Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Bézier Curve–Based Quantitative Eyelid Symmetry Analysis and Artificial Intelligence–Assisted Photographic Assessment After Upper Eyelid Surgery: A Prospective Study

Muzaffer Said Guler

said.guler@icloud.com

University of Health Sciences, Kanuni Sultan Süleyman Training and Research Hospital

Fatma Savur

Sağlık Bilimleri Üniversitesi

said kemal cengiz

Yıldız Technical University

Gülistan Oyur

Sağlık Bilimleri Üniversitesi

Fatma Selin Kaya

Sağlık Bilimleri Üniversitesi

Kubra Serefoglu Cabuk

Research Article

Keywords: Upper eyelid surgery, Bézier curves, bilateral symmetry, margin reflex distance, artificial intelligence, photographic assessment

Posted Date: November 23rd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11526076/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Background

To investigate the potential of Bézier curve–based quantitative analyses of bilateral eyelid margin and contour-shape symmetry for assessing surgical outcomes after upper eyelid surgery, and to examine their associations with assessments by experienced oculoplastic surgeons and artificial intelligence (AI)–generated photographic outcome scores.

Methods

This prospective, single-center study included 67 patients who underwent Müller muscle–conjunctival resection (MMCR; n = 26), external levator resection (n = 20), or bilateral functional upper eyelid blepharoplasty (n = 21). Preoperative and 6-month postoperative photographs were analyzed using pupil-centered and eyelid-apex height–equalized Bézier curve alignments. Mean absolute error (MAE), mean squared error (MSE), root mean squared error (RMSE), and coefficient of determination (R^2) were calculated. Associations with the assessments of two oculoplastic surgeons and AI-generated photographic scores were examined using Spearman correlation analysis.

Results

In the pupil-centered analysis, MAE decreased from 1.1 ± 0.8 to 0.8 ± 0.6 and RMSE decreased from 1.2 ± 0.8 to 0.9 ± 0.6, whereas R^2 increased from − 1.5 ± 4.2 to 0.2 ± 1.2 (all $p < 0.001$). Changes in the eyelid-apex height–equalized contour analysis were not significant (all $p > 0.05$). Postoperative pupil-centered R^2 correlated positively with surgeons' overall outcome ratings ($\rho = 0.543$) and AI-generated scores ($\rho = 0.570$; both $p < 0.001$). The overall correlation between surgeon and AI scores was $\rho = 0.760$ ($p < 0.001$).

Conclusions

Pupil-centered Bézier curve–based measures of bilateral eyelid margin symmetry showed significant postoperative improvement and were associated with surgical outcome assessments by experienced oculoplastic surgeons. AI-assisted photographic scores were also positively correlated with surgeons' assessments. These findings provide a promising basis for developing more comprehensive and standardized approaches that combine quantitative symmetry analysis with AI-assisted photographic evaluation to complement conventional margin reflex distance-1 (MRD1) measurement in the assessment of upper eyelid surgical outcomes.

Background

In the evaluation of upper eyelid surgery performed for ptosis or dermatochalasis, aesthetic features such as eyelid contour and bilateral symmetry are important in addition to functional improvement. Müller muscle–conjunctival resection (MMCR) and levator resection for ptosis aim to correct eyelid height, whereas blepharoplasty performed for pseudoptosis and superior visual field restriction due to dermatochalasis aims to remove redundant skin. Although these procedures have different surgical indications, evaluation of their outcomes may benefit from consideration of eyelid contour and interocular symmetry alongside eyelid height and functional gains [1–3].

In routine clinical practice, the success of upper eyelid surgery is frequently assessed by subjective review of preoperative and postoperative photographs, whereas objective measurements are often limited to margin reflex distance-1 (MRD1). However, a single-point measurement of eyelid height cannot adequately characterize the entire upper eyelid contour or differences in shape between the two eyes. Eyelid contour symmetry, curvature, and natural appearance are important aspects of the surgical outcome, yet they are difficult to quantify using conventional methods. Consequently, a need remains for objective and reproducible approaches to eyelid contour assessment [4, 5].

Recent advances in digital image analysis and mathematical modeling have enabled new approaches to quantitative characterization of the upper eyelid contour. In particular, Bézier curves allow mathematical representation of the upper eyelid margin and quantification of interocular contour similarity using error and goodness-of-fit metrics [6–8]. Such measurements may facilitate preoperative-to-postoperative comparisons of bilateral contour symmetry and complement observer-dependent assessments with objective data. Nevertheless, because bilateral contour similarity alone does not capture every dimension of surgical success, quantitative measurements should be interpreted alongside clinical evaluations.

Advances in artificial intelligence (AI) and computerized image analysis have also enabled new approaches to photographic evaluation of surgical outcomes [9–12]. Multimodal AI models capable of interpreting visual inputs can generate ratings based on visible features, including eyelid height, symmetry, and the natural appearance of the eyelid contour, from paired preoperative and postoperative images. Examining how such ratings relate to quantitative contour measurements and assessments by experienced surgeons may inform the future development of more comprehensive, standardized, multimodal approaches to surgical outcome assessment.

The aim of this prospective study was to quantify bilateral upper eyelid contour symmetry from standardized preoperative and 6-month postoperative photographs using Bézier curves in patients undergoing MMCR, levator resection, or upper eyelid blepharoplasty, and to investigate associations between these quantitative measurements, subjective ratings by experienced oculoplastic surgeons, and AI-generated photographic outcome scores. By jointly examining conventional clinical measurements, Bézier-derived contour symmetry metrics, and AI-assisted photographic assessments, we sought to characterize relationships among these approaches and provide a basis for future multimodal image-based systems for surgical outcome evaluation.

Methods

Study design and participants

This prospective, single-center observational study included 67 patients who underwent upper eyelid surgery at the Oculoplastic Surgery Clinic of Başakşehir Çam and Sakura City Hospital between February 1, 2024, and February 1, 2025. All procedures were performed by a single experienced surgeon. According to clinical indications, 26 patients underwent unilateral MMCR, 20 underwent unilateral levator resection, and 21 underwent bilateral upper eyelid blepharoplasty.

Surgical selection was based on levator function, the response to the phenylephrine test, and the presence of dermatochalasis. Appropriate patients exhibiting an increase of at least 1 mm in MRD1 after administration of 2.5% phenylephrine into the superior conjunctival fornix underwent MMCR, whereas patients with an insufficient response underwent levator resection. Patients with dermatochalasis-related superior visual field restriction documented by computerized visual field testing underwent bilateral functional upper eyelid blepharoplasty. Surgical approaches were selected according to clinical indications; participants were not randomly allocated to treatment groups.

MMCR was performed using a posterior transconjunctival approach with resection of the Müller muscle and conjunctiva. Levator resection was performed through a standard external approach, accessing the levator aponeurosis through the upper eyelid crease and adjusting eyelid height by appropriate resection and tarsal fixation. For blepharoplasty, redundant upper eyelid skin was excised. No patients underwent an additional simultaneous upper eyelid procedure.

Patients with traumatic, neurogenic, myogenic, or mechanical ptosis; orbital, palpebral, neurological, or systemic conditions that could affect eyelid contour; strabismus; significant corneal or ocular surface disease; ocular or eyelid surgery within the preceding 6 months; or intraoperative complications were excluded.

Standardized photography

Standardized frontal photographs were obtained preoperatively and at postoperative month 6. Head position was standardized using the chin and forehead supports of a slit-lamp biomicroscope. Images were captured under comparable lighting conditions, and a 10-mm-diameter reference marker was placed over the glabella for calibration. Before analysis, identifying information was removed and the photographs were cropped to include the periocular region.

Quantitative upper eyelid symmetry analysis using Bézier curves

Right and left upper eyelid contours were manually modeled with Bézier curves on preoperative and postoperative photographs using ImageJ (National Institutes of Health, Bethesda, MD, USA), and pupil

centers were marked. Contours traced by Said Kemal Cengiz (SKC) were exported to MATLAB R2025a. Each upper eyelid contour was represented by 1,000 (x,y) coordinate points.

The left upper eyelid contour was kept fixed, whereas the right contour was mirrored across the y-axis. Bilateral symmetry was assessed using two distinct alignment approaches.

In the first approach (pupil-centered analysis), the pupil centers of the two eyes were aligned before comparison of the upper eyelid contours. This approach preserved differences in eyelid height and quantified bilateral eyelid margin symmetry by jointly reflecting the relative vertical positions and geometric shapes of the eyelid contours.

In the second approach (eyelid-apex height–equalized analysis), the lower-positioned upper eyelid contour was translated only in the vertical direction until the apices of the two contours reached the same height. By equalizing apex height, this approach assessed geometric contour-shape similarity independently of the relative vertical displacement between the eyelids.

For both analyses, only curve segments within the common horizontal coordinate range after alignment were compared. Contour segments outside the overlapping range were excluded; the medial and lateral contour endpoints were not artificially forced to coincide. Matched vertical coordinates at corresponding x positions were used to calculate mean absolute error (MAE), mean squared error (MSE), root mean squared error (RMSE), and the coefficient of determination (R^2) [7].

The following mathematical expressions were used:

$$MAE = (1/n)\ \Sigma\ |y_i - \hat{y}_i|$$

$$MSE = (1/n)\ \Sigma\ (y_i - \hat{y}_i)^2$$

$RMSE = \sqrt{MSE}$

$$R^2 = 1 - [\Sigma(y_i - \hat{y}_i)^2 / \Sigma(y_i - \bar{y})^2]$$

Here, y_i and $\hat{y}_i$ denote the matched vertical coordinates of the two compared curves, $\bar{y}$ denotes the mean vertical coordinate of the reference curve, and n is the number of matched points. Lower MAE, MSE, and RMSE values and higher R^2 values indicate greater geometric similarity between bilateral contours.

MRD1 was additionally calculated from calibrated photographs as the vertical distance between the pupil-center level and the upper eyelid margin.

Representative preoperative and postoperative photographs illustrating the analytical workflow and the two alignment approaches are presented in Fig. 1.

Subjective assessments by surgeons

Preoperative and 6-month postoperative photographs were independently evaluated by two experienced oculoplastic surgeons (KŞÇ and FS). Overall upper eyelid symmetry, contour symmetry, and overall surgical outcome were graded on a five-point Likert scale ranging from 1 (poor) to 5 (excellent). Interobserver agreement was assessed using the intraclass correlation coefficient (ICC), and the mean of the two surgeons' ratings was used in the statistical analyses.

AI-assisted photographic assessment

ChatGPT-4 (OpenAI, San Francisco, CA, USA; April–August 2025) was used for photographic assessment of postoperative cosmetic outcomes. For each patient, anonymized preoperative and 6-month postoperative photographs were evaluated in two independent sessions using the same standardized prompt.

The model was instructed to generate a cosmetic outcome score ranging from 0 to 100 based on visible improvement in upper eyelid height (40%), bilateral eyelid height symmetry (30%), and contour symmetry and natural appearance (30%). One integer score was obtained per session, and the mean of the two ratings was recorded as the final AI score. The model had no access to clinical data, surgeons' assessments, or quantitative Bézier analysis results.

AI scores were treated as photographic observer ratings based on prespecified visual criteria rather than independently validated morphometric measurements.

Statistical analysis

Statistical analyses were performed using IBM SPSS Statistics version 28 (IBM Corp., Armonk, NY, USA). Continuous variables were summarized as mean ± standard deviation and median (minimum–maximum, interquartile range), and categorical variables as counts and percentages. Normality was assessed using the Kolmogorov–Smirnov test.

Preoperative and postoperative MRD1 measurements and Bézier symmetry metrics were compared using paired-samples t tests or Wilcoxon signed-rank tests, according to their distributional characteristics. Between-group comparisons were performed using one-way analysis of variance or the Kruskal–Wallis test, with Bonferroni adjustment for pairwise comparisons when appropriate.

Associations between quantitative eyelid margin and contour-shape symmetry metrics and surgeon or AI scores were examined using Spearman rank correlation coefficients. Analyses stratified by surgical procedure were also conducted. The patient was the unit of analysis for bilateral symmetry measurements. Statistical significance was defined as $p \leq 0.05$.

Ethics approval

This prospective study was approved by the Clinical Research Ethics Committee of Başakşehir Çam and Sakura City Hospital (January 17, 2024; approval No. 19) and was conducted in accordance with the Declaration of Helsinki.

Results

Patient characteristics and surgical outcomes

A total of 67 patients, with a mean age of 52.4 ± 18.2 years (range, 5–82 years), were included. Forty-eight patients (71.6%) were female and 19 (28.4%) were male. Twenty-six patients (38.8%) underwent unilateral MMCR, 20 (29.9%) underwent unilateral levator resection, and 21 (31.3%) underwent bilateral upper eyelid blepharoplasty. Overall, 88 upper eyelids were operated on.

The change in MRD1 (ΔMRD1) differed significantly among procedures ($p < 0.001$): 1.2 ± 1.4 mm in the levator resection group, 0.8 ± 0.8 mm in the MMCR group, and 0.2 ± 0.6 mm in the blepharoplasty group. In pairwise comparisons, the increase in MRD1 after levator resection was significantly greater than that after blepharoplasty (Bonferroni-adjusted $p < 0.001$).

MRD1 changes were calculated using the operated eye for unilateral procedures and the mean measurements from both eyes for bilateral blepharoplasty in the statistical comparison.

Quantitative eyelid margin and contour-shape symmetry analyses

In the bilateral eyelid margin symmetry analysis using pupil-centered Bézier curves, postoperative MAE, MSE, and RMSE values decreased significantly from their preoperative values, whereas R^2 increased significantly (all $p < 0.001$).

In contrast, the contour-shape symmetry analysis conducted after equalizing the apex heights of the two contours solely by vertical translation showed no statistically significant preoperative-to-postoperative changes in MAE, MSE, RMSE, or R^2 (all $p > 0.05$). Detailed results of both quantitative analyses are provided in Table 1.

Table 1
Preoperative and postoperative quantitative upper eyelid symmetry metrics

Parameter	Preoperative	Postoperative, month 6	p
Pupil-centered eyelid margin symmetry			
MAE	1.1 ± 0.8	0.8 ± 0.6	< 0.001
MSE	2.2 ± 3.2	1.2 ± 2.0	< 0.001
RMSE	1.2 ± 0.8	0.9 ± 0.6	< 0.001
R^2	−1.5 ± 4.2	0.2 ± 1.2	< 0.001
Eyelid-apex height–equalized contour symmetry			
MAE	0.6 ± 0.2	0.6 ± 0.2	0.359
MSE	0.7 ± 0.4	0.7 ± 0.4	0.659
RMSE	0.8 ± 0.2	0.8 ± 0.2	0.499
R^2	0.2 ± 0.6	0.3 ± 0.4	0.315

Data are presented as mean ± standard deviation. MAE, mean absolute error; MSE, mean squared error; RMSE, root mean squared error; R^2, coefficient of determination.

Surgeon assessments and AI-generated scores

The mean subjective eyelid symmetry rating assigned by the experienced oculoplastic surgeons improved from 2.2 ± 0.9 preoperatively to 3.5 ± 1.0 postoperatively, and the mean contour symmetry rating increased from 2.3 ± 1.0 to 3.5 ± 1.0 (both $p < 0.001$). The mean postoperative overall surgical outcome rating was 3.6 ± 1.1.

High interobserver agreement was observed for the surgeons' overall surgical outcome ratings (ICC = 0.82; 95% confidence interval, 0.70–0.89; $p < 0.001$).

AI-generated cosmetic outcome scores had a mean of 67.4 ± 19.9, a median of 72, and a range of 18–92. The surgeons' overall surgical outcome ratings were positively correlated with AI scores (Spearman $\rho = 0.760$; $p < 0.001$).

Associations between quantitative symmetry metrics and surgeon or AI assessments

Postoperative pupil-centered eyelid margin symmetry metrics showed statistically significant negative correlations between MAE, MSE, and RMSE and both the surgeons' overall surgical outcome ratings and AI-generated scores. R^2 was positively correlated with both assessment methods (all $p < 0.001$).

When associations with the surgeons' direct assessments of eyelid contour were examined, postoperative pupil-centered R^2 showed a positive correlation with postoperative contour symmetry ratings (Spearman $\rho = 0.607$; $p < 0.001$).

By contrast, none of the contour-shape symmetry metrics obtained after equalization of eyelid-apex height showed a statistically significant correlation with the surgeons' overall outcome ratings or AI scores (Table 2).

Table 2
Correlations of postoperative quantitative symmetry metrics with surgeon and AI outcome scores

Parameter	Surgeon score, ρ	p	AI score, ρ	p
Pupil-centered eyelid margin symmetry				
MAE	−0.497	< 0.001	−0.528	< 0.001
MSE	−0.502	< 0.001	−0.523	< 0.001
RMSE	−0.502	< 0.001	−0.523	< 0.001
R^2	0.543	< 0.001	0.570	< 0.001
Eyelid-apex height–equalized contour symmetry				
MAE	0.165	0.186	0.140	0.258
MSE	0.057	0.647	0.033	0.792
RMSE	0.057	0.647	0.033	0.792
R^2	0.132	0.292	0.079	0.526

ρ, Spearman rank correlation coefficient; AI, artificial intelligence. The surgeon score represents the mean rating assigned by two experienced oculoplastic surgeons.

Subgroup analyses by surgical procedure

Procedure-specific analyses showed that the associations between postoperative pupil-centered eyelid margin symmetry metrics and surgeon or AI outcome scores were most pronounced in the MMCR

group. In this group, R^2 was strongly and positively correlated with the surgeons' overall outcome ratings ($\rho = 0.794$; $p < 0.001$) and AI scores ($\rho = 0.760$; $p < 0.001$).

In the levator resection group, R^2 was also positively correlated with the surgeons' outcome ratings ($\rho = 0.531$; $p = 0.016$) and AI scores ($\rho = 0.500$; $p = 0.025$). MAE, MSE, and RMSE showed negative associations with outcome scores in both surgical groups.

In contrast, no statistically significant correlations were identified between postoperative quantitative eyelid margin symmetry metrics and either the surgeons' or AI outcome scores in the bilateral blepharoplasty group. Procedure-specific correlations between pupil-centered Bézier symmetry metrics and surgeon or AI-generated outcome scores are summarized in Table 3.

Table 3
Procedure-specific correlations between pupil-centered Bézier metrics and outcome scores

Surgical group / parameter	Surgeon, ρ	p	AI, ρ	p
MMCR (n = 26)				
MAE	−0.789	< 0.001	−0.714	< 0.001
MSE	−0.807	< 0.001	−0.720	< 0.001
RMSE	−0.807	< 0.001	−0.720	< 0.001
R^2	0.794	< 0.001	0.760	< 0.001
Levator resection (n = 20)				
MAE	−0.529	0.016	−0.518	0.019
MSE	−0.497	0.026	−0.477	0.033
RMSE	−0.497	0.026	−0.477	0.033
R^2	0.531	0.016	0.500	0.025
Blepharoplasty (n = 21)				
MAE	−0.236	0.303	−0.347	0.123
MSE	−0.230	0.315	−0.343	0.128
RMSE	−0.230	0.315	−0.343	0.128
R^2	0.298	0.189	0.388	0.082

ρ, Spearman rank correlation coefficient; AI, artificial intelligence. Only pupil-centered contour analyses are presented in this table. Because MSE and RMSE are related by a monotonic transformation, they may yield identical Spearman rank correlation coefficients. The patient was the unit of analysis.

Discussion

Outcomes following upper eyelid surgery are generally assessed in clinical practice by examination, the surgeon's visual judgment, and single-point measures of eyelid height such as MRD1. However, the overall geometry of the eyelid margin and interocular symmetry cannot be adequately characterized by one linear measurement [4–6]. In this prospective study, we jointly evaluated Bézier curve–derived quantitative symmetry metrics, expert assessments, and AI-assisted photographic ratings in patients undergoing MMCR, levator resection, or functional upper eyelid blepharoplasty. Our findings showed measurable postoperative improvement in pupil-centered bilateral eyelid margin symmetry, and quantitative metrics were associated with surgeons' postoperative assessments. These results provide clinical evidence supporting the potential use of mathematical parameters derived from standardized photographs as complementary measures in surgical outcome assessment beyond conventional MRD1 measurement.

Efforts to quantify upper eyelid contour have progressed from the multiple radial midpupil lid distances described by Milbratz et al. to continuous contour modeling using Bézier curves [4–6]. Huelin et al. demonstrated the utility of Bézier curves for characterizing different geometric features of the upper eyelid contour, whereas Golbert et al. showed that contour similarity between the two eyelids can be quantified [5, 6]. Şerefoğlu Çabuk et al. reported associations between mathematical metrics, including MAE, MSE, RMSE, and particularly R^2, and experts' subjective symmetry ratings [7]. Consistent with these observations, postoperative pupil-centered R^2 in our study was positively correlated with surgeons' overall outcome ratings (Spearman $\rho = 0.543$; $p < 0.001$), whereas MAE and RMSE were negatively associated with higher surgeon ratings ($\rho=-0.497$ and $\rho=-0.502$, respectively; both $p < 0.001$). Studies by Ayıntap and Aytoğan on regional contour symmetry following levator surgery and by Bilici et al. on eyelid elevation at different margin locations likewise underscore that surgical outcomes encompass geometric characteristics extending beyond a single central distance measurement [8, 13].

The two alignment approaches used in this study are also important for interpreting what the quantitative metrics represent. Pupil-centered alignment compares relative eyelid height and contour shape together; postoperative MAE, MSE, and RMSE decreased, whereas R^2 increased (all $p < 0.001$). In contrast, when the contour apices were brought to the same height solely by vertical translation, the resulting shape-focused metrics showed no significant group-level preoperative-to-postoperative changes (all $p > 0.05$). This divergence suggests that improvements in bilateral similarity after surgery may be more apparent in metrics sensitive to eyelid height differences; however, a direct causal conclusion about the contribution of height changes cannot be drawn. Likewise, a nonsignificant result does not establish that contour morphology remained unchanged. Both approaches quantify geometric similarity between bilateral curves under different alignment conditions rather than the anatomical normality of either individual eyelid or overall surgical success.

The procedure-specific findings further highlight the importance of clinical indications. Mean MRD1 change, calculated using the operated eye, was 1.2 ± 1.4 mm after levator resection, 0.8 ± 0.8 mm after

MMCR, and 0.2 ± 0.6 mm after bilateral blepharoplasty. Because procedure selection was based on phenylephrine response, levator function, and the underlying anatomical condition, the groups were neither randomized nor equivalent in their baseline characteristics; our findings should therefore not be interpreted as evidence of superiority of one procedure over another. Şerefoğlu Çabuk et al. quantitatively compared symmetry after unilateral ptosis surgery and reported improvements in eyelid height and symmetry with both techniques [14]. In a multicenter prospective MMCR study, Leung et al. evaluated eyelid height success and bilateral height symmetry as distinct outcomes, emphasizing that quantitative measures may reflect different aspects of clinical results [15]. In a meta-analysis of seven studies, Karam et al. found no significant differences in postoperative MRD1 or undercorrection rates between MMCR and external levator advancement [16]. Likewise, the stronger correlations observed for some measures in our MMCR group do not establish surgical superiority, given the small subgroups and differing indications.

The inclusion of isolated bilateral blepharoplasty was consistent with our aim of evaluating different dimensions of upper eyelid surgical outcomes. Functional blepharoplasty for dermatochalasis primarily aims to relieve the mechanical obstruction of the superior visual field by removing redundant skin rather than raising the eyelid margin; consequently, its results may not be adequately captured by MRD1 or eyelid margin contour alone [1–3]. The absence of significant correlations between pupil-centered Bézier metrics and either surgeon or AI outcome scores in the blepharoplasty group should not be interpreted as evidence of a lack of surgical efficacy. The study by Er Bilir and Eroğul on automated morphometric measurements after blepharoplasty similarly indicates that different anatomical outcomes may be assessed together [12]. Moreover, our AI photographic scoring rubric assigned a combined weight of 70% to eyelid height and bilateral height symmetry and may therefore not fully capture the reduction in redundant skin or improvement in visual field following blepharoplasty. These observations reinforce the need to assess outcomes appropriate to each surgical indication alongside quantitative contour analysis.

The strong positive rank correlation between surgeons' overall outcome ratings and AI-generated photographic scores (Spearman $\rho = 0.760$; $p < 0.001$) suggests that structured assessments of photographic appearance may relate to expert clinical impressions. This finding warrants further investigation of semiautomated or multimethod frameworks that may eventually contribute to postoperative outcome assessment. However, Spearman correlation does not demonstrate absolute agreement between scores or establish their interchangeability. The general-purpose multimodal model used in this study was neither trained specifically for eyelid surgery nor independently calibrated as a morphometric system; rather, it functioned as a photographic observer applying prespecified criteria. Thus, our findings support investigation of relationships among different assessment methods, not substitution of AI ratings for quantitative Bézier measurements or clinical judgment.

Image-based oculoplastic assessment is evolving rapidly. Sun et al. investigated a deep-learning approach to predicting appearance after blepharoptosis surgery, and Nam et al. developed a system for automated measurement of MRD1, MRD2, and eyelid length [9, 10]. Using automated contour analysis in

540 healthy individuals, Shao et al. examined age- and sex-related morphological characteristics and suggested that more detailed eyelid-shape quantification may facilitate objective preoperative and postoperative assessment in oculoplastic surgery [11]. The AI-driven eyeball exposure rate analysis described by Lee et al. also illustrates the potential value of two-dimensional parameters beyond a single linear measurement [17]. Collectively, these studies demonstrate that eyelid height, contour, and exposure can be quantified as complementary visual components. Rather than developing a dedicated automated segmentation system, our study examined associations between Bézier-derived mathematical symmetry metrics and surgeon and AI photographic ratings within the same prospective cohort.

The limitations of general-purpose multimodal models in aesthetic assessment should also be considered. Haider et al. reported promising findings for the interpretation of qualitative facial features but identified limitations in estimating precise quantitative proportions [18]. In a 2026 study evaluating paired photographs before and after botulinum toxin treatment, Güler et al. highlighted variability in the assessment of regional aesthetic improvement [19]. Although these studies cannot be directly extrapolated to upper eyelid surgery, they support treating AI outputs as exploratory tools that may provide additional information alongside expert assessments. Accordingly, our findings concerning AI-based photographic scoring should be viewed as preliminary evidence of a potentially useful adjunct to postoperative assessment; reproducibility, external validation in different populations, and evaluation criteria tailored to surgical indications will require further investigation.

Strengths of this study include its prospective design, standardized photographic protocol, 6-month follow-up, independent assessments by two oculoplastic surgeons, and concurrent evaluation of Bézier-derived mathematical error and goodness-of-fit metrics alongside photographic ratings in the same patients. Nevertheless, its single-center design, sample of 67 patients, small procedure-specific subgroups, and differing clinical targets of the interventions limit generalizability. Because the contours were traced manually by a single investigator, intraobserver and interobserver reproducibility of the contour measurements was not separately tested. Obtaining AI scores in two independent sessions does not, by itself, establish rating stability. Patient-reported outcomes, postoperative visual field gains, and the independent clinical contribution of these quantitative metrics beyond MRD1 were not evaluated. Multiple correlations and small-subgroup analyses should be interpreted as exploratory.

Conclusions

In conclusion, bilateral upper eyelid symmetry metrics calculated from Bézier curves on standardized photographs enabled quantitative characterization of postoperative changes and investigation of their associations with expert ratings. The relationship between AI-assisted photographic scores and expert assessments further suggests avenues for extending this research approach. Rather than constituting a validated standalone measure of success that replaces MRD1, our findings provide a foundation for future development of more standardized, reproducible, multimethod systems that assess upper eyelid surgical outcomes across complementary quantitative and clinical dimensions.

Abbreviations

AI
artificial intelligence
ICC
intraclass correlation coefficient
MAE
mean absolute error
MMCR
Müller muscle–conjunctival resection
MRD1
margin reflex distance-1
MSE
mean squared error
R^2
coefficient of determination
RMSE
root mean squared error.

Declarations

Meeting Presentation:

This material has not been presented at any national or international meeting.

Financial Support:

No financial support was received for this study.

Artificial Intelligence–Assisted Technologies Statement:

A large language model–based artificial intelligence system, ChatGPT-4 (OpenAI, San Francisco, CA, USA), was used as part of the study methodology to generate standardized cosmetic success scores from de-identified preoperative and postoperative periocular photographs. The AI system had no access to clinical data, surgeon ratings, or quantitative metric outputs. No AI-assisted technology was used to generate or modify the scientific content of the manuscript outside the described study methodology.

Proprietary Interest Statement:

The authors have no proprietary or financial interest in any product, device, software, or technology described in this manuscript.

Conflict of Interest Statement:

The authors declare no conflicts of interest.

Ethics approval

This prospective study was approved by the Clinical Research Ethics Committee of Başakşehir Çam and Sakura City Hospital (January 17, 2024; approval No. 19) and was conducted in accordance with the Declaration of Helsinki.

Consent to participate

Written informed consent to participate in the study was obtained from all 67 participants or, in the case of minors, from their parents or legal guardians.

Consent for publication

Written informed consent for the publication of the de-identified preoperative and postoperative periocular photographs presented in Figure 1 was obtained from the patient depicted.

Availability of data and materials

The de-identified datasets used and analyzed during the current study are available from the corresponding author upon reasonable request, subject to applicable ethical and patient confidentiality requirements.

Competing interests

The authors declare that they have no competing interests.

Funding

No external funding was received for this study.

Authors' contributions

MSG conceived and designed the study, coordinated data collection, interpreted the findings, and drafted the original manuscript. FS performed all surgical procedures and, together with KSC, independently evaluated the clinical and photographic surgical outcomes. SKC performed Bézier curve tracing and MATLAB-based quantitative eyelid contour analyses. GO and FSK contributed to manuscript writing,

translation, and English-language editing. All authors critically reviewed and approved the final manuscript.

Acknowledgements

Not applicable.

References

1. Cahill KV, Bradley EA, Meyer DR, Custer PL, Holck DE, Marcet MM, et al. Functional indications for upper eyelid ptosis and blepharoplasty surgery: a report by the American Academy of Ophthalmology. Ophthalmology. 2011;118(12):2510–7. 10.1016/j.ophtha.2011.09.029. . PMID:22019388.
2. Kim KK, Granick MS, Baum GA, Beninger F, Cahill KV, Donnelly KC, et al. American Society of Plastic Surgeons evidence-based clinical practice guideline: eyelid surgery for upper visual field improvement. Plast Reconstr Surg. 2022;150(2):e419–34. 10.1097/PRS.0000000000009329. . PMID:35895522.
3. Hollander MHJ, Contini M, Pott JWP, Vissink A, Schepers RH, Jansma J. Functional outcomes of upper eyelid blepharoplasty: a systematic review. J Plast Reconstr Aesthet Surg. 2019;72(2):294–309. PMID:30528286.
4. Milbratz GH, Garcia DM, Guimarães FC, Cruz AAV. Multiple radial midpupil lid distances: a simple method for lid contour analysis. Ophthalmology. 2012;119(3):625–8. PMID:22197435.
5. Huelin FJ, Sales-Sanz M, Ye-Zhu C, Cruz AAV, Muñoz-Negrete FJ, Rebolleda G. Bézier curves as a total approach to measure the upper lid contour: redefining clinical outcomes in palpebral surgery. Br J Ophthalmol. 2023;107(1):6–11. 10.1136/bjophthalmol-2021-319666. . PMID:34340977.
6. Golbert MB, Garcia DM, Akaishi PMS, Cruz AAV. Upper eyelid contour symmetry measurement with Bézier curves. Arq Bras Oftalmol. 2020;83(1):28–32. 10.5935/0004-2749.20200002. . PMID:31531548.
7. Serefoglu Cabuk K, Cengiz SK, Guler MG, Topcu H, Cetin Efe A, Ulas MG, et al. Chasing the objective upper eyelid symmetry formula; R^2, RMSE, POC, MAE, and MSE. Int Ophthalmol. 2024;44:303. 10.1007/s10792-024-03157-y. . PMID:38954051.
8. Ayintap E, Aytogan H. Comparison of eyelid contour symmetry after ptosis repair via centralized versus temporally biased levator suturing techniques. BMC Ophthalmol. 2025;25:486. 10.1186/s12886-025-04318-y. . PMID:40859207.
9. Sun Y, Huang X, Zhang Q, Lee SY, Wang Y, Jin K et al. A fully automatic postoperative appearance prediction system for blepharoptosis surgery with image-based deep learning. Ophthalmol Sci. 2022;2(3):100169. 10.1016/j.xops.2022.100169. PMID:36245755.

10. Nam Y, Song T, Lee J, Lee JK. Development of a neural network-based automated eyelid measurement system. Sci Rep. 2024;14:1202. 10.1038/s41598-024-51838-6. . PMID:38216653.
11. Shao J, Cao J, Wang C, Xu P, Lou L, Ye J. Automatic measurement and comparison of normal eyelid contour by age and gender using image-based deep learning. Ophthalmol Sci. 2024;4(5):100518. 10.1016/j.xops.2024.100518. PMID:38881605.
12. Er Bilir A, Eroğul Ö. Artificial intelligence-assisted quantitative evaluation of upper eyelid blepharoplasty outcomes in dermatochalasis: methodological validation and clinical correlation. BMC Ophthalmol. 2026;26:162. 10.1186/s12886-026-04693-0. . PMID:41761106.
13. Bilici S, Harbigil-Sever T, Ugurbas SH. Digital analysis of unilateral ptosis repair: external levator advancement vs. Müller's muscle conjunctival resection. Arq Bras Oftalmol. 2024;87(3):e20230028. 10.5935/0004-2749.2023-0028. . PMID:38537045.
14. Serefoglu Cabuk K, Asik Nacaroglu S, Ozturk Karabulut G, et al. Muller muscle conjunctival resection or external levator advancement; a quantitative comparison of symmetry in unilateral ptosis. Eur J Ophthalmol. 2022;32(4):2125–32. 10.1177/11206721211049057. . PMID:34704509.
15. Leung VC, El-Khazen Dupuis J, Ashraf DC, Idowu OO, Massicotte E, Vagefi MR, et al. Müller muscle conjunctival resection: a multicentered prospective analysis of surgical success. Ophthalmic Plast Reconstr Surg. 2023;39(3):226–31. 10.1097/IOP.0000000000002292. . PMID:36356179.
16. Karam M, Alsaif A, Abul A, Alkhabbaz A, Alotaibi A, Shareef E, et al. Muller's muscle conjunctival resection versus external levator advancement for ptosis repair: systematic review and meta-analysis. Int Ophthalmol. 2023;43(7):2563–73. 10.1007/s10792-023-02633-1. . PMID:36692699.
17. Lee B, Xu L, Oh SH, Ha Y, Kwon H, Lee KC, et al. AI-driven eyeball exposure rate (EER) analysis: a useful tool for assessing ptosis surgery effectiveness. PLoS ONE. 2025;20(3):e0319577. 10.1371/journal.pone.0319577. . PMID:40132041.
18. Haider SA, Prabha S, Gomez-Cabello CA, Borna S, Genovese A, Trabilsy M, et al. Facial analysis for plastic surgery in the era of artificial intelligence: a comparative evaluation of multimodal large language models. J Clin Med. 2025;14(10):3484. 10.3390/jcm14103484. . PMID:40429479.
19. Güler I, Kraus A, Grieb G, Stelling H. Evaluation of aesthetic outcomes following botulinum toxin treatment using multimodal large language models: a paired before-and-after analysis. Aesthet Surg J Open Forum. 2026;8:ojag112. 10.1093/asjof/ojag112. . PMID:42405004.

Figures

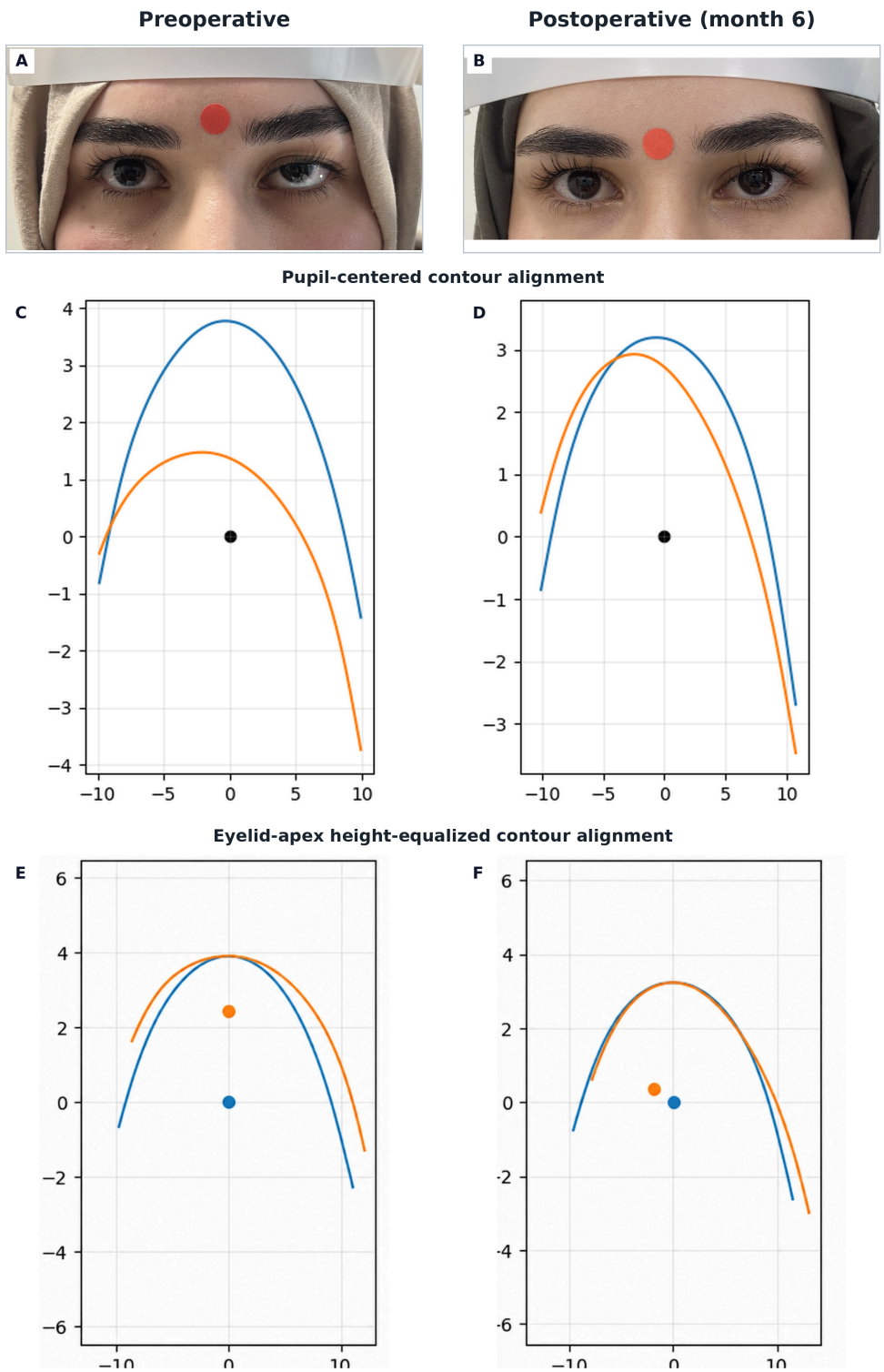

Figure 1

Two Bézier curve alignment approaches before and after upper eyelid surgery. (A, B) Frontal photographs of a representative patient before surgery and at postoperative month 6. (C, D) Bilateral eyelid margin analysis after alignment of pupil centers, preserving the relative vertical positions of the contours. (E, F) Shape-focused analysis after bringing contour apices to the same height by vertical translation alone.

Blue and orange curves represent the two compared eyelid contours; markers indicate the reference coordinates used for alignment. Photographs and plots depict the same patient.