Original Article

"Utility of Smartphone-based Three-dimensional Surface Imaging for Digital Facial Anthropometry" ☆

Robin Hartmann [a], Felix Nieberle [a], Christoph Palm [b,c], Vanessa Brébant [d], Lukas Prantl [d], Reinald Kuehle [e], Torsten E. Reichert [a], Juergen Taxis [a,#], Tobias Ettl [a,#,*]

[a] *University Hospital Regensburg, Clinic and Polyclinic for Oral and Maxillofacial Surgery, Franz-Josef-Strauß-Allee 11, 93053, Regensburg, Germany*

[b] *Regensburg Medical Image Computing (ReMIC), Ostbayerische Technische Hochschule Regensburg (OTH Regensburg), Galgenbergstr. 32, 93053, Regensburg*

[c] *Regensburg Center of Biomedical Engineering (RCBE), OTH Regensburg and Regensburg University, Galgenbergstr. 32, 93053, Regensburg*

[d] *University Center of Plastic, Aesthetic, Hand and Reconstructive Surgery, University Hospital Regensburg, Franz-Josef-Strauß-Allee 11, 93053 Regensburg, Germany*

[e] *University of Heidelberg, Department of Oral and Maxillofacial Surgery, Im Neuenheimer Feld 400, 69120, Heidelberg, Germany*

ARTICLE INFO

ABSTRACT

Background: The utilization of three-dimensional (3D) surface imaging for facial anthropometry is a significant asset for patients undergoing maxillofacial surgery. Notably, there have been recent advancements in smartphone technology that enable 3D surface imaging.

In this study, anthropometric assessments of the face were performed using a smartphone and a sophisticated 3D surface imaging system.

Methods: 30 healthy volunteers (15 females and 15 males) were included in the study. An iPhone 14 Pro (Apple Inc., USA) using the

☆ Clinical trial registration number: DRKS00033007

* Corresponding author: Univ.-Prof. Dr. Dr. Tobias Ettl, M.D. D.M.D, Department of Oral and Maxillofacial Surgery, University Hospital Regensburg, Germany. University Hospital Regensburg Franz-Josef-Strauß-Allee 11, 93053 Regensburg, Tel: +49941944-6329.

E-mail address: tobias.ettl@ukr.de (T. Ettl).

# These authors contributed equally to the present work.

application 3D Scanner App (Laan Consulting Corp., USA) and the Vectra M5 (Canfield Scientific, USA) were employed to create 3D surface models. For each participant, 19 anthropometric measurements were conducted on the 3D surface models. Subsequently, the anthropometric measurements generated by the two approaches were compared. The statistical techniques employed included the paired t-test, paired Wilcoxon signed-rank test, Bland–Altman analysis, and calculation of the intraclass correlation coefficient (ICC).

Results: All measurements showed excellent agreement between smartphone-based and Vectra M5-based measurements (ICC between 0.85 and 0.97). Statistical analysis revealed no statistically significant differences in the central tendencies for 17 of the 19 linear measurements. Despite the excellent agreement found, Bland–Altman analysis revealed that the 95% limits of agreement between the two methods exceeded ±3 mm for the majority of measurements.

Conclusion: Digital facial anthropometry using smartphones can serve as a valuable supplementary tool for surgeons, enhancing their communication with patients. However, the proposed data suggest that digital facial anthropometry using smartphones may not yet be suitable for certain diagnostic purposes that require high accuracy.

Introduction

Digital facial anthropometry (DFA) is a valuable measure for therapy and outcome assessment in patients undergoing maxillofacial surgery.[1–3] In this method, measurements are performed on three-dimensional (3D) surface models (SMs) of the face, enabling the individual assessment of parameters such as symmetry or changes in facial morphology.[1–5] SMs are usually created by standard 3D surface imaging systems such as the 3dMD system (*3dMD, USA*), Vectra systems (*Canfield Scientific, USA*), or Artec systems (*Artec Group, Luxembourg, Luxembourg*).[6–8] Numerous studies have validated these systems for clinical use in DFA.[4,6,7,9–11]

A novelty in the field is smartphone use for DFA. In this approach, smartphones are used to generate precise SMs of facial soft tissue, which are subsequently employed for anthropometric assessments.[12–14]

There is a paucity of protocols available for validating smartphone-based DFA. Rudy et al. used the iPhone X (*Apple Inc., USA*) with the application ScandyPro (*Scandy Inc., USA*), along with the Vectra H1 (*Canfield Scientific, USA*), to obtain surface information from 16 participants.[12] Besides other evaluations, they conducted anthropometric measurements between 10 anatomical landmarks. They concluded that the iPhone X offers accurate and precise 3D scanning. Chong et al. developed an investigational application for the iPad/iPhone (*Apple Inc., USA*) and enrolled 20 subjects to acquire surface information.[13] Among other assessments, they performed 21 digital anthropometric measurements, which showed excellent correlations with direct measurements. They concluded that their 3D facial scanning system allowed patients to capture 3D images on their own. Akan et al. enrolled 26 individuals and performed alongside other evaluations, seven linear and three angular measurements, comparing the iPhone X and the 3dMD 3D imaging system. The researchers concluded that images captured with a smartphone can be used to record and evaluate 3D soft tissue changes.[14]

Notably, previous validation of smartphone-based DFA has primarily relied on unsophisticated protocols for anthropometry or investigational applications. Consequently, both the validity and the reli-

ability of these studies may be compromised. The absence of a current gold standard for smartphone-based DFA limits the validity of previous investigations.

To address these limitations, a study was conducted comparing smartphone-based DFA with a previously established protocol that utilizes a sophisticated 3D imaging system, considered the gold standard for digital anthropometry. The iPhone 14 Pro (*Apple Inc., USA*), using the application 3D Scanner App V2.1.2 (*Laan Consulting Corp., USA*), and the Vectra M5 (*Canfield Scientific, USA*) were used to generate SMs of the face. A set of standardized anthropometric measurements was conducted based on these models, which were then compared.

This comparison allows the evaluation of the accuracy of digital anthropometric measurements with smartphones and determines their potential as an alternative method for digital anthropometry. By filling this knowledge gap, valuable insights were gained to be provided for future research and clinical applications, contributing to the advancement of smartphone-based DFA.

Material and methods

Study protocol

This trial was designed as a monocentric prospective study. It was conducted at the Department of Oral and Maxillofacial Surgery at the University Hospital Regensburg, Germany. Prior to participant recruitment, the study received approval from the local ethics committee (23-3400-101). The proposed method was then tested on a group of 30 healthy adult students who were enrolled at the University of Regensburg. Individuals with a recent history of craniofacial surgery, maxillofacial trauma, or substantial skeletal deformities were not included in the study.

The participants were invited to undergo an anthropometric examination. After receiving an explanation of the procedure, 15 specific landmarks were identified and marked using a white eyeliner. The anthropometry protocol used had been previously described by Othman et al.[11] To maintain consistency, participants were positioned on a stool in a relaxed posture, with their heads in a natural position and under standardized lighting conditions. To eliminate any potential interference, participants were instructed to remove any glasses, jewelry, and makeup. Additionally, they were asked to wear a hairband to minimize any disturbances caused by their hair. Subsequently, 3D SMs of the face were generated, which were used to conduct a set of anthropometric measurements.

3D surface imaging

To capture the participants' surface information, the iPhone 14 Pro using the application 3D Scanner App V2.1.2 and the Vectra M5 System using the software Vectra Analysis Module (VAM) (*Canfield Scientific, USA*) were utilized.

The Vectra M5 is a stationary and highly reliable surface imaging system.[15] It uses stereophotogrammetry for surface imaging. The technique involves arranging multiple pictures as a stereo pair and merging the intersections between different images into a 3D model.[16] Stereophotogrammetry is a well-established method that has been widely used for 3D surface imaging.[17,18] It is considered the gold standard for DFA.[19] Previous studies have proved its effectiveness in capturing the face, making it valuable for surgical planning and evaluation in maxillofacial surgery.[18,20] The Vectra M5 comes with the software VAM for SM analysis.[21] The software allows anthropometric measurements.

The 3D Scanner App is based on the Object Capture API by Apple (*Apple Inc., USA*). It provides a LiDAR-Mode, a Pointcloud-Mode, a TrueDepth-Mode, and a Photo-Mode. In this study, the Photo-Mode was used to capture the patients' surface information. According to the Laan Consulting Corp., the Photo-Mode combines photogrammetry and LiDAR (light detection and ranging) to estimate object scale. LiDAR involves time-of-flight measurements, which determine the time taken for an object to travel a distance.[22] Combining both technologies helps to create an accurate 3D SM that can be used for anthropometric measurements. Once the 3D SM is completed, users can edit it or share it in various formats. The 3D Scanner App features a measurement tool for conducting linear measurements. It is complemented by the 3D Scanner App software designed for MacOS (Apple Inc., USA), which facilitates point cloud processing and mesh generation.

Fig. 1. *Landmarks*: Appearance of a 28-year-old male participant; the iPhone 14 Pro (*Apple Inc., USA*) using the application 3D Scanner App V2.1.2 (*Laan Consulting Corp., USA*) was used to create the surface model; (1) soft tissue nasion (N), (2) pronasale (PRN), (3) subnasale (SN), (4) labrale superius (LS), (5) stomion (STO), (6) labrale inferius (LI), (7) soft tissue gnathion (GN), (8) alare (AL) (L), (9) alare (AL) (R), (10) subalare (SBAL) (L), (11) subalare (SBAL) (R), (12) christa philtri (CPH) (L), (13) christa philtri (CPH) (R), (14) cheilion (CH) (L), and (15) cheilion (CH) (R).

Landmarks

Prior studies have emphasized the significance of accurately labeling landmarks for anthropometric measurements.[23] Thus, following previous research, landmarks were manually marked using a removable white eyeliner.[11,24]

The following landmarks were marked unilaterally: (1) soft tissue nasion (N), (2) pronasale (PRN), (3) subnasale (SN), (4) labrale superius (LS), (5) stomion (STO), (6) labrale inferius (LI), and (7) soft tissue gnathion (GN).

The following landmarks were marked bilaterally left (L) and right (R): (8) and (9) alare (AL), (10) and (11) subalare (SBAL), (12) and (13) christa philtri (CPH), (14), and (15) cheilion (CH).

Figure 1 provides an overview of all landmarks.
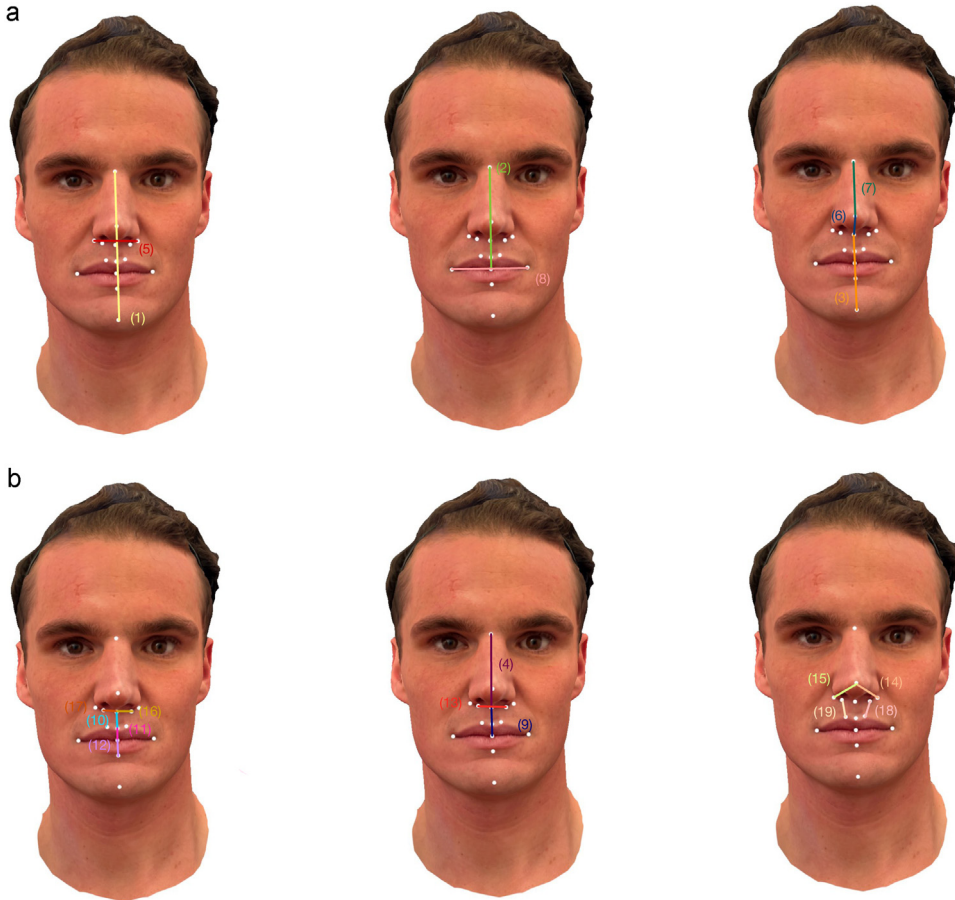
Anthropometric measurements

The following measurements were performed: (1) N – GN, (2) N – STO, (3) SN – GN, (4) N – SN, (5) AL – AL, (6) SN – PRN, (7) N – PRN, (8) CH – CH, (9) SN – STO, (10) SN – LS, (11) LS – STO, (12) LI – STO, (13) SBAL – SBAL, (14) AL – PRN (L), (15) AL – PRN (R), (16) SBAL – SN (L), (17) SBAL – SN (R), (18) SBAL – CPH (L), and (19) SBAL – CPH (R).

A medical doctor digitally conducted all anthropometric measurements in the study using the 3D Scanner App measurement tool and the VAM software. Measurements based on the Vectra M5 were conducted on the computer using the VAM, while smartphone-based measurements were conducted using the smartphone application. Digital anthropometry, as used in the study, refers to the process that uses SMs to perform anthropometric assessments.[1] The physician digitally selected the midpoints of the landmarks in concordance to direct measurements. The software then measured the metric linear distance between the manually selected vertices.

Figure 2(a) and (b) provides an overview of all anthropometric measurements. Figure 3 juxtaposes the smartphone-based and the Vectra-based approach for anthropometric measurements.

Statistical analysis

IBM SPSS 29 (*SPSS Inc., USA*) was used for statistical analysis. Normal distribution of each measurement was assessed using the Kolmogorov–Smirnov Test. Normality was assumed for measurements

Fig. 2. (a) and (b) *Anthropometric measurements*: Appearance of a 28-year-old male participant; the iPhone 14 Pro (*Apple Inc., USA*) using the application 3D Scanner App V2.1.2 (*Laan Consulting Corp., USA*) was used to create the surface model; (1) N – GN, (2) N – STO, (3) SN – GN, (4) N – SN, (5) AL – AL, (6) SN – PRN, (7) N – PRN, (8) CH – CH, (9) SN – STO, (10) SN – LS, (11) LS – STO, (12) LI – STO, (13) SBAL – SBAL, (14) AL – PRN (L), (15) AL – PRN (R), (16) SBAL – SN (L), (17) SBAL – SN (R), (18) SBAL – CPH (L), and (19) SBAL – CPH (R). The figure exemplifies linear measurements on 3D SMs. For illustrational purposes, the measurements on 3D SMs were depicted in 2D pictures.

(1) to (11) and (14) to (19). Consequently, for these 17 measurements, a paired t-test was conducted to examine any systematic bias between the two methods. However, measurements (12) and (13) did not exhibit normal distribution. For these two variables, a Wilcoxon signed-rank test for paired samples was used to evaluate any systematic bias between the methods. The consistency between measurements was evaluated using the intraclass correlation coefficient (ICC). The correlation was evaluated according to Cicchetti et al. using the following guidelines for interpretation: less than 0.40 – poor, between 0.40 and 0.59 – fair, between 0.60 and 0.74 – good, and between 0.75 and 1.00 – excellent.[25] To evaluate the agreement between the measurements, Bland–Altman analyses were conducted. An acceptable total deviation of ±3.0 mm was deemed appropriate, considering any limit of agreement beyond 3 mm at a 95% confidence level as clinically unacceptable. Table 1 provides an overview of the means, maxima, and minima of the two approaches.
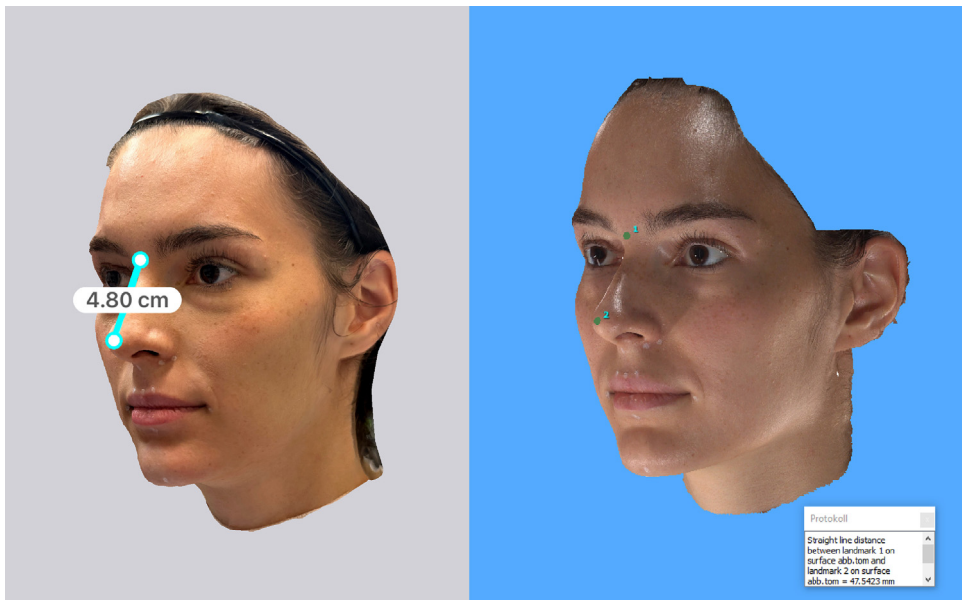
Fig. 3. *Presentation of the two methods:* Appearance of a 21-year-old female participant; the iPhone 14 Pro (*Apple Inc., USA*) using the application 3D Scanner App V2.1.2 (*Laan Consulting Corp., USA*) was used to perform measurement (7) (N – PRN) on the surface model on the left side; the Vectra M5 (*Canfield Scientific, USA*) using the software VAM (*Canfield Scientific, USA*) was used to perform measurement (7) (N – PRN) on the right side.

Table 1
Descriptive Statistics: Values in millimeters (mm) for Vectra M5-based and smartphone-based measurements (1) to (19); IBM SPSS 29 was used for data analysis.

Descriptive Statistics						
Variables	Vectra M5			Smartphone		
	Minimum	Maximum	Mean	Minimum	Maximum	Mean
(1) N – GN	109.5	136.2	118.6	106.1	128.2	118.9
(2) N – STO	69.2	87.3	79.2	67.3	85.3	78.7
(3) SN – GN	52.4	73.3	61.7	51.3	76.7	62.3
(4) N – SN	48.1	68.7	58.4	47.5	65.5	57.9
(5) AL – AL	28.0	40.2	34.9	29.6	41.4	35.3
(6) SN – PRN	15.6	25.0	20.3	15.7	24.9	20.5
(7) N – PRN	42.1	60.0	51.2	40.7	60.3	50.9
(8) CH – CH	46.9	60.9	55.3	46.4	63.0	54.5
(9) SN – STO	14.9	28.1	21.6	17.1	26.6	21.4
(10) SN – LS	7.0	13.7	9.9	5.4	13.4	10.0
(11) LS – STO	9.1	18.6	12.1	9.8	17.2	11.9
(12) LI – STO	8.2	17.2	12.8	8.2	18.3	13.4
(13) SBAL – SBAL	14.2	31.8	20.6	14.9	37.5	20.7
(14) AL – PRN (L)	20.6	39.2	33.6	29.5	39.3	34.0
(15) AL – PRN (R)	28.4	40.0	33.8	29.6	39.4	34.2
(16) SBAL – SN (L)	8.8	15.3	11.6	8.5	15.4	11.8
(17) SBAL – SN (R)	7.0	15.5	11.5	6.8	16.6	11.6
(18) SBAL – CPH (L)	9.0	16.8	13.1	9.3	16.7	13.1
(19) SBAL – CPH (R)	8.0	16.6	13.0	7.7	17.6	13.2

Table 2

Intraclass correlation coefficient (ICC) and Bland–Altman analysis: ICC and Bland–Altman 95% limits of agreement between the two methods; IBM SPSS 29 was used for data analysis.

Intraclass correlation coefficient (ICC) and Bland-Altman analysis						
	ICC			Bland-Altman		
Variables	ICC	95% Confidence Interval		Mean bias	95% Confidence Intervall	
		Upper	Lower		Upper	Lower
(1) N – GN	.9	.95	.8	-.2	7.0	-7.4
(2) N – STO	.96	.98	.9	.5	4.4	-3.3
(3) SN – GN	.95	.98	.89	-.6	3.6	-4.7
(4) N – SN	.96	.98	.92	.5	4.0	-3.0
(5) AL – AL	.93	.97	.86	-.4	2.7	-3.5
(6) SN – PRN	.9	.95	.8	-.2	2.3	-2.7
(7) N – PRN	.97	.99	.94	.3	3.6	-3.0
(8) CH – CH	.95	.98	.9	.7	4.1	-2.7
(9) SN – STO	.92	.96	.82	.2	2.8	-2.4
(10) SN – LS	.91	.96	.81	-.2	1.8	-2.2
(11) LS – STO	.92	.96	.84	.2	2.1	-1.7
(12) LI – STO	.92	.96	.84	-.7	1.5	-2.8
(13) SBAL – SBAL	.95	.97	.88	-.1	3.6	-3.7
(14) AL – PRN (L)	.87	.94	.72	-.4	4.0	-4.7
(15) AL – PRN (R)	.92	.96	.82	-.4	2.5	-3.4
(16) SBAL – SN (L)	.85	.93	.68	-.3	2.2	-2.8
(17) SBAL – SN (R)	.93	.97	.84	-.2	2.0	-2.3
(18) SBAL – CPH (L)	.94	.97	.87	.1	1.9	-1.8
(19) SBAL – CPH (R)	.9	.95	.78	-.2	2.2	-2.7

Results

Patient demographics

The cohort included 15 males and 15 females. Their mean age was M = 24 years ($SD = \pm2.3$), the mean height M = 176 cm ($SD = \pm8$ cm), the mean weight M = 69.6 kg ($SD = \pm14.0$ kg), and the mean BMI M = 22.5 ($SD = \pm3.6$).

Accuracy of measurements in the mid-face region

The values for measurements (1) to (3) showed excellent correlation with ICC values from 0.9 to 0.96 (Table 2). T-test for paired samples revealed that the mean values for measurements (1) to (3) did not differ significantly between Vectra-based (1) (M = 118.6); (2) (M = 79.2); (3) (M = 61.7) and smartphone-based (1) (M = 118.9); (2) (M= 78.7); (3) (M = 62.3) measurements (*t-test for paired samples; (1) p = 0.73, (2) p = 0.16, (3) p = 0.15, n = 30*) (Table 3). Nevertheless, the Bland–Altman analysis revealed clinically unacceptable disparities (>3 mm) between the two methods, as indicated by the 95% limits of agreement exceeding ±3.00 mm (Table 2). Measurement (1) showed the highest difference between the two methods with Bland–Altman 95% limits of agreement reaching from 7.0 mm to -7.4 mm. Figure 4 shows the Bland–Altman plots for the mid-face region.

Accuracy of measurements in the nose region

All measurements of the nose region (4) to (7) demonstrated excellent correlation with ICC values ranging from 0.9 to 0.97 (Table 2). The t-test for paired samples indicated that the mean values for measurements (4) to (7) did not exhibit a significant difference between Vectra-based (4) (M = 58.4); (5) (M = 34.9); (6) (M = 20.3); (7) (M = 51.2) and smartphone-based (4) (M = 57.9); (5) (M = 35.3); (6) (M = 20.5); (7) (M = 50.9) measurements (*t-test for paired samples; (4) p = 0.14, (5) p = 0.19, (6) p = 0.47, (7) p = 0.36, n = 30*) (Table 3). Measurements (4), (5), and (7) showed clinically unacceptable

Table 3

T-test for paired samples: Comparison of Vectra M5-based and smartphone-based measurements (1)–(11) and (14)–(19); IBM SPSS 29 was used for data analysis.

T-test for paired samples					
Variables	Paired Differences Mean	Std. Deviation	95% Confidence Interval of the Difference		p
			Lower	Upper	
(1) N – GN	-.2	3.7	-1.6	1.1	.73
(2) N – STO	.5	2.0	-.2	1.3	.16
(3) SN – GN	-.6	2.1	-1.4	.2	.15
(4) N – SN	.5	1.8	-.2	1.2	.14
(5) AL – AL	-.4	1.6	-1.0	.2	.19
(6) SN – PRN	-.2	1.3	-.6	.3	.47
(7) N – PRN	.3	1.7	-.3	.9	.36
(8) CH – CH	.7	1.7	.1	1.4	.03
(9) SN – STO	.2	1.3	-.3	.7	.45
(10) SN – LS	-.2	1.0	-.5	.2	.39
(11) LS – STO	.2	1.0	-.2	.5	.29
(14) AL – PRN (L)	-.4	2.2	-1.2	.4	.35
(15) AL – PRN (R)	-.4	1.5	-1.0	.1	.13
(16) SBAL – SN (L)	-.3	1.3	-.8	.2	.22
(17) SBAL – SN (R)	-.2	1.1	-.6	.2	.40
(18) SBAL – CPH (L)	.1	.9	-.3	.4	.73
(19) SBAL – CPH (R)	-.2	1.3	-.7	.2	.30

limits of agreement between the two methods with Bland–Altman 95% limits of agreement exceeding ±3.00 mm. The only acceptable value for the nose region was measurement (6) with Bland–Altman 95% limits of agreement from 2.3 mm to -2.7 mm (Table 2). Figure 5 shows the Bland–Altman plots for the nose region.

Accuracy of measurements in the orolabial region

The orolabial region (8) to (19) showed the highest agreement for both methods. Eight out of 12 measurements were found within the clinically acceptable Bland–Altman 95% limits of ±3.00 mm. However, measurements (8), (13), (14), and (15) exceeded the ±3.00 mm limits. For measurement (8), values reached from 4.1 mm to -2.7 mm; for measurement (13), from 3.6 mm to -3.7 mm; for measurement (14), from 4.0 mm to -4.7 mm; and for measurement (15), from 2.5 mm to -3.4 mm (Table 2). The t-test for paired samples was conducted for measurements (8)–(11) and (14)–(19). For measurements (9)–(11) and (14)–(19) mean values did not exhibit a significant difference between Vectra-based measurements and smartphone-based measurements (Table 3). A significant difference was found in measurement (8) between Vectra-based (M = 55.3) and smartphone-based (M = 54.5) measurements (*t-test for paired samples; p = .03, n = 30*) (Table 3). A Wilcoxon signed-rank test for paired samples was used to compare values for measurements (12) and (13). The values for measurement (12) did differ significantly between Vectra-based (Median = 12.7) and smartphone-based (Median = 13.5) measurements (*Wilcoxon signed-rank test for paired samples; p = <.001, n = 30*) (Table 4). However, values for measurement (13) did not differ significantly between Vectra-based (Median = 19.7) and smartphone-based (Median = 19.9) measurements (*Wilcoxon signed-rank test for paired samples; p = .46, n = 30*) (Table 4). All measurements showed excellent correlation between the two methods. ICC values ranged from 0.85 to 0.95 (Table 2). Figure 6(a) and (b) shows the Bland–Altman plots for the orolabial region.

Overall accuracy

Most of the measurements demonstrated excellent agreement between the two methods, as supported by high ICC values and comparable central tendencies. However, the majority of measurements
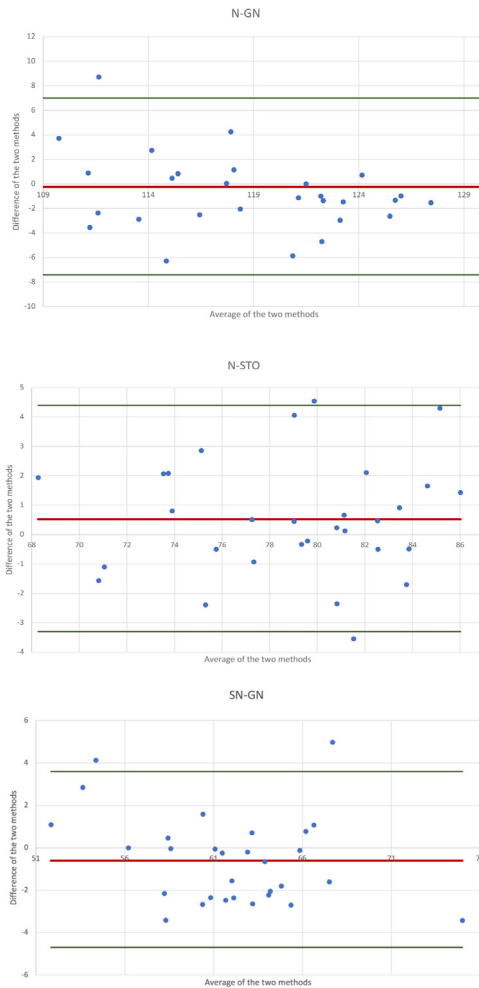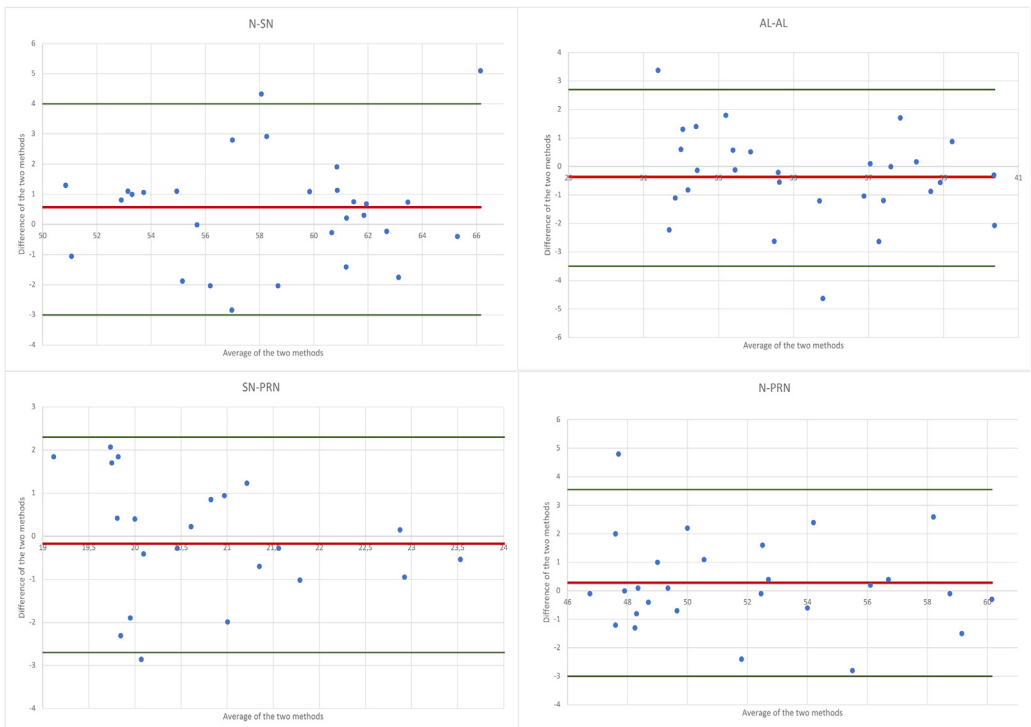
Fig. 4. *Bland–Altman plots of the mid-face region*: Measurements (1) to (3) in millimeters (mm); MS Excel was used to create the illustration.

Table 4
Wilcoxon Signed Rank Test for paired samples: Median values in Millimeters (mm) and Wilcoxon Signed Rank Test for paired samples for measurements (12) and (13); Comparison of Vectra M5-based and smartphone-based measurements (12) and (13); IBM SPSS 29 was used for data analysis.

Median values and Wilcoxon Signed Rank Test for paired samples				
Variables	(12) LI – STO		(13) SBAL – SBAL	
	Vectra M5	Smartphone	Vectra M5	Smartphone
Median	12.7	13.5	19.7	19.9
Z	-2.81		-.74	
Asymp. Sig. (2-tailed)	<.001		.46	

Fig. 5. *Bland–Altman plots of the nose region*: Measurements (4) to (7) in millimeters (mm); MS Excel was used to create the illustration.
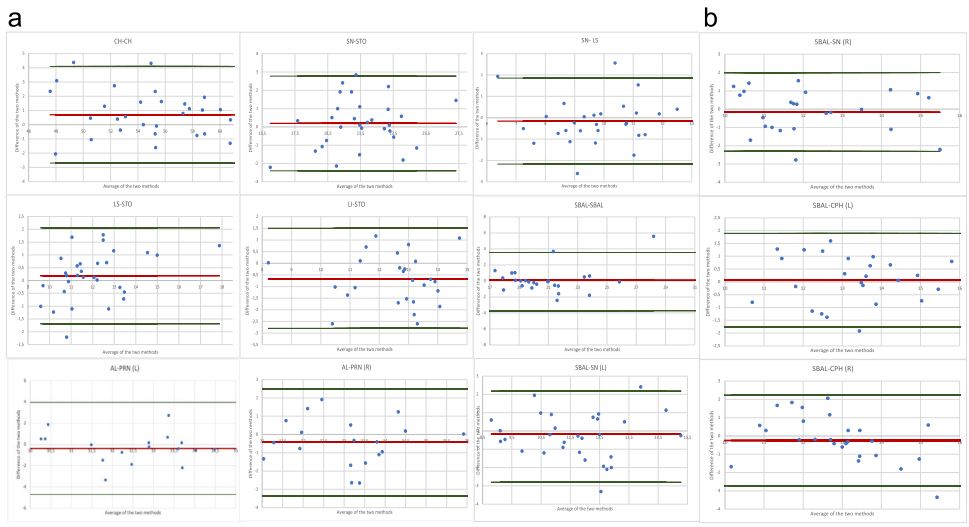
Fig. 6. (a) *Bland–Altman plots of the orolabial region*: Measurements (8) to (16) in millimeters (mm); MS Excel was used to create the illustration. (b) *Bland–Altman plots of the orolabial region*: Measurements (17) to (19) in millimeters (mm); MS Excel was used to create the illustration.

displayed Bland–Altman 95% limits of agreement that surpassed the limits of ±3.00 mm. This implies that smartphone-based DFA may be inappropriate for specific diagnostic purposes that demand a high level of accuracy. Nevertheless, it may be used as a supplementary tool for clinicians and patients. Some values, particularly in the orolabial region, displayed Bland–Altman 95% limits of agreement that did not exceed ±3.00 mm.

Discussion

This study has certain limitations that necessitate further discussion. In this trial, a fully smartphone-based approach was compared to a fully Vectra-based technique under realistic circumstances, with the aim of providing physicians with valuable insights into future considerations for implementation. One limitation of this study is that anthropometric measurements were conducted manually using two different systems: the 3D Scanner App and the Vectra VAM. Both approaches involve performing measurements manually. Although this approach aimed to replicate realistic conditions, such as those encountered in telemedicine, performing measurements manually may affect the measurements' accuracy. Prior studies have, therefore, introduced methods for automated anthropometry.[26,27] However, the field of smartphone-based DFA lacks automation, which may help to overcome this limitation.

Despite technological advancements, landmark detection remains a challenge in smartphone-based DFA. In the present study, a smartphone-based approach that involved selecting landmarks using a smartphone was employed. While this method allows untrained users to perform anthropometry, accurately identifying landmarks can be difficult. According to Farkas et al., the reliability of anthropometric evaluations relies on both the examiner's expertise in measurements and their ability to identify anatomical landmarks accurately.[28] To enhance visibility and accuracy, a white eyeliner was used to mark the landmarks on the participants' faces, and a clinician reduced the landmark diameter to a minimum. However, in this study landmarks involving the orolabial region, such as the stomion (5) and the labrale inferius (6), as well as the left and right cheilion (14) and (15), were challenging to identify as they are highly influenced by facial expressions. These observations are supported by a study performed by Aynechi et al. on landmarking in anthropometry.[24] In their trial, the researchers investigated the impact of landmark labeling on the accuracy and precision of an indirect facial anthropometric technique. They obtained 18 facial anthropometric measurements from a sample of 10 adults using the 3dMDface system. These measurements were taken with both labeled and unlabeled landmarks and were compared to measurements obtained through direct anthropometry. The authors observed that modifications in facial expressions during image capturing can introduce potential errors due to the inherent difficulty of maintaining a natural facial expression.[24] This could potentially account for the observed differences in measurements (8) (cheilion to cheilion) and (12) (labrale inferius to stomion).

These limitations emphasize the need for further research and improvements in smartphone-based DFA to address the challenges associated with landmark detection accuracy and to enhance its applicability for untrained users. By implementing automation in DFA, it may be possible to overcome at least some of these challenges.

Furthermore, critically examining the findings of the study regarding the agreement of the orolabial region is crucial. Due to the lack of a gold standard for DFA, comparing this study's findings is challenging. Nonetheless, previous studies have evaluated the disparities in the orolabial region using similar assessments. In their comparative study, Othman et al. evaluated the Vectra M5 System alongside direct anthropometric measurements, comparing 19 linear facial measurements obtained from 37 cleft patients. They found that for 8 out of 12 linear measurements in the orolabial region, the 95% limits of agreement did not meet clinically acceptable standards.[11] However, the present study's findings are consistent with previous trials on smartphone-based digital anthropometry in the orolabial region. Chong et al. conducted a comprehensive study comparing anthropometric measurements obtained from an iPhone/iPad-based system with direct measurements, specifically examining 21 anthropometric measurements including the orolabial region. Their results revealed that the majority of the measurements in the orolabial region exhibited disparities within 95% Bland–Altman limits of 2 mm.[13] Akan et al. performed a study comparing anthropometric measurements conducted on SMs

obtained by the iPhone X and the 3dMD 3D imaging system. They found no statistically significant differences in vermillion height and mouth width between the two systems.[14] Nevertheless, in the present trial, the absolute metric values for orolabial measurements were lower compared to other regions, such as the mid-face region. Although a clinically acceptable agreement was observed using the Bland–Altman method for 8 of 12 measurements, the result may be attributed to the generally low metric values in this region.

A drawback of this investigation is that volunteers were evaluated by a single examiner only once. Further studies should investigate the inter- and intra-rater reliability of smartphone-based DFA. Nevertheless, the data provided in this study contribute to the progress in the clinical application of smartphone-based DFA by offering insights into the precision of the method.

In addition, it is notable that the ± 3 mm agreement used in the present study to assess the Bland–Altman 95% limits of agreement may differ from that in previous investigations, which have used a ± 2 mm agreement for validation of stereophotogrammetry systems.[11,29] It is essential to recognize that the quality of meshes generated by smartphones may not yet be on par with the sophisticated scan systems used in previous investigations. Previous studies have found that smartphone-based scans show less accuracy for 3D surface imaging when compared with sophisticated scan systems.[30,31] The implications of this in terms of clinical applications are still unclear.

Earlier investigations have defined differences in SMs greater than 3 mm as clinically relevant.[32] This approach acknowledges that slight differences within the range of volumetric fluctuations due to circadian rhythm changes in the face are unlikely to have a significant impact on clinical decision-making.[32] Therefore, in the present study, an agreement of ± 3 mm was used to define the method as clinically acceptable.

When considering the use of smartphone-based SMs for DFA, it is crucial to consider factors that may influence the method's accuracy. Previous investigations have found that the accuracy of SMs generated by smartphones may vary between different smartphone applications.[33] This variation could potentially account for the reported inaccuracy. Furthermore, various elements, such as the quality of the smartphone's lens system and the inherent measurement error of manually performing measurements, must be considered as factors influencing the accuracy of smartphone-based DFA.

Finally, it is necessary to consider the context in which the proposed method is used. As smartphones continue to evolve and improve their capabilities, future research may warrant reassessing the acceptable limits of agreement for smartphone-based DFA.

Further research is required to evaluate the extent of inaccuracy in the reproduction of anatomically complex surfaces of smartphone-based surface imaging and the potential impact on clinical applications. It is important to consider these limitations when applying DFA using smartphones in clinical settings where accuracy is critical.

Conclusion

The proposed data suggest that DFA using smartphones may not yet be suitable for certain diagnostic purposes requiring high accuracy. Today, it can serve as a complementary tool to facilitate communication between patients and clinicians. To enhance the accuracy of the method, additional automation may be necessary. The technology of DFA using smartphones demonstrates the potential to become a powerful tool in the future, and maxillofacial surgeons are at the forefront of implementing this technology.

Conflict of interest

The authors declare that they have no conflicts of interest.

Acknowledgments

Department of Oral and Maxillofacial Surgery Heidelberg (University Hospital Heidelberg), and the Department of Oral and Maxillofacial Surgery Regensburg (University Hospital Regensburg) for their collaboration and support.

Ethical approval

The institutional Ethics Committee of the Medical Faculty of the University of Regensburg (23-3400-101), Germany approved the study. All procedures performed in this study were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

Patient consent for photo publication

We hereby affirm that informed consent has been obtained from the patient(s) or their legal guardian(s) for the use of their photographs in our manuscript. This consent encompasses all forms of media, including print, digital, and electronic formats, as part of the publication process. The patient(s) or their legal guardian(s) have been thoroughly informed about the nature of the publication, the context in which the photographs will be used, and the potential audience. They have been reassured that their anonymity will be preserved and that no personal or identifiable information will be disclosed. The consent was obtained in a manner that adheres to ethical standards and respects the rights and dignity of the patient(s).

Author contributions

RH: project development, data collection, data analysis and interpretation, manuscript writing; FN: data analysis and interpretation, critical revision of the article; CP: data analysis and interpretation, critical revision of the article; VB: data analysis and interpretation, critical revision of the article; LP: data analysis and interpretation, critical revision of the article; RK: data analysis and interpretation, critical revision of the article; TR: project development, data analysis and interpretation, critical revision of the article, final approval; JT: project development, data collection, data analysis and interpretation, critical revision of the article, manuscript writing, final approval; TE: project development, data collection, data analysis and interpretation, critical revision of the article, manuscript writing, final approval.

Funding

Informed consent

Informed consent was obtained from every participant in this study.

References

1. Jayaratne YSN, Zwahlen RA. Application of digital anthropometry for craniofacial assessment. *Craniomaxillofacial Trauma & Reconstruction*. 2014;7:101–107. doi:10.1055/s-0034-1371540.
2. Wong JY, Oh AK, Ohta E, Hunt AT, Rogers GF, Mulliken JB, et al. Validity and reliability of craniofacial anthropometric measurement of 3D digital photogrammetric images. *The Cleft Palate-Craniofacial Journal*. 2008;45:232–239. doi:10.1597/06-175.
3. Heike CL, Cunningham ML, Hing AV, Stuhaug E, Starr JR. Picture perfect? Reliability of craniofacial anthropometry using three-dimensional digital stereophotogrammetry: Plastic and Reconstructive Surgery 2009;124:1261–72. https://doi.org/10.1097/PRS.0b013e3181b454bd.
4. Ghoddousi H, Edler R, Haers P, Wertheim D, Greenhill D. Comparison of three methods of facial measurement. *International Journal of Oral and Maxillofacial Surgery*. 2007;36:250–258. doi:10.1016/j.ijom.2006.10.001.
5. Koban KC, Perko P, Li Z, Xu Y, Giunta RE, Alfertshofer MG, et al. 3D anthropometric facial imaging - a comparison of different 3D scanners. *Facial Plastic Surgery Clinics of North America*. 2022;30:149–158. doi:10.1016/j.fsc.2022.01.003.
6. Düppe K, Becker M, Schönmeyr B. Evaluation of facial anthropometry using three-dimensional photogrammetry and direct measuring techniques. *Journal of Craniofacial Surgery*. 2018;29:1245–1251. doi:10.1097/SCS.0000000000004580.

7. de Menezes M, Rosati R, Ferrario VF, Sforza C. Accuracy and reproducibility of a 3-dimensional stereophotogrammetric imaging system. *Journal of Oral and Maxillofacial Surgery*. 2010;68:2129–2135. doi:10.1016/j.joms.2009.09.036.

8. Modabber A, Peters F, Kniha K, Goloborodko E, Ghassemi A, Lethaus B, et al. Evaluation of the accuracy of a mobile and a stationary system for three-dimensional facial scanning. *Journal of Cranio-Maxillofacial Surgery*. 2016;44:1719–1724. doi:10.1016/j.jcms.2016.08.008.

9. Naini FB, Akram S, Kepinska J, Garagiola U, McDonald F, Wertheim D. Validation of a new three-dimensional imaging system using comparative craniofacial anthropometry. *Maxillofac Plast Reconstr Surg*. 2017;39:23. doi:10.1186/s40902-017-0123-3.

10. Savoldelli C, Benat G, Castillo L, Chamorey E, Lutz J-C. Accuracy, repeatability and reproducibility of a handheld three-dimensional facial imaging device: the Vectra H1. Journal of Stomatology. *Oral and Maxillofacial Surgery*. 2019;120:289–296. doi:10.1016/j.jormas.2019.03.012.

11. Othman SA, Saffai L, Wan Hassan WN. Validity and reproducibility of the 3D VECTRA photogrammetric surface imaging system for the maxillofacial anthropometric measurement on cleft patients. *Clin Oral Invest*. 2020;24:2853–2866. doi:10.1007/s00784-019-03150-1.

12. Rudy HL, Wake N, Yee J, Garfein ES, Tepper OM. Three-dimensional facial scanning at the fingertips of patients and surgeons: accuracy and precision testing of iPhone X three-dimensional scanner. *Plastic & Reconstructive Surgery*. 2020;146:1407–1417. doi:10.1097/PRS.0000000000007387.

13. Chong Y, Liu X, Shi M, Huang J, Yu N, Long X. Three-dimensional facial scanner in the hands of patients: validation of a novel application on iPad/iPhone for three-dimensional imaging. *Ann Transl Med*. 2021;9:1115 –1115. doi:10.21037/atm-21-1620.

14. Akan B, Akan E, Şahan AO, Kalak M. Evaluation of 3D face-scan images obtained by stereophotogrammetry and smartphone camera. *International Orthodontics*. 2021;19:669–678. doi:10.1016/j.ortho.2021.08.007.

15. De Stefani A, Barone M, Hatami Alamdari S, Barjami A, Baciliero U, Apolloni F, et al. Validation of Vectra 3D imaging systems: a review. *IJERPH*. 2022;19:8820. doi:10.3390/ijerph19148820.

16. Douglas TS. Image processing for craniofacial landmark identification and measurement: a review of photogrammetry and cephalometry. *Computerized Medical Imaging and Graphics*. 2004;28:401–409. doi:10.1016/j.compmedimag.2004.06.002.

17. Gibelli D, Cappella A, Dolci C, Sforza C. 3D surface acquisition systems and their applications to facial anatomy: let's make a point. *Italian Journal of Anatomy and Embryology*. 2020:422–431. doi:10.13128/IJAE-11671.

18. Gibelli D, Dolci C, Cappella A, Sforza C. Reliability of optical devices for three-dimensional facial anatomy description: a systematic review and meta-analysis. *International Journal of Oral and Maxillofacial Surgery*. 2020;49:1092–1106. doi:10.1016/j.ijom.2019.10.019.

19. Othman SA, Majawit LP, Wan Hassan WN, Wey MC, Mohd Razi R. Anthropometric study of three-dimensional facial morphology in Malay adults. *PLoS ONE*. 2016;11:e0164180. doi:10.1371/journal.pone.0164180.

20. Quinzi V, Polizzi A, Ronsivalle V, Santonocito S, Conforte C, Manenti RJ, et al. Facial scanning accuracy with stereophotogrammetry and smartphone technology in children: a systematic review. *Children*. 2022;9:1390. doi:10.3390/children9091390.

21. Santander P, Quast A, Hubbert J, Horn S, Meyer-Marcotty P, Küster H, et al. Three-dimensional head shape acquisition in preterm infants - translating an orthodontic imaging procedure into neonatal care. *Early Human Development*. 2020;140:104908. doi:10.1016/j.earlhumdev.2019.104908.

22. Vogt M, Rips A, Emmelmann C. Comparison of iPad Pro®'s LiDAR and TrueDepth capabilities with an industrial 3D scanning solution. *Technologies*. 2021;9:25. doi:10.3390/technologies9020025.

23. Gibelli D, Cappella A, Bertozzi F, Sala D, Sitta S, Tasso DR, et al. Three-dimensional facial anthropometric analysis with and without landmark labelling: is there a real difference? *Journal of Craniofacial Surgery*. 2022;33:665–668. doi:10.1097/SCS.0000000000007687.

24. Aynechi N, Larson BE, Leon-Salazar V, Beiraghi S. Accuracy and precision of a 3D anthropometric facial analysis with and without landmark labeling before image acquisition. *The Angle Orthodontist*. 2011;81:245–252. doi:10.2319/041810-210.1.

25. Cicchetti DV. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology n.d.

26. Hartmann R, Weiherer M, Schiltz D, Seitz S, Lotter L, Anker A, et al. A novel method of outcome assessment in breast reconstruction surgery: comparison of autologous and alloplastic techniques using three-dimensional surface imaging. *Aesth Plast Surg*. 2020. doi:10.1007/s00266-020-01749-4.

27. Hartmann R, Weiherer M, Schiltz D, Baringer M, Noisser V, Hösl V, et al. New aspects in digital breast assessment: further refinement of a method for automated digital anthropometry. *Arch Gynecol Obstet*. 2021;303:721–728. doi:10.1007/s00404-020-05862-2.

28. Farkas LG. Accuracy of anthropometric measurements: past, present, and future. *The Cleft Palate-Craniofacial Journal*. 1996;33:10–22. doi:10.1597/1545-1569_1996_033_0010_aoampp_2.3.co_2.

29. Dindaroğlu F, Kutlu P, Duran GS, Görgülü S, Aslan E. Accuracy and reliability of 3D stereophotogrammetry: a comparison to direct anthropometry and 2D photogrammetry. *The Angle Orthodontist*. 2016;86:487–494. doi:10.2319/041415-244.1.

30. Elbashti M, Sumita Y, Aswehlee A, Seelaus R. Smartphone application as a low-cost alternative for digitizing facial defects: is it accurate enough for clinical application? *Int J Prosthodont*. 2019;32:541–543. doi:10.11607/ijp.6347.

31. Amornvit P, Sanohkan S. The accuracy of digital face scans obtained from 3D scanners: an in vitro study. *IJERPH*. 2019;16:5061. doi:10.3390/ijerph16245061.

32. Thurzo A, Strunga M, Havlínová R, Reháková K, Urban R, Surovková J, et al. Smartphone-based facial scanning as a viable tool for facially driven orthodontics? *Sensors*. 2022;22:7752. doi:10.3390/s22207752.

33. Van Lint L, Christiaens L, Stroo V, Bila M, Willaert R, Sun Y, et al. Accuracy comparison of 3D face scans obtained by portable stereophotogrammetry and smartphone applications. *J Med Biol Eng*. 2023;43:550–560. doi:10.1007/s40846-023-00817-9.