Evaluation of Changes in Facial Attractiveness and Estimated Facial Age After Blepharoplasty with an Artificial Intelligence Algorithm

Yunus Balel

1 Department of Oral and Maxillofacial Surgery, Faculty of Dentistry, Tokat Gaziosmanpaşa University, Tokat, Turkey
2 T.R. Ministry of Health, Sivas Oral and Dental Health Hospital, Sivas, Turkey

ABSTRACT

Objective: The aim of this study is to evaluate the effects of blepharoplasty operation on facial attractiveness and estimated facial age with an artificial intelligence-based algorithm over pre- and post-treatment facial photographs. In addition, it is aimed to make a comparison by reviewing the observable changes according to gender and operation type (upper, lower, combined).

Methods: Preoperative and postoperative photos of patients who underwent open access and copyright-free blepharoplasty operation on social media platforms (instagram and youtube) were collected. The photos were evaluated by an artificial intelligence algorithm trained to estimate facial age and evaluate facial attractiveness.

Results: A total of 541 patients, of which 454 (83.92%) were female and 87 (16.08%) were male. When all patients were evaluated without subgrouping, there was a -1.91±3.35 years younger face age and 0.43±0.64 point increase in facial attractiveness (p<0.005).

Conclusion: In this study, the effects of blepharoplasty on facial attractiveness and apparent age are presented with quantitative data. In addition, it has been concluded that artificial intelligence can be used in scoring the apparent age and facial attractiveness after blepharoplasty.

Keywords: Blepharoplasty; Artificial intelligence; Facial age; Facial attractiveness

INTRODUCTION

Blepharoplasty is a surgical procedure that aims to regulate the eyelids aesthetically and functionally. This procedure is usually performed to eliminate aesthetic problems such as puffiness, wrinkles and under-eye bags [1]. Blepharoplasty can also help to eliminate functional disorders of the eyelids [2], and in general, the primary purpose of individuals who have this procedure is to improve facial aesthetics [3,4].

Studies that made various anthropometric measurements of the structures of the eye were previously available in the literature[5,6], and numerous studies[3,4,7,8] have been conducted to evaluate the aesthetic results of blepharoplasty. The main evaluation criteria in these studies were related to the evaluation of professional observers or patient satisfaction. Aesthetic evaluations performed by doctors/surgeons are based on predetermined ideal facial features and rules such as the golden ratio [9]. However, aesthetic evaluations made by professionals based on these rules cannot fully reflect the aesthetic perception of the society [10,11]. Generally, in the aesthetic evaluation made in the society, how attractive the appearance of the person is measured. However, the factor of subjectivity also plays an important role in aesthetic evaluation. Subjectivity can change according to one’s own views, preferences, and values.
Therefore, one person’s aesthetic evaluation result may differ from another person’s aesthetic evaluation result. Subjectivity can emerge from various aspects in aesthetic evaluation [12]. For example, there may be features that one person may find beautiful, and features that another person may not find beautiful. Subjectivity is also affected by factors such as gender, ethnicity, age and cultural values [13]. Therefore, the management of the subjectivity factor is important in aesthetic evaluation. In this way, it can be ensured that the person makes an aesthetic evaluation according to his own views and preferences and is open to the opinions of other people.

Advances in artificial intelligence (AI) have enabled these algorithms to be used in more and more fields [9]. Estimation of facial age and measurement of facial attractiveness are among these areas. In these algorithms, artificial intelligence uses the datasets it has learned by scanning various faces and makes an attractiveness assessment in line with these data [14–16]. Quantitative data such as ideal face ratios can be used to train datasets, while their combination with data from human evaluators can also be used. In this way, in the evaluations to be made, the beauty rules taught by dictation in certain proportions and the subjectivity factor can be used together. In addition, interpretation-free and reproducible results can be obtained.

The aim of this study is to evaluate the effects of blepharoplasty operation on facial attractiveness and estimated facial age with an artificial intelligence-based algorithm over pre- and post-treatment facial photographs. In addition, it is aimed to make a comparison by reviewing the observable changes according to gender and operation type (upper, lower, combined). The hypothesis of this study is that blepharoplasty provides rejuvenation and an increase in attractiveness in patients. In addition, it is thought that these parameters are affected differently according to the type of operation and gender.

**Main Points:**
- Rejuvenation was observed in 80.4% of the patients.
- There was an increase in attractiveness in 80.5% of the patients.
- There was a higher increase in attractiveness in women than in men.
- There was no statistically significant difference in attractiveness change between the types of surgery

**MATERIALS AND METHODS**

**Obtaining the Sample**

In this cross sectional cohort study, it was aimed to obtain a large and multinational sample size. The patient photos, which were open access and copyright-free on social media platforms (instagram and youtube), were the most appropriate source for this purpose. The obtained photographic data has been used only for artificial intelligence evaluation and has not been used or reproduced for any other purpose.

Search settings for Instagram: #blepharoplasty keyword (hashtag), for Youtube: ‘blepharoplasty before and after’ words were used and the search was carried out on 20.10.2022. All posts in Instagram posts and the first 1000 posts according to relevance in Youtube posts were evaluated.

Inclusion Criteria: (I) Patients undergoing blepharoplasty, (II) Patients with preoperative and postoperative full-face photos shared, (III) Specifying which type of blepharoplasty was performed, upper, lower or combined (upper+lower).

Exclusion Criteria: (I) Shipments that do not specify the type of surgery (II). Covering the patient’s face by any method (III). Photographs without a full face photograph and showing only the eye area (IV). Presence of make-up on the patient’s face (V). Smile on the patient’s face emotional expression such as irritability (VI). Redness, bruising, swelling on the face (VII). Combination of another procedure (dermal filler, face lift, etc.) with blepharoplasty. The photographs were evaluated by an oral and maxillofacial surgeon. After applying the inclusion and exclusion criteria, appropriate photographs were used for the study.

**Artificial Intelligence Algorithm**

Built a deep Convolutional Neural Network (CNN) for multi-class age classification and attractiveness determination.

For age estimation, the model used a combination of three different datasets as training data and was trained from scratch. These datasets are the APPA-REAL dataset [9], the UTKFace dataset [17], and the IMDB-WIKI dataset [18]. Images were preprocessed by the researcher using proprietary facial recognition software and the model was evaluated in both an extended test set and Adience benchmark. In the test set, the model achieved a categorical accuracy of 51%. Age assessment gave a score between 0 and 100.
For the attractiveness score, the model was further fine-tuned using the Chicago Face Dataset [19] trained on a custom dataset from the BLINQ dating app [20]. The attractiveness rating gave scores from 1 (least attractive) to 10 (most attractive).

In order to prevent the artificial intelligence algorithm from being affected by the background, the background in the photographs of the patients was blackened using Adobe® Photoshop® CS6 software. Then, preoperative and postoperative photographs were evaluated by artificial intelligence in terms of apparent age and facial attractiveness, and the results were recorded.

**Calculating Sample Size**

In order to determine the sample size for the study, a pilot study was conducted with 30 patients. The objective of the pilot study was to evaluate the change in patients' apparent age as assessed by AI after blepharoplasty. The mean apparent age of the patients before surgery was 40.23±8.52, and the mean postoperative age was 38.45±8.95. Using these data, the effect size was calculated as 0.2035930. With an alpha margin of error of 0.05 and a power of 95%, the minimum required sample size was determined to be 263. However, in order to enhance the statistical power of the study, the researcher aimed to obtain as many samples as possible. The calculations were performed using G*Power software (latest ver. 3.1.9.7; Heinrich-Heine-Universität Düsseldorf, Düsseldorf, Germany).

**Statistical Analysis**

Descriptive analyzes were made to give information about the general characteristics of the study groups. In the representation of variables, continuous variables are in the form of mean ± standard deviation; categorical variables are indicated as n (%). In examining the normality of the variables, Kurtosis and Skewness values were examined and values between -1.5 and +1.5 were accepted as having a normal distribution [21]. Paired Samples T-Test was used to examine the change of quantitative variables before and after surgery. Repeated measures one-way analysis of variance (ANOVA) tests were used while evaluating the relationships between the data of quantitative variables and the study groups (gender-operation type). Linear regression analysis was performed to measure the relationship between two quantitative variables. In determining the degree of statistical significance, if the p value is less than 0.05, it was considered significant. IBM SPSS Version 26.0 package software (IBM Statistical Package for the Social Sciences(SPSS)Version26, SPSS inc., IBM Co., Somers, NY) was used for statistical calculations.

**RESULTS**

A total of 33,000 posts were evaluated on Instagram, while the first 1,000 posts were evaluated on Youtube, and a total of 34,000 posts were evaluated. A total of 541 patients, of which 454 (83.92%) were female and 87 (16.08%) were male, who met the inclusion and exclusion criteria from these posts were included in the study. The descriptive values obtained as a result of the evaluations made by the artificial intelligence algorithm are listed in Table 1.

A statistically significant rejuvenation and increase in attractiveness were observed after blepharoplasty operation (p<0.05) (Table 2). Rejuvenation was observed in 80.4% of the patients and an increase in attractiveness in 80.5%. In subgroup comparisons: Women compared to men, Combined blepharoplasty compared to upper and lower blepharoplasty, lower blepharoplasty compared to upper blepharoplasty provided more rejuvenation (p<0.05). There was a higher increase in attractiveness in women than in men (p<0.05). There was no statistically significant difference in attractiveness change between the types of surgery (p=0.169).

Linear regression analysis was used to identify two possible relationships. The first model (Figure 1) shows that age is related to post-operative attractiveness change (Regression coefficient: 0.135 [95% CI: -0.015; -0.004]; F:9.931; p = 0.02). According to the results of this model, it can be expected that there will be more aesthetic increase after surgery in younger patients. In the second model (Figure 2) shows the preoperative aesthetic score was not associated with the change in attractiveness after surgery (Regression coefficient: 0.059 [95% CI: -0.079; 0.014]; F:1.901; p = 0.169).

**DISCUSSION**

The evaluation to be made after any cosmetic surgery on the face shows how successful the results of the procedure are and whether the procedure has achieved its goals. Most of the blepharoplasty patients expect an increase in facial aesthetics [22]. The two most important parameters of facial aesthetics are facial attractiveness and facial age. To the best of our knowledge, this is the first study to evaluate the effect of blepharoplasty on facial attractiveness and facial age with artificial intelligence algorithms.

Due to the natural features of a person's face, it is difficult to accurately predict age based on an image. Facial aging variation
is complex and unique to a particular individual and many external factors such as lifestyle and climate [23]. Therefore, two individuals of the same chronological age may appear at different ages. In addition, there is an ordinal relationship and correlation between age tags. Age 40 is closer to 10 than 35, making age estimation more difficult compared to a problem where there is no correlation between grades [24]. Although face perception is controlled by a special region of the brain by the human visual system, age estimation through facial aesthetics can be influenced by individual, cultural and social experiences, but there is no personal information-based intervention with the interpretation of computer and artificial intelligence software applications, because only special algorithms are used [25,26].

As a result of the analysis, it was seen that blepharoplasty provided a rejuvenation in women (Mean: 2.07; p<0.001) and men (Mean: 1.03; p=0.005). The effect of blepharoplasty on facial age has been evaluated in a limited number of studies. One of these studies, Bater et al [4], evaluated the results of blepharoplasty with a questionnaire study and reported that blepharoplasty provided 1.04 years of rejuvenation, similar to our present study. From this point of view, it can be said that blepharoplasty in the appropriate indication is promising for individuals who want a younger appearance.

Table 1. The gender and type of surgery of the patients who underwent blepharoplasty, and the apparent facial age and facial attractiveness scores obtained as a result of the evaluation of the photographs by artificial intelligence.

<table>
<thead>
<tr>
<th></th>
<th>Before surgery (T0)</th>
<th>After Surgery (T1)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Apparent age (years) Mean±SD</td>
</tr>
<tr>
<td>All</td>
<td>541 (100.0%)</td>
<td>39.78±9.23</td>
</tr>
<tr>
<td>Females</td>
<td>454 (83.92%)</td>
<td>38.87±9.04</td>
</tr>
<tr>
<td>Males</td>
<td>87 (16.08%)</td>
<td>44.52±8.75</td>
</tr>
</tbody>
</table>

Subgroups by type of surgery

<table>
<thead>
<tr>
<th></th>
<th>Upper</th>
<th>Lower</th>
<th>Combined (Upper+Lower)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>116 (21.44%)</td>
<td>184 (34.01%)</td>
<td>241 (44.54%)</td>
</tr>
<tr>
<td></td>
<td>39.24±9.27</td>
<td>36.57±8.56</td>
<td>42.49±8.89</td>
</tr>
<tr>
<td></td>
<td>4.93±1.19</td>
<td>4.91±1.23</td>
<td>4.89±1.12</td>
</tr>
<tr>
<td></td>
<td>38.19±9.30</td>
<td>34.70±8.73</td>
<td>40.13±8.54</td>
</tr>
<tr>
<td></td>
<td>5.47±1.33</td>
<td>5.33±1.34</td>
<td>5.34±1.25</td>
</tr>
</tbody>
</table>

Table 2. The change in facial attractiveness scores and facial appearance perceived by artificial intelligence after surgery, and comparison between and within groups.

<table>
<thead>
<tr>
<th></th>
<th>Apparent Age</th>
<th>Attractiveness</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean Difference (T1-T0) Mean±SD</td>
<td>Impact of therapy</td>
</tr>
<tr>
<td></td>
<td></td>
<td>*p</td>
</tr>
<tr>
<td>All</td>
<td>-1.91±3.35</td>
<td>0.000</td>
</tr>
<tr>
<td>Subgroups by gender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Females</td>
<td>-2.07±3.33</td>
<td>0.000</td>
</tr>
<tr>
<td>Males</td>
<td>-1.03±3.36</td>
<td>0.005</td>
</tr>
<tr>
<td>Subgroups by type of surgery</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Upper</td>
<td>-1.05±3.48</td>
<td>0.002</td>
</tr>
<tr>
<td>Lower</td>
<td>-1.87±2.96</td>
<td>0.000</td>
</tr>
<tr>
<td>Combined (Upper+Lower)</td>
<td>-2.35±3.50</td>
<td>0.000</td>
</tr>
</tbody>
</table>

* Paired Samples Test
† Repeated Measures ANOVA (Greenhouse-Geisser)

T0: Before surgery T1: After Surgery
Figure 1. Scatter diagram with changes in attractiveness caused by the blepharoplasty plotted against preoperative age. The black dotted line represents the mean improvement of attractiveness, the red line represents linear regression, and the blue lines represent the 95% confidence interval.

Figure 2. Scatter diagram with changes in attractiveness caused by the blepharoplasty plotted against preoperative facial attractiveness. The black dotted line represents the mean improvement of attractiveness, the red line represents linear regression, and the blue lines represent the 95% confidence interval.
Evaluation of facial attractiveness with quantitative parameters consisting of the ratio of several facial distances to each other may not always give accurate results [27,28]. This is a complex assessment influenced by many factors that depend on the evaluator and the evaluator. The artificial intelligence algorithm used in this study included a combination of certain face proportions along with the dataset from the BLINQ dating app [20], which included more than 17 million evaluations of more than 13,000 face images. Also, Chicago Face Dataset [19] was used to increase accuracy. Thanks to this artificial intelligence algorithm, human and computer comments are combined and the results of the surgery are evaluated in an objective and reproducible way. When the results of present study were evaluated in the facial attractiveness parameter, there was an increase in attractiveness after blepharoplasty in women (Mean: 0.47; p=0.000) and men (Mean: 0.22; p=0.001). The compatibility of the results of the present study with the results of previous studies evaluating the effect of blepharoplasty on facial attractiveness highlights the appropriateness and usefulness of artificial intelligence-based scoring.

By evaluating the results of different operation types, whether upper, lower or combined blepharoplasty was applied, rejuvenation in the perceived age in all three types of surgery also increased the attractiveness. Although combined blepharoplasty provided more rejuvenation than other types of surgery, the increase in attractiveness was less than that of other types of surgery. The model given in Figure 1 may be useful to explain this situation. Although there is no significant difference in preoperative age score between the operation type groups, patients who need combined blepharoplasty are generally older than patients who need only upper or only lower blepharoplasty. The model in Figure 1 emphasizes that the increase in attractiveness after surgery is greater in younger patients. However, it is controversial to what extent the clinical reflection of the small differences between the increases in attractiveness score after upper, lower and combined blepharoplasty can be discerned by the human eye.

Patcas et al. [9] evaluated the changes in apparent age and facial attractiveness after orthognathic surgery with an algorithm similar to the artificial intelligence algorithm in the presented study. In this study, they reported that changes in attractiveness were associated with baseline attractiveness, not age at baseline. The findings of the presented study contradict the findings of the study of Patcas et al.[9] while the increase in attractiveness score in the blepharoplasty patient population was not affected by baseline attractiveness, more attractiveness increased in younger patients. This situation may have arisen for two reasons. First, the blepharoplasty patient population is older than the orthognathic surgery patient population. Secondly, the changes made in the jaws may be perceived differently than the changes made around the eyes. New studies are needed to determine this situation clearly.

The use of artificial intelligence to evaluate clinical outcomes is becoming more and more common nowadays. Obtaining objective and reproducible results and continuous self-education of the algorithm by learning new information helps clinicians and patients. Creating simulation photographs by estimating the postoperative patient image and evaluating these photographs by artificial intelligence will help the clinician to provide realistic information to the patient and to ensure that the patient has realistic expectations. It will also reduce the gap between the patient’s level of aesthetic expectation and the level of aesthetics the clinician can offer.

**Limitations**

Despite all these advantages, there are also situations where artificial intelligence is disadvantageous. While the patient who applies for an increase in facial aesthetics may think that the surgical correction of the area that he thinks is the problem area is very important and very valuable, it is unlikely that this situation can be fully represented by artificial intelligence algorithms. In addition, the attractiveness evaluation based on the artificial intelligence algorithm used in this study was carried out by training the data obtained from a dating platform. However, although it is an algorithm obtained by analyzing more than 17 million attractiveness assessments [20], training these algorithms with much more data input will provide much more inclusive results.

The source of the data evaluated in this study is social media. This poses a potential risk of bias. Each surgeon uses different techniques in operations and our patient population was multinational. In addition, although images with make-up and photoshop were excluded from the study during selection, it is still not possible to guarantee that this potential risk of bias is eliminated.

**CONCLUSIONS**

In conclusion, with the artificial intelligence evaluation made
In this study, it was concluded that blepharoplasty provides an increase in facial attractiveness and rejuvenation at the estimated facial age. Combined blepharoplasty provides more rejuvenation than only upper and only lower blepharoplasty. Blepharoplasties applied in younger patients provide more aesthetic increase than those applied in older patients. In addition, artificial intelligence has shown promising performance in the evaluation of blepharoplasty results and this article is expected to guide future studies.

**Funding:** None

**Informed Consent:** In this study, which was carried out in line with the Declaration of Helsinki, the data were obtained from open access and copyright-free social media posts.

**Conflict of Interest:** No conflict of interest.

**Authors’ Contributions:** Conception; Design; Supervision; Fundings; Materials; Data Collection and/or Processing; Analysis and/or Interpretation; Literature Review; Writing; Critical Review: YB.

**Ethical Approval:** This study, which was evaluated by the Clinical Research Ethics Committee at Tokat Gaziosmanpaşa University with registration number 2023-KAEK-036, was determined to be exempt from requiring ethics committee approval and was documented as such.

**REFERENCES**


How to Cite: