OXFORD UNIVERSITY PRESS | Aesthetic Surgery Journal Open Forum

Comparative Analysis of Facial Rejuvenation Techniques Using Artificial Intelligence

Journal:	*Aesthetic Surgery Journal Open Forum*
Manuscript ID	Draft
Manuscript Type:	Original Article
Portal Keywords:	High SMAS < FACE, SMASectomy < FACE, SMAS plication < FACE
Portal Keywords 2:	artificial intellegence, facial analysis, machine learning
Custom keywords:	
Subject Collection:	Facial Surgery

Abstract:

Abstract:
Introduction:
Aesthetic facial surgeries historically rely on subjective analysis in determining success. This subjectivity limits our ability to objectively compare and track surgical outcomes due to the inherent biases within the examination. This study examines the use of the artificial intelligence software: FaceReader, to objectively compare three aesthetic surgical techniques for facial rejuvenation.

Methods:
32 patients who underwent facial rejuvenation surgery with concomitant procedures (Chart 1) between 01/01/2015 and 12/31/2017 were identified.10 patients underwent a SMAS plication facelift (Group A), 7 had a SMASectomy facelift (Group B), and 15 had a high SMAS facelift (Group C). Neutral repose (no expressed emotions) images pre- and post-operatively (average >3 months) were analyzed using the FaceReader software. The software tracks nearly 500 key locations on the face to accurately measure 28 action units and 7 cardinal emotions within the face. The action units and emotion are assessed for presence and intensity of functioning.

Results:
Across the procedures only Group C, the high SMAS facelift, experienced a change in emotional expression and action unit functioning (Graph 1). Post-operatively, 11/15 Group C patients' experienced activation of the lip corner puller AU, increasing in average intensity from 0% to 18.7%. This action unit pattern correlated with an average increase in detected happy emotion from 1.03% to 13.17% (p=0.008). Conversely, the average angry emotion detected decreased from 14.66% to 0.63% (p=0.032). Group A experienced a decrease in happiness by 0.84% and a decrease in anger by 6.87% (P>>0.1). Group B had an increase in happiness by 0.77% and an increase in anger by 1.91% (P>>0.1). Both Group A and B did not show any discernable action unit patterns.

	Conclusion: This study provides the first proof of concept for the use of a machine learning software application to objectively compare aesthetic surgical outcomes in facial rejuvenation. This study demonstrates the objective increased efficacy in high SMAS facelift facial rejuvenation over other compared techniques. Future applications of this software include large patient population assessment of various facial surgeries to uncover which have the greatest impact on functioning and emotional expression.

Group (n)	Avg Age	Sex	Avg Follow-up (months)	Facelift Type	Concomitant Procedures (% of total)							
					Browlift	Blepharoplasty	Necklift	Skin Revitalizing	Lip Aug.	Fat Grafting	Lipectomy	Chin Aug.
A (10)	70	9F 1M	>4.5	SMAS Plication	0	40% (4/10)	70% (7/10)	0	0	50% (5/10)	20% (2/10)	0
B (7)	69	7F	>2	SMAS-ectomy	29% (2/7)	29% (2/7)	0	43% (3/7)	29% (2/7)	71% (5/7)	0	0
C (15)	57	11F 4M	>3	High SMAS	73% (11/15)	53% (8/15)	0	0	0	87% (13/15)	0	20% (3/15)

Chart 1. Description of the facelift type and type of concomitant procedures categorized into groups: SMAS plication facelift (Group A), SMASectomy facelift (Group B), and high SMAS facelift (Group C).

655x282mm (38 x 38 DPI)

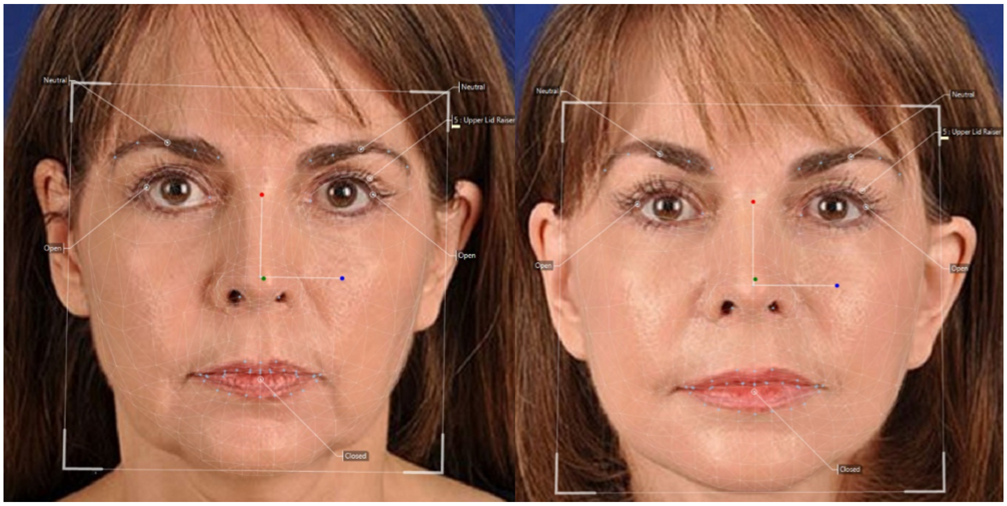

Figure 2. Representative images of the FaceReader analysis. The facial images are overlayed with a virtual mesh with labelled action units with their respective functioning. The patient images are preoperative on the left and post-operative on the right (3 months post-operatively). The patient is a 49-year-old female who underwent high SMAS facelift with lateral temporal endoscopic browlift, bilateral canthopexy, and fat transfer to the lower eye lids, midface face, jawline, chin, and upper and lower lip.

665x331mm (38 x 38 DPI)

Figure 1. Pre- and post-operative repose images. Figures A and B are exemplary images from Group A, C and D are images from Group B, and E and F are from Group C. Images A, C, and E show the preoperative images while images B, D, and F show post-operative images. The patient in Image A and B (3 months post-operatively) is an 89-year-old female who underwent a SMAS plication with fat grafting and a blepharoplasty. The patient in Image C and D (4 months post-operatively) is a 78-year-old female who underwent a SMASectomy facelift with fat grafting, dermabrasion, and lip augmentation. The patient in Image E and F (3 months post-operatively) is a 49-year-old female who underwent a high SMAS facelift with browlift, canthopexy, and fat grafting.

294x570mm (38 x 38 DPI)

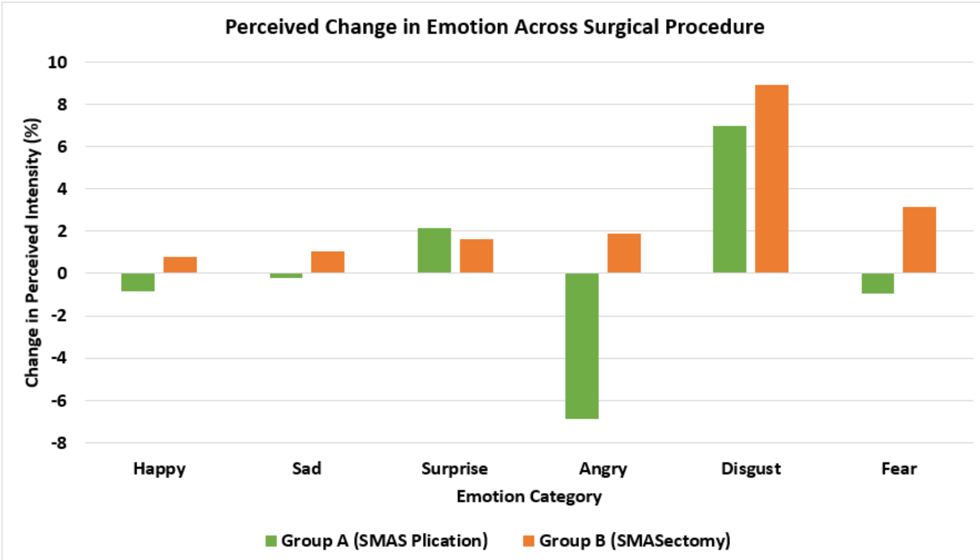

Chart 2. Graphical representation of the change in perceived emotional expression at neutral repose after surgery relative to before. Inflections below the x-axis represent decreases after surgery while positive inflections represent increases after surgery.

600x346mm (38 x 38 DPI)

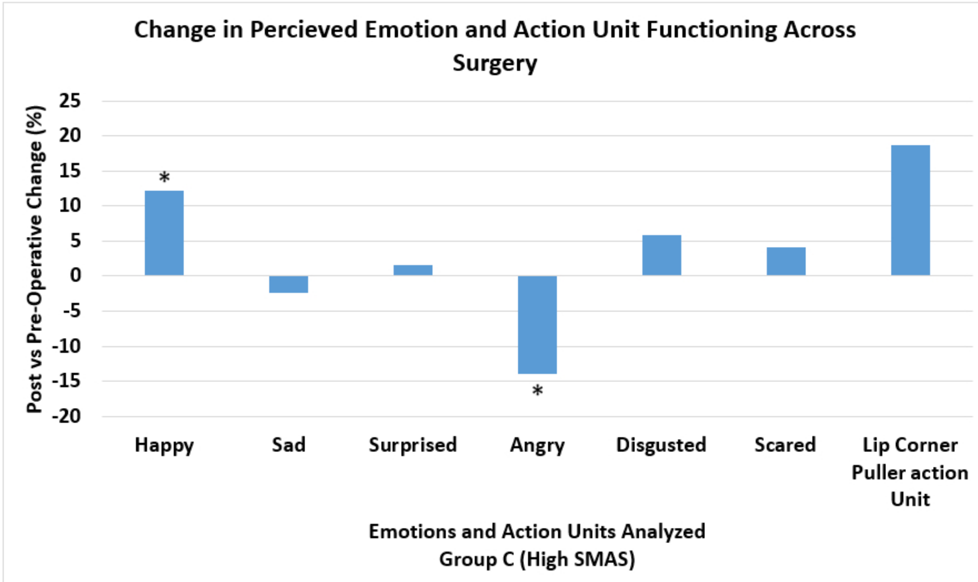

Chart 3. Outputted analysis of the perceived emotions and action unit of interest post-operatively relative to preoperatively. Inflections below the x-axis represent decreases after surgery while positive inflections represent increases after surgery. An Asterix denotes a statistically significant p value <0.05.

600x360mm (38 x 38 DPI)

Abstract:

Introduction:

Aesthetic facial surgeries historically rely on subjective analysis in determining success. This subjectivity limits our ability to objectively compare and track surgical outcomes due to the inherent biases within the examination. This study examines the use of the artificial intelligence software: FaceReader, to objectively compare three aesthetic surgical techniques for facial rejuvenation.

Methods:

32 patients who underwent facial rejuvenation surgery with concomitant procedures (Chart 1) between 01/01/2015 and 12/31/2017 were identified.10 patients underwent a SMAS plication facelift (Group A), 7 had a SMASectomy facelift (Group B), and 15 had a high SMAS facelift (Group C). Neutral repose (no expressed emotions) images pre- and post-operatively (average >3 months) were analyzed using the FaceReader software. The software tracks nearly 500 key locations on the face to accurately measure 28 action units and 7 cardinal emotions within the face. The action units and emotion are assessed for presence and intensity of functioning.

Results:

Across the procedures only Group C, the high SMAS facelift, experienced a change in emotional expression and action unit functioning (Graph 1). Post-operatively, 11/15 Group C patients' experienced activation of the lip corner puller AU, increasing in average intensity from 0% to 18.7%. This action unit pattern correlated with an average increase in detected happy emotion from 1.03% to 13.17% ($p=0.008$). Conversely, the average angry emotion detected decreased from 14.66% to 0.63% ($p=0.032$). Group A experienced a decrease in happiness by 0.84% and a decrease in anger by 6.87% ($P>>0.1$). Group B had an increase in happiness by 0.77% and an increase in anger by 1.91% ($P>>0.1$). Both Group A and B did not show any discernable action unit patterns.

Conclusion:

This study provides the first proof of concept for the use of a machine learning software application to objectively compare aesthetic surgical outcomes in facial rejuvenation. Future applications of this software include large patient population assessment of various facial surgeries to uncover which have the greatest impact on emotional expression and facial function.

Introduction:

Emotional expression has been a fundamental aspect of human communication and social connection throughout our species' evolution.[1] The practical importance of emotional expression for human interaction is accompanied by the aesthetic value humans place on these emotions; as seen by the implementation of surgeries such as facial rejuvenation.[2] Universal emotional expression by means of facial movements have been subdivided into seven categories: sadness, happiness, anger, neutrality, surprise, fear, and disgust.[3,4] These emotions have been systematically linked to the functioning of facial muscle action units through the Facial Action Coding System originally developed by Ekman and Friesen in 1978.[3,4] This system has been used as a means for analyzing facial expressions and the emotions they emit.[4-6]

Throughout the past century, surgical interventions have been employed for altering apparent undesired emotional expression from facial aging.[2,6,7] When reviewing clinical practice literature, it becomes apparent that the lack of standardization in surgical techniques plagues the facial surgery profession.[2,5] The utilization and adoption of so many varying surgical techniques is partially a lack of objective pre- and postoperative analysis.[2,5] Without an objective measure of surgical outcomes, techniques will continue to vary and patient outcomes will remain at the discretion of the patient, the operating team, and their colleagues.[2] Subjective analysis is a common but unstandardized practice for determining the success and efficacy of surgical interventions and is subject to many biases.[2] These outcome measures remain crucial from the evaluation of a successful surgery, however, objective emotional expression as a form of unbiased evaluation of outcome can be an additional valuable measure of success.

The presented research comparatively assesses the techniques of facial rejuvenation via high SMAS, SMAS plication, and SMAS-ectomy facelift with concomitant procedures using artificial intelligence (see figure 1). These patients often present expressing that their faces appear tired, sad, or angry, even when in a neutral repose. In this paper, we examine the use of artificial intelligence and deep learning software as a novel objective measure for comparing surgical outcomes in facial rejuvenation patients. The aim of this research is the examination of the efficacy of three facelift techniques and the exploration of a universal objective measure of surgical outcomes with respect to emotional expression.

Methods

I an attempt to demonstrate the ability to use artificial intelligence to compare surgical techniques, we examined facial rejuvenation patients at a tertiary medical center with stratification by procedure type.

All patients who underwent facial rejuvenation surgery (high SMAS, SMAS plication, and SMAS-ectomy facelift with possible concomitant procedures (see figure 1)) between January of 2015 and December of 2017 were retrospectively identified. The identified patients were then subcategorized by their facelift type for evaluating the differences between surgical techniques. The groups were defined into the SMAS plication (Group A), SMAS-ectomy (Group B), and high SMAS facelift (Group C) categories. The surgeries were completed by three surgeons over a 2-year span. Two surgeons completed the surgeries in Groups A and B while 1 surgeon completed the Group C surgeries. We excluded those patients without postoperative photos, those who did not want their photos shown, ones with postoperative photos without resolution of bruising, those with concomitant diagnoses such as facial paralysis, and patients who underwent additional complex head and face reconstructive procedures.

We obtained pre- and post-operative images in repose for all patients (n=32), for a total of 64 images. The post-operative photos were taken on follow-up with slight variation between the groups in the average length of follow-up (figure 1). Photos were obtained using a Canon XH-A1S 3CCD HDV Camcorder positioned 1.5 meters away from the patient. Images were analyzed using a commercially available facial expression recognition software package (FaceReader™, Noldus Information Technology BV, Wageningen, Netherlands). The data generated from the software reflected the proportion of each of the seven cardinal emotions and the associated functioning of the action units. The software's capability to classify facial expressions was achieved by training an artificial neural network, using more than 10,000 images that were manually annotated by trained experts.[8-10] The system assesses the movements of more than 500 facial landmarks on each face to perform the classification.

The emotion results outputted by FaceReader are in percentages with the total equaling 100%. Therefore, a 5-point increase in an emotion across the surgery would represent a change in perceived emotion from a value of n% to a value of n+5%. The emotions of most keen interest were happiness, sadness, surprise, and anger, due to the aim of the surgeries being a reduction in negative emotion with an increase in appearing happy and youthful without the creation of a surprised complexion. The action unit function intensity is reported as a value of 0 through 5 with 0 being undetected and 5 being maximal functioning. Therefore, a 1-point increase in functioning across the surgery would represent a change of intensity of 20%. Action units involving the lower half of the face were of most keen interest due to their direct change from the facelift surgery.

Differences between variables were assessed for variation within the samples before comparative analysis using dependent and independent T-tests. A value of $p < 0.05$ was considered statistically significant.

The authors have no disclosures with regards to the production of this original article.

Results

A total of 32 facial rejuvenation patients (27 females, 5 males) were included in the study after exclusion requirements were met. All patients were Caucasian of Fitzpatrick type III or lower with an average age of 63.7 years. Exemplary images of the patient images pre- and postoperatively are provided in Figure 2, with a FaceReader analysis examples shown in Figure 3.

Group A and Group B:

Among the Group A (n=10) and Group B (n=7) cohorts, there was no observed significant change across the surgery in either perceived emotional expression or action unit functioning (Figure 4.). Among the average emotional changes in Group A, happiness decreased from 3.57% to 2.73%, sadness decreased from 4.8% to 4.61%, surprise increased from 7.35% to 9.5%, and anger decreased from 8.69% to 1.82% (no values observed to be statistically significant). The average emotional changes for Group B were observed as an increase in happiness from 0% to 0.77%, an increase in sadness from 12% to 13.1%, an increase in surprise from 0.96% to 2.57%, and an increase in anger from 1.86% to 3.77% (no values observed to be statistically significant). Evaluation of the action units provided no discernable patterns in either group.

Group C:

The unique discernable pattern in emotion for Group C was an increase in average overall perceived happiness from 1.03% to 13.17% ($p<0.01$) (Figure 5). Conversely, the average angry emotion detected decreased from 14.66% to 0.63% ($p=0.03$) (Figure 5). Other analyzed emotions fluctuated; however, these were determined to be statistically insignificant.

Specific analysis of the facial action units uncovered a correlative finding to the emotional analysis. Preoperatively, there was no activation of the lip corner puller action unit in any of the patients on software analysis. After high SMAS facelift, 11/15 patients had activation of the lip corner puller action unit; ranging in intensity from 1/5 to 3/5 (Figure 5). This was the most distinct and nearly universal change in the muscle action units. No other discernible patterns were observed across the examined cohort.

Discussion

Our report demonstrates the use of artificial intelligence in the objective comparison of emotion and facial functioning across three groups of facial rejuvenation techniques.

The repose image analysis showed that the surgical interventions varied in their ability to enact significant change within the face. In particular, facial rejuvenation aims at reversing facial wearing and thus being able to measure these subjective outcomes will ultimately inform surgical decision making.

Our examination of repose images of patients pre- and post- facial rejuvenation demonstrates an increased efficacy with the high SMAS facelift technique of rejuvenation. Within Groups A and B, it was observed that emotions and action unit functioning fluctuated across the procedure, but no discernable pattern was seen. This result illustrates the power of the software to more objectively distinguish differences in outcome across whole patient group analysis.

On the other hand, Group C patients undergoing rejuvenation with a high SMAS facelift, were shown to have a significant increase in happiness and a concurrent decrease in anger. This analysis of emotion between groups shows the objective form of comparison for a previously subjective critique of post-operative outcomes.

Unfortunately, the study has considerable limitations within the comparative groups making these findings biased. In particular the groups differ substantially in a few categories. Demographically, for example, Group A and B had an average age nearly a decade older than Group C. Younger patients generally respond better to rejuvenation making the comparative analysis biased. As well, Group B had less than half the average follow-up time compared to Group A. This may impact findings despite rhytidectomy resulting in its final aesthetic impact on the face relatively quickly after the procedure has taken place.

As well, it is expected that concomitant procedures will be performed during rejuvenation but within our examined patients, the 8 observed concomitant procedures were not uniformly distributed. In conjunction, multiple surgeons completed the cases which can skew their concomitant procedure recommendations as well as their variability in surgical efficacy. Finally, all the examined patients were of a low Fitzpatrick skin type which limits the generalizability to patients with darker skin types.

With an understanding of the limitations of the analyzed groups we may be better equipped to evaluate the use of artificial intelligence in their objective comparison. The emotion results outlined above show the capability of the software for detecting differences in emotion post-operation. Furthering the software's utility, is the action unit analysis which can show correlative findings with emotion or separate muscle functioning analysis.

Within Group C, post-operatively it was shown that there was a significant increase in the functioning of the lip corner puller action unit, a trend unobserved in the other surgical groups. The high SMAS facelift raises the SMAS in the vector of the zygomaticus major[6,11-15], this results in the observed increase in its perceived functioning. The increase in happiness correlates with the understanding that the lip corner puller action unit, the action unit of zygomaticus major, is crucial for the formation of a smile[16,17]. The correlated increase in both happiness and the lip corner puller action unit functioning shows internal validity for the correct objective analysis by the artificial intelligence. It may be posited that the high-SMAS facelift is better equipped at visualizing the plane and vector of the zygomaticus major,[11] relative to the other two techniques, especially with variability in the angle of vector of zygomaticus major.[11,] However, it is accepted that general elevation along an acute angle relative to the oral commissure will result in rejuvenation. [6,11-15] Once again, this support is to show how internal validity of the artificial intelligence software for analysis as the many potential confounders may be the reason for the differential findings between the groups.

Our aim was to examine the available methods of postoperative examination in hopes to have a standard method of assessing surgeries and informing patient decision making. If a subset of surgical techniques is deemed the standard, techniques should be examined and compared objectively so that patients' satisfaction can only increase through ensuring the most effective surgery is provided. Previous attempts at this goal have included patient satisfaction surveys, perception of the patients by other people, quality of life measurements, anthropometric measurements, and three-dimensional digitization of landmarks.[7,16,18] Additional objective measures have been developed incorporating technology such as eye tracking for facial procedures.[5,19]

One major point of contention for the use of these methods is the innate biases such as subconscious perception and contextual information that impact their results. [9,10] For examining aesthetic surgeries, it is possible to say that the patient's satisfaction is the most crucial post-operative analysis.[6] However, assessing the success of a surgical technique requires objective evaluation which the previously mentioned techniques lack. Our favorability falls on the artificial intelligence software, Facereader, which removes theses biases through its developmental process and extensive testing. It utilizes the pioneered Facial Action Coding System by Ekman and Friesen in 1978 to develop a comparative accuracy of 80%.[3,18,20] In conjunction with the removed biases and variability, our research demonstrates that the software is able to quantify the subjective measures in a standardized way allowing for direct comparison of surgical techniques. Due to our limitations, direct group conclusions can not be drawn and future work will require more homogenized cohorts for examining surgical efficacy relative to other techniques.

Conclusion

Due to the ease and availability of the artificial intelligence software from Noldus Information Technology, its efficacy must be assessed for measuring aesthetic outcomes in facial surgery. The outputted numerical values for historically subjective measures like perceived emotions provides a reliable and comparable measure which can avoid the inter and intraobserver biases. Removal of these biases with the direct production of quantitative results created the avenue for direct surgical comparison. This technique will aid in informing patients for the rational for technique selection as well as how it may impact their complexion. Especially as surgical techniques become more abundant, it will be crucial for surgeons to be able to assess and compare techniques to ensure the greatest chance of success for their patients.

References

1. Jablonka E, Ginsburg S, Dor D. The co-evolution of language and emotions. *Philos Trans R Soc Lond B Biol Sci.* 2012;367(1599):2152-2159.
2. Barrett DM, Gerecci D, Wang TD. Facelift Controversies. *Facial Plast Surg Clin North Am.* 2016;24(3):357-366.
3. Ekman P, Friesen W. *Facial Action Coding System: A Technique for the Measurement of Facial Movement.* Consulting Psychologists Press; 1978.
4. Clark EA, Kessinger J, Duncan SE, et al. The Facial Action Coding System for Characterization of Human Affective Response to Consumer Product-Based Stimuli: A Systematic Review. *Front Psychol.* 2020;11:920.
5. Asaad M, Dey JK, Al-Mouakeh A, et al. Eye-Tracking Technology in Plastic and Reconstructive Surgery: A Systematic Review. *Aesthetic Surgery Journal.* 2020.
6. Charafeddine AH, Drake R, McBride J, Zins JE. Facelift: History and Anatomy. *Clin Plast Surg.* 2019;46(4):505-513.
7. Reilly MJ, Tomsic JA, Fernandez SJ, Davison SP. Effect of facial rejuvenation surgery on perceived attractiveness, femininity, and personality. *JAMA Facial Plast Surg.* 2015;17(3):202-207.
8. K Scherer PE. Handbook of Methods in Nonverbal Behavior Research. *Cambridge Univ Press; Cambridge, UK.* 1982.
9. Kehrer A, Engelmann S, Bauer R, et al. The nerve supply of zygomaticus major: Variability and distinguishing zygomatic from buccal facial nerve branches. *Clin Anat.* 2018;31(4):560-565.
10. van der Schalk J, Hawk ST, Fischer AH, Doosje B. Moving faces, looking places: validation of the Amsterdam Dynamic Facial Expression Set (ADFES). *Emotion (Washington, DC).* 2011;11(4):907-920.
11. Jacono AA, Bryant LM, Alemi AS. Optimal Facelift Vector and its Relation to Zygomaticus Major Orientation. Aesthet Surg J. 2020 Mar 23;40(4):351-356. doi: 10.1093/asj/sjz114. PMID: 30997513.
12. Jacono AA, Ransom ER. Patient-specific rhytidectomy: finding the angle of maximal rejuvenation. Aesthet Surg J. 2012;32(7):804-813.
13. Marten TJ. Facelift. Planning and technique. Clin Plast Surg. 1997;24(2):269-308.
14. Marten TJ. High SMAS facelift: combined single flap lifting of the jawline, cheek, and midface. Clin Plast Surg. 2008;35(4):569-603, vi.
15. Connell BF, Marten TJ. The trifurcated SMAS flap: threepart segmentation of the conventional flap for improved results in the midface, cheek, and neck. Aesthetic Plast Surg. 1995;19(5):415-420.
16. Ching S, Thoma A, McCabe RE, Antony MM. Measuring outcomes in aesthetic surgery: a comprehensive review of the literature. *Plast Reconstr Surg.* 2003;111(1):469-480; discussion 481-462.
17. Boonipat T, Asaad M, Lin J, Glass GE, Mardini S, Stotland M. Using Artificial Intelligence to Measure Facial Expression following Facial Reanimation Surgery. *Plast Reconstr Surg.* 2020;146(5):1147-1150.
18. Chauhan N, Warner JP, Adamson PA. Perceived age change after aesthetic facial surgical procedures quantifying outcomes of aging face surgery. *Arch Facial Plast Surg.* 2012;14(4):258-262.
19. Boonipat T, Abu-Ghname A, Charaffadine A, Fleming KD, Bite U, Stotland MA. Objective Outcomes in Upper Blepharoplasty. *Plastic and Reconstructive Surgery – Global Open.* 2019;7(8S-1):137-138.

20. Skiendziel T, Rösch AG, Schultheiss OC. Assessing the convergent validity between the automated emotion recognition software Noldus FaceReader 7 and Facial Action Coding System Scoring. *PLOS ONE.* 2019;14(10):e0223905.

Figure Legends

Chart 1. Description of the facelift type and type of concomitant procedures categorized into groups: SMAS plication facelift (Group A), SMASectomy facelift (Group B), and high SMAS facelift (Group C).

Figure 1. Pre- and post-operative repose images. Figures A and B are exemplary images from Group A, C and D are images from Group B, and E and F are from Group C. Images A, C, and E show the preoperative images while images B, D, and F show post-operative images. The patient in Image A and B (3 months post-operatively) is an 89-year-old female who underwent a SMAS plication with fat grafting and a blepharoplasty. The patient in Image C and D (4 months post-operatively) is a 78-year-old female who underwent a SMASectomy facelift with fat grafting, dermabrasion, and lip augmentation. The patient in Image E and F (3 months post-operatively) is a 49-year-old female who underwent a high SMAS facelift with browlift, canthopexy, and fat grafting.

Figure 2. Representative images of the FaceReader analysis. The facial images are overlayed with a virtual mesh with labelled action units with their respective functioning. The patient images are preoperative on the left and post-operative on the right (3 months post-operatively). The patient is a 49-year-old female who underwent high SMAS facelift with lateral temporal endoscopic browlift, bilateral canthopexy, and fat transfer to the lower eye lids, midface face, jawline, chin, and upper and lower lip.

Chart 2. Graphical representation of the change in perceived emotional expression at neutral repose after surgery relative to before. Inflections below the x-axis represent decreases after surgery while positive inflections represent increases after surgery.

Chart 3. Outputted analysis of the perceived emotions and action unit of interest post-operatively relative to preoperatively. Inflections below the x-axis represent decreases after surgery while positive inflections represent increases after surgery. An Asterix denotes a statistically significant p value <0.05.