;86(2):420-428.
- Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159-174.
- Fleiss JL, Cohen J. The equivalence of weighted kappa and the intraclass correlation coefficient as measure of reliability. *Educ Psychol Measurement*. 1973;33(3):613-619.
- Lemperle G, Anderson R, Knapp TR. An index for quantitative assessment of lip augmentation. *Aesthet Surg J*. 2010;30(3):301-310.

# ASAPS Las Vegas 2016 Aesthetic Symposium

STATE OF THE ART IN FACIAL REJUVENATION

January 14–16, 2016

The Cosmopolitan Hotel of Las Vegas  
Las Vegas, Nevada



THE AMERICAN SOCIETY FOR  
AESTHETIC PLASTIC SURGERY, INC.

[www.surgery.org/lasvegas2016](http://www.surgery.org/lasvegas2016)

AMA PRA Category 1 Credit™ available

- Hands-on Cadaver Lab—Sold Out Every Year!
- Intimate Learning Environment, International Faculty
- CME Credit Available—See Website for Details