Original Article

Three-dimensional facial scanner in the hands of patients: validation of a novel application on iPad/iPhone for three-dimensional imaging

Yuming Chong[1,2#^], Xinyu Liu[2#], Mai Shi[3], Jiuzuo Huang[1], Nanze Yu[1], Xiao Long[1,4^]

[1]Department of Plastic and Aesthetic Surgery, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, China; [2]Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, China; [3]School of Medicine, Tsinghua University, Beijing, China; [4]State Key Laboratory of Complex Severe and Rare Diseases, Peking Union Medical College Hospital, Chinese Academy of Medical Science and Peking Union Medical College, Beijing, China

Contributions: (I) Conception and design: Y Chong, X Long; (II) Administrative support: X Long; (III) Provision of study materials or patients: J Huang, N Yu, X Long; (IV) Collection and assembly of data: Y Chong, X Liu, M Shi; (V) Data analysis and interpretation: Y Chong, X Liu; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

Correspondence to: Xiao Long. Department of Plastic and Aesthetic Surgery, State Key Laboratory of Complex Severe and Rare Diseases, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, China.
Email: longxiao@pumch.cams.cn.

Background: Three-dimensional (3D) photography plays an important role in surgical planning and postoperative evaluation. Commercial 3D facial scanners are expensive, and they require patients to come to the clinics for 3D photography. To solve this problem, we developed an iPad/iPhone application to enable patients to capture 3D images of themselves on their own. This study aimed to evaluate the validity and reproducibility of this novel imaging system.

Methods: 3D images were taken on 20 volunteers using the novel imaging system. Twenty-one anthropometric parameters were measured using calipers (direct measurement) and 3D photographs (3D photogrammetry). The results were compared to assess the accuracy and bias of 3D photogrammetry. The reproducibility was evaluated by testing intra- and interobserver reliabilities. Furthermore, 3D virtual models obtained by the novel imaging system and Vectra H1 camera were compared by performing heat map analysis.

Results: The 3D photogrammetric results showed excellent correlations with direct measurements. Most anthropometric parameters did not show statistically significant differences between the two methods. The 95% limits of agreement exceeded 2 mm in some parameters, especially those with large numbers, although their relative error measurements were very small. Intra- and interobserver reliabilities were high enough to ensure good reproducibility. The comparison of 3D models obtained by the novel imaging system and Vectra H1 camera showed that the mean distance and the mean RMS were 0.08 and 0.67 mm, respectively.

Conclusions: The novel 3D facial scanning system is validated to enable patients to take 3D images on their own. The imaging quality of the subnasale region needs further improvement. Future clinical applications include surgical planning, postoperative evaluation, and early diagnosis of diseases that affect facial appearance.

Keywords: Three-dimensional imaging (3D imaging); iPad/iPhone application; anthropometry; validity; reproducibility

^ ORCID: Yuming Chong, 0000-0001-9307-3046; Xiao Long, 0000-0003-0136-2508.

Page 2 of 14

Chong et al. Validation of a novel 3D imaging system

Introduction

Anthropometry on three-dimensional (3D) photography, or 3D photogrammetry, plays an important role in surgical planning and postoperative evaluation. Compared with direct measurement, the quick capture replaces the time-consuming process of caliper measurement to reduce the impact of patient movement (1), and the 3D virtual models can be archived for further analysis. Measuring on two-dimensional (2D) images does not allow the evaluation of surface area or volumetric change (2). Meanwhile, it has been argued that 2D photogrammetry was less precise and subject to lighting variation (3,4). In recent years, the value of 3D photography has been widely recognized as increasingly more studies have employed this technology to evaluate surgical outcomes (5-7).

As 3D facial scanning continues to gain popularity in aesthetic and craniofacial surgery, surgeons and patients begin to acknowledge the limitations of commercial 3D scanning devices. First, although the portability of 3D facial scanners improved in recent years and some portable devices have been validated to have high reliability, they are designed for surgeons but not patients and they require familiarity with capture technology (8). Patients must come to the clinics and seek help from professionals for 3D scanning. Second, commercial 3D facial scanners are expensive, and they are sold together with additional replaceable analytical software. For these reasons, there is an urgent need to design a user-friendly, inexpensive facial scanning system that allows patients to take 3D images by themselves.

To address this problem, the authors developed a novel iPad/iPhone application for 3D facial scanning. It allows patients to take 3D images of themselves on their own. The application, if it works well, will benefit both patients and doctors. Patient compliance for follow-up is expected to improve as they do not need to travel to the hospital, and doctors can better monitor their surgical outcomes to make proper decisions.

A novel device should be validated before its clinical application. Several 3D imaging systems are currently in use, and they have been validated to be accurate. Previous studies showed that the 3dMDface digital photogrammetry system (3dMD Inc., Atlanta, GA, USA) was valid and reliable and produced anthropometric measurements with a high degree of agreement with direct measurement and the Genex 3D imaging system (Genex Technologies, Inc. Kensington, MD, USA) (9-12). The Vectra imaging system (Canfield Scientific, Inc., NJ, USA) was also validated to be accurate and reproducible in normal adults and patients with various diseases (13-15). This study aimed to examine the validity and reproducibility of this novel imaging system. We present the following article in accordance with the MDAR reporting checklist (available at https://dx.doi.org/10.21037/atm-21-1620).

Methods

Study sample

Twenty young volunteers aging from 18 to 30 without any kind of craniofacial deformity were enrolled in this study. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). Ethical approval was provided by the Institutional Review Board of Peking Union Medical College Hospital (No. ZS-1243). All volunteers signed informed consent and agreed on their anthropometric data to be used for analysis.

3D imaging system

The authors collaboratively developed a 3D imaging application called "MeiXuan" that works on iPad/iPhone (Apple Inc., Cupertino, CA, USA) that supports FaceID technology. The customized software application was developed based on an open-source development kit (16). Some modification was made to the algorithm to improve imaging accuracy. Briefly, the iPad/iPhone camera collects point cloud data, and the software converts the point cloud to a mesh. The color and texture are further integrated into the mesh to complete a 3D image. A user-friendly interface with a voice prompt was designed to facilitate accurate scanning.

When taking a 3D image, the user holds the iPad/iPhone by hand and follows the voice prompt to rotate his/her head (*Figure 1A,B,C*). The system provides feedback on whether the capturing is successful in seconds. Any improper operation could result in low-quality 3D virtual models that do not support further photogrammetry. In this study, all 3D images were captured in the same room. The volunteers set right under an electronic light to guarantee there was no deep shadow on the face. In real practice, the user will be reminded by the application that the 3D scanning should be done in a bright room and that deep shadow on the face should be avoided.

Data collection

Eighteen anthropometric landmarks were used in this study (*Figure 2*) (17). Twenty-one anthropometric parameters
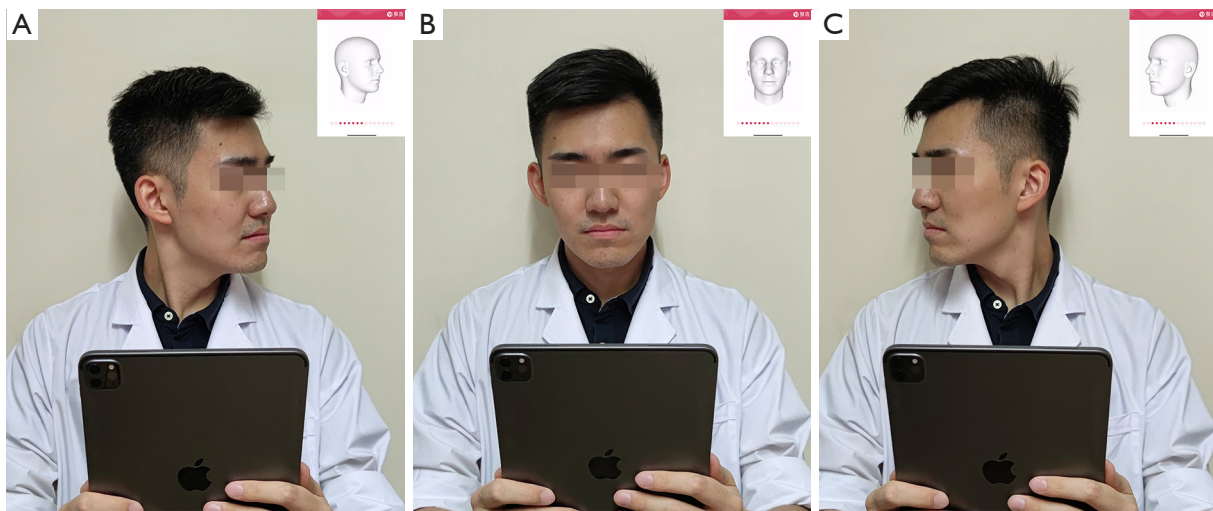
Figure 1 The process of 3D image capturing. The volunteer followed the voice prompt to turn his head to the left (A), turn his head to look forward (B), and finally turn his head to the right (C). The interface of the application at each step is shown in the top right corner.
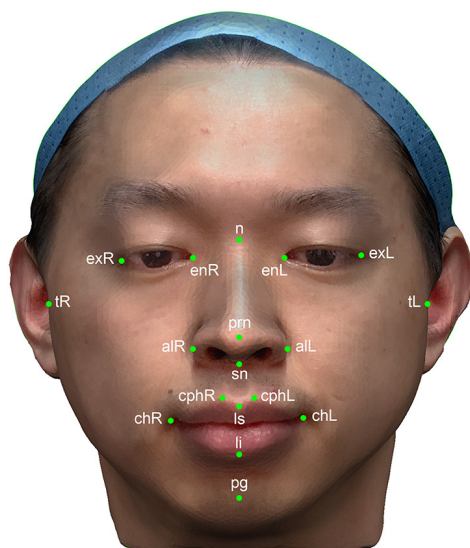
Figure 2 Illustration of facial anthropometric landmarks used in this study. n, soft tissue nasal; enR, right endocanthion; enL, left endocanthion; exR, right exocanthion; exL, left exocanthion; tR, right tragion; tL, left tragion; prn, pronasale; alR, right alare; alL, left alare; sn, subnasale; cphR, right crista philtra; cphL, left crista philtra; ls, labiale superius; chR, right chelion; chL, left chelion; li, labrale inferius; pg, pogonion.

were generated from these landmarks and categorized into three groups. The six anthropometric parameters in the periorbital region included the right ocular width (en-exR),

left ocular width (en-exL), intercanthal width (en-en), outer canthal width (ex-ex), right orbito-tragial depth (t-exR), and left orbito-tragial depth (t-exL). The nine anthropometric parameters in the nasal region included the nose dorsum length (n-prn), nose height (n-sn), distance between the nasion and right alare (n-alR), distance between the nasion and left alare (n-alL), nose width (al-al), right upper third facial depth (t-nR), left upper third facial depth (t-nL), right maxillary depth (t-snR), and left maxillary depth (t-snL). The six anthropometric parameters in the orolabial region included the cutaneous upper lip height (sn-ls), vermillion height (ls-li), philtrum width (cph-cph), mouth width (ch-ch), right lower third face depth (t-pgR) and left lower third face depth (t-pgL). All parameters were measured both on volunteers and on 3D virtual models.

The 3D virtual models could be accessed from the background database and exported in OBJ file format. They were further imported into Geomagic Wrap 2017 (Geomagic, Inc., Research Triangle Park, NC, USA) for measurement. The 3D virtual models were measured twice independently. The first author (Y Chong) manually identified the 18 landmarks on each 3D virtual model and took linear measurements between landmarks. The landmarks were then removed for the next set of 3D photogrammetry. Another author (X Liu) followed the same protocol and measured all parameters again.

Direct measurement served as the gold standard. In this study, a digital caliper was used to measure the distance between two landmarks. Measurements were recorded

Page 4 of 14

Chong et al. Validation of a novel 3D imaging system

to the nearest 0.01 mm. One author (Y Chong) took the first set of measurements of the 20 volunteers. Another author (X Liu) took the second set of measurements. Two measurements on the same volunteer were no less than an hour apart.

Comparison with Vectra H1 imaging system

The 3D virtual models of eight volunteers generated by this novel iPad/iPhone imaging system were further compared with their 3D models obtained by Vectra H1 camera, a common handheld 3D facial scanner that has been validated by previous studies (13,14). Registration of the 3D models by two imaging systems of the same volunteer was performed using best-fit alignment. The best-fit alignment command moves one 3D model to share physical space with another via a two-phase best-fit algorithm. Then, the distance between two 3D virtual models was visualized by generating a color-coded heat map. Average distance, root mean square (RMS), and the average maximum distance were recorded.

Statistical analyses

Statistical analyses were performed using SPSS software version 23.0 (IBM, New York, USA). The differences between the measurements of the two methods were expressed as mean absolute difference (MAD) and relative error measurement (REM).

Validity, depicting how close 3D photogrammetric results are with direct measurements, could be defined as accuracy and bias (4). For accuracy, we compared the average of two sets of 3D photogrammetric measurements with the average of two sets of direct measurements. The Shapiro-Wilk W test confirmed data normality. The Pearson's product-moment correlation coefficient (PPMCC) was calculated. For bias, a paired t-test was used to assess whether the measurements of the two methods had statistically significant differences. The Bland-Altman analysis was further performed to assess the agreement between the 3D photogrammetric and direct measurements. According to Othman *et al.*, Bland-Altman 95% limits of agreement over 2 mm were regarded as clinically unacceptable (15).

Reproducibility indicates whether repeated measurements yielded consistent results. For intraobserver reliability, five 3D virtual models were randomly selected. One author measured them again. The two sets of measurements of the five photographs from the same

author were 1 week apart to avoid memory bias. The intraclass correlation coefficient (ICC) was used to evaluate the agreement between two sets of measurements. For interobserver reliability, another five 3D virtual models were randomly selected and measured again by a different author. ICC was used to compare the 3D photogrammetric measurements of two independent observers. An ICC value over 0.8 suggested good reliability.

To reduce type I error, the level of statistical significance was set at 0.0024 (0.05/21) after Bonferroni correction.

Results

Accuracy

The results of the direct and 3D photogrammetric measurements are presented in *Table 1*. The PPMCC (r) was higher than 0.90 for all parameters, showing that the 3D measurement was positively correlated with the direct measurement. The MADs were less than 1 mm in 15 of 21 parameters. No parameter had a MAD over 1.50 mm. The REMs ranged from 0.85% to 3.25%.

Bias

In the periorbital region, statistically significant difference of the two methods were shown on ex-ex (t=–4.935, P<0.001). Bland-Altman analysis manifested good correlation between the two methods on en-exR, en-exL, en-en, and ex-ex, while 95% limits of agreement exceeded 2 mm for t-exR (95% limits of agreement: –2.24 to 2.20), and t-exL (95% limits of agreement: –2.82 to 2.12) (*Figure 3*). In the nasal region, statistically significant difference of the two methods were shown on t-nR (t=–4.876, P<0.001), t-snR (t=–3.824, P=0.001), and t-snL (t=–8.287, P<0.001). Bland-Altman analysis manifested good correlation between the two methods on n-prn, al-al, and n-alL, while 95% limits of agreement exceeded 2 mm for n-sn (95% limits of agreement: –2.51 to 1.22), n-alR (95% limits of agreement: –2.14 to 1.03), t-nR (95% limits of agreement: –3.10 to 0.88), t-nL (95% limits of agreement: –2.10 to 2.22), t-snR (95% limits of agreement: –3.59 to 1.41), and t-snL (95% limits of agreement: –3.03 to 0.08) (*Figure 4*). In the orolabial region, two methods showed no statistically significant difference on the six parameters. Bland-Altman analysis manifested good correlation between the two methods on sn-ls, cph-cph, ch-ch, and ls-li, while 95% limits of agreement exceeded 2 mm for t-pgR (95% limits

Table 1 Accuracy and bias of the novel imaging system

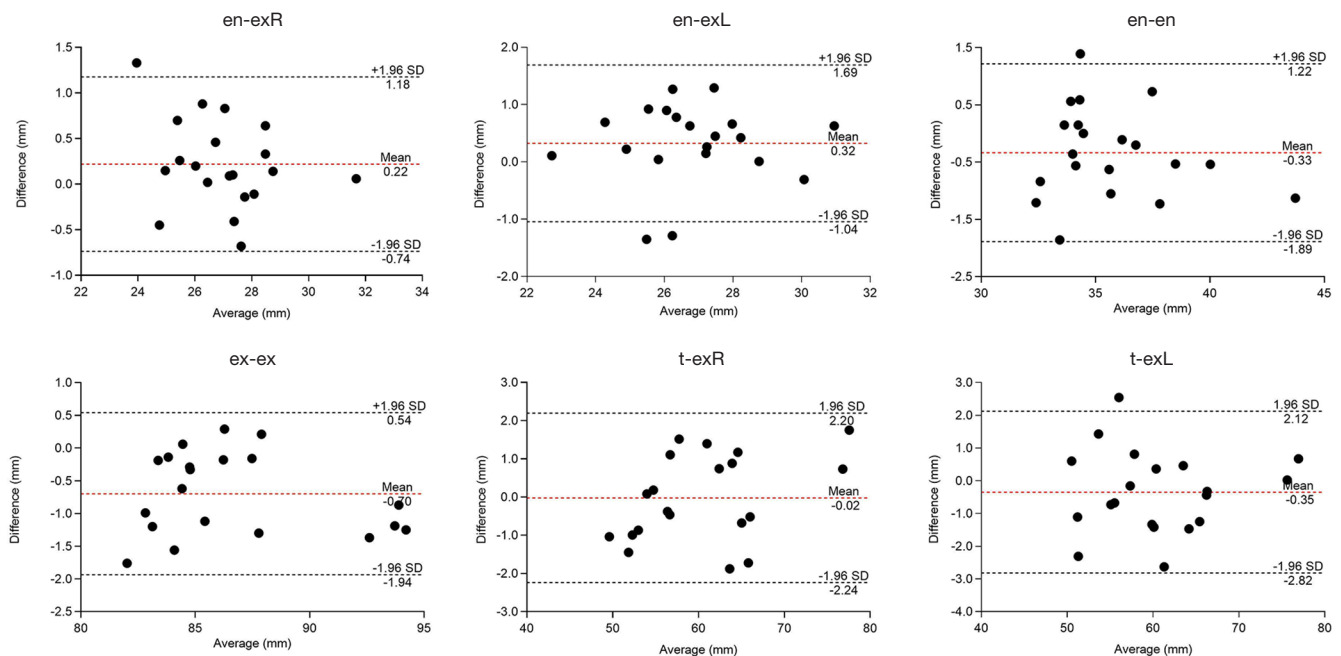| Parameters	Direct measurement		3D photogrammetry		MAD	REM (%)	PPMCC	Paired t-test	
	Mean	SD	Mean	SD				t	P value
Periorbital region									
en-exR	27.10	1.68	26.88	1.82	0.40	1.52	0.964	2.022	0.057
en-exL	26.95	1.95	26.63	1.92	0.62	2.33	0.935	2.081	0.051
en-en	35.49	2.74	35.83	2.86	0.69	1.95	0.961	−1.888	0.074
ex-ex	86.31	3.87	87.01	3.99	0.75	0.87	0.987	−4.935	<0.001*
t-exR	60.48	7.93	60.50	7.50	0.98	1.61	0.991	−0.084	0.934
t-exL	60.25	7.38	60.59	7.34	1.04	1.79	0.985	−1.234	0.232
Nasal region									
n-prn	36.60	4.27	36.67	4.31	0.69	1.95	0.982	−0.356	0.726
n-sn	43.44	4.53	44.08	4.52	0.88	2.03	0.978	−3.035	0.007
al-al	29.59	2.81	29.58	2.82	0.41	1.37	0.985	0.051	0.96
n-alR	38.98	4.53	39.53	4.45	0.75	1.98	0.984	−3.047	0.007
n-alL	39.70	4.57	39.77	4.46	0.64	1.67	0.983	−0.372	0.714
t-nR	99.46	7.70	100.56	7.18	1.27	2.05	0.993	−4.876	<0.001*
t-nL	99.82	7.19	99.76	7.02	0.81	1.75	0.988	0.243	0.811
t-snR	105.54	7.62	106.62	7.03	1.41	1.35	0.988	−3.824	0.001*
t-snL	104.82	7.02	106.30	6.94	1.47	1.41	0.994	−8.287	<0.001*
Orolabial region									
sn-ls	13.05	2.23	12.90	2.05	0.44	3.25	0.950	0.968	0.345
cph-cph	10.70	1.80	10.81	1.67	0.33	3.19	0.974	−1.202	0.244
ch-ch	44.64	4.71	44.69	4.85	0.44	1.00	0.995	−0.449	0.658
ls-li	14.19	1.45	13.97	1.57	0.35	2.55	0.975	2.737	0.013
t-pgR	121.46	9.06	121.61	8.59	1.04	0.85	0.991	−0.509	0.616
t-pgL	120.96	9.18	120.94	8.40	1.40	1.14	0.985	0.047	0.963

*, paired t-test reported statistical significance after Bonferroni correction. Mean, SD, and MAD are recorded in millimeters. SD, standard deviation; MAD, mean absolute difference; REM, relative error measurement; PPMCC, Pearson's product-moment correlation coefficient; en-exR, right ocular width; en-exL, left ocular width; en-en, intercanthal width; ex-ex, outer canthal width; t-exR, right orbito-tragial depth; t-exL, left orbito-tragial depth; n-prn, nose dorsum length; n-sn, nose height; al-al, nose width; n-alR, distance between the nasion and right alare; n-alL, distance between the nasion and left alare; t-nR, right upper third facial depth; t-nL, left upper third facial depth; t-snR, right maxillary depth; t-snL, left maxillary depth; sn-ls, cutaneous upper lip height; cph-cph, philtrum width; ch-ch, mouth width; ls-li, vermillion height; t-pgR, right lower third face depth; t-pgL, left lower third face depth.

of agreement: −2.66 to 2.36) and t-pgL (95% limits of agreement: −3.37 to 3.41) (*Figure 5*).

Reproducibility

The results of intraobserver reliability are shown in *Table 2*.

All parameters were in good agreement between the two sets of measurements. The results of the interobserver reliability are shown in *Table 3*. The two independent observers reported consistent results on all parameters except for n-sn. Excellent intra- and interobserver reliabilities demonstrated good reproducibility of the novel

Page 6 of 14

Chong et al. Validation of a novel 3D imaging system

Figure 3 Bland-Altman analysis of parameters in the periorbital region. Bland-Altman plots of en-exR, en-exL, en-en, ex-ex, t-exR, and t-exL. The X axis is the average of direct and 3D photogrammetric measurements. The Y axis is the difference of these two methods. en-exR, right ocular width; en-exL, left ocular width; en-en, intercanthal width; ex-ex, outer canthal width; t-exR, right orbito-tragial depth; t-exL, left orbito-tragial depth.

3D imaging system.

Heat map analysis

The results of heat map analysis following superimposition of the 3D virtual models obtained by this novel iPad/iPhone facial scanner and Vectra H1 imaging system of each volunteer are shown in *Table 4* and *Figure 6*. The mean RMS of all 3D model pairs was 0.67 mm. The mean distance of all 3D model pairs was 0.08 mm. The average positive and negative maximum distances of all 3D model pairs were 3.96 and –3.83 mm, respectively.

Discussion

It has been demonstrated that handheld 3D imaging systems are becoming increasingly more popular (18). The authors have been working on the clinical application of 3D anthropometry and agreed that the portability of a 3D imaging system brought great convenience to both clinicians and patients (19,20). The authors worked with engineers to develop a novel iPad/iPhone application for

3D photography. The imaging system was proven by this study to be accurate and reliable for clinical use. Our newly developed application allows patients to capture 3D images of themselves on their own, which provides patients with great convenience and less chance of hospital-acquired infection.

Importance of validation study

A validity study is important before the clinical application of any novel device. Specific to a 3D imaging system, a device should be tested for its accuracy, bias, and reproducibility with direct measurement, which is often regarded as the gold standard. The handheld Vectra H1 device is one of the most common portable imaging systems and has been comprehensively validated for its accuracy and reliability. Camison *et al.* measured 136 linear distances with both portable Vectra H1 and 3dMDface systems (13). Most of the differences were within a 1-mm threshold, which indicated that the Vectra H1 is accurate enough for clinical use. Gibelli *et al.* further compared the portable Vectra H1 imaging system with the static Vectra M3 system
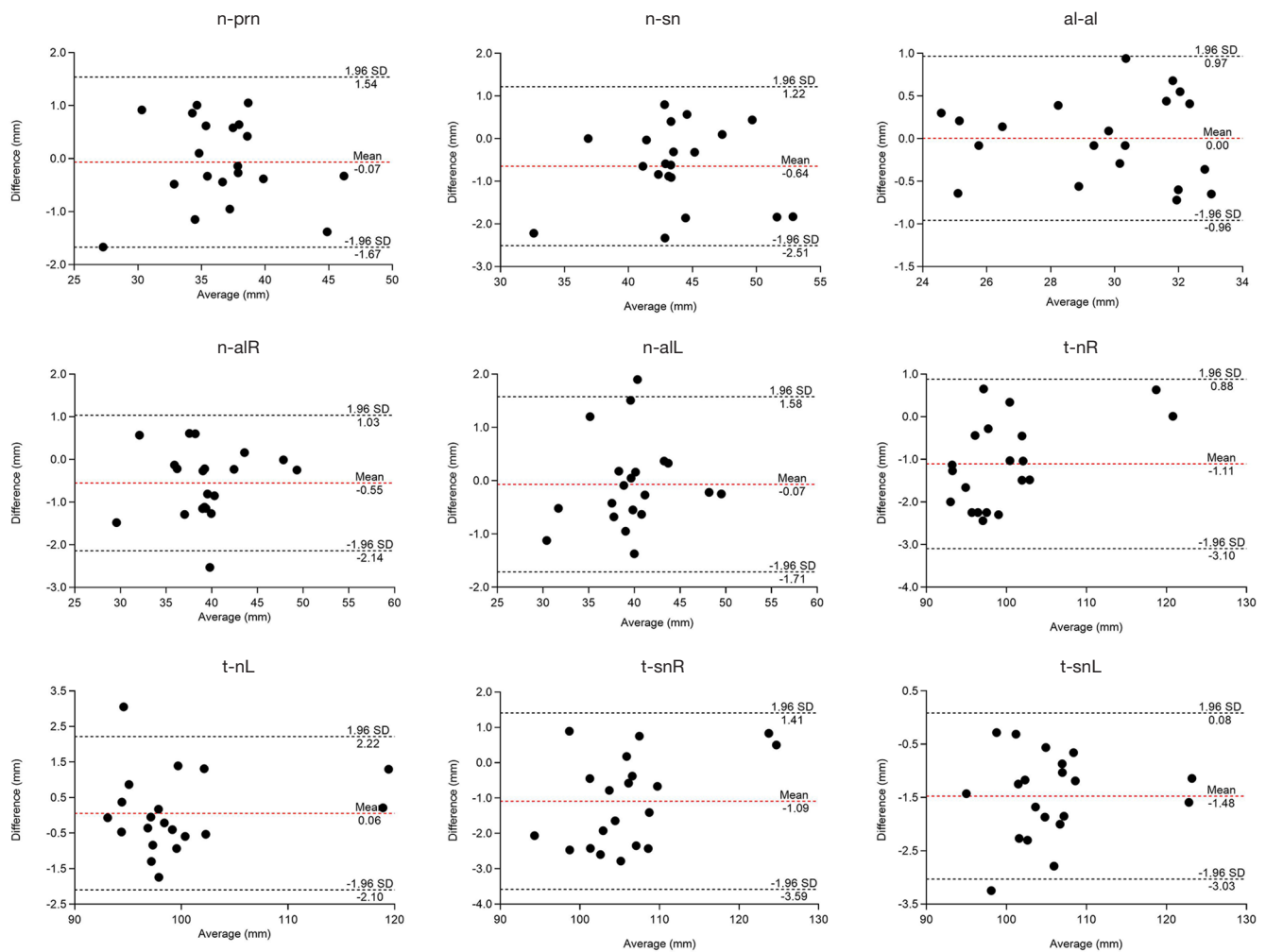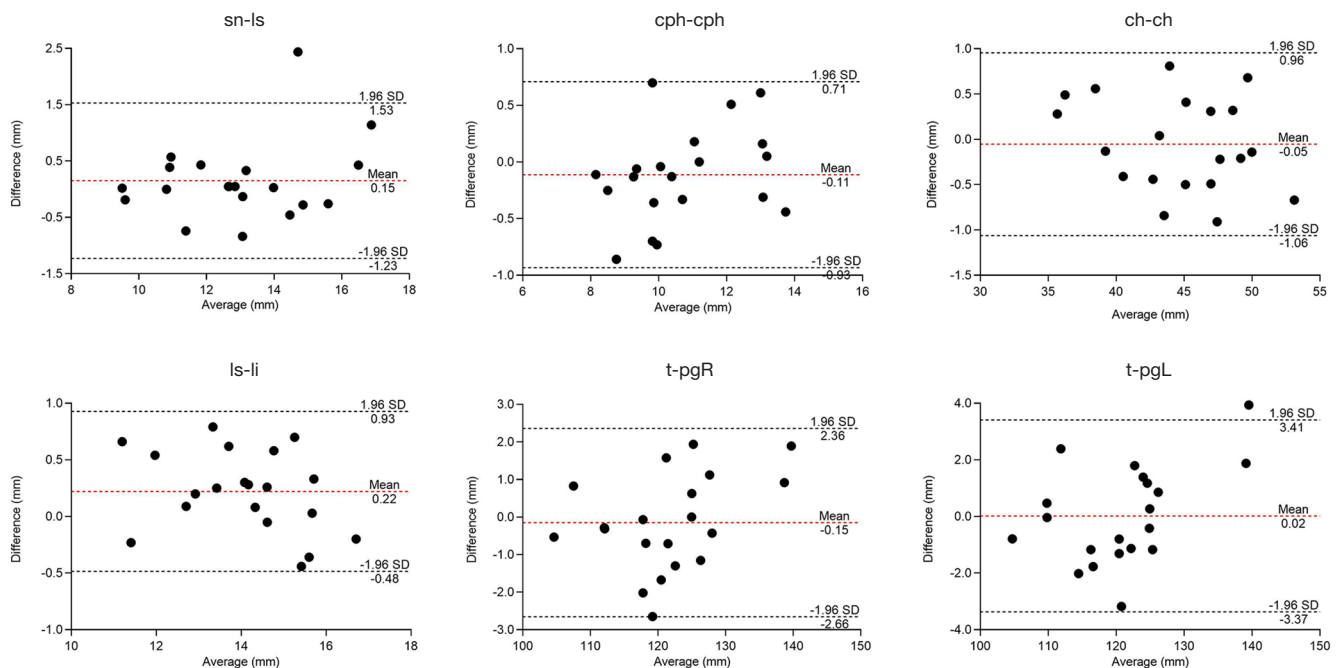
Figure 4 Bland-Altman analysis of parameters in the nasal region. Bland-Altman plots of n-prn, n-sn, al-al, n-alR, n-alL, t-nR, t-nL, t-snR, and t-snL. The X axis is the average of direct and 3D photogrammetric measurements. The Y axis is the difference of these two methods. n-prn, nose dorsum length; n-sn, nose height; al-al, nose width; n-alR, distance between the nasion and right alare; n-alL, distance between the nasion and left alare; t-nR, right upper third facial depth; t-nL, left upper third facial depth; t-snR, right maxillary depth; t-snL, left maxillary depth.

from the same company in a larger population and found that the portable device was reliable in the measurement of linear distances, angles, and surface area (14). Koban *et al.* investigated the validation of two other handheld devices [Sense 3D Scanner (3D Systems Inc., Rock Hill, SC, USA) and Artec Eva (Artec 3D Inc., Luxembourg, USA)] against a nonportable 3D imaging system (Vectra XT) and found that only one portable device could yield accurate facial scanning images (21). Apart from normal adults, a 3D imaging system has also been verified for its reliability in patients with craniofacial conditions (10,15,22-24), on cadaver heads (25),

and on mannequin heads (26).

Accuracy comparison with other 3D imaging systems

In this study, we measured several anthropometric parameters in the periorbital, nasal, and orolabial regions. The parameters included horizontal distances, vertical distances, and facial depth parameters to comprehensively evaluate 3D virtual models in horizontal, vertical, and sagittal dimensions. In the periorbital region, the MADs of en-en, ex-ex, and en-ex were all less than 1 mm. The MAD

Page 8 of 14

Chong et al. Validation of a novel 3D imaging system

Figure 5 Bland-Altman analysis of parameters in the orolabial region. Bland-Altman plots of sn-ls, cph-cph, ch-ch, ls-li, t-pgR, and t-pgL. The X axis is the average of direct and 3D photogrammetric measurements. The Y axis is the difference of these two methods. sn-ls, cutaneous upper lip height; cph-cph, philtrum width; ch-ch, mouth width; ls-li, vermillion height; t-pgR, right lower third face depth; t-pgL, left lower third face depth.

of t-exL was slightly greater than 1 mm (MAD =1.04 mm, REM =1.79%). Bland-Altman analysis indicated that t-exR and t-exL had 95% limits of agreement over 2 mm. However, considering that the REMs of these parameters were small (1.61%, and 1.79%, respectively) and the paired *t*-test did not show statistical significance, the difference between direct measurement and photogrammetric measurement could be regarded as acceptable.

In the nasal region, the MADs of n-prn, n-sn, al-al, n-alR, and n-alL did not exceed 1 mm, and the paired *t*-test did not show statistical significance between the two methods. It has been concluded that measurements on 3D virtual models should undergo an adjustment of around 1.35 mm (15,27). The MADs of t-snR and t-snL were 1.41 and 1.47 mm, respectively, slightly over the threshold value. The paired *t*-test also confirmed statistical significance between the two methods. Bland-Altman analysis further showed that their 95% limits of agreement were –3.59 to 1.41 mm and –3.03 to 0.08 mm, respectively. However, the REMs of the two parameters were just around 1.4%, indicating that the mean difference took up a very small proportion of the original value; thus, no adjustment was needed. Notably, the 95%

limit of agreement for n-sn was –2.51 to 1.22 mm. The lower bound exceeded 2 mm and was higher than the value reported by Othman *et al.* (–1.95 mm) (15) and Dindaroğlu *et al.* (–1.16 mm) (11). Three parameters related to the subnasale showed clinically unacceptable discrepancies, possibly because the nose bridge partially blocks the light shining on the subnasale, which may affect the digital identification of the subnasale landmark. To solve a similar problem, some studies advocated for facial landmark labeling ahead of image capture to improve accuracy (13,15), while other reports demonstrated that marking landmarks before imaging did not influence the accuracy and precision of the 3D imaging system (28). For our imaging system, placement of landmarks before imaging is not practical as the patients will take photographs on their own without help from the professionals. Patients should not be asked to place landmarks on their faces on their own. Bias on the placement of landmarks should be a systematic error. This is the reason why we did not draw landmarks on the volunteers before imaging in this validation study.

In the orolabial region, all linear distances expected for facial depth parameters had small MADs and discrepancies

Table 2 Results of intraobserver reliability

Parameters	MAD	REM (%)	ICC			P value
			ICC values	95% CI		
				Lower	Upper	
Periorbital region						
en-exR	0.59	2.19	0.907	0.360	0.990	0.006
en-exL	0.36	1.35	0.855	0.143	0.984	0.015
en-en	0.79	2.17	0.979	0.817	0.998	<0.001
ex-ex	0.69	0.77	0.989	0.901	0.999	<0.001
t-exR	1.51	2.46	0.985	0.866	0.998	<0.001
t-exL	1.96	3.19	0.954	0.633	0.995	0.002
Nasal region						
n-prn	1.44	3.72	0.963	0.693	0.996	0.001
n-sn	0.95	2.01	0.965	0.710	0.996	0.001
al-al	0.70	2.43	0.916	0.409	0.991	0.005
n-alR	1.01	2.40	0.959	0.665	0.996	0.001
n-alL	1.22	2.89	0.939	0.539	0.994	0.003
t-nR	1.14	1.07	0.988	0.888	0.999	<0.001
t-nL	1.21	1.21	0.977	0.801	0.998	<0.001
t-snR	0.76	0.72	0.995	0.952	0.999	<0.001
t-snL	1.12	1.08	0.985	0.867	0.998	<0.001
Orolabial region						
sn-ls	0.48	3.53	0.949	0.598	0.995	0.002
cph-cph	0.54	4.75	0.923	0.445	0.992	0.004
ch-ch	1.73	3.92	0.943	0.558	0.994	0.002
ls-li	0.79	6.20	0.963	0.695	0.996	0.001
t-pgR	1.51	1.20	0.980	0.824	0.998	<0.001
t-pgL	2.15	1.81	0.974	0.774	0.997	0.001

MAD is recorded in millimeters. MAD, mean absolute difference; REM, relative error measurement; ICC, intraclass correlation coefficient; CI, confidence interval; en-exR, right ocular width; en-exL, left ocular width; en-en, intercanthal width; ex-ex, outer canthal width; t-exR, right orbito-tragial depth; t-exL, left orbito-tragial depth; n-prn, nose dorsum length; n-sn, nose height; al-al, nose width; n-alR, distance between the nasion and right alare; n-alL, distance between the nasion and left alare; t-nR, right upper third facial depth; t-nL, left upper third facial depth; t-snR, right maxillary depth; t-snL, left maxillary depth; sn-ls, cutaneous upper lip height; cph-cph, philtrum width; ch-ch, mouth width; ls-li, vermillion height; t-pgR, right lower third face depth; t-pgL, left lower third face depth.

less than 2 mm according to Bland-Altman analysis, similar to what previous studies have found (10,15). The discrepancies of t-pgR and t-pgL were larger than 2 mm, although the paired *t*-test did not report statistical significance. In the real world, direct measurement of the lower one-third of

the face is not easy because slight facial expression changes cannot be completely avoided (15). Spending too long on measurement will only increase facial movement and bring more discomfort to the volunteers. As a result, the authors only measured six parameters in the orolabial region to

Page 10 of 14

Chong et al. Validation of a novel 3D imaging system

Table 3 Results of interobserver reliability

Parameters	MAD	REM (%)	ICC			P value
			ICC values	95% CI		
				Lower	Upper	
Periorbital region						
en-exR	0.30	1.09	0.994	0.943	0.999	<0.001
en-exL	0.33	1.26	0.988	0.892	0.999	<0.001
en-en	1.21	3.51	0.884	0.258	0.987	0.010
ex-ex	0.66	0.79	0.970	0.745	0.997	0.001
t-exR	0.94	1.61	0.997	0.974	1.000	<0.001
t-exL	1.45	2.43	0.960	0.670	0.996	0.001
Nasal region						
n-prn	0.91	2.54	0.841	0.093	0.982	0.018
n-sn	0.92	2.13	0.737	−0.185	0.969	0.047
al-al	0.98	3.33	0.910	0.378	0.990	0.006
n-alR	0.94	2.50	0.829	0.054	0.981	0.021
n-alL	0.60	1.59	0.944	0.569	0.994	0.002
t-nR	0.46	0.48	0.989	0.896	0.999	<0.001
t-nL	1.46	1.47	0.807	−0.013	0.978	0.026
t-snR	1.27	1.22	0.948	0.594	0.994	0.002
t-snL	0.95	0.91	0.918	0.420	0.991	0.005
Orolabial region						
sn-ls	0.41	3.20	0.959	0.667	0.996	0.001
cph-cph	0.67	5.71	0.899	0.323	0.989	0.007
ch-ch	1.65	3.64	0.864	0.178	0.985	0.013
ls-li	0.44	2.95	0.970	0.744	0.997	0.001
t-pgR	0.53	0.43	0.996	0.966	1.000	<0.001
t-pgL	1.34	1.13	0.969	0.741	0.997	0.001

MAD is recorded in millimeters. MAD, mean absolute difference; REM, relative error measurement; ICC, intraclass correlation coefficient; CI, confidence interval; en-exR, right ocular width; en-exL, left ocular width; en-en, intercanthal width; ex-ex, outer canthal width; t-exR, right orbito-tragial depth; t-exL, left orbito-tragial depth; n-prn, nose dorsum length; n-sn, nose height; al-al, nose width; n-alR, distance between the nasion and right alare; n-alL, distance between the nasion and left alare; t-nR, right upper third facial depth; t-nL, left upper third facial depth; t-snR, right maxillary depth; t-snL, left maxillary depth; sn-ls, cutaneous upper lip height; cph-cph, philtrum width; ch-ch, mouth width; ls-li, vermillion height; t-pgR, right lower third face depth; t-pgL, left lower third face depth.

minimize the error caused by soft tissue movement.

### *The emerging trend of 3D scanning on smartphones*

The value of low-cost 3D scanning technology on mobile devices such as smartphones and iPads has been recognized in recent years. A number of applications are now available on iPhone/iPads and Android phones. This increases the popularity and accessibility of 3D scanning. 3D facial models obtained by mobile devices can be used for facial

Table 4 Results of heat map analysis of 3D virtual models generated by the novel imaging system and Vectra H1 imaging system

Volunteer	Average distance	RMS	Maximum distance	
			Positive	Negative
1	0.11	0.66	3.09	−2.50
2	0.10	0.67	3.79	−2.67
3	0	0.64	3.61	−2.40
4	0.01	0.64	3.57	−6.45
5	0.09	0.70	6.03	−5.68
6	0.10	0.75	3.08	−3.01
7	0.13	0.70	4.79	−5.00
8	0.08	0.60	3.72	−2.92
Mean	0.08	0.67	3.96	−3.83

Average distances, RMS, and maximum distances are recorded in millimeter. RMS, root mean square.
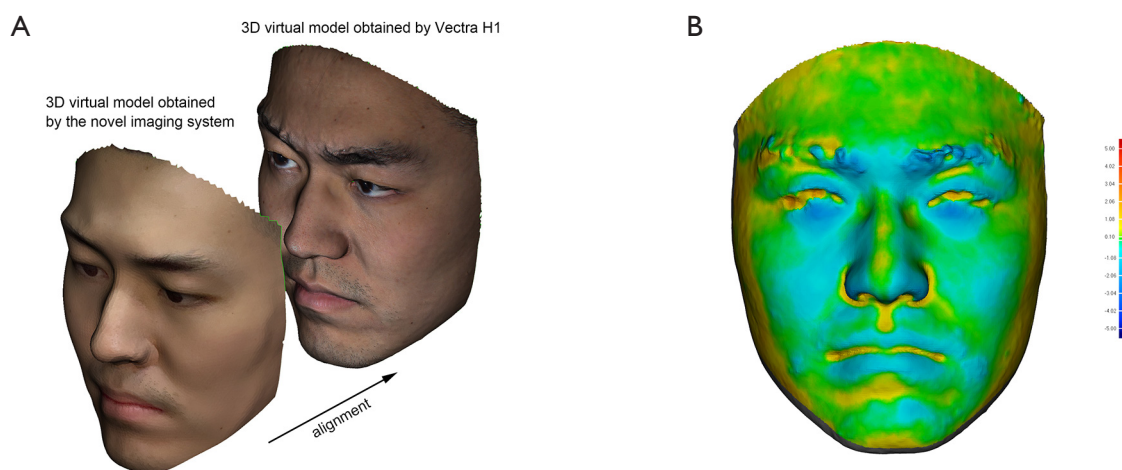
Figure 6 Heat map analysis. (A) Superimposition of 3D virtual models obtained by the novel imaging system and Vectra H1 camera via best-fit alignment; (B) results of heat map analysis of one volunteer. The color-coded bar on the right represents the deviation between two models.

recognition, augmented reality application scenarios, and clinical practice. A recent study by Rudy *et al.* tested the accuracy and precision of an application called ScandyPro on iPhone X for 3D facial scanning (8). The study compared 16 pairs of 3D virtual models captured by iPhone X and Vectra H1 and found that the average RMS between the two methods was 0.44±0.10 mm, which was comparable to the average RMS of the present study. Another study developed and validated a smartphone photogrammetry framework for 3D scanning of the external ear (29). The authors concluded that novice operators could rapidly learn

to do accurate and reliable 3D scans of the external ear after watching instructional videos.

Clinical relevance

3D facial scanning has been intensively used in the preoperative planning and postoperative evaluation of plastic and aesthetic surgery, oral and maxillofacial surgery, and other craniofacial surgery. It outperforms direct measurement and 2D photogrammetry because it provides accurate facial information in minutes and supports facial

Page 12 of 14

Chong et al. Validation of a novel 3D imaging system

depth measurement and volumetric analysis. 3D virtual models can be used for simulation of surgical effect, which allows surgeons and patients to see how surgery changes facial appearance intuitively. Our 3D scanning system has more value in postoperative evaluation. Traditional 3D scanning devices require patients to come to the clinics for 3D photography. However, patients who install our application on their mobile devices can take 3D images by themselves at home. This saves them much time and effort.

Additionally, our device can be further applied in the early diagnosis of diseases that affect facial appearance. For example, genetic disorders such as Turner syndrome and Down syndrome all have remarkable facial changes. In recent years, studies have applied facial recognition algorithms in the diagnosis of these diseases (30,31). Thanks to the technological advance in 3D facial recognition, diagnosis based on 3D virtual models has the potential to yield better accuracy. Our application can be bundled with analytical software that helps detect diseases that affect appearance and facial skin.

Limitations

There are some limitations to this study. First, although iPad/iPhones are common personal electronics, they are not accessible to every patient. Our solution is not available for every patient. Second, the image color and texture of the novel imaging system are not as precise as those of the commonly used 3D imaging devices such as Vectra H1 and 3dMDface. In the future, we will continue to work on algorithms and this novel application will be recommended to the patients in our center.

Conclusions

To reduce the cost of 3D facial scanning and enable patients to capture 3D images on their own, the authors developed a novel iPad/iPhone application for 3D imaging. This study comprehensively evaluated the validity and reproducibility of the novel imaging system. With good portability, high accessibility, and a user-friendly interface, the system shows fairly good accuracy for clinical use. Efforts will be made to improve the imaging quality of the subnasale region in the future.

Acknowledgments

Footnote

Reporting Checklist: The authors have completed the MDAR reporting checklist. Available at https://dx.doi.org/10.21037/atm-21-1620

Data Sharing Statement: Available at https://dx.doi.org/10.21037/atm-21-1620

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://dx.doi.org/10.21037/atm-21-1620). YC and XL collaborated with engineers from Beijing MeiXuan 3D Technology Co., Ltd. in the development of the MeiXuan application. The other authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). Ethical approval was provided by the Institutional Review Board of Peking Union Medical College Hospital (No. ZS-1243). All volunteers signed informed consent and agreed on their anthropometric data to be used for analysis.

References

1. Weinberg SM, Scott NM, Neiswanger K, et al. Digital three-dimensional photogrammetry: evaluation of anthropometric precision and accuracy using a Genex 3D camera system. Cleft Palate Craniofac J 2004;41:507-18.

2. Jacobs RA; Plastic Surgery Educational Foundation DATA Committee. Three-dimensional photography. Plast Reconstr Surg 2001;107:276-7.

3. Farkas LG, Bryson W, Klotz J. Is photogrammetry of the face reliable? Plast Reconstr Surg 1980;66:346-55.

4. Wong JY, Oh AK, Ohta E, et al. Validity and reliability of craniofacial anthropometric measurement of 3D digital photogrammetric images. Cleft Palate Craniofac J 2008;45:232-9.

5. Meier JD, Glasgold RA, Glasgold MJ. 3D photography in the objective analysis of volume augmentation including fat augmentation and dermal fillers. Facial Plast Surg Clin North Am 2011;19:725-35, ix.

6. Miller TR. Long-term 3-Dimensional Volume Assessment After Fat Repositioning Lower Blepharoplasty. JAMA Facial Plast Surg 2016;18:108-13.

7. Jacono AA, Bryant LM, Ahmedli NN. A Novel Extended Deep Plane Facelift Technique for Jawline Rejuvenation and Volumization. Aesthet Surg J 2019;39:1265-81.

8. Rudy HL, Wake N, Yee J, et al. Three-Dimensional Facial Scanning at the Fingertips of Patients and Surgeons: Accuracy and Precision Testing of iPhone X Three-Dimensional Scanner. Plast Reconstr Surg 2020;146:1407-17.

9. Aldridge K, Boyadjiev SA, Capone GT, et al. Precision and error of three-dimensional phenotypic measures acquired from 3dMD photogrammetric images. Am J Med Genet A 2005;138A:247-53.

10. Heike CL, Cunningham ML, Hing AV, et al. Picture perfect? Reliability of craniofacial anthropometry using three-dimensional digital stereophotogrammetry. Plast Reconstr Surg 2009;124:1261-72.

11. Dindaroğlu F, Kutlu P, Duran GS, et al. Accuracy and reliability of 3D stereophotogrammetry: A comparison to direct anthropometry and 2D photogrammetry. Angle Orthod 2016;86:487-94.

12. Weinberg SM, Naidoo S, Govier DP, et al. Anthropometric precision and accuracy of digital three-dimensional photogrammetry: comparing the Genex and 3dMD imaging systems with one another and with direct anthropometry. J Craniofac Surg 2006;17:477-83.

13. Camison L, Bykowski M, Lee WW, et al. Validation of the Vectra H1 portable three-dimensional photogrammetry system for facial imaging. Int J Oral Maxillofac Surg 2018;47:403-10.

14. Gibelli D, Pucciarelli V, Cappella A, et al. Are Portable Stereophotogrammetric Devices Reliable in Facial Imaging? A Validation Study of VECTRA H1 Device. J Oral Maxillofac Surg 2018;76:1772-84.

15. 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 Investig 2020;24:2853-66.

16. ARKit. Available online: https://developer.apple.com/documentation/arkit

17. Farkas LG. Anthropometry of the Head and Face. 2nd edition. New York, NY, USA: Raven Press, 1994.

18. Knoops PG, Beaumont CA, Borghi A, et al. Comparison of three-dimensional scanner systems for craniomaxillofacial imaging. J Plast Reconstr Aesthet Surg 2017;70:441-9.

19. Chong Y, Dong R, Liu X, et al. Stereophotogrammetry to reveal age-related changes of labial morphology among Chinese women aging from 20 to 60. Skin Res Technol 2021;27:41-8.

20. Chong Y, Li J, Liu X, et al. Three-dimensional anthropometric analysis of eyelid aging among Chinese women. J Plast Reconstr Aesthet Surg 2021;74:135-42.

21. Koban KC, Perko P, Etzel L, et al. Validation of two handheld devices against a non-portable three-dimensional surface scanner and assessment of potential use for intraoperative facial imaging. J Plast Reconstr Aesthet Surg 2020;73:141-8.

22. Ort R, Metzler P, Kruse AL, et al. The Reliability of a Three-Dimensional Photo System- (3dMDface-) Based Evaluation of the Face in Cleft Lip Infants. Plast Surg Int 2012;2012:138090.

23. Tse R, Booth L, Keys K, et al. Reliability of nasolabial anthropometric measures using three-dimensional stereophotogrammetry in infants with unrepaired unilateral cleft lip. Plast Reconstr Surg 2014;133:530e-542e.

24. Zhao YJ, Xiong YX, Wang Y. Three-Dimensional Accuracy of Facial Scan for Facial Deformities in Clinics: A New Evaluation Method for Facial Scanner Accuracy. PLoS One 2017;12:e0169402.

25. Fourie Z, Damstra J, Gerrits PO, et al. Evaluation of anthropometric accuracy and reliability using different three-dimensional scanning systems. Forensic Sci Int 2011;207:127-34.

26. Metzler P, Sun Y, Zemann W, et al. Validity of the 3D VECTRA photogrammetric surface imaging system for cranio-maxillofacial anthropometric measurements. Oral Maxillofac Surg 2014;18:297-304.

27. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;1:307-10.

28. Aynechi N, Larson BE, Leon-Salazar V, et al. Accuracy and precision of a 3D anthropometric facial analysis with and without landmark labeling before image acquisition. Angle Orthod 2011;81:245-52.
29. Nightingale RC, Ross MT, Cruz RLJ, et al. Frugal 3D scanning using smartphones provides an accessible framework for capturing the external ear. J Plast Reconstr Aesthet Surg 2021. [Epub ahead of print]. doi: 10.1016/j.bjps.2021.03.131.
30. Pan Z, Shen Z, Zhu H, et al. Clinical application of an automatic facial recognition system based on deep learning for diagnosis of Turner syndrome. Endocrine 2021;72:865-73.
31. Qin B, Liang L, Wu J, et al. Automatic Identification of Down Syndrome Using Facial Images with Deep Convolutional Neural Network. Diagnostics (Basel) 2020;10:487.