doi:10.36849/JDD.6548

 ORIGINAL ARTICLE

Safety and Effectiveness of a Novel Hyaluronic Acid Gel for Lip Augmentation

Jeffrey Adelglass MD,[a] Javier Alonso-Llamazares MD PhD,[b] Jeremy Fenton MD,[c] Michael H. Gold MD,[d] Joel Schlessinger MD,[e] Stacy R. Smith MD[f]

[a]Skintastic, Plano, TX
[b]International Dermatology Research, Inc. Miami, FL
[c]Schweiger Dermatology, PLLC New York, NY
[d]Tennessee Clinical Research Center, Nashville, TN
[e]Skin Specialists, PC, Omaha, NE
[f]California Dermatology & Clinical Research Institute Encinitas, CA

ABSTRACT

Objective: To compare the safety and efficacy of a novel hyaluronic acid injectable gel with 0.3% lidocaine (test device) with that of a commercially available injectable hyaluronic acid gel with 0.3% lidocaine (comparator) for lip augmentation.

Methods: Eligible patients (n = 158) with an overall score of very thin (n = 0) or thin (n = 1) on a 5-point Lip Fullness Grading Scale (LFGS) participated in the double-blind, randomized, multicenter study. Efficacy was assessed periodically over 6 months on a per protocol (PP) population (definitive) and a modified intent-to-treat (mITT) population (supportive).

Results: In the PP population, the mean change from baseline (day 56) in LFGS score was 1.52 for the test device and 1.53 for the comparator. This 56-day change was the primary efficacy endpoint. The 95% confidence interval (CI) limits for the mean difference in scores (test device minus comparator) were -0.33 and 0.31. In the mITT population, the corresponding 95% CI limits were -0.26 and 0.31. In both populations, the lower limits, -0.33 and -0.26, were higher than the prespecified -0.50, indicating that the test device was non-inferior to comparator. The adverse event profile was similar between the treatment groups. Ninety-three percent of patients treated with test device considered themselves improved, much improved, or very much improved at day 168 compared to 82% of those treated with comparator. The corresponding investigator improvement ratings were 100% and 76%, respectively.

Conclusion: For lip augmentation, the efficacy and safety of the test device is non-inferior to comparator.

J Drugs Dermatol. 2022;21(1):13-20. doi:10.36849/JDD.6548

INTRODUCTION

Full lips are associated with a youthful appearance, while thinning of the visible red lip indicates natural aging. Lip augmentation is a procedure designed to increase vermilion height, create pout (effacement), soften perioral lines and wrinkles, add volume, and reduce excess visible dentition, all to improve the dimensional relationship of the lips to the face. Injectable hyaluronic acid (HA) gel fillers are well-established for augmenting lips and correcting perioral rhytids.[1,2]

The objective of the present study was to compare the safety and efficacy profile of a novel HA injectable gel with 0.3% lidocaine (test device) with that of a commercially available injectable HA gel with 0.3% lidocaine (comparator) for lip augmentation. The test device consists of small, spherical, and uniform particles designed to facilitate optimal integration in the treated area, slow and predictable breakdown, and ease of injection.[3] The comparator is commercially available for (1) submucosal implantation with lip augmentation and (2) dermal implantation with correction of perioral rhytids in adults over the age of 21 years.

MATERIALS AND METHODS

Study Design

Qualified patients enrolled in the double-blind, randomized, controlled, multicenter, 6-month study of patients seeking lip augmentation. Patients were randomized 1:1 to treatment with either test device (Revanesse® Lips+, Prollenium Medical Technologies, Inc., Aurora, ON, Canada) or comparator (Restylane Silk, Galderma Laboratories, L.P., Fort Worth, TX). The test device is a clear, colorless gel in 1.0 mL pre-filled syringes with 25 mg/mL of stabilized HA and lidocaine 0.3% w/w. Comparator is a clear, colorless gel in 1.0 mL pre-filled syringes formulated to a concentration of 20 mg/mL of stabilized HA and lidocaine 0.3% w/w. The treating investigator was unblinded and the evaluating investigator was blinded to treatments administered.

Informed consent, medical history, physical examination, demographics, concomitant medications, vision evaluation, pregnancy tests, treatment administration, and diary dispensation were obtained or conducted on day 1. Patient visits to evaluate efficacy and safety occurred at week 0 (baseline, treatment) and days 28 (touch-up if necessary), 56, 84, and 168 of the study.

For patients requiring touch-up, telephone inquiries for safety were made on days 3,14, 33, 44, 112, and 140. Safety was also assessed by monitoring adverse events (AEs) at all visits. Lip function, sensation, texture, firmness, symmetry, and lip movement/function were evaluated, as well as investigator ease of use, swelling, and patient satisfaction.

Patients

Eligible patients (n = 158, men or non-pregnant or non-breastfeeding women) were over 21 years of age and had an overall score of very thin or thin on the Lip Fullness Grading Scale (LFGS). The LFGS is a 5-point photonumeric rating scale to quantify 3 dimensional fullness of the lip (0 = very thin, 1 = thin, 2 = moderately thick, 3 = thick, and 4 = full).[4] Patients sought ≥1-point improvement in overall LFGS score. Females of childbearing potential had a negative pregnancy test on day 1 and agreed to use adequate contraception during the study period. All patients provided written informed consent to participate in the study.

Exclusion Criteria

Major grounds for exclusion were pregnancy; current lactation; lip scars, tattoos, adornments, or facial hair; abnormal lip function, sensation, symmetry, or mass formation; dentures or device covering upper palate; dentofacial or maxillofacial deformities; recent plastic surgery or permanent facial implants, semi-permanent dermal filler treatment in lower face, or cosmetic procedures of the face, neck, or lips; recent use of anti-wrinkle products for lips or perioral region; current treatment with anticoagulants or related agents; recent use of systemic corticosteroids or immunosuppressive medications; current regimen of lidocaine or structurally related agent; active inflammation, infection, cancerous or pre-cancerous lesion, or unhealed wound on the face; susceptibility to scar formation; porphyria; active herpes labialis; impaired cardiac, liver, or renal function; uncontrolled undiagnosed disease; and severe cardiac disease.

Populations

To evaluate efficacy and safety, specific populations were identified for analysis. Efficacy analysis was performed on a per protocol (PP) group and a modified intent-to-treat (mITT) group and while safety and ease of use were evaluated on an as-treated (AT) population.

Procedure

Randomized patients were blindfolded before and during injections for optimal lip augmentation. Correction was to 100% of desired volume effect without overcorrection or overfilling. Diaries were dispensed with instructions to record the (1) extent (mild, moderate, severe) of bruising, redness, swelling, pain, tenderness, itching, and impact on lip function and lip sensation and (2) other AEs for two weeks after treatment. Phone calls were made to assess safety, protocol, and AEs. On days 28, 56, 84, and 168, LFGS, in-office assessments of Patient Global Aesthetic Improvement (pGAI), Investigator Global Aesthetic Improvement (iGAI), safety, and AE were performed.

Injection Technique

Antegrade linear threading and retrograde linear threading techniques were used to inject upper and lower lips of approximately half of patients. Serial puncture was used most often for perioral injections. Most lip injections were submucosal while most perioral injections were into the mid or superficial dermis.

Injection Volume

Typical usage for each treatment was specific to site and amount of augmentation or rhytids correction desired. Per US clinical studies, the maximum volume allowed per treatment was 1.5 mL per lip (1.5 for upper, 1.5 for lower) and 1.0 mL for perioral rhytid correction.

The mean volume of filler for the upper lip was 0.731 mL for the test device and 0.833 ml for the comparator on day 1 and 0.422 and 0.513 mL, respectively, on day 30. The mean volume used for the lower lip was 0.700 mL for the test device and 0.793 ml for the comparator on day 1 and 0.453 and 0.526 mL, respectively, on day 30. The mean volume used for the perioral areas was 0.769 ml for the test device and 0.797 ml for the comparator on day 1 and 0.650 and 0.813 ml, respectively, on day 30.

Assessments

Comparisons between test device and comparator for gender,

TABLE 1.

Patient and Investigator Global Aesthetic Improvement (pGAI, iGAI) Scale	
1	Worse – the appearance is worse than the original condition
2	No change – the appearance is the same as the original condition
3	Improved – obvious improvement in appearance from the initial condition, a touch up might further improve the result
4	Much improved – marked improvement in appearance from the initial condition, but not completely optimal; a touch-up might slightly improve the result
5	Very much improved – optimal cosmetic result

ethnicity, age, body mass index (BMI), and Fitzpatrick skin type were made using the AT (as treated) population and tested for significance, using $P<0.05$ as the cutoff level. Efficacy was evaluated by the LFGS, Perioral Lines at Rest Severity Scale (POL),[5] pGAI scale (Table 1), and (4) iGAI scale (Table 1). The POL is a validated 4-point scale for the most severe perioral line with the patient's mouth at rest.

The LFGS was also used to assess treatment efficacy in the upper and lower lips separately at each visit.

Primary Efficacy

The primary efficacy endpoint was change from baseline to day 56 in overall LFGS with both lips together. The goal was to disprove the null hypothesis, that test device was inferior to comparator by more than 0.50 units. Two-sided hypothesis testing was conducted for all inferential analyses with *P*-values <0.05 considered statistically significant.

It was assumed that the LFGS data were normally distributed. The 95% confidence interval (CI) for the difference in mean values between the treatment groups (test device minus comparator) was calculated. If the lower limit of the CI was above the pre-specified non-inferiority limit of -0.50, the null hypothesis was rejected to support the claim that test device was non-inferior to comparator. These calculations were performed for both the PP and mITT populations. The results for the PP population were considered as definitive and those for the mITT population as supportive.

Secondary Efficacy

Endpoints were (1) percent of responders with at least a 1-grade increase from baseline on the overall LFGS of both lips together on day 56, (2) percent of responders that achieved ≥1-point improvement [decrease in severity] from baseline on the overall POL severity scale of both lips together on day 84, (3) the change from baseline to day 84 in overall LFGS of both lips together, and (4) change from baseline to day 168 in overall LFGS of both lips together.

Other efficacy variables were (1) pGAI, iGAI, and swelling at each visit, (2) percent of patients with treatment success (upper and lower lips separately) at day 56 where responders had ≥1-grade increase from baseline on the LFGS post augmentation, (3) percent of responders (upper and lower lips separately) on the POL severity scale at day 84, where responders had ≥1-point improvement from baseline, and (4) patient satisfaction with the lips visual analog scale (0 = very unsatisfied to 100 = very satisfied).

Safety

For the AT population, treatment-emergent AEs (TEAEs), serious AEs, and AEs of special interest (vision changes or AEs events due to embolic or ischemic cause) were monitored. Other safety assessments included lip function, sensation, texture, firmness, symmetry, and movement/function, all evaluated by the blinded evaluating investigator before injection and at follow-up visits. Lip function was assessed by the patient's ability to sip liquid through a straw. Lip sensation was assessed by (1) the monofilament test, a patient's ability to feel the sensation of a 0.4G monofilament at 3 points on the upper lip and 3 points on the lower lip; and (2) the cotton wisp test, a patient's ability to feel the sensation of a cotton wisp at 3 points on the upper lip and 3 points on the lower lip. Lip texture, firmness, and symmetry were assessed as normal or abnormal.

Lip movement/function was evaluated by the patient's ability to pucker lips, blow, and pronounce words that began with the letter "W" such as water, work, week, and wind.

Ease of Use

Both devices were evaluated for ease of use (0 = not easy, 10 = most easy).

RESULTS

Of the 158 randomized patients, 141 (89.2%) completed the study. The most frequent reasons for discontinuation were withdrawal of consent in the test device group (n = 6, 7.5%) and lost to follow-up in the comparator group (n = 7, 9.0%).

Demographics

Demographics of the test device and comparator groups were compared using the AT population. Comparisons for gender, ethnicity, age, body mass index (BMI), and Fitzpatrick skin type showed that differences between the two groups were not significant except for ethnicity (P=0.044) and age (P=0.048), which were of borderline significance.

Primary Efficacy

In the PP population, the mean change from baseline to day 56 in overall LFGS with both lips together was 1.52 for the test device and 1.53 for comparator (Figure 1). The lower limit of the 95% CI for test device minus comparator was -0.33, which is higher than the pre-specified non-inferiority margin of -0.50.

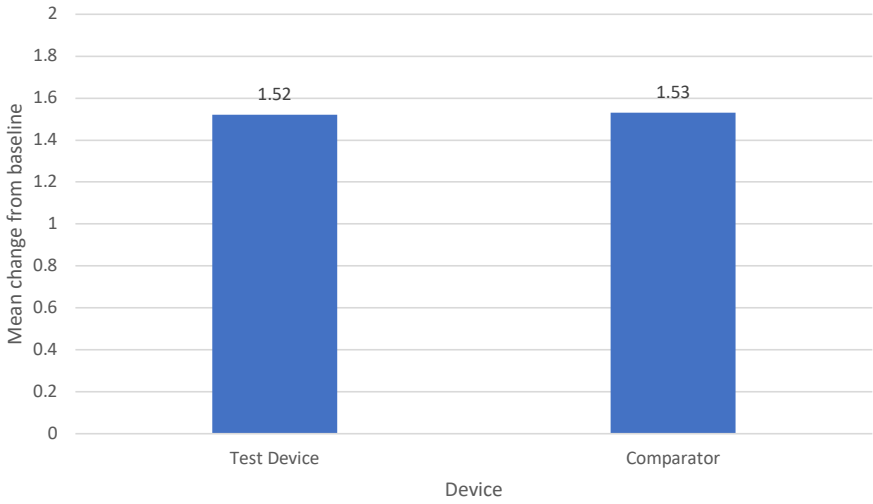
In the mITT population, the lower limit of the 95% CI test device minus comparator was -0.26, which is also higher than -0.50.

Secondary Efficacy

In the following comparisons, numerical results are limited to those of the PP population due to spatial constraints. Results for the mITT population were similar in all cases.

The percent of PP patients with at least a 1-grade increase from baseline on the overall LFGS of both lips together on day 56 was 90.7% with test device and 92.7% with comparator.

FIGURE 1. Mean change from baseline to day 56 in overall Lip Fullness Grading Scale (LFGS) with both lips together in the per-protocol (PP) population for test device and comparator.



The percent of PP patients that achieved ≥1-point improvement [decrease in severity] from baseline on the overall POL severity scale of both lips together on day 84 was 66.7% with test device and 57.1% with comparator.

The change from baseline to day 84 in overall LFGS with both lips together in the PP population was 1.37 with test device and 1.42 with comparator (Figure 2).

The change from baseline to day 168 in overall LFGS of both lips together in the PP population was 1.00 with test device and 0.93 with comparator.

pGAI

In the PP population, the proportion of patients much improved or very much improved was greatest at day 56 for both groups (81% test device, 76% comparator) and lowest on day 168 (65% test device, 44% comparator). On day 84, the proportions improved, much improved, or very much improved for test device vs comparator were 96% vs 89%, respectively, and on day 168 were 93% vs 82%, respectively. The results for each category (improved, much improved, very much improved) are shown in Figures 3 through 5.

iGAI

For the PP population, the proportions of patients much improved or very much improved was greatest at day 56 for both groups (78% test device, 78% comparator) and lowest at day 168 (46% test device, 40% comparator). At day 28, the proportions of patients much improved or very much improved were 59% and

FIGURE 2. Mean change from baseline to day 84 in Overall Lip Fullness Grading Scale (LFGS) with both lips together in the per-protocol (PP) population for test device and comparator.

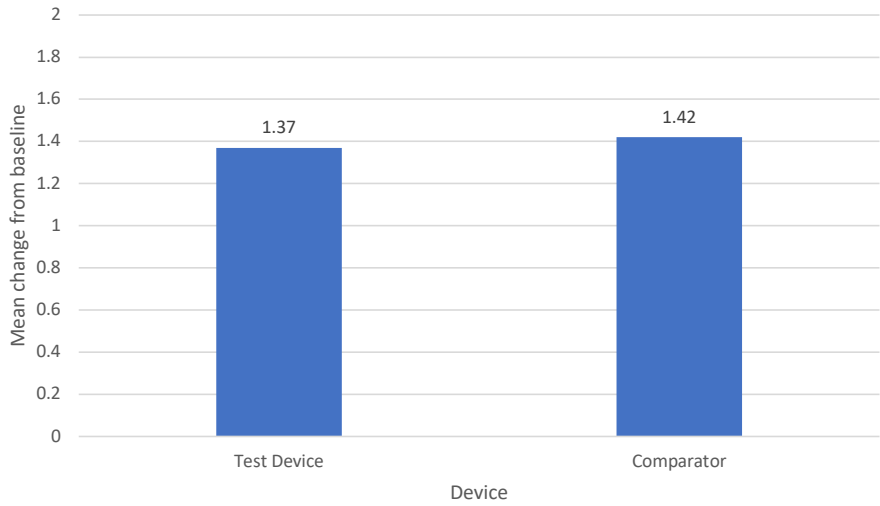


FIGURE 3. Patient Global Aesthetic Improvement (pGAI) of the per-protocol PP population on day 56 for test device and comparator.

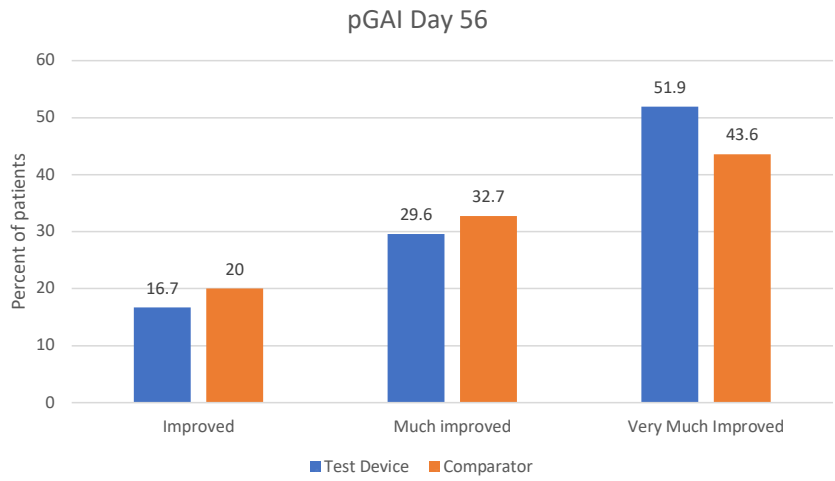


FIGURE 4. Patient Global Aesthetic Improvement (pGAI) of the per-protocol (PP) population on day 84 for test device and comparator.

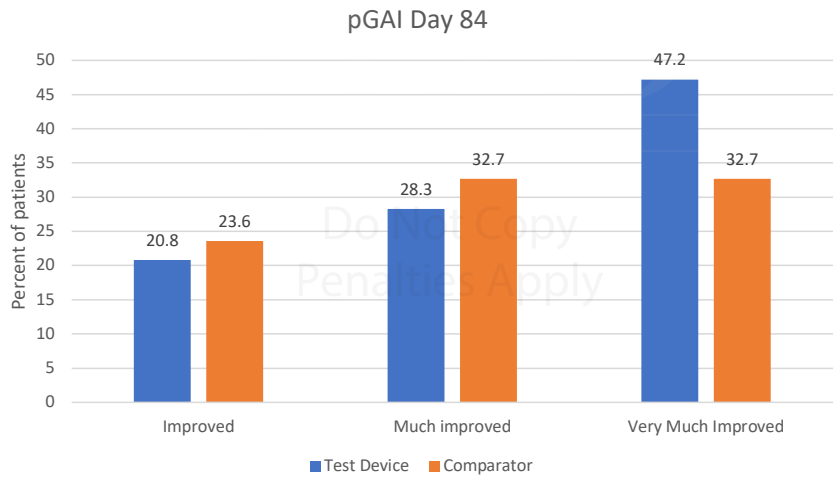


FIGURE 5. Patient Global Aesthetic Improvement (pGAI) of the per-protocol (PP) population on day 168 for test device and comparator.

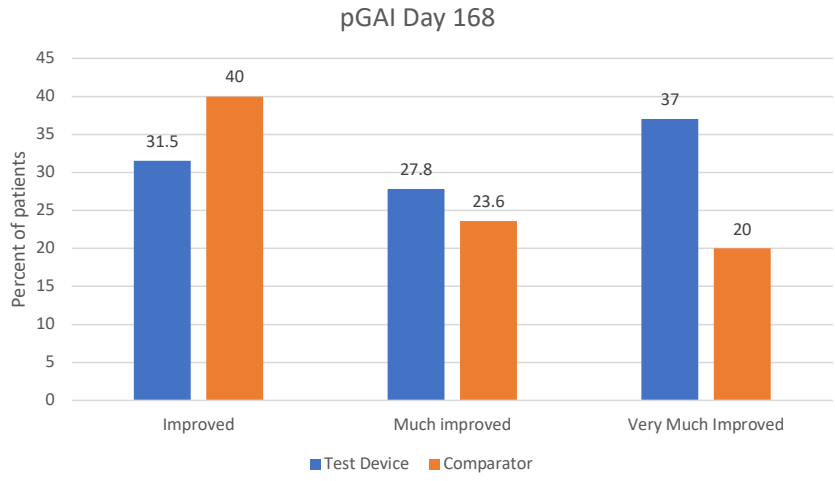


FIGURE 6. Investigator Global Aesthetic Improvement (iGAI) for the per protocol (PP) population on day 56 for test device and comparator.

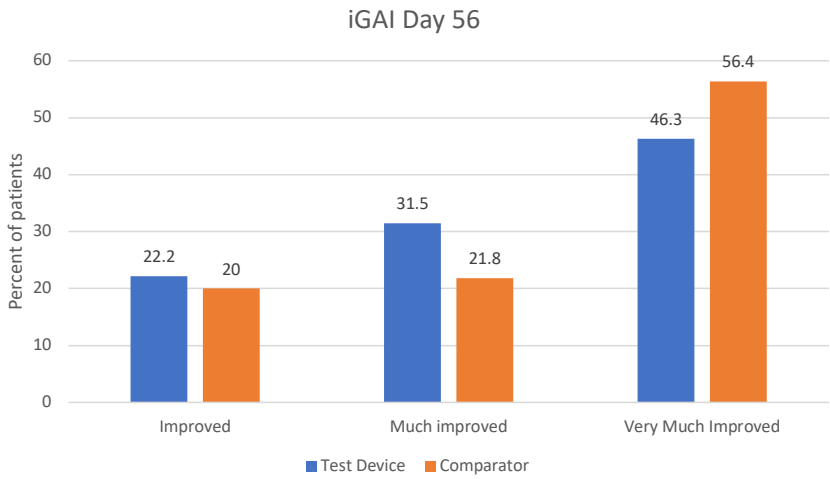


FIGURE 7. Investigator Global Aesthetic Improvement (iGAI) for the per protocol (PP) population on day 84 for test device and comparator.

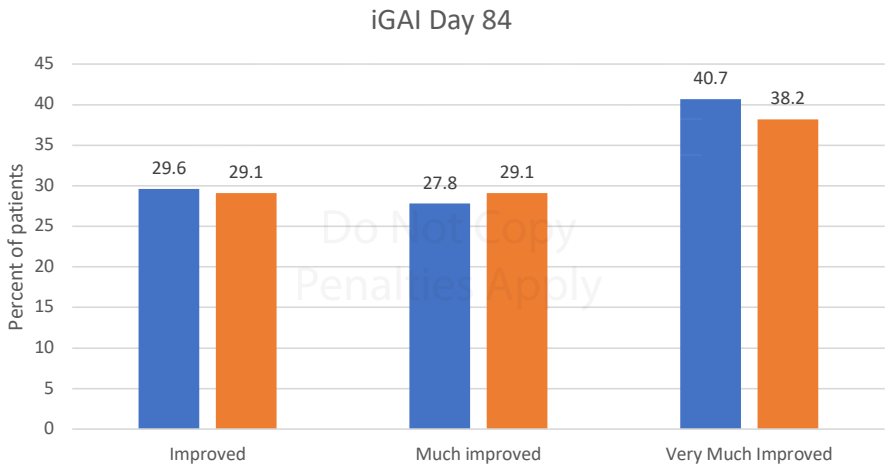


FIGURE 8. Investigator Global Aesthetic Improvement (iGAI) for the per protocol (PP) population on day 168 for test device and comparator.

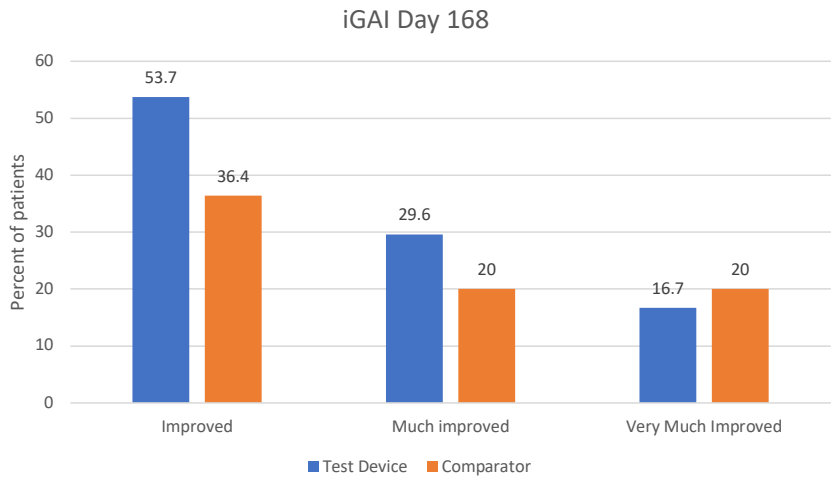
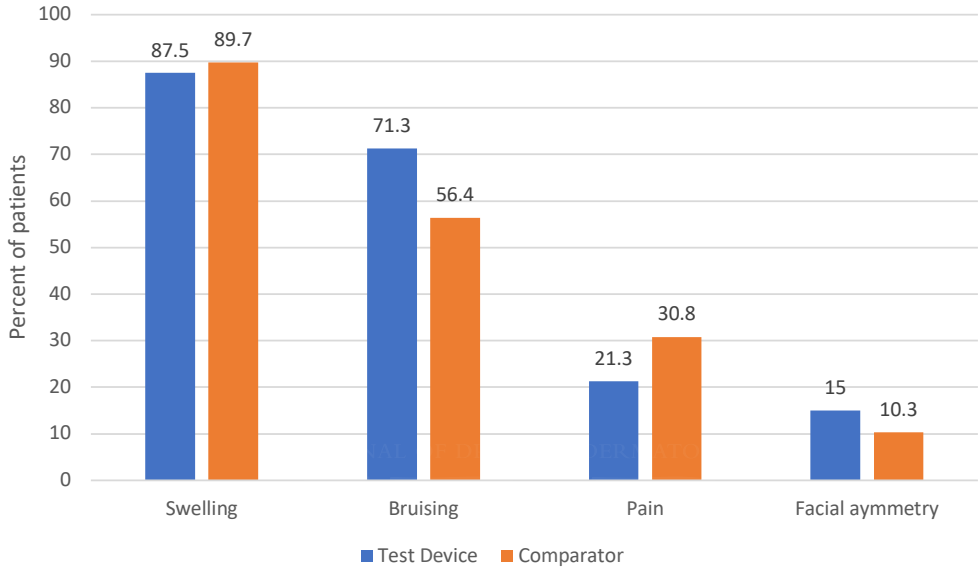


FIGURE 9. Treatment-emergent adverse events (TEAEs) appearing most frequently after treatment with the test device and comparator. Swelling, bruising, and pain were observed at the injection site.



75% for test device and comparator, respectively. On day 168 the proportions improved, much improved, or very much improved were 100% for the test device and 76% for comparator). The results for each category (improved, much improved, very much improved) are shown in Figures 6 through 8.

Swelling

For the PP population, minimal swelling was reported on day 28 for 4 (7.4%) patients with test device and on 1 (1.8%) patient with comparator. On day 56 the only subsequent swelling was minimal for 1 patient with comparator.

Upper and Lower Lips Separately

The percent of patients with treatment success at day 56 (where success was defined as achieving a ≥1-grade increase from baseline on the LFGS) was 94.4% with test device and 89.1% with comparator for the upper lips, and 92.6% with test device and 87.3% with comparator for the lower lips.

On day 84, the percent of responders on the POL severity scale (responder defined as a patient with a ≥ 1-point improvement [decrease in severity] from baseline), was 66.7% with test device and 61.9% with comparator for the upper lips, and 47.6% with test device and 42.9% with comparator for the lower lips.

Satisfaction

Mean Patient Satisfaction with Lips on the visual analog scale at each scheduled visit ranged from 76.3 to 84.4 in the test device group and from 72.8 to 86.3 in the comparator group. The mean rating was highest on day 84 in the test device group (84.4) and on day 56 in the comparator group (86.3).

Safety

For the AT population, the AE profile was similar between the treatment groups. Most patients (93.8% test device, 96.2% comparator) had TEAEs (excluding vascular injections/visual events) with the most frequent being injection site swelling, injection site bruising, injection site pain, and facial asymmetry (Figure 9). Most TEAEs were reported as mild or moderate in intensity.

Two patients in each treatment group had AEs of special interest that involved vascular injections/visual events, which were not related to treatment. One serious AE was reported, breast cancer stage II. Two of the AEs of special interest, retinal detachment (test device) and facial paralysis (Bell's palsy, comparator), were originally reported as AEs and subsequently elevated to serious AEs.

No patient discontinued the study due to a TEAE. One patient in the test device group had treatment interrupted/discontinued due to injection-site TEAEs, which were treated and resolved, and the patient completed the study. One patient discontinued the study due to pregnancy.

Regarding lip function and sensation, all patients were able to sip liquid through a straw, feel sensation of a 0.4G monofilament, and feel sensation of a cotton wisp at all visits. All except 2 patients (1 in each group) had normal lip texture and all except 1 patient (test device) had normal lip firmness as at all visits.

Their ability to pucker lips, blow with lips, and pronounce words that began with "W" were also normal. Lip symmetry was

mildly abnormal in 5.4% of patients in the test device group and 3.2% in the comparator group prior to injection at day 1. At subsequent visits, 4.1% to 11.1% in the test device group and 0% to 8.6% in the comparator group had abnormal lip symmetry that was generally mild.

Ease of Use

On a numeric scale from 0 = not easy to 10 = most easy, the median rating for each treatment group in the AT population at both day 1 and day 28 was 8.00; the range was 0 to 10 in the test device group and 6 to 10 in the comparator group.

DISCUSSION

The results of the current study demonstrate the durability of the augmentation provided by the test device and its non-inferiority to the comparator device. The data from the GAIS scoring, on which patients and physicians noted significant differences from baseline at all time points, shows the persistence of clinical effects.

The primary efficacy endpoint was change from baseline to day 56 in overall LFGS of both lips together. If the lower limit of the 95% CI for treatment difference in means (test device minus comparator) was on or above 0.50, the null hypothesis was to be rejected to support the claim that the test device is non inferior to comparator. This was achieved with a 95% CI of (0.33) for the definitive analysis in the PP population and was supported by the results (-0.26) for the mITT population. Both treatments showed high rates of treatment success across the secondary and other efficacy endpoints, as well as high rates of patient satisfaction with their lips.

The test device was well tolerated and the AE profiles of the two devices were similar. As expected, injection-site TEAEs were the most frequent AEs. Most were anticipated, mild or moderate in severity, resolved promptly. No persistent nodules, masses, or significant asymmetry were noted during the study. Treatment due to injection-site TEAEs was interrupted in only 1 patient and no patient discontinued the study due to a TEAE.

The success of the test device may be attributed to the manufacturing process which includes an advanced crosslinking process (thixofix) that promotes links between different HA polymer chains and minimizes less effective links on parts of the same chain. The resulting level of crosslinking inhibits natural degradation of filler which prolongs longevity in the treated area. The thixofix process may also minimize elongation of the upper lip over time due to edema.[6]

The strengths of the present study include the large number of patients, the use of three populations for comparisons, and the comprehensive evaluation of lip function after treatment with both devices. The encouraging results warrant additional studies to further evaluate the clinical benefits achievable by the test device.

CONCLUSION

For lip augmentation, the efficacy and safety of the test device is non-inferior to comparator.

DISCLOSURES

Dr. Smith is a consultant to Prollenium Medical Technologies, Inc. The other authors have no conflicts of interest to disclose.

REFERENCES

1. Beer K, Glogau RG, Dover JS, et al. A randomized, evaluator-blinded, controlled study of effectiveness and safety of small particle hyaluronic acid plus lidocaine for lip augmentation and perioral rhytides. *Dermatol Surg.* 2015;41(Suppl 1):S127–S136.
2. Brody-Camp S, Raggio BS. Lip Implants. 2020 Jun 18. In: *StatPearls [Internet].* Treasure Island (FL): StatPearls Publishing; 2021 Jan–.
3. Gold MH, Baumann LS, Clark CP III, Schlessinger J. A multicenter, double-blinded, randomized, split-face study of the safety and efficacy of a novel hyaluronic acid gel for the correction of nasolabial folds. *J Drugs Dermatol.* 2018;17(1):66-73.
4. Carruthers A, Carruthers J, Hardas B, et al. A validated lip fullness grading scale. *Dermatol Surg.* 2008;34(Suppl 2): S161-S166.
5. Cohen JL, Thomas J, Paradkar D, et al. An interrater and intrarater reliability study of 3 photographic scales for the classification of perioral aesthetic figures. *Dermatol Surg.* 2014;40:663-670.
6. Woodward J, Ranjit-Reeves R, Katz DF, et al. Comparing water absorption of Food and Drug Administration-approved hyaluronic acid fillers. *Dermatol Surg.* 2021;47(9):1237-1242.

AUTHOR CORRESPONDENCE

Stacy R. Smith MD

E-mail:..... srs@cdcrinstitute.com