Periorbital Biometric Measurements using ImageJ Software: Standardisation of Technique and Assessment of Intra- and Interobserver Variability

R. Rajyalakshmi, Winston D. Prakash, Mohammad Javed Ali, Milind N. Naik

Department of Ophthalmic Plastic Surgery, LV Prasad Eye Institute, Hyderabad, Telangana, India

Abstract

Purpose: To assess the reliability and repeatability of periorbital biometric measurements using ImageJ software and to assess if the horizontal visible iris diameter (HVID) serves as a reliable scale for facial measurements. **Methods:** This study was a prospective, single-blind, comparative study. Two clinicians performed 12 periorbital measurements on 100 standardised face photographs. Each individual's HVID was determined by Orbscan IIz and used as a scale for measurements using ImageJ software. All measurements were repeated using the 'average' HVID of the study population as a measurement scale. Intraclass correlation coefficient (ICC) and Pearson product-moment coefficient were used as statistical tests to analyse the data. **Results:** The range of ICC for intra- and interobserver variability was 0.79–0.99 and 0.86–0.99, respectively. Test-retest reliability ranged from 0.66–1.0 to 0.77–0.98, respectively. When average HVID of the study population was used as scale, ICC ranged from 0.83 to 0.99, and the test-retest reliability ranged from 0.83 to 0.96 and the measurements correlated well with recordings done with individual Orbscan HVID measurements. **Conclusion:** Periorbital biometric measurements using ImageJ software are reproducible and repeatable. Average HVID of the population as measured by Orbscan is a reliable scale for facial measurements.

Keywords: Facial measurements, ImageJ, periorbital biometry

Introduction

Accurate orbitofacial biometric measurements are of vital importance to the ophthalmic and facial plastic surgeon in the management, surgical planning and evaluation of outcomes.[1,2] Clinical biometric measurements are most commonly obtained using a millimetre ruler, with the examiner sitting at the eye level of the patient. Boboridis *et al*. have reported 'modest and clinically acceptable' inter- and intraobserver variability of eyelid measurements using this technique in patients with various clinical conditions.[3] Clinical measurement techniques, though simple and practical for day-to-day patient care, do have limitations such as observer variables, parallax, variable illumination and patient movement. Moreover, clinical measurements are limited only to linear measurements and more complex parameters such as area and volume cannot be measured. When performing clinical trials with utmost scientific rigor, a more objective and reproducible method is desirable.

In an attempt to improvise on the clinical measurement technique, several digital photographic techniques have applied facial and periorbital measurements. The photographic technique has several potential advantages. It provides standard illumination and working distance, allows for masking of observers, and permits rapid transfer of images between institutions, thus facilitating data storage and archiving.[1] The digital photographs can also be further analysed using various software programs.

Address for correspondence: Dr. Milind Naik,
Department of Ophthalmic Plastic Surgery, LV Prasad Eye Institute,
LV Prasad Marg, Banjara Hills, Hyderabad - 500 034, Telangana, India.
E-mail: milind@drmilindnaik.com

Access this article online

Quick Response Code:

Website: www.jcasonline.com

DOI: 10.4103/JCAS.JCAS_61_17

How to cite this article: Rajyalakshmi R, Prakash WD, Ali MJ, Naik MN. Periorbital biometric measurements using ImageJ software: Standardisation of technique and assessment of intra- and interobserver variability. J Cutan Aesthet Surg 2018;10:130-5.

ImageJ software is a reliable measurement tool available from the National Institute of Health, Bethesda, Maryland, USA (http://rsb.info.nih.gov/ij).[4] It has been successfully used as a measurement tool in various subspecialties of medicine.[5-8] ImageJ has been used for periorbital and facial measurements.[9,10] Although ImageJ provides an objective measure of periorbital biometry, its inter- and intraobserver variability in the measurement of clinically significant periorbital parameters is not known. In this study, therefore, we aimed to determine the intra- and interobserver variability of periorbital biometric measurements when measured using ImageJ software.

Another difficulty while performing objective periorbital and facial measurements is the lack of a reliable measurement scale. Having a millimetre ruler within the frame of the photograph can provide measurement scale but is not always practical. Horizontal visible iris diameter when measured by a reliable instrument like Orbscan can provide a good measurement scale for facial measurements.[11] The human corneal diameter (horizontal visible iris diameter) has fairly stable dimensions for a given population and can be thought of as a naturally available measurement scale for all facial photographs. In this study, therefore, we also evaluated if the average corneal diameter of the population can be used as a reliable measurement scale.[12]

Procedure

This was a prospective, observational comparative study involving 100 consecutive normal controls who visited the Department of Ophthalmic Plastic Surgery at LV Prasad Eye Institute, Hyderabad, India.

Study objectives

The study had two objectives: First, to evaluate the intra- and interobserver variability of periorbital measurements as performed using ImageJ software. Second, to assess if the mean horizontal visible iris diameter (HVID) of the study population can be reliably used as a reference scale for these measurements.

Pilot study to determine sample size

Since there were no such studies reported in the literature, a pilot study was conducted to arrive at a sample size. For the pilot study, a standardised face photograph of one individual was taken after obtaining an informed consent. Five recordings of the HVID of the right eye were obtained, and an average was taken as the reference measurement. The photograph was uploaded in ImageJ software as a TIFF format image at 300 dots per inch (dpi) resolution. Margin reflex distance was measured 10 times by two investigators (MNN and RR). With this data, a sample size was calculated with a standard deviation of 0.16 mm to obtain an 80% power of the study at 95% confidence interval so as to have a precision of 0.1 mm in the measurements. A sample size of 100 individuals was arrived at.

Study protocol

We recruited 100 normal controls presenting to the Department of Ophthalmic Plastic Surgery, fulfilling the inclusion and exclusion criteria as shown in Table 1.

The study was approved by the Institute Review Board and Ethics Committee of LV Prasad Eye Institute, Hyderabad, India.

The purpose and the procedure of the study were explained to all the participants, and a written informed consent was obtained. Apart from complete ocular examination, each individual underwent following two evaluations: measurement of HVID and standardised face photographs.

Measurement of horizontal visible iris diameter

HVID was measured in the right eye of each individual using Orbscan (Bausch and Lomb Zyoptix Orbscan IIz Anterior Segment Analyzer). Five readings were obtained for the right eye of each individual [Figure 1b]. In case of poor fixation or increased blink and errors, the Orbscan reading was aborted and all the five readings were repeated. The average HVID of the right eye was then calculated and entered into patient data sheet.

Face photographs

Face photographs of 100 consecutive individuals were taken after explaining the purpose, the procedure and obtaining the informed consent for the study. Images were taken in a standardised manner with a Nikon D2X Digital SLR Camera with indirect lighting from a distance of 45 cm. All the images were uniformly cropped to include trichion to menton vertically and tragus to tragus horizontally. Image size was 6 inches (height) by 4.5 inches (width), stored at 300dpi in TIFF format.

Table 1: Inclusion and exclusion criteria used to enrol individuals in the study

Inclusion criteria
Age >18 years
Informed consent for face photograph
Study individuals with normal slit lamp and indirect ophthalmoscope eye examination
Study individuals with refractive error ranging from +4D to −6D
Exclusion criteria
Prior ocular, orbital or maxillofacial surgery
Prior history of facial trauma
Obvious facial asymmetry syndromes
Contact lens wearers
Study individuals using any topical medications
Study individuals with refractive error beyond the +4D–−6D range
Poor quality photographs

Figure 1: Each photograph was imported in Adobe Photoshop 7, magnified to ×600 and the centre of both the pupils was manually marked (a). Horizontal visible iris diameter was measured in the right eye of each individual using Orbscan (b). Image imported into ImageJ software and horizontal visible iris diameter was marked with a line tool. This line was set at the horizontal visible iris diameter value for that patient (as measured by Orbscan in mm) and taken as the reference for biometric measurement. This step is called 'setting the scale' (c). Once the scale is set, the measurements were recorded (d)

Biometric measurements

A code number (between 1 and 100) was randomly assigned to the face photographs. Two physicians were asked to measure predefined periocular and facial biometric measurements on each photograph. Each physician repeated the whole set of measurements, to obtain a final set of two measurements.

Each photograph was imported in Adobe Photoshop 7, magnified to ×600 and the centre of both the pupils was manually marked using the 'pencil tool' set at 1-pixel size [Figure 1a]. This was performed to enable all measurements along the mid-pupillary plane. The photographs were then opened in ImageJ. In each photograph, HVID was then manually marked using the line tool. HVID value for that patient (as measured by Orbscan) was taken as the reference measurement for the respective photograph. This step is called 'setting the scale' [Figure 1c]. Once the scale is set, the measurements were recorded [Figure 1d]. The biometric measurements that were evaluated are given in Table 2.

All measurements were recorded on a sheet against the coded number of the photograph. Two investigators (MNN and RR) performed the measurements twice (measurement 1 and measurement 2) at two different times. The two observers were masked to each other's readings. Another masked observer (KS) performed the analysis of the measurements.

Statistical analysis

Data were analysed by calculating intraclass correlation coefficient (ICC), Pearson product-moment correlation coefficient and Bland–Altman plots.

Outcome measures

The first component of the study involved biometric measurements based on individual HVID of each individual. Intraobserver variability for each measurement was calculated for each physician by comparing measurement 1 to measurement 2. Interobserver variability for each measurement was calculated by comparing the average of each measurement of one physician to the other.

The second component of the study involved biometric measurements based on the 'mean' HVID of all individuals. All biometric measurements were repeated using the mean horizontal corneal diameter (11.34 mm). The

Table 2: Periorbital biometrics that was measured and the definition used while measuring the same on photographs

Measurement	Definition
VPFH	Distance between upper and the lower eyelid margin in the mid-pupillary plane
HPFH	Distance between medial and lateral canthus (internal)
MRD1	Distance between upper eyelid margin to the central corneal light reflex in the mid-pupillary plane
MRD2	Distance between lower eyelid margin and central corneal light reflex in the mid-pupillary plane
MBH	Distance between medial canthus to the lower limit of hair-bearing eyebrow
CBH	Distance between upper eyelid margin to the lower limit of hair-bearing eyebrow in the mid-pupillary plane
LBH	Distance between lateral canthus to the lower limit of hair-bearing eyebrow (lower most eyebrow lash along the contour of the eyebrow)
LFH	Distance between upper eyelid margins to the eyelid fold in mid-pupillary plane (when multiple folds were present, the highest eyelid fold was measured)
IMCD	Distance between both internal canthi
IPD	Distance between the mid-points of both pupils
COC distance	Distance between lateral canthus to the lateral oral commissure

VPFH: Vertical palpebral fissure height, HPFH: Horizontal palpebral fissure height, MRD1: Margin reflex distance-1, MRD2: Margin reflex distance-2, MBH: Medial brow height, CBH: Central brow height, LBH: Lateral brow height, LFH: Lid fold height, IMCD: Intermedial canthal distance, IPD: Interpupillary distance, COC: Canthus to oral commissure

intra- and interobserver variability was again calculated in a similar manner.

Results

A total of 111 individuals fulfilling the inclusion criteria were enrolled in the study. Eleven photographs were excluded due to the poor quality, which could have precluded accurate analysis. The remaining 100 face photographs were analysed by two observers (MNN and RR) in a masked manner. The patient demographics and the HVID values as pleasured by Orbscan are given in Table 3.

For measurements representing part I of the study (each patient's own HVID reading was utilised for measurement), the mean and the standard deviation of each parameter (two measurements; both intra- and interobserver) is depicted in Table 4. The ICC values for the intra- and interobserver variability are depicted in Table 5. The test-retest reliability was measured using Pearson product-moment correlation coefficient and is depicted in Table 6.

The intraobserver test-retest reliability varied from 66 to 100. The value was least for medial brow height (0.66) and highest for canthus to commissure (1.00). All the other parameters ranged above 0.90, indicating good reliability.

Table 3: Demographics of 100 individuals who participated in the study

Parameter	Value
Age, mean (range)	25.6 years (18–78 years)
Male:female	53:47
Mean HVID (Orbscan II)	11.34 mm (10.7–12.7 mm)

HVID: Horizontal visible iris diameter

The interobserver test-retest reliability varied from 0.77 to 0.97. The value was least for medial brow height (0.77), and highest for canthus to commissure distance on the left (0.98), indicating good reliability. Bland–Altman scatter plots were generated for each measurement to look for the agreement between the two observers [Table 7].

The differences (investigator-1 minus investigator-2) are plotted on Y-axis and the average value of the two investigators is plotted on X-axis. Mean difference (bias) is depicted along the central line and 2 standard deviations are shown above and below the mean bias line. These values (mean + 2 standard deviation) are known as 'limits of agreement'. The plots depicted good agreement between the two observers in all the measurements.

The ICC values for the intraobserver variability for the measurements using the 'mean' HVID of 100 individuals (11.34 mm) as the scale ranged from 0.75 to 0.99 and the test-retest reliability for the same ranged from 0.62 to 0.94 [Table 8].

Discussion

The inter- and intraobserver variability in the measurement of clinically significant periorbital parameters are not known. This study aimed to determine the intra- and interobserver variability in the periorbital biometric measurements performed using ImageJ software.

ICC values were >90 for both intra- and interobserver measurements and ranged from 0.75 to 0.99 for measurements using the 'mean' HVID value. The test-retest reliability for the intra- and interobserver measurements was good (>90%). For the measurements using mean HVID, it ranged from 0.62 to 0.97. This indicates that the periorbital biometric measurements using ImageJ software are reproducible and repeatable. When the same measurements were repeated using the mean HVID of the study population (11.34 mm), the measurements were still comparable to the previous measurements. This indicates that the 'mean' HVID of the population (as measured by Orbscan) is a reliable scale for facial measurements. The latter inference has the potential to allow reliable facial measurements if the mean HVID of the population is known.

The ICC and test-retest reliability values were low for brow height as compared to the other measurements. The arrangement of the eyebrow hair is sparse, and less

Table 4: The mean readings and *P* values for intra- and interobserver measurements of various biometric parameters

Parameter	Intraobserver measurements		Interobserver measurements	
	Reading (1 and 2)	*P**	Reading (1 and 2)	*P**
HPFH	28.6 ± 1.9	0.687	29.0 ± 1.9	<0.0001
	28.6 ± 1.9		28.6 ± 1.9	
VPFH	10.1 ± 1.2	0.646	10.3 ± 1.0	0.018
	10.1 ± 1.2		10.1 ± 1.2	
MRD1	3.7 ± 0.7	0.172	3.8 ± 0.7	0.462
	3.8 ± 0.7		3.8 ± 0.7	
CBH	9.8 ± 2.6	0.927	9.2 ± 2.6	<0.0001
	9.7 ± 2.7		9.7 ± 2.7	
LBH	14.8 ± 3.0	0.599	14.1 ± 2.9	<0.0001
	14.7 ± 3.0		14.7 ± 3.0	
MBH	15.0 ± 2.5	0.371	14.6 ± 3.1	0.025
	14.8 ± 2.8		14.8 ± 2.8	
IMCD	32.9 ± 3.1	0.968	32.4 ± 3.0	<0.0001
	32.9 ± 3.1		32.9 ± 3.1	
IPD	63.9 ± 4.4	0.613	63.8 ± 4.5	0.771
	63.9 ± 4.6		63.9 ± 4.6	
COCR	68.2 ± 4.8	0.024	68.3 ± 4.8	0.479
	68.2 ± 4.8		68.2 ± 4.8	
COCL	67.3 ± 4.9	<0.0001	67.4 ± 4.9	0.168
	67.5 ± 4.9		67.5 ± 4.9	
MRD2	6.4 ± 0.9	0.701	6.4 ± 0.9	0.878
	6.3 ± 0.9		6.3 ± 0.9	

VPFH: Vertical palpebral fissure height, HPFH: Horizontal palpebral fissure height, MRD1: Margin reflex distance-1, MRD2: Margin reflex distance-2, MBH: Medial brow height, CBH: Central brow height, LBH: Lateral brow height, IPD: Interpupillary distance, COCR: Canthus to oral commissure right, COCL: Canthus to oral commissure left, IMCD: Inter medial canthal distance

*$P < 0.05$%

Table 5: Intraclass correlation coefficient values for intra- and interobserver variability

Parameter	ICC	
	Intraobserver	Interobserver
MRD1	0.97 (0.96–0.98)	0.98 (0.97–0.99)
MRD2	0.95 (0.93–0.97)	0.97 (0.96–0.98)
VPFH	0.97 (0.96–0.98)	0.96 (0.95–0.97)
HPFH	0.99 (0.98–1.00)	0.99 (0.98–1.00)
IPD	0.93 (0.90–0.96)	0.94 (0.91–0.96)
CBH	0.95 (0.92–0.96)	0.97 (0.95–0.98)
MBH	0.79 (0.69–0.86)	0.86 (0.80–0.90)
LBH	0.99 (0.98–1.00)	0.99 (0.98–1.00)
LFH	0.99 (0.98–1.00)	0.99 (0.98–1.00)
IMCD	0.99 (0.98–1.00)	0.99 (0.98–1.00)
COCL	0.99 (0.98–1.00)	0.99 (0.98–1.00)
COCR	0.99 (0.99–1.00)	0.98 (0.97–0.99)

ICC: Intraclass correlation coefficient, MRD1: Margin reflex distance-1, MRD2: Margin reflex distance-2, IPD: Interpupillary distance, CBH: Central brow height, MBH: Medial brow height, LBH: Lateral brow height, LFH: Lid fold height, COCR: Canthus to oral commissure right, COCL: Canthus to oral commissure left, IMCD: Inter medial canthal distance, VPFH: Vertical palpebral fissure height, HPFH: horizontal palpebral fissure height

Table 6: Test-retest reliability values for all biometric parameters

Parameter	Test-retest reliability	
	Intraobserver	Interobserver
MRD1	0.95	0.94
MRD2	0.91	0.88
VPFH	0.94	0.86
HPFH	0.98	0.91
IPD	0.88	0.78
CBH	0.90	0.91
MBH	0.66	0.77
LBH	0.99	0.95
LFH	0.98	0.97
IMCD	1.00	0.97
COCL	1.00	0.98
COCR	1.00	0.92

MRD1: Margin reflex distance-1, MRD2: Margin reflex distance-2, IPD: Interpupillary distance, CBH: Central brow height, MBH: Medial brow height, LBH: Lateral brow height, LFH: Lid fold height, COCR: Canthus to oral commissure right, COCL: Canthus to oral commissure left, IMCD: Inter medial canthal distance, VPFH: Vertical palpebral fissure height, HPFH: Horizontal palpebral fissure height

distinct a landmark, leading to variability. Further, in some individuals, eyebrow hair may not extend more medially, so as to be in line with the medial canthus, which was our measurement reference. Measurement of the lateral canthus was also difficult in some cases, due to lateral lash ptosis that precluded accurate identification of the lateral canthus.

Table 7: Scatter plot for Bland–Altman analysis for data of margin reflex distance-1 parameter

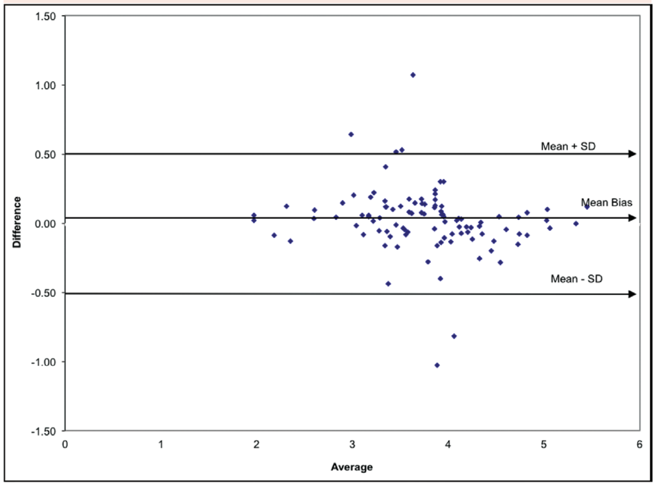

Table 8: Intersession intraclass correlation coefficient and test-retest reliability of different measurements by using the 'mean' horizontal visible iris diameter as a scale

Parameter	ICC	Test-retest reliability
HPFH	0.90 (0.86–0.94)	0.84
VPFH	0.97 (0.96–0.98)	0.94
MRD1	0.98 (0.97–0.99)	0.96
CBH	0.99 (0.97–0.98)	0.97
LBH	0.96 (0.94–0.97)	0.92
MBH	0.96 (0.94–0.97)	0.93
IMCD	0.75 (0.63–0.83)	0.62
IPD	0.87 (0.81–0.92)	0.79
COCR	0.90 (0.84–0.93)	0.82
COCL	0.75 (0.63–0.83)	0.67
MRD2	0.86 (0.80–0.91)	0.78

MRD1: Margin reflex distance-1, MRD2: Margin reflex distance-2, CBH: Central brow height, MBH: Medial brow height, IPD: Interpupillary distance, ICC: Intraclass correlation coefficient, COCR: Canthus to oral commissure right, COCL: Canthus to oral commissure left, IMCD: Inter medial canthal distance, VPFH: Vertical palpebral fissure height, HPFH: horizontal palpebral fissure height

Our study for the first time outlines a standardised approach to periorbital measurement using ImageJ software and HVID. In conclusion, our study found that reliable periorbital biometric measurements can be obtained using ImageJ software. The intra- and interobserver variability is low. We also found that the mean HVID of the population can be a reliable scale, against which other facial measurements can be made.

Financial support and sponsorship

Nil.

Conflicts of interest

None.

References

1. Edwards DT, Bartley GB, Hodge DO, Gorman CA, Bradley EA. Eyelid position measurement in graves' ophthalmopathy: Reliability of a photographic technique and comparison with a clinical technique. Ophthalmology 2004;111:1029-34.
2. Guyot L, Dubuc M, Richard O, Philip N, Dutour O. Comparison between direct clinical and digital photogrammetric measurements in patients with 22q11 microdeletion. Int J Oral Maxillofac Surg 2003;32:246-52.
3. Boboridis K, Assi A, Indar A, Bunce C, Tyers AG. Repeatability and reproducibility of upper eyelid measurements. Br J Ophthalmol 2001;85:99-101.
4. Holmes JM, Hohberger GG, Leske DA. Photographic and clinical techniques for outcome assessment in sixth nerve palsy. Ophthalmology 2001;108:1300-7.
5. Dello SA, van Dam RM, Slangen JJ, van de Poll MC, Bemelmans MH, Greve JW, *et al*. Liver volumetry plug and play: Do it yourself with imageJ. World J Surg 2007;31:2215-21.
6. Wilson WD, Valet AS, Andreotti RF, Green-Jarvis B, Lyshchik A, Fleischer AC, *et al*. Sonographic quantification of ovarian tumor vascularity. J Ultrasound Med 2006;25:1577-81.
7. Irving BA, Weltman JY, Brock DW, Davis CK, Gaesser GA, Weltman A, *et al*. NIH imageJ and slice-O-matic computed tomography imaging software to quantify soft tissue. Obesity (Silver Spring) 2007;15:370-6.
8. Carmona R, Macías D, Guadix JA, Portillo V, Pérez-Pomares JM, Muñoz-Chápuli R, *et al*. A simple technique of image analysis for specific nuclear immunolocalization of proteins. J Microsc 2007;225:96-9.
9. Mancini R, Taban M, Lowinger A, Nakra T, Tsirbas A, Douglas RS, *et al*. Use of hyaluronic acid gel in the management of paralytic lagophthalmos: The hyaluronic acid gel "gold weight". Ophthal Plast Reconstr Surg 2009;25:23-6.
10. Cruz AA, Coelho RP, Baccega A, Lucchezi MC, Souza AD, Ruiz EE, *et al*. Digital image processing measurement of the upper eyelid contour in graves disease and congenital blepharoptosis. Ophthalmology 1998;105:913-8.
11. Salouti R, Nowroozzadeh MH, Zamani M, Ghoreyshi M, Salouti R. Comparison of horizontal corneal diameter measurements using Galilei, EyeSys and Orbscan II systems. Clin Exp Optom 2009;92:429-33.
12. Hashemi H, KhabazKhoob M, Yazdani K, Mehravaran S, Mohammad K, Fotouhi A, *et al*. White-to-white corneal diameter in the Tehran eye study. Cornea 2010;29:9-12.