OPEN FORUM

Facial Surgery

Artificial Intelligence in Surgical Evaluation: A Study of Facial Rejuvenation Techniques

Nathan S.D. Hebel, BS; Thanapoom Boonipat, MD; Jason Lin, MD; Daniel Shapiro, MD; and Uldis Bite, MD

Aesthetic Surgery Journal Open Forum 2023, 1–7

https://doi.org/10.1093/asjof/ojad032
www.asjopenforum.com

OXFORD UNIVERSITY PRESS

Abstract

Background: Aesthetic facial surgeries historically rely on subjective analysis in determining success; this limits objective comparison of surgical outcomes.

Objectives: This case study exemplifies the use of an artificial intelligence software on objectively analyzing facial rejuvenation techniques with the aim of reducing subjective bias.

Methods: Retrospectively, all patients who underwent facial rejuvenation surgery with concomitant procedures from 2015 to 2017 were included ($n = 32$). Patients were categorized into Groups A to C: Group A—10 superficial musculoaponeurotic system (SMAS) plication facelift ($n = 10$), Group B—SMASectomy facelift ($n = 7$), and Group C—high SMAS facelift ($n = 15$). Neutral repose images preoperatively and postoperatively (average >3 months) were analyzed using artificial intelligence for emotion and action unit alterations.

Results: Postoperatively, Group A experienced a decrease in happiness by 0.84% and a decrease in anger by 6.87% ($P >> .1$). Group B had an increase in happiness by 0.77% and an increase in anger by 1.91% ($P >> .1$). Both Group A and Group B did not show any discernable action unit patterns. In Group C, the lip corner puller AU increased in average intensity from 0% to 18.7%. This correlated with an average increase in detected happiness from 1.03% to 13.17% ($P = .008$). Conversely, the average detected anger decreased from 14.66% to 0.63% ($P = .032$).

Conclusions: This study provides the first proof of concept for the use of a machine learning software application to objectively assess various aesthetic surgical outcomes in facial rejuvenation. Due to limitations in patient heterogeneity, this study does not claim one technique's superiority but serves as a conceptual foundation for future investigation.

Level of Evidence: 4

Editorial Decision date: March 22, 2023; online publish-ahead-of-print March 28, 2023.

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

Dr Hebel is a medical student, Mayo Clinic Alix School of Medicine, Rochester, MN, USA. Dr Boonipat is the plastic surgery chief resident, Dr Shapiro is the director of resident aesthetic surgery training, and Dr Bite is a plastic surgeon, Division of Plastic Surgery, Mayo Clinic, Rochester, MN, USA. Dr Lin is a plastic surgery resident, Division of Plastic Surgery, St. Louis University, St. Louis, MI, USA.

Corresponding Author:
Dr Uldis Bite, Division of Plastic Surgery, Mayo Clinic, 200 First Street SW, Rochester, MN 55905, USA.
E-mail: bite.uldis@mayo.edu

Presented at: The Aesthetic Meeting in April 2022, San Diego, CA, USA.

Table. 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)

Group (*n*)	Avg age	Sex	Avg follow-up (months)	Facelift type	Concomitant procedures (%total)							
					Browlift	Blepharoplasty	Necklift	Skin revitalizing	Lip augmentation	Fat grafting	Lipectomy	Chin augmentation
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)

SMAS, superficial musculoaponeurotic system.

facial rejuvenation.[2] Universal emotional expression by means of facial movements have been subdivided into 7 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 are partially a lack of objective preoperative 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 in 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 assesses the techniques of facial rejuvenation through a high superficial musculoaponeurotic system (SMAS), SMAS plication, and SMAS-ectomy facelift using objective artificial intelligence outputs of emotion and action unit functioning. These patients often present expressing that their faces appear tired, sad, or angry, even when in a neutral repose. Our aim was to use artificial intelligence to determine if the software can observe a reversal in this facial wearing. The foundation provided by this case study of artificial intelligence efficacy in surgical evaluation and comparison will further the investigation into improved postoperative assessments.

METHODS

In 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. With IRB approval by the Mayo Clinic and written consent provided by patients for the use of their images in evaluation, all patients who underwent facial rejuvenation surgery (high SMAS, SMAS plication, and SMAS-ectomy facelift, with possible concomitant procedures; Table) between January 2015 and December 2017 were retrospectively identified. The identified patients were then subcategorized by their facelift type for evaluation of the 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 3 surgeons over a 2-year span. Two surgeons completed the surgeries in Groups A and B, while 1 surgeon completed Group C surgeries. We excluded those patients without postoperative photographs, those who did not want their photographs shown, those with concomitant diagnoses such as facial paralysis, and patients who underwent additional complex head and face reconstructive procedures.

We obtained preoperative and postoperative repose affect images for all patients ($n = 32$), for a total of 64 images. The postoperative photographs were taken on follow-up with slight variation between the groups in the average length of follow-up (Table). Photographs were obtained using a Canon XH-A1S 3CCD HDV Camcorder (Ota City, Tokyo, Japan) positioned 1.5 m away from the patient. Full-face images were analyzed using a commercially available facial expression recognition software package (FaceReader, Noldus Information Technology BV, Wageningen, the Netherlands). The data generated from the software reflected the proportion of each of the 7 cardinal emotions and the associated functioning of the facial action units.[8-10]

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 $P < .05$ was considered statistically significant. The authors have no disclosures with regard to the production of this original article.

RESULTS

A total of 32 facial rejuvenation patients (27 females and 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 preoperatively and postoperatively are provided in Figure 1, with a FaceReader analysis examples shown in Figure 2. No complications requiring reoperation occurred.

Group A and Group B

Demographically, Group A had an average age of 70 (range, 56-89) with a male-to-female ratio of 1:9. Group B was similar with an average age of 69 (range, 56-78) with a male-to-female ratio of 0:7. No complications requiring reoperation occurred. Among Group A ($n = 10$) and Group B ($n = 7$), there was no observed significant change across the surgery in either perceived emotional expression or action unit functioning (Figure 3). 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). As exemplified in Figure 1A and B, 1 patient from Group A was found to have 0% happiness before surgery and 12.1% happiness post surgery. For the other emotions, surprise increased from 0% to 5.7%, and sadness and anger were not detected at either time point. This was a slightly larger net effect in terms of happiness and surprise relative to the whole group. 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). In Figure 1C and D, the

Figure 1. Preoperative and postoperative repose images. Group A: exemplary (A) preoperative and (B) 3-month postoperative images of an 89-year-old female who underwent a SMAS plication with fat grafting and a blepharoplasty. Group B: (C) preoperative and (D) 4-month postoperative images of a 78-year-old female who underwent a SMASectomy facelift with fat grafting, dermabrasion, and lip augmentation. Group C: (E) preoperative and (F) 3-month postoperative images of a 49-year-old female who underwent a high SMAS facelift with browlift, canthopexy, and fat grafting. SMAS, superficial musculoaponeurotic system.

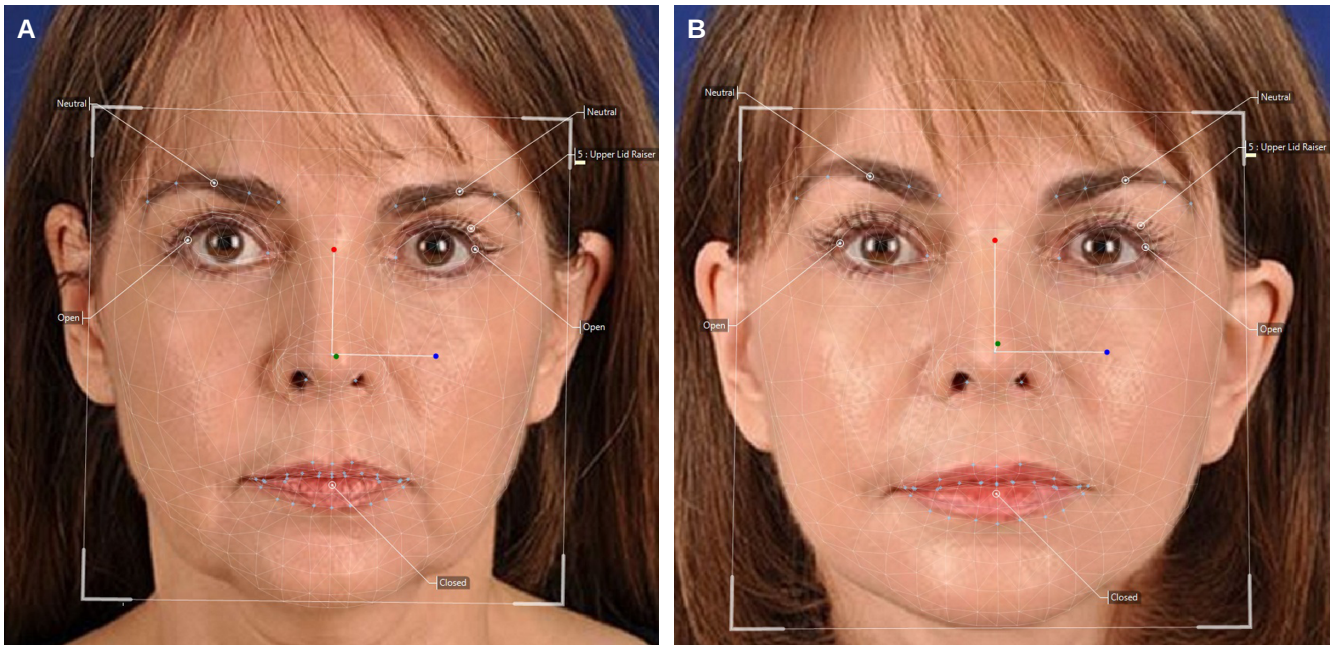

Figure 2. Representative images of the FaceReader (Noldus Information Technology BV, Wageningen, the Netherlands) analysis. The facial images are overlayed with a virtual mesh with labeled action units with their respective functioning. A 49-year-old female patient who underwent high SMAS facelift with a lateral temporal endoscopic browlift, bilateral canthopexy, and fat transfer to the lower eyelids, midface face, jawline, chin, and upper and lower lip is shown (A) preoperatively and (B) 3 months postoperatively. SMAS, superficial musculoaponeurotic system.

exemplary patient was found to have no detected happiness or anger before or after surgery. After the procedure, sadness had decreased from 15.5% to 0% and surprise had increased from 0% to 7.6%. These findings were similar to the whole cohort in terms of the minimal to no discernable change in happiness or anger but did show larger effects in surprise and sadness relatively. Evaluation of the action units provided no discernable patterns in either group.

Group C

Demographically, Group C varied from the other groups with an average age of 56 (range, 46-68) with a male-to-female ratio of 4:11. No complications requiring reoperation occurred.

In terms of emotion, the discernable pattern in Group C was an increase in average overall perceived happiness from 1.03% to 13.17% ($P < .01$) (Figure 4). Conversely, the average angry emotion detected decreased from 14.66% to 0.63% ($P = .03$) (Figure 4). Other analyzed emotions fluctuated; however, these were determined to be statistically insignificant. The exemplified patient in Figure 1E and F did not show detected happiness but did have a reduction in anger after the surgery from 9.9% to 9.4%. There was an increase in surprise and fear for this patient from 0% to 7.6% and 12.6%, respectively.

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 a high SMAS facelift, 11 of the 15 patients had activation of the lip corner puller action unit, ranging in intensity from 1/5 to 3/5 (Figure 4). This was the most distinct and nearly universal change in the muscle action units. No other discernible patterns were observed across the examined cohort. The findings of this group had been previously published in a short communication.[11]

DISCUSSION

Our report demonstrates the use of artificial intelligence in the objective analysis of emotion and facial functioning in 3 groups of facial rejuvenation techniques. These objective values of emotional expression and action unit functioning create a foundation for comparing and contrasting surgical techniques in the future. Our reported findings are significantly tempered by the heterogeneity between the groups. The findings are intended to function as a case study for the utility of artificial intelligence instead of functioning as a direct superiority analysis so of the 3 techniques.

Within our 3 groups, the artificial intelligence software reliably outputted values of emotional expression and the corresponding action unit functioning. Within Group C, postoperatively, it was shown that there was a significant

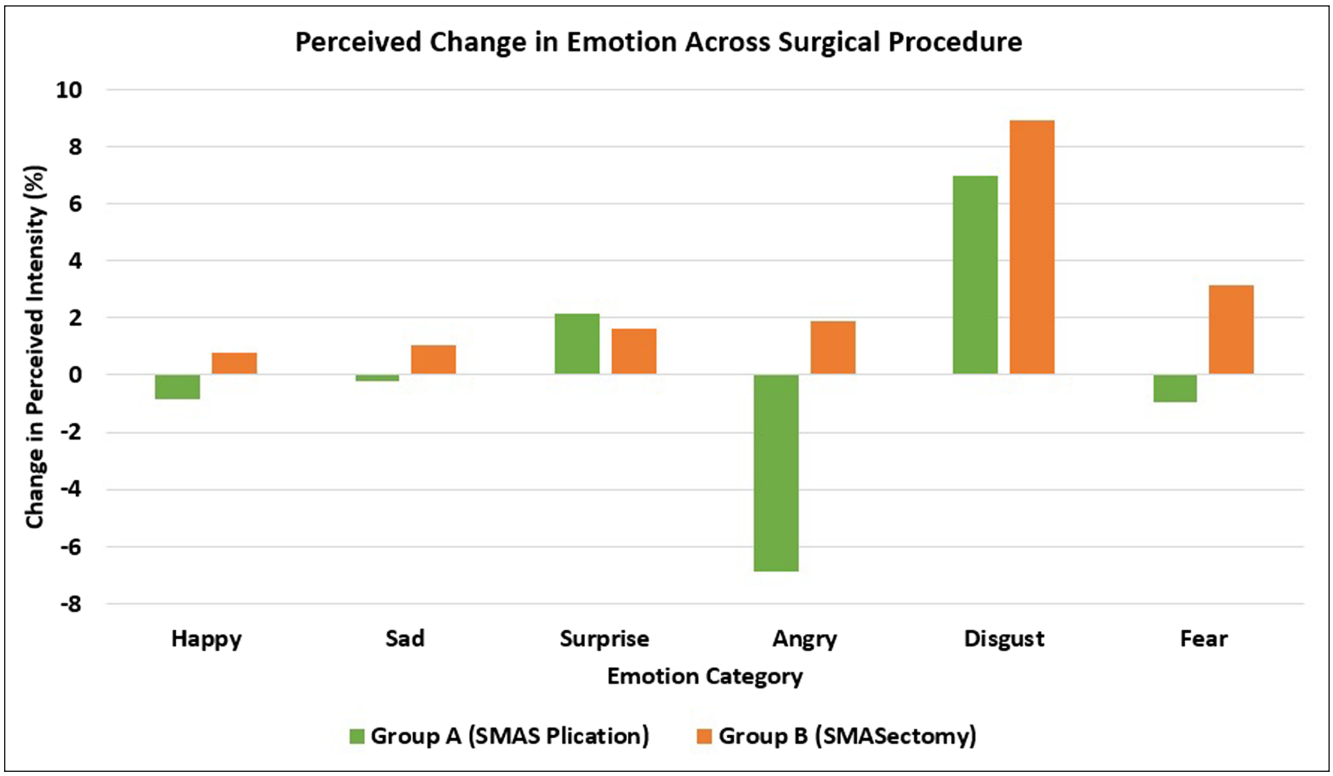

Figure 3. Graphical representation of the change in perceived emotional expression at neutral repose after surgery relative to before. Inflections below the *x*-axis represent decreases after surgery, while positive inflections represent increases after surgery. SMAS, superficial musculoaponeurotic system.

increase in the functioning of the lip corner puller action unit,[11] a trend unobserved in the other surgical groups. As well, Group C had significant increases in perceived happiness in the face.[11]

These findings are unsurprising as the high SMAS facelift raises the SMAS in the vector of the zygomaticus major[6,12-16] resulting in the observed increase in the lip corner puller action unit 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.[17,18] As well, the perception of anger and happiness should be inversely correlated by the nature of human emotion. Group C had also observed significant decreases in perceived anger. The correlated increase in both happiness and the lip corner puller action unit functioning with a concurrent decrease in anger shows internal validity for the correct objective analysis by the artificial intelligence.

However, it is accepted that general elevation along an acute angle relative to the oral commissure will result in rejuvenation.[6,12-16] All 3 examined techniques function along the vector of the zygomaticus major and should result in a similar pattern, albeit with potential variability by surgical technique. Within Groups A and B, it was observed that emotions and action unit functioning fluctuated across the procedure, but no discernable pattern was seen.

These varying output patterns by groups are impacted by the nature of the cohort and the limitations of the study. In particular, the groups differ substantially in a few categories. Demographically, Groups A and B had an average age over 10 years older than Group C. Younger patients generally respond better to rejuvenation perhaps in part due to lifestyle or intrinsic physiologic differences making the comparative analysis biased. As well, Group B had less than half the average follow-up time compared to Group A. Although we were not examining long-term outcomes, and facial rejuvenation impact should be observed after swelling and bruising have resolved in 4 to 6 weeks, the variation presents confounding factors. Finally, the ratio of males to females in the study was not equivalent further introducing potential bias.

For the studied groups, a large confounding factor was the concomitant procedures. These are to be expected in facial rejuvenation surgery but they were not uniformly observed in the cohorts. The techniques such as browlift and chin augmentation may result in more dramatic, and unknown, alterations in emotional expression, whereas variability in fat grafting can subtly alter the facelift procedures. Concomitant procedures impact on the patients is exacerbated by the 3 primary surgeons on the

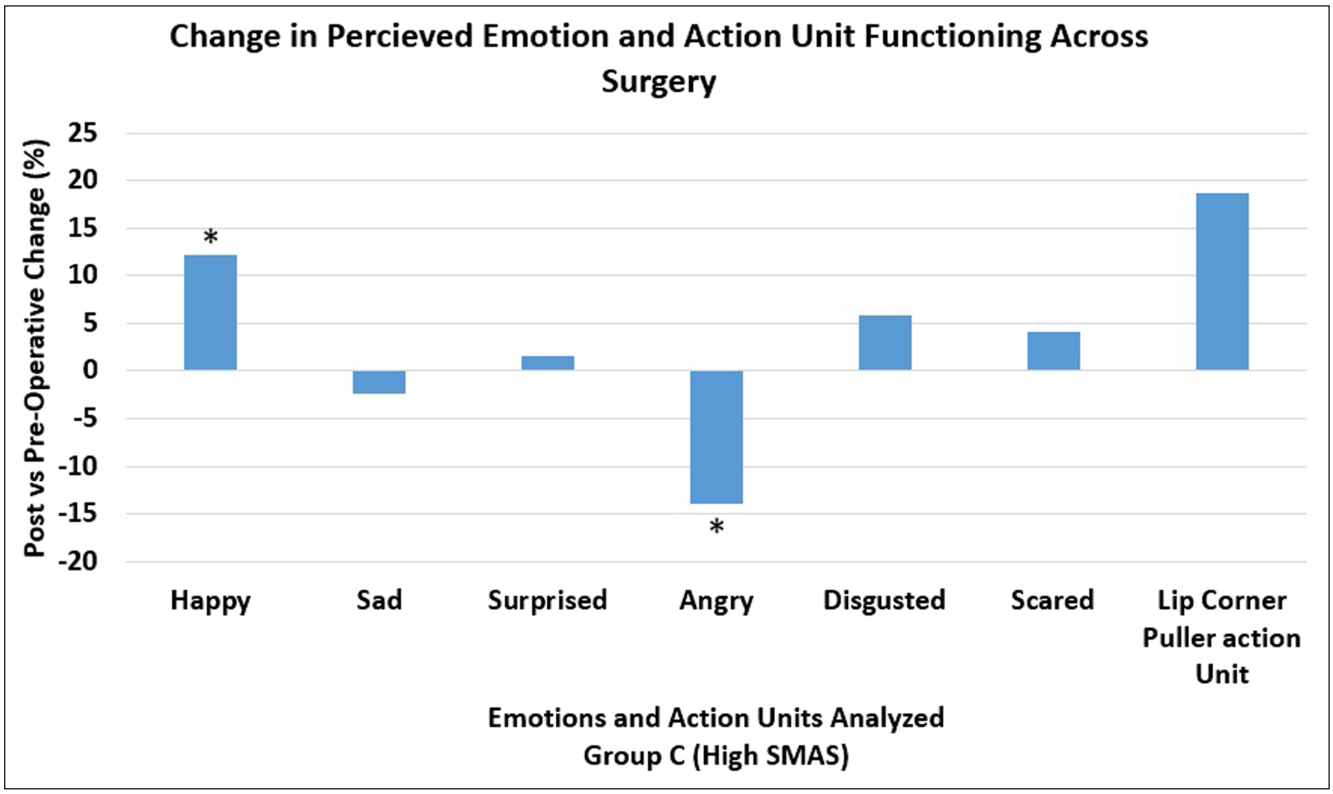

Figure 4. Outputted analysis of the perceived emotions and action unit of interest postoperatively relative to preoperatively. Inflections below the *x*-axis represent decreases after surgery, while positive inflections represent increases after surgery. An asterisk denotes a statistically significant *P* value <.05. SMAS, superficial musculoaponeurotic system.

cases which have alterations in technique and skill. In combination with their personal choice for a certain technique, the nonstandardized approach limits claims made about specific facelift techniques.

With limitations being acknowledged, it is evident that this case study cannot make claims as to the superiority of one facelift technique over the other. However, the outputted values demonstrate how in a homogenous population, surgical techniques can be assessed and compared for differential efficacy. In particular, facial rejuvenation aims at reversing facial wearing and thus being able to measure these historically subjective outcomes will ultimately inform surgical 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 by 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, 3-dimensional digitization of landmarks, and eye-tracking techniques.[5,7,17,19,20]

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 postoperative analysis[6]; however, this lacks the objectivity required for comparison. Our favorability falls on the artificial intelligence software, Facereader, which removes these biases through its developmental process and extensive testing. The software's capability to classify facial expressions was achieved by training an artificial neural network, using more than 10,000 images of males and females of varying races that were manually annotated by trained experts. The system assesses the movements of more than 500 facial landmarks on each face to perform the classification. It utilizes the pioneered Facial Action Coding System by Ekman and Friesen in 1978 to develop a comparative accuracy of 80%.[3,19,21] 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 cannot be drawn and future work will require more homogenized cohorts for examining surgical efficacy relative to other techniques.

CONCLUSIONS

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 provide a reliable and comparable measure that can avoid the inter- and intra-observer 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 of the rationale 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.

Disclosures

The authors declared no potential conflicts of interest with respect to the research, authorship, and publication of this article.

Funding

The authors received no financial support for the research, authorship, and publication of this article, including payment of the article processing charge.

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. doi: 10.1098/rstb.2012.0117
2. Barrett DM, Gerecci D, Wang TD. Facelift controversies. *Facial Plast Surg Clin North Am*. 2016;24(3):357-366. doi: 10.1016/j.fsc.2016.03.012
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. doi: 10.3389/fpsyg.2020.00920
5. Asaad M, Dey JK, Al-Mouakeh A, et al. Eye-tracking technology in plastic and reconstructive surgery: a systematic review. *Aesthet Surg J*. 2020;40(9):1022-1034. doi: 10.1093/asj/sjz328
6. Charafeddine AH, Drake R, McBride J, Zins JE. Facelift: history and anatomy. *Clin Plast Surg*. 2019;46(4):505-513. doi: 10.1016/j.cps.2019.05.001
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. doi: 10.1001/jamafacial.2015.0158
8. Ekman P, Scherer KR. *Handbook of Methods in Nonverbal Behavior Research*. Cambridge University Press; 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. doi: 10.1002/ca.23044
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*. 2011;11(4):907-920. doi: 10.1037/a0023853
11. Boonipat T, Hebel N, Zhu A, Lin J, Shapiro D. Using artificial intelligence to analyze emotion and facial action units following facial rejuvenation surgery. *J Plast Reconstr Aesthet Surg*. 2022;75(9):3628-3651. doi: 10.1016/j.bjps.2022.08.007
12. Jacono AA, Bryant LM, Alemi AS. Optimal facelift vector and its relation to zygomaticus major orientation. *Aesthet Surg J*. 2020;40(4):351-356. doi: 10.1093/asj/sjz114
13. Jacono AA, Ransom ER. Patient-specific rhytidectomy: finding the angle of maximal rejuvenation. *Aesthet Surg J*. 2012;32(7):804-813. doi: 10.1177/1090820X12455826
14. Marten TJ. Facelift. Planning and technique. *Clin Plast Surg*. 1997;24(2):269-308. doi: 10.1016/S0094-1298(20)31065-8
15. Marten TJ. High SMAS facelift: combined single flap lifting of the jawline, cheek, and midface. *Clin Plast Surg*. 2008;35(4):569-603, vi–vii. doi: 10.1016/j.cps.2008.04.002
16. Connell BF, Marten TJ. The trifurcated SMAS flap: three-part segmentation of the conventional flap for improved results in the midface, cheek, and neck. *Aesthetic Plast Surg*. 1995;19(5):415-420. doi: 10.1007/BF00453874
17. 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):469-482. doi: 10.1097/01.PRS.0000036041.67101.48
18. 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. doi: 10.1097/PRS.0000000000007251
19. 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. doi: 10.1001/archfacial.2011.1561
20. Boonipat T, Abu-Ghname A, Charaffadine A, Fleming KD, Bite U, Stotland MA. Objective outcomes in upper blepharoplasty. *Plast Reconstr Surg Global Open*. 2019;7(8S-1):137-138. doi: 10.1097/01.GOX.0000585028.07793.16
21. 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. doi: 10.1371/journal.pone.0223905