ISAPS®

ORIGINAL ARTICLES

OCULOPLASTIC

Protocol Establishment and Reliability Verification of Three-Dimensional Digital Stereophotogrammetry in Lower Eyelid Anthropometry

Yarong Chi[1] · Yuyan Yang[2] · Lin Jin[2] · Yuming Chong[2] · Jiuzuo Huang[2] ·
Nanze Yu[2] · Xiao Long[2]

4

Abstract

Background Three-dimensional (3D) stereophotogrammetry has been widely used in anthropometry for both medical and esthetic purposes. However, no studies have assessed its reliability on measuring the lower eyelid by 3D imaging. This study aimed to establish a standardized 3D anthropometric protocol for lower eyelid region and validate its reliability.

Methods Fifty-eight participants (116 eyes) were recruited with mean age of 39.14 ± 11.25 years. Two sets of VECTRA 3D images were taken for each subject, and each set of images was individually measured twice by two raters. Twenty-seven landmarks were identified in the lower eyelid region, and then 19 linear, 4 curvilinear, 7 angular and 2 areal metrics were assessed for intrarater, interrater and intramethod reliability.

Results High reliability was found in this 3D imaging-based lower eyelid anthropometry. The mean absolute difference within 2 intrarater measurement were 0.22 and 0.08 units, the technical errors measurement were 0.31 and 0.15 units, the relative errors of measurement were 0.90% and 0.31%, the relative technical errors of measurement were 1.53% and 0.64%, and the intra-group correlation coefficient was 0.99 and 0.99. The results for interrater measurement were 0.53 units, 0.59 units, 2.94%, 3.41% and 0.96, and for intramethod measurement were 0.71 units, 0.77 units, 4.12%, 4.05% and 0.95 units, respectively.

Conclusion 3D stereophotogrammetry is reliable for lower eyelid anthropometry. The standardized protocol can be further applied for many purposes such as lower eyelid aging evaluation, surgical related assessment and periocular rejuvenation plan.

Level of Evidence IV This journal requires that authors assign a level of evidence to each article. For a full description of these Evidence-Based Medicine ratings, please refer to the Table of Contents or the online Instructions to Authors www.springer.com/00266.

Keywords Three-dimensional stereophotogrammetry · Lower eyelid anthropometry · Protocol establishment · Reliability

Yarong Chi and Yuyan Yang have contributed equally to this article.

✉ Jiuzuo Huang
   hjz1983@126.com

✉ Nanze Yu
   yunanze@pumch.cn

✉ Xiao Long
   pumclongxiao@126.com

1   Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

2   Department of Plastic and Reconstructive Surgery, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Dongcheng District, Beijing 100730, China

Introduction

The lower eyelid area, including the pretarsal roll, tear trough, lid-cheek junction, and eyelid bag, has long been nonnegligible for the aesthetics and rejuvenation of the face. According to previous studies, from the fourth decade of life on, the restriction of the orbicularis retaining ligament (ORL) and tear trough ligament on swollen orbital septum fat leads to the gradual appearance of the eyelid pouch, tear trough and lid-cheek junction, which further leads to an aged and fatigued appearance [1–5]. Moreover,

the pretarsal roll, consisting of the orbicularis oculi muscle, is widely believed by Easterners to make the eyes more attractive [6]. Therefore, such operations as blepharoplasty and pretarsal roll injection are often used to achieve facial rejuvenation. A complete evaluation method of the lower eyelid region would greatly help obtain satisfactory surgical results.

Three-dimensional (3D) surface imaging has been widely used in recent years for the morphological assessment of facial parts, such as periocular aging and lower eyelid sag [7–10]. Compared with traditional 2D photography, 3D stereophotogrammetry provides a more complete and accurate collection of soft tissue features, allowing researchers to more finely assess facial structure, feature differences between populations, and soft-tissue changes before and after operation [11–14]. Furthermore, instead of lower eyelid evaluation scales, which can only give us subjective grades, or computerized tomography (CT), magnetic resonance imaging (MRI) and other imaging techniques, which are costly and easily affected by body position, 3D stereophotogrammetry overcomes such obstacles with its accuracy, convenience and noninvasiveness [15, 16].

Unfortunately, the existing periocular evaluation systems that use 3D imaging analysis either have a single index—a basic indicator such as the height and width of the eyelid fissure—or they focus on the upper eyelid region or only analyze the age-related changes of lower eyelid fulness, and no effective 3D morphological assessment system has been proposed for the lower eyelid area [7, 10, 17, 18]. For this reason, the periocular anthropometric assessment we established first describes the characteristics of the lower eyelid region in detail through metrics such as lines, curvatures, angles and areas, making it the first 3D-imaging system assessment of lower eyelid region aging and the blepharoplasty effect.

In this study, we proposed a practical protocol for lower eyelid morphology assessment with 3D stereophotogrammetry, introduced various novel periocular landmarks, and validated the intrarater, interrater and intramethod reliability of our anthropometry system through the measurement of linear distances, curvatures, angles and area metrics.

Methods

Patients and Recruitment

Fifty-eight volunteers (9 males, 49 females, 116 eyes) were recruited at the Department of Plastic and Reconstructive Surgery, Peking Union Medical College Hospital, from September 2021 to January 2023. They were aged 23 to 61

(39.14±11.25) years. Obvious pretarsal roll and eyelid bags were observed in all patients collected. Patients with congenital facial deformity, pathologies, recent or previous trauma, periorbital surgery and botulinum toxin injection were excluded. All patients signed the informed consent. This study was performed in line with the Declaration of Helsinki and was approved by the Institutional Review Board of the Peking Union Medical College Hospital (Reference Number K2502).

Equipment and Three-Dimensional Image Acquisition

The VECTRA H1-270 handheld camera (Canfield Scientific, Inc., Parsippany, NJ, USA), a three-dimensional stereophotogrammetry system, was used for patient image collection. All participants were required to remove all make-up and fully expose their eyebrows and forehead by pulling back their hair and removing glasses and jewelry before 3D image acquisition. During the image acquisition, patients were asked to sit in a fixed position, maintain a neutral facial position and expression, and look straight ahead. An experienced operator used the VECTRA H1-270 camera to capture photos of each patient at a fixed distance and three different angles (right oblique, frontal and left oblique) exactly according to the manufacturer's instructions. The lens height in the camera was adjusted to ensure that it was even with the middle point of the bilateral medial canthus of the patient before shooting. Subsequently, the 3D models of the patient's faces were automatically generated with a resolution of 0.95 mm and exported in the OBJ format for picture processing. Two sets of 3D images (Capture 1 and Capture 2) were taken for each patient at an interval of more than 45 minutes with recalibration of the VECTRA device. Geomagic Wrap 2021 (Geomagic, Inc., Research Triangle Park, NC, USA) was finally selected for further facial landmark identification, metric measurement and analysis.

Landmark Identification and Metric Measurement

A coordinate system is essential for 3D model evaluation. Taking the middle point between the left and right endocanthions as the origin, Camper's plane was determined by the left tragus, right tragus, and midpoint of the right and left ala. Then, the horizontal plane was obtained after turning up 7.5° and translating to pass through the origin. The sagittal plane was then defined as the plane perpendicular to both Camper's plane and the horizontal plane. Finally, the vertical plane was set to pass through the origin and be perpendicular to both the sagittal plane and the horizontal plane.

Table 1 Definition of 27 anthropometric periocular landmarks

Landmarks	Definition
En	Endocanthion, inner commissure of the palpebral fissure
Ex	Exocanthion, outer commissure of the upper and lower eyelash roots of palpebral fissure
Pc	Pupillary center
Mcl	Medial corneoscleral limbus point horizontal to Pc
Lcl	Lateral corneoscleral limbus point horizontal to Pc
LpMcl	Point vertical to Mcl at the lower palpebral margin
LpPc	Point vertical to Pc at the lower palpebral margin
LpLcl	Point vertical to Lcl at the lower palpebral margin
LpMm	Midpoint of LpMcl and En
LpMl	Midpoint of LpLcl and Ex
Pro	Origin of the pretarsal roll
PrMm	Point vertical to LpMm at the pretarsal roll
PrMcl	Point vertical to Mcl at the pretarsal roll
PrPc	Point vertical to Pc at the pretarsal roll
PrLcl	Point vertical to Lcl at the pretarsal roll
PrMl	Point vertical to LpMl at the pretarsal roll
PrEx	Point vertical to Ex at the pretarsal roll
TtEn	Point vertical to En at the tear trough
TtMm	Point vertical to LpMm at the tear trough
TtMCL	Point vertical to Mcl at the tear trough
TtPc	Point vertical to Pc at the tear trough
LjLCL	Point vertical to Lcl at the lid-cheek junction
LjML	Point vertical to LpMl at the lid-cheek junction
LjEx	Point vertical to Ex at the lid-cheek junction
UpPc	Point vertical to Pc at the upper palpebral margin
EbPc	Point vertical to Pc at the inferior margin of eyebrows
Na	Nasion, the lowest point of the nasal bridge

A total of 27 periocular landmarks are described in Table 1, which were then digitally identified and marked following the standard protocol shown in Fig. 1. In the front view, the pupillary center (Pc), endocanthion (En), exocanthion (Ex), medial corneoscleral limbus (Mcl) and lateral corneoscleral limbus (Lcl) were first marked because of their easy recognition. Then, the vertical corresponding points of these five points on the lower palpebral margin were marked. The midpoints of En and Mcl and Ex and Lcl were labeled for a more detailed

Fig. 1 The 27 three-dimensional anthropometric landmarks of the periocular region. 26 landmarks were indicated on the frontal view, and the 3 landmarks were shown on the lateral view (Pc and Ex represented the same points on both frontal view and lateral view).
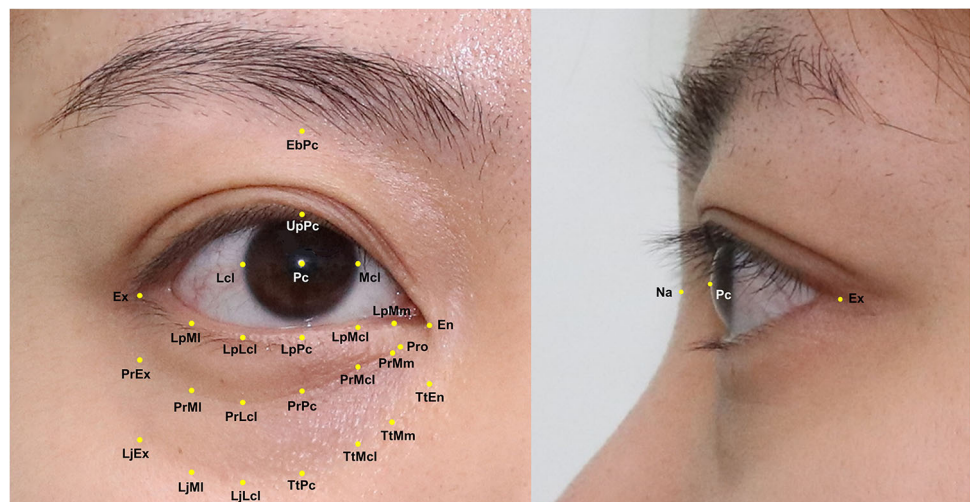
description. With reference to the point of the lower palpebral margin, the vertical corresponding points on the pretarsal roll, tear trough and lid-cheek junction were identified on their lower margin in consideration of their width. The vertical correspondence of the pupillary center at the upper eyelid margin and the lower eyebrow margin was also marked for the overall evaluation of the orbital part. As for the lateral view, the nasion point was labeled to describe lateral orbital features for different subjects. On the basis of the 27 landmarks mentioned above, 19 linear distances, 4 curvatures, 7 angles and 2 areas were measured (Table 2 and Fig. 2), of which 2 linear distances were accessed in the lateral view.

Intrarater, Interrater and Intramethod Reliability Assessment

The first author (Y. C., Rater 1) finished two measurement sessions at an interval of more than 24 hours for each set of 3D pictures, and the second author (L. J., Rater 2) performed two measurements on the first set of 3D pictures at an interval of at least 24 hours. From their results, intrarater, interrater and intramethod reliability were assessed (Fig. 3). Intrarater reliability referred to the uniformity of two measurements of the same image performed by the same rater. Interrater reliability was the comparison of the second measurements of the same image performed by Rater 1 vs. Rater 2. Intramethod reliability meant the similarity between the second measurements of two different sets of images carried out by the same rater.

Data Analysis

Intrarater, interrater and intramethod reliability were assessed by 7 statistics (Table 3) [18]. Through the intraclass correlation coefficient (ICC), the reliability between repeated measures was evaluated, a value close to 1 indicating high reliability and close to 0 indicating low reliability. Generally, the ICC value is divided into 3 categories with thresholds of 0.4 and 0.75, where a value greater than 0.75 indicates excellent reliability. The mean absolute difference (MAD) and technical error of measurement (TEM) shared the same acceptable error threshold values of 1 unit (millimeter) for linear and curvature metrics, 2 units (degrees) for angular metrics and 4 units (square millimeters) for areal metrics according to the relatively small or large magnitudes of periocular and lower eyelid measurements. The relative technical error of measurement (REM) and relative TEM (%TEM) were also calculated to provide an estimate of variation considering the influence of measurement size on reliability assessment. Five reliability categories were defined for REM and %TEM with reference to previous studies: < 1%, excellent; 1 to 3.9%,

very good; 4 to 6.9%, good; 7 to 9.9%, moderate; and > 10%, poor [17]. Total TEM and relative % total TEM were also calculated to negate the influence of the positive correlation between total TEM and sample size.

All 7 statistics mentioned above and the means of the measurements with their standard deviations (SDs) were entered into the software Microsoft Excel 2019 (Microsoft Corp., Redmond, WA, USA) after all measurements were finished, for further processing. ICCs were calculated and significance tests run using SPSS version 25 software (IBM Corp., Armonk, NY, USA). The normality of all measurements was tested by the Kolmogorov–Smirnov test first; then, the paired-sample t test was employed for normally distributed measurements, and the Wilcoxon signed-rank test was performed for nonnormally distributed paired measurements. A P value of < 0.05 was considered statistically significant.

Results

Baseline Data and General Results

Fifty-eight healthy volunteers were recruited (49 females, 9 males). The two sexes had similar ages (male 38.22 ± 9.90 years, female 39.30 ± 11.49 years). The mean value and standard deviation (SD) of each measurement are shown in Table 4. The reliability assessment was divided into three parts, namely, intrarater, interrater and intramethod, and a total of seven evaluation indicators were calculated: MAD, TEM, REM, %TEM, total TEM, % total TEM and ICC (Supplementary Figs. 1, 2, 3, 4, and 5 and Table 5, respectively). On the whole, most of the measurements showed relatively high reliability, as the ICC was more than 0.95, the MAD and TEM values were within the specified normal ranges, and the REM and %TEM were less than 7%. No REM or %TEM estimates exceeded 10%, which was the threshold for poor reliability. In addition, the vast majority of statistics calculated indicated a higher reliability in intrarater assessment than in interrater and intramethod measurements.
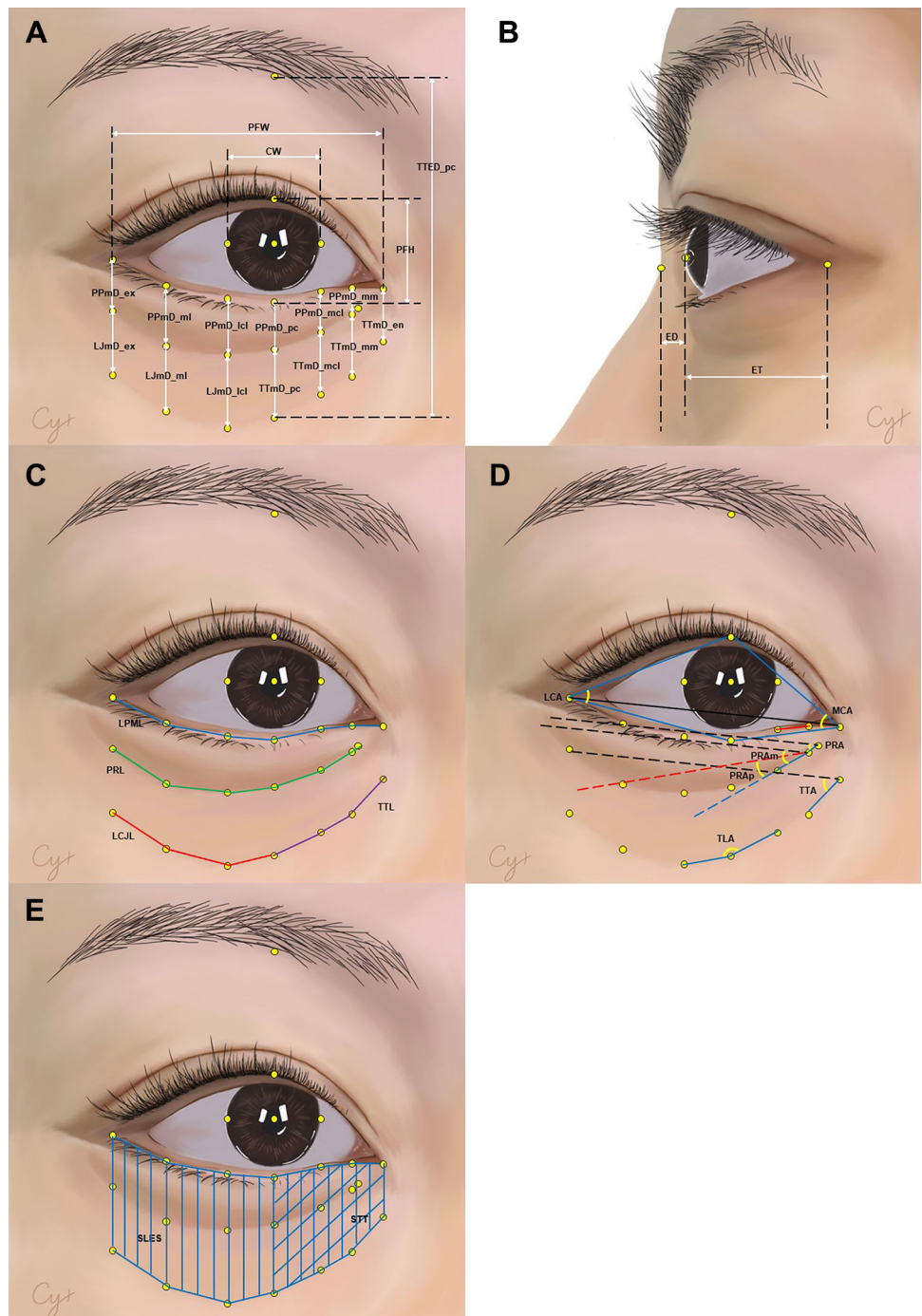
Intrarater Reliability

All of the ICC estimates in the two intrarater measurements exceeded 0.95, with 17 of them reaching 1.00, which indicated high intrarater correlation. All intrarater MAD and TEM estimates were less than the self-set limits, which were 1 unit for linear and curvature metrics, 2 units for angles and 4 units for areal measurements. In particular, the linear and curvature MAD estimates were less than 0.2 units. As for REM estimates, 10 linear estimates (CW, TTED_pc, PFW, LJmD_lcl, TTmD_pc, TTmD_mcl, PFH,

Table 2 List of 32 periocular measurements

Definition	Abbreviation	Landmarks
Linear distances		
Frontal view		
Palpebral fissure width	PFW	LpPc-UpPc
Palpebral fissure height	PFH	En-Ex
Cornea width	CW	Mcl-Lcl
Pretarsal roll-palpebral margin distance (medial corneoscleral limbus)	PPmD_mcl	PrMcl-LpMcl
Pretarsal roll-palpebral margin distance (pupillary center)	PPmD_pc	PrPc-LpPc
Pretarsal roll-palpebral margin distance (lateral corneoscleral limbus)	PPmD_lcl	PrLcl-LpLcl
Pretarsal roll-palpebral margin distance (midpoint of En and Mcl)	PPmD_mm	PrMm-LpMm
Pretarsal roll-palpebral margin distance (midpoint of Ex and Lcl)	PPmD_ml	PrMl-LpMl
Pretarsal roll-palpebral margin distance (exocanthion)	PPmD_ex	PrEx-Ex
Tear trough-palpebral margin distance (endocanthion)	TTmD_en	TtEn-En
Tear trough-palpebral margin distance (medial corneoscleral limbus)	TTmD_mcl	TtMcl-LpMcl
Tear trough-palpebral margin distance (pupillary center)	TTmD_pc	TtPc-LpPc
Lid cheek junction-palpebral margin distance (lateral corneoscleral limbus)	LJmD_lcl	LjLcl-LpLcl
Lid cheek junction-palpebral margin (exocanthion)	LJmD_ex	LjEx-Ex
Tear trough-palpebral margin distance (midpoint of En and Mcl)	TTmD_mm	TtMm-LpMm
Lid cheek junction-palpebral margin (midpoint of Ex and Lcl)	LJmD_ml	LjMl-LpMl
Tear trough-Eyebrows distance (pupillary center)	TTED_pc	TtPc-EbPc
Lateral view		
Eye depth (Z difference of Na and Pc)	ED	Na(z)-Pc(z)
Eye thickness (Z difference of Pc and En)	ET	Pc(z)-En(z)
Curvatures		
Lower palpebral margin length	LPML	En-LpMm-LpMcl-LpPc-LpLcl-LpMl-Ex
Pretarsal roll length	PRL	Pro-PrMm-PrMcl-PrPc-PrLcl-PrMl-PrEx
Tear trough length	TTL	TtEn-TtMm-TtMcl-TtPc
Lid cheek junction length	LCJL	TtPc-LjLcl-LjMl-LjEx
Angles		
Medial canthal angle	MCA	LEPc-En-UEPc
Letral canthal angle	LCA	LEPc-Ex-UEPc
Pretarsal roll angle	PRA	Pro-PrMm-horizontal line of En-Ex
Tear trough angle	TTA	TtMm-TtEn-horizontal line of En-Ex
Tear trough-lid cheek junction angle	TLA	TtMcl-TtPc-TtLcl
Pretarsal roll angle (medial)	PRAm	PrMcl-PrMm-horizontal line of En-Ex
Pretarsal roll angle (palpebral)	PRAp	PrMcl-PrMm-horizontal line of LpMm-LpMcl
Areas		
Surface area between tear trough and lower palpebral margin	STT	En-LpMm-LpMcl-LpPc-TtPc-TtMcl-TtMm-TtEn-En
Surface area of lower eyelid sag	SLES	En-LpMm-LpMcl-LpPc-LpLcl-LpMl-Ex-LjEx-LjMl-LjLcl-TtPc-TtMcl-TtMm-TtEn-En

Fig. 2 Schematic drawing of each metrics of lower eyelid region. **A**, **B** showed 19 linear distances of the protocol in both frontal and side view. **C** showed 4 curvatures of the protocol. **D** showed angles of the protocol, in which the dotted lines were parallel with the solid lines in the same color. **E** showed 2 areal metrics of the protocol.

LJmD_ml, ED, LJmD_ex), 3 curvature estimates (PRL, LPML, LCJL), 5 angular estimates (TLA, MCA, TTA, LCA, PRAm) and both area estimates (SLES, STT) reached excellent reliability, with intrarater reliability values less than 1%. The remaining 9 linear estimates, 1 curvature estimate and 2 angular estimates all had very good reliability (1–3.9%). A total of 12 intrarater %TEM values had excellent reliability (< 1%), including 3 linear

estimates (CW, TTED_pc, PFW), 2 curvature estimates (PRL, LPML), 5 angular estimates (TLA, TTA, MCA, LCA, PRAm) and 2 area estimates (STT, SLES). The other 13 linear estimates, 2 curvature estimates and 2 angular estimates had very good reliability, and only 2 linear estimates (ET, TTmD_en) had good reliability, with values of 4%-6.9%.

Fig. 3 Schematic of intrarater, interrater and intramethod reliability assessment.
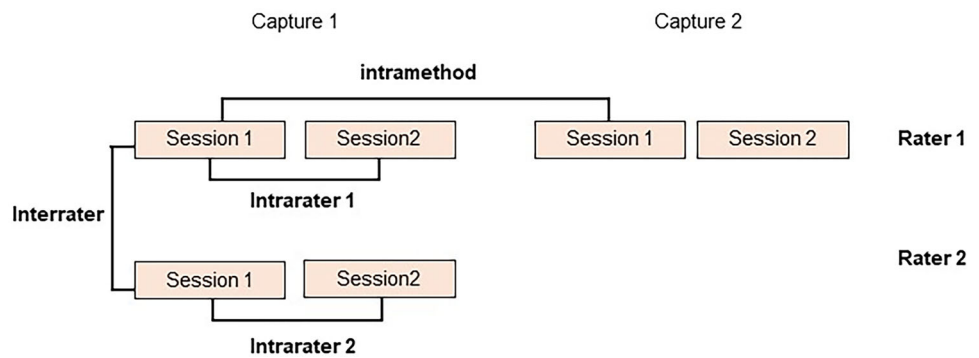

Table 3 List of statistics for reliability evaluation

Statistic	Equation
Intraclass correlation coefficient (ICC)	$B/(B+W)$
Mean absolute difference (MAD)	$\lvert X1 - X2 \rvert$
Technical error of measurement (TEM)	$\sqrt{(\sum D^2/2N)}$
Relative error measurement (REM)	$(MAD/X3)\times 100$
Relative TEM (% TEM)	$(TEM/X3)\times 100$
Total TEM	$\sqrt{\left(\frac{TEM(intra1)^2 + TEM(intra2)^2}{2}\right) + TEM(inter)^2}$
Relative total TEM (% total TEM)	$(\%\ total\ TEM/X3)\times 100$

B, Between-measurement variance; W, Within measurement variance; D, Difference between measurements; N, Number of eyes or subjects measured; X1, Mean for rater 1 (session 1, or session 2 of capture 1); X2, Mean for rater 2 (session 2, or session 2 of capture 2); X3, Grand mean; TEM (intra1), Intrarater TEM for the first rater; TEM (intra2), Intrarater TEM for the second rater; TEM (inter), Interrater TEM between the two raters

Interrater Reliability

The majority of interrater ICC estimates were no less than 0.95. The exceptions were PFW, PFH, LJmD_ml, LCJL and PRAm, whose ICCs ranged from 0.92 to 0.94, and only two of the ICCs were below 0.90 (0.83 for CW, 0.88 for PRL). All interrater MAD estimates were within the self-set limits, as were all TEM estimates except a linear metric, LJmD_ex, whose TEM value was 1.10 units. Similarly, nearly all intrarater measurements had high reliabilities in terms of total TEM, except LJmD_ex, which had a total TEM value of 1.12%. Furthermore, of the % total TEM values, only one linear distance (PPmD_mcl) failed to meet good reliability, with a value of 8.05%.

Among the interrater REM estimates, only 1 angle (TLA) and 2 areal measurements had REM of less than 1%; 10 linear distances, all 4 curvatures and the remaining 6 angular measurements had very good reliability, with REM between 1% and 3.9%. Eight other linear distances (LJmD_ex, TTmD_en, TTmD_mm, PPmD_ex, PPmD_lcl, PPmD_pc, PPmD_mm, PPmD_ml) had good reliability (4%-6.9%). The last linear distance (PPmD_mcl) had the worst REM of 7.56%, which was still in the moderate reliability category. Of the %TEM values, similarly, 1 angle and 2 areal estimates achieved excellent reliability, with %TEM of less than 1%; 8 linear estimates (TTED_pc, PFW, LJmD_lcl, LJmD_ml, CW, PFH, ED, ET), 4 curvatures and 6 angles had very good reliability (1–3.9%); 10 other linear estimates reached good reliability, with %TEM ranging from 4 to 6.9%; and one linear distance had the worst %TEM of 7.52%, which was considered moderate reliability.

Intramethod Reliability

Twenty-three intramethod measurements had ICC estimates of no less than 0.95, the other seven intramethod ICC estimates were between 0.90 and 0.94, and only 2 ICC estimates were below 0.90 (PFH, ICC=0.81; PRL, ICC=0.88). All intramethod MAD and TEM estimates were within our defined limits. In terms of intramethod REM, only 1 angle (TLA) and 2 areal measurements reached excellent reliability, with values below 1%. Three linear distances (PFW, TTED_pc, CW), 2 curvatures

Table 4 Mean and standard deviations (SDs) of all measurements

Variable	Rater 1								Rater 2			
	Capture 1				Capture 2				Capture 1			
	Session 1		Session 2		Session 1		Session 2		Session 1		Session 2	
	μ	SD	μ	SD	μ	SD	μ	SD	μ	SD	μ	SD
PFW	29.45	1.47	29.41	1.50	29.20	1.49	29.19	1.54	29.46	1.57	29.33	1.52
PFH	11.74	1.11	11.77	1.10	11.56	1.03	11.59	1.14	11.81	1.05	11.77	1.07
CW	11.56	0.73	11.56	0.74	11.53	0.64	11.54	0.67	11.53	0.77	11.55	0.77
PPmD_mcl	3.82	0.88	3.80	0.86	3.90	0.95	3.91	0.97	3.80	0.89	3.83	0.90
PPmD_pc	4.68	1.07	4.67	1.04	4.74	1.01	4.74	1.00	4.67	1.05	4.70	1.04
PPmD_lcl	5.13	1.21	5.10	1.19	5.30	1.19	5.27	1.20	5.13	1.25	5.21	1.19
PPmD_mm	5.55	1.42	5.53	1.39	5.72	1.38	5.70	1.38	5.54	1.39	5.55	1.39
PPmD_ml	5.89	1.58	5.86	1.54	6.18	1.55	6.16	1.54	5.81	1.59	5.85	1.53
PPmD_ex	6.04	1.88	6.02	1.86	6.20	1.87	6.20	1.87	6.12	1.92	6.08	1.91
TTmD_en	8.16	1.99	8.10	2.05	8.02	2.06	8.01	2.07	8.15	2.12	8.18	2.09
TTmD_mcl	11.70	2.12	11.61	2.16	11.71	2.17	11.71	2.17	11.70	2.03	11.71	2.10
TTmD_pc	13.66	2.19	13.63	2.20	13.71	2.19	13.71	2.19	13.66	2.15	13.72	2.17
LJmD_lcl	14.91	2.28	14.90	2.28	15.10	2.26	15.09	2.26	14.94	2.24	14.97	2.26
LJmD_ex	15.62	2.40	15.60	2.40	15.62	2.44	15.61	2.44	15.62	2.36	15.61	2.39
TTmD_mm	16.25	2.50	16.25	2.49	16.34	2.54	16.33	2.54	16.13	2.39	16.14	2.39
LJmD_ml	16.88	3.27	16.87	3.26	17.05	3.10	17.05	3.10	16.80	3.05	16.79	3.05
TTED_pc	39.13	3.57	39.22	3.52	38.77	3.49	38.76	3.50	39.17	3.24	39.18	3.25
ED	6.11	2.23	6.15	2.21	6.19	2.26	6.19	2.26	6.06	2.21	6.09	2.22
ET	7.14	1.46	7.09	1.49	7.06	1.48	7.06	1.48	7.19	1.37	7.17	1.88
LPML	31.98	1.73	31.93	1.73	31.72	1.71	31.73	1.73	32.00	1.81	31.94	1.82
PRL	31.30	1.92	31.29	1.94	30.83	1.88	30.83	1.89	31.30	2.08	31.24	2.06
TTL	16.05	1.89	16.09	1.94	16.30	1.92	16.29	1.91	16.12	1.87	16.16	1.85
LCJL	16.96	1.55	16.92	1.59	16.81	1.71	16.80	1.70	17.10	1.63	16.97	1.62
MCA	44.94	4.95	44.95	4.94	44.74	4.96	44.72	4.95	45.10	4.97	45.02	4.93
LCA	35.17	3.68	35.12	3.73	35.36	3.66	35.35	3.66	35.11	3.73	35.15	3.77
PRA	45.81	7.63	45.75	7.67	45.70	7.73	45.64	7.77	46.02	8.04	45.88	7.98
TTA	52.14	6.93	52.13	6.93	51.99	6.86	52.02	6.90	52.24	7.05	52.07	6.99
TLA	158.61	8.40	158.50	8.42	158.41	8.46	158.31	8.56	158.71	8.49	158.66	8.44
PRAm	34.98	4.92	34.94	4.93	34.85	5.05	34.77	5.10	35.01	4.98	34.82	4.94
PRAp	16.19	5.76	16.17	5.76	16.19	5.82	16.14	5.81	16.18	5.78	16.15	5.79
STT	199.83	49.26	199.96	49.32	200.07	49.35	200.09	49.30	200.55	49.35	200.43	49.31
SLES	490.90	78.14	491.39	77.84	491.21	78.16	490.85	78.20	491.51	77.88	491.03	78.02

(LPML, PRL) and 6 remaining angular measurements were in the very good reliability category (REM ranges from 1 to 3.9%). Eleven linear distances and 2 curvatures had good reliability, with REM ranging from 4 to 6.9%. The other 5 linear measurements (ED, ET, PPmD_ex, PPmD_pc and PPmD_mcl) had moderate reliability, with a maximum REM of 7.65%. As for %TEM, 1 angle (TLA) and 2 areal estimates had excellent reliability, with %TEM of less than 1%. Three linear measurements (PFW, TTED_pc, CW), 2 curvatures (LPML, PRL) and the other 6 angles had %TEM between 1 and 3.9%. Fifteen lines and 2 curvatures had good reliability, with %TEM ranged from 4 to 6.9%. Only 1 linear estimate (PPmD_mcl) had moderate reliability (%TEM=7.09%).

Table 5 Intrarater, interrater and intramethod intraclass correlation coefficient (ICC) and mean differences of all measurements

Variable	Intrarater 1			Intrarater 2			Interrater			Intramethod		
	ICC (CI 95%)	Dmean	p value	ICC (CI 95%)	Dmean	p value	ICC (CI 95%)	Dmean	p value	ICC (CI 95%)	Dmean	p value
PFW	0.97 (0.96–0.98)	0.04	0.326	0.97 (0.96–0.98)	0.01	0.787	0.92 (0.88–0.95)	0.25	0.002	0.93 (0.89–0.95)	−0.01	0.907
PFH	0.98 (0.97–0.99)	−0.03	0.132	0.95 (0.93–0.97)	−0.03	0.627*	0.94 (0.90–0.96)	0.18	<0.001	0.81 (0.71–0.87)	−0.08	0.388
CW	1.00 (0.99–1.00)	0.01	0.969*	0.99 (0.98–0.99)	−0.02	0.039*	0.84 (0.77–0.90)	0.04	0.003*	0.92 (0.87–0.94)	0.03	0.242*
PPmD_mcl	0.99 (0.98–0.99)	0.02	0.300*	0.98 (0.98–0.99)	−0.01	0.545*	0.95 (0.92–0.97)	−0.08	0.014*	0.95 (0.93–0.97)	0.02	0.764*
PPmD_pc	0.99 (0.98–0.99)	0.01	0.960	1.00 (0.99–1.00)	0.01	0.767	0.97 (0.96–0.98)	−0.07	0.068	0.95 (0.93–0.97)	0.01	0.941
PPmD_lcl	0.99 (0.98–0.99)	0.03	0.041	1.00 (0.99–1.00)	0.03	0.006	0.98 (0.95–0.99)	−0.17	<0.001	0.96 (0.94–0.98)	0.01	0.986
PPmD_mm	1.00 (0.99–1.00)	0.02	0.550*	1.00 (0.99–1.00)	0.02	0.011*	0.97 (0.95–0.98)	−0.17	<0.001*	0.97 (0.96–0.98)	0.01	0.669*
PPmD_ml	1.00 (0.99–1.00)	0.03	0.130*	1.00 (0.99–1.00)	0.03	0.022*	0.97 (0.92–0.98)	−0.29	<0.001*	0.97 (0.96–0.98)	0.07	0.487*
PPmD_ex	0.99 (0.99–0.99)	0.01	0.911*	1.00 (1.00–1.00)	0.01	0.381*	0.99 (0.98–0.99)	−0.16	<0.001*	0.98 (0.97–0.99)	−0.08	0.184*
TTmD_en	0.97 (0.95–0.98)	0.06	0.243	1.00 (1.00–1.00)	0.01	0.052	0.97 (0.96–0.98)	0.14	0.038	0.97 (0.96–0.98)	0.01	0.851
TTmD_mcl	0.97 (0.95–0.98)	0.06	0.243	1.00 (1.00–1.00)	0.01	0.052	0.97 (0.96–0.98)	0.14	0.037	0.97 (0.96–0.98)	0.01	0.851
TTmD_pc	0.99 (0.99–0.99)	0.03	0.282	0.99 (0.99–1.00)	−0.01	0.710	0.95 (0.93–0.97)	−0.05	0.580	0.93 (0.90–0.95)	−0.01	0.975
LJmD_lcl	0.99 (0.99–1.00)	0.02	0.482	1.00 (1.00–1.00)	0.01	0.044	0.95 (0.92–0.96)	−0.18	0.083	0.95 (0.92–0.97)	−0.03	0.793
LJmD_ex	0.99 (0.99–1.00)	0.02	0.379	0.98 (0.97–0.99)	0.01	0.006	0.98 (0.98–0.99)	0.01	0.983	0.96 (0.95–0.98)	−0.01	0.992
TTmD_mm	0.99 (0.99–1.00)	0.01	0.792	0.99 (0.98–0.99)	1	0.352	0.98 (0.98–0.99)	−0.09	0.196	0.96 (0.94–0.97)	0.12	0.218
LJmD_ml	0.99 (0.99–1.00)	0.01	0.806	0.99 (0.98–1.00)	0.01	0.541	0.94 (0.91–0.96)	−0.17	0.273	0.96 (0.94–0.97)	0.08	0.548
TTED_pc	0.99 (0.99–1.00)	−0.08	0.199*	0.98 (0.98–0.99)	0.01	0.117*	0.98 (0.97–0.99)	0.37	<0.001*	0.97 (0.95–0.98)	−0.04	0.718*
ED	0.99 (0.98–0.99)	−0.04	0.083	0.99 (0.98–0.99)	0.01	0.131	1.00 (0.99–1.00)	−0.09	0.008	0.98 (0.97–0.99)	0.04	0.470
ET	0.97 (0.96–0.98)	0.05	0.135	0.97 (0.96–0.98)	−0.01	0.599	0.98 (0.97–0.99)	0.08	0.041	0.94 (0.92–0.96)	−0.05	0.474
LPML	0.98 (0.97–0.98)	0.46	0.223	0.98 (0.97–0.99)	−0.01	0.605	0.95 (0.91–0.97)	0.26	0.001	0.93 (0.90–0.95)	−0.19	0.836
PRL	0.99 (0.98–0.99)	0.01	0.763	0.99 (0.99–0.99)	0.01	0.544	0.88 (0.80–0.93)	0.47	<0.001	0.88 (0.82–0.92)	0.01	0.930*
TTL	0.96 (0.94–0.97)	−0.40	0.490	0.97 (0.96–0.98)	0.01	0.147	0.96 (0.93–0.97)	−0.25	0.002	0.91 (0.87–0.94)	−0.07	0.496
LCJL	0.98 (0.97–0.98)	0.05	0.159	1.00 (1.00–1.00)	0.01	0.219	0.93 (0.90–0.95)	0.15	0.075	0.90 (0.85–0.93)	−0.14	0.158
MCA	1.00 (1.00–1.00)	−0.01	0.931	0.99 (0.98–0.99)	0.02	0.053	0.99 (0.99–1.00)	0.20	0.017	0.99 (0.98–0.99)	−0.16	0.152
LCA	0.99 (0.99–1.00)	0.05	0.186	0.98 (0.98–0.99)	0.01	0.310	0.98 (0.98–0.99)	−0.18	0.037	0.97 (0.96–0.98)	0.06	0.591
PRA	0.99 (0.99–0.99)	0.05	0.431	0.98 (0.98–1.00)	0.05	0.073	1.00 (1.00–1.00)	0.11	0.222	0.99 (0.99–1.00)	−0.22	0.131
TTA	1.00 (0.99–1.00)	0.01	0.825	0.99 (0.98–0.99)	−0.03	0.465	0.99 (0.99–1.00)	0.15	0.128	0.98 (0.98–0.99)	−0.10	0.557
TLA	0.99 (0.99–1.00)	0.11	0.033*	0.98 (0.98–0.99)	0.09	0.064*	0.98 (0.98–0.99)	0.20	0.008*	0.99 (0.99–0.99)	−0.10	0.738*
PRAm	0.99 (0.98–0.99)	0.04	0.374	0.99 (0.98–1.00)	0.08	0.031	0.93 (0.90–0.95)	0.13	0.130	0.96 (0.94–0.97)	−0.04	0.645
PRAp	1.00 (0.99–1.00)	0.02	0.559	0.99 (0.99–1.00)	0.05	0.022	0.99 (0.99–1.00)	0.01	0.971	0.99 (0.99–0.99)	0.01	0.919
STT	1.00 (1.00–1.00)	−0.13	0.143*	0.99 (0.99–1.00)	−0.02	0.683*	0.99 (0.99–0.99)	−0.25	0.238*	0.99 (0.99–1.00)	−0.73	0.001*
SLES	1.00 (1.00–1.00)	−0.49	0.149	0.99 (0.98–0.99)	0.36	0.011	0.97 (0.96–0.98)	−0.31	0.142	0.96 (0.94–0.97)	−0.60	0.294

*p values are from Wilcoxon signed rank tests for paired data, and the rest ones are from paired-sample t tests

Discussion

Considering the marked individual differences in the lower eyelid region, previous lower eyelid assessments were mostly conducted through grading scales, which are convenient but subjective and inaccurate. As for rare quantitative evaluation, orbital grayscale analysis can only provide limited information, such as the difference in the grayscale value between the lower eyelid fold and the surrounding area [19, 20]. In addition, imaging operations such as CT and MRI are cumbersome and expensive, and so far there has been no accurate evaluation model of lower eyelid soft tissue that can be widely promoted. Lo and Lin provided a comprehensive overview of the application of 3D imaging in craniomaxillofacial surgery, highlighting the numerous advantages offered by 3D photography, including minimal invasiveness, high repeatability and ease of use [21]. Fan et al. applied a 3D imaging system to the measurement of upper eyelid area and volume, finding that it was reliable in area measurement despite doubts about the accuracy of volume measurement [22]. Our study is the first periorbital morphology evaluation system focusing on the lower eyelid region by using 3D stereophotogrammetry. Referring to the opinions of Guo et al., 58 subjects with both pretarsal roll and eyelid bags were recruited, 27 periorbital landmarks were set up, and 32 linear distances, 4 curvatures, 7 angular metrics and 2 areas were calculated to describe the characteristics of the lower eyelid region in detail. Through the assessment of many statistical indices, such as MAD, TEM, REM, %TEM, total TEM, % total TEM and ICC, the protocol was found to have superb intrarater, interrater and intramethod reliability.

Referring to previous research and considering the enormous difference in our measured values, we defined 1 unit for linear and curvature metrics, 2 units for angles and 4 units for areas as the precision error limits for MAD and TEM evaluation. Fortunately, almost all measurements were within this range, except the TEM of interrater LJmD_ex, which slightly exceeded it (TEM = 1.10). The high reliability of the mean values of MAD, TEM, REM, % TEM and ICC were confirmed in rater 1 (0.22 unit, 0.31 unit, 0.90%, 1.53%, 0.99) and rater 2 (0.08 unit, 0.15 unit, 0.31%, 0.64%, 0.99). Such consistency also extended to the interrater measures (0.53 units, 0.59 units, 2.94%, 3.41%, 0.96) and intramethod measures (0.71 units, 0.77 units, 4.12%, 4.05%, 0.95). Among them, the highest reliability was found within the intrarater measurements and the lowest but still acceptable accuracy in intramethod measurements, in line with Guo et al. [18]. Moreover, a sharp decrease in reliability was found between intrarater measurements and interrater and intramethod measurements, while 32 of the estimates reached excellent intrarater

reliability, but the number was 6 for both interrater and intramethod reliability. This may be explained by the fact that the tear trough, lid-cheek junction and pretarsal roll all had a certain width rather than being characterized by a single line. Although their lower margins had been uniformly marked, there were still differences in judgment between raters. For the differences between methods, apart from the errors caused by equipment recalibration, patient position and facial expression, which were certainly included, errors of intrarater measurement were also introduced because two sets of landmark identification were carried out at different times. However, Gibelli D. et al. found intramethod differences by premarking 50 landmarks on patients' faces, which effectively eliminated intrarater errors [23].

In terms of specific indicators, it was found that when MAD and TEM were small, % TEM and REM tended to be large, and vice versa, which may be because the relative errors of indicators with smaller values were larger relative to the value itself [18, 24]. Comparing this study with Guo et al.'s, a total of 3 linear distances, 1 curvature and 2 angles were repeated, namely, PFW, PFH, CW, LPML, MCA and LCA. However, the mean value differences of PFH, CW, and LCA were within the self-set limits (this study vs. Guo's study, 11.61 mm vs. 11.71 mm, 11.55 mm vs. 12.11 mm, 35.21° vs. 36.70°), while others were not (PFW 29.34 mm vs. 30.62 mm, LPML 31.88 mm vs. 34.44 mm, MCA 44.91° vs. 39.42°). These discrepancies may be caused by the different races enrolled in the studies: there are longer eyes in Caucasians and rounder eyes in Asians [18].

The linear distances with low accuracy were concentrated in the pretarsal roll and the LJmD_ex, where the lower margin was difficult to judge precisely because of the fine wrinkles. However, the related curvatures and angles still met the self-defined ranges, indicating the same judgment of the curve shapes between different raters. Furthermore, since the boundary of areas did not involve pretarsal roll, both area metrics showed high reliability in all of the evaluation statistics, which confirmed our conclusion. In addition, some estimates of eyes had high accuracy, such as PFW, PFH and CW, which also illustrated the high reliability of this protocol.

There is no denying the significance of volume measurement in evaluating the outcomes of lower eyelid surgery. We did not include volume measures in this article for two primary reasons. First, 3D photography primarily focuses on capturing surface information. Although certain software can measure volumes of marked areas, it still relies on surface labeling within the images. Fan et al. also highlighted that volume measurements obtained through 3D photography are not highly reliable [22]. Second, considering the complex anatomical structure and

nonplanar bottom surface of the lower eyelid region, such internal information cannot be adequately captured by 3D photography alone, making it challenging to accurately evaluate lower eyelid volume solely through 3D imaging. For precise measurement of lower eyelid volume, we suggest combining CT, MRI and ultrasound measurements to obtain more accurate data.

This study proposed a standardized method for measuring the lower eyelid region by using 3D photography and confirmed its reliability and accuracy. The new imaging technology can thus be applied to many other clinical situations. Before performing lower blepharoplasty, the standardized method can be used to evaluate the severity of tear trough deformity and eyelid bags so that operators can facilitate more individualized surgical planning for each patient. Meanwhile, the therapeutic or aesthetic effects can be further assessed by such a protocol. Additionally, the attractiveness of pretarsal rolls can be better defined with such measurements, and changes in the lower eyelid region caused by aging or certain diseases can be easily identified and diagnosed through the establishment of a normative 3D anthropometric database of the lower eyelid region.

This study had some limitations. First, only 58 participants were included, of whom more than 80% were female. This sample size was relatively small, making it difficult to generalize the results. Second, the inevitable use of flash in 3D photography not only weakened the visibility of the margins of the pretarsal roll, tear trough and lid-cheek junction after exposure but also made it hard for the participants to maintain exactly the same expression in two different captures because the eyes are extremely sensitive to light. These edges can still be accurately identified and marked, but the small changes in expression inevitably led to the larger intramethod error compared to the intrarater and interrater errors. Third, no other validation was conducted in this study. A further comparison of 3D stereophotogrammetry to 2D photogrammetry or direct measurement may be best to confirm its validity and reliability. In summary, further studies are still necessary for more reliable and universal evaluation of the lower eyelid region.

Conclusion

This is the first periocular anthropometry system concentrating on the lower eyelid region using a 3D digital stereophotogrammetry technique. It proved highly reliable and repeatable. It takes advantage of the convenience of 3D imaging and can be widely applied to lower eyelid aging assessment, surgical design, therapeutic effect evaluation and periocular rejuvenation.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Ethical Approval This study was approved by the Institutional Review Board of the Peking Union Medical College Hospital (Reference Number K2502) and conducted in accordance with the Declaration of Helsinki.

Informed Consent All patients provided written consent for the use of their 3D facial images.

References

1. Chen S, Ma H, Song T et al (2022) Anatomical characterization of the tear trough and the lid-cheek junction: an iodine staining technique based on high-resolution micro-computed tomography. Plast Reconstr Surg 149(4):646e–654e. https://doi.org/10.1097/PRS.0000000000008964

2. Haddock NT, Saadeh PB, Boutros S et al (2009) The tear trough and lid/cheek junction: anatomy and implications for surgical correction. Plast Reconstr Surg 123(4):1332–1340. https://doi.org/10.1097/PRS.0b013e31819f2b36

3. Wong CH, Hsieh MKH, Mendelson B (2012) The tear trough ligament: anatomical basis for the tear trough deformity. Plast Reconstr Surg 129(6):1392–1402. https://doi.org/10.1097/PRS.0b013e31824ecd77

4. Hussain SN, Mangal S, Goodman GJ (2019) The Tick technique: a method to simplify and quantify treatment of the tear trough region. J Cosmet Dermatol 18(6):1642–1647. https://doi.org/10.1111/jocd.13169

5. Mojallal A, Cotofana S (2017) Anatomy of lower eyelid and eyelid-cheek junction. Ann Chir Plast Esthet 62(5):365–374. https://doi.org/10.1016/j.anplas.2017.09.007

6. Cheng CY, Hu S, Chang SL et al (2022) Comprehensive evaluation of the lower eyelid aging process among Asian women. Dermatol Surg 48(6):653–658. https://doi.org/10.1097/DSS.0000000000003425

7. Camp MC, Wong WW, Filip Z et al (2011) A quantitative analysis of periorbital aging with three-dimensional surface imaging. J Plast Reconstr Aesthet Surg 64(2):148–154. https://doi.org/10.1016/j.bjps.2010.04.037

8. Donath AS, Glasgold RA, Meier J et al (2010) Quantitative evaluation of volume augmentation in the tear trough with a hyaluronic acid-based filler: a three-dimensional analysis. Plast Reconstr Surg 125(5):1515–1522. https://doi.org/10.1097/PRS.0b013e3181d70317

9. Jodeh DS, Curtis H, Cray JJ et al (2018) Anthropometric evaluation of periorbital region and facial projection using three-dimensional photogrammetry. J Craniofac Surg 29(8):2017–2020. https://doi.org/10.1097/SCS.0000000000004761

10. Chong Y, Li J, Liu X et al (2021) Three-dimensional anthropometric analysis of eyelid aging among Chinese women. J Plast Reconstr Aesthet Surg 74(1):135–142. https://doi.org/10.1016/j.bjps.2020.08.010

11. Guo Y, Hou X, Rokohl AC et al (2020) Reliability of periocular anthropometry: a comparison of direct, 2-dimensional, and 3-dimensional techniques. Dermatol Surg 46(9):e23–e31. https://doi.org/10.1097/DSS.0000000000002243

12. De Stefani A, Barone M, Hatami Alamdari S et al (2022) Validation of vectra 3D imaging systems: a review. Int J Environ Res Public Health. https://doi.org/10.3390/ijerph19148820

13. Fan W, Guo Y, Hou X et al (2022) Validation of the portable next-generation VECTRA H2 3D imaging system for periocular anthropometry. Front Med (Lausanne) 9:833487. https://doi.org/10.3389/fmed.2022.833487

14. Liu J, Rokohl AC, Guo Y et al (2021) Reliability of stereophotogrammetry for area measurement in the periocular region. Aesthet Plast Surg 45(4):1601–1610. https://doi.org/10.1007/s00266-020-02091-5

15. Chen YS, Tsai TH, Wu ML et al (2008) Evaluation of age-related intraorbital fat herniation through computed tomography. Plast Reconstr Surg 122(4):1191–1198. https://doi.org/10.1097/PRS.0b013e318185d370

16. Chen J, Wu Y, Wang Y et al (2022) Objective assessment of the effectiveness of the transconjunctival fat-repositioning technique using magnetic resonance imaging and three-dimensional surface imaging. Facial Plast Surg Aesthet Med 24(5):382–384. https://doi.org/10.1089/fpsam.2021.0195

17. Gibelli D, Cappella A, Bertozzi F et al (2022) Three-dimensional facial anthropometric analysis with and without landmark labelling: Is there a real difference? J Craniofac Surg 33(2):665–668. https://doi.org/10.1097/SCS.0000000000007687

18. Guo Y, Rokohl AC, Schaub F et al (2019) Reliability of periocular anthropometry using three-dimensional digital stereophotogrammetry. Graefe's Arch Clin Exp Ophthalmol 257(11):2517–2531. https://doi.org/10.1007/s00417-019-04428-6

19. Jo SJ, Kim HS, Park JT et al (2018) Assessment of age- and sex-related changes in baggy lower eyelids using a novel objective image analysis method: orbital gray scale analysis. J Cosmet Dermatol 17(5):874–880. https://doi.org/10.1111/jocd.12423

20. Yang S, Kim BR, Kim HS et al (2019) A quantitative, objective method for evaluating the surgical outcomes of baggy eyelid correction using orbital gray scale analysis. J Cosmet Dermatol 18(6):1814–1820. https://doi.org/10.1111/jocd.12922

21. Lo L-J, Lin H-H (2023) Applications of three-dimensional imaging techniques in craniomaxillofacial surgery: a literature review. Biomed J. https://doi.org/10.1016/j.bj.2023.100615

22. Fan W, Rokohl AC, Maus J et al (2023) Evaluation of the portable next-generation VECTRA H2 3D imaging system for measuring upper eyelid area and volume. Aesthet Surg J. https://doi.org/10.1093/asj/sjad136

23. Gibelli D, Pucciarelli V, Cappella A et al (2018) Are portable stereophotogrammetric devices reliable in facial imaging? A validation study of VECTRA H1 device. J Oral Maxillofac Surg Off J Am Assoc Oral Maxillofac Surg 76(8):1772–1784. https://doi.org/10.1016/j.joms.2018.01.021

24. Andrade LM, Rodrigues da Silva AMB, Magri LV et al (2017) Repeatability study of angular and linear measurements on facial morphology analysis by means of stereophotogrammetry. J Craniofac Surg 28(4):1107–1111. https://doi.org/10.1097/SCS.0000000000003554