UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

A clinimetric assessment of the validity and reliability of 3D technology for scar surface area measurement

Doomen, M.C.H.A.; Rijpma, D.; Pijpe, A.; Meij-de Vries, A.; Niessen, F.B.; Karaoglu, S.; de Vet, H.C.W.; Gevers, T.; van Zuijlen, P.P.M.

DOI
10.1016/j.burns.2022.12.008

Publication date
2023

Document Version
Final published version

Published in
Burns

License
Article 25fa Dutch Copyright Act (https://www.openaccess.nl/en/policies/open-access-in-dutch-copyright-law-taverne-amendment)

Link to publication

Citation for published version (APA):
Doomen, M. C. H. A., Rijpma, D., Pijpe, A., Meij-de Vries, A., Niessen, F. B., Karaoglu, S., de Vet, H. C. W., Gevers, T., & van Zuijlen, P. P. M. (2023). A clinimetric assessment of the validity and reliability of 3D technology for scar surface area measurement. *Burns*, *49*(3), 583-594. https://doi.org/10.1016/j.burns.2022.12.008

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)

Download date:14 Nov 2026

BURNS 49 (2023) 583–594

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/burns

A clinimetric assessment of the validity and reliability of 3D technology for scar surface area measurement

M.C.H.A. Doomen [a,b,c,i,1], D. Rijpma [a,c,i,*,1], A. Pijpe [a,c,i], A. Meij-de Vries [a,d,e], F.B. Niessen [c], S. Karaoglu [f], H.C.W. de Vet [g], T. Gevers [f], P.P.M. van Zuijlen [a,c,e,h,i]

[a] Burn Center, Red Cross Hospital, Beverwijk, the Netherlands
[b] Association of Dutch Burn Centers, Beverwijk 1941 AJ, the Netherlands
[c] Amsterdam UMC location Vrije Universiteit Amsterdam, Department of Plastic Reconstructive and Hand Surgery, De Boelelaan 1117, Amsterdam, Netherlands
[d] Department of Surgery, Red Cross Hospital, Beverwijk, the Netherlands
[e] Amsterdam UMC location University of Amsterdam, Paediatric Surgical Centre, Emma Children's Hospital, Meibergdreef 9, Amsterdam, Netherlands
[f] 3DUniversum, 1098 XH Amsterdam, the Netherlands
[g] Amsterdam UMC location Vrije Universiteit Amsterdam, Epidemiology and Data Science, De Boelelaan1117, Amsterdam, the Netherlands
[h] Department of Plastic, Reconstructive & Hand Surgery, Red Cross Hospital, Beverwijk, the Netherlands
[i] Amsterdam Movement Sciences, Tissue Function and Regeneration, Amsterdam, the Netherlands

ARTICLE INFO

Article history:
Accepted 19 December 2022

Keywords:
Scars
Surface area
Measurement instrument
(3D) depth sensor
Clinimetric evaluation

ABSTRACT

Introduction: The quality of scars has become an important outcome of burn care. Objective scar assessment through scar surface area measurement enables quantification of scar formation and evaluation of treatment efficacy. 3D technology has proven valid and reliable but often remains cumbersome, expensive, and time-consuming. 3D technology with depth sensors on mobile devices has become available and might surpass these limitations. This study provides a clinimetric assessment of the validity and reliability of a 3D system with a depth sensor for scar surface area measurement.

Methods: A technology involving a depth sensor mounted on a mobile device was used. Images and analyses were made with a custom-made software application. A standardized one-keyframe image capturing procedure was followed. To assess validity, stickers with predefined dimensions (8.01 cm^2 – 77.70 cm^2) were imaged in a single observer setting on various body parts of healthy volunteers. To assess reliability, hypertrophic scars, keloids, and normotrophic scars were imaged and rated by two observers independently. Data are

* Correspondence to: Burn Center, Red Cross Hospital, Vondellaan 13, 1942 LE Beverwijk, the Netherlands.
E-mail addresses: drijpma@rkz.nl.cc (D. Rijpma), ppmvanzuijlen@me.com (P.P.M. van Zuijlen).
[1] Shared first author, equal contribution

https://doi.org/10.1016/j.burns.2022.12.008
0305-4179/© 2022 Elsevier Ltd and ISBI. All rights reserved.

expressed as mean (+/-SD), Coefficient of Variation (CV), Intraclass Correlation Coefficients (ICC), and Limits of Agreements (LoA).

Results: Eighty stickers placed on 20 healthy volunteers showed validity with CV between 0.62%– 1.67% for observer A and 0.75%– 1.19% for observer B. For the reliability study, 69 scars on 36 patients were included. Mean scar surface area ranged from 0.83 cm^2 to 155.59 cm^2. Mean scar surface area measurement was 13.83 cm^2 (SD 23.06) for observer A and 13.59 cm^2 (SD 23.31) for observer B. Adjusted interobserver CV for trained observers is estimated as 5.59%, with corresponding LoA = 0 ± 0.15 x mean surface area. Interobserver ICCs were 0.99–1.00.

Conclusion: This 3D technology with a depth sensor for measuring scar surface area provides valid and reliable data and thereby surpasses expensive and time-consuming 3D cameras.

© 2022 Elsevier Ltd and ISBI. All rights reserved.

1. Introduction

Burn injuries are considered to be one of the most common types of traumatic injuries [1,2]. Throughout the last decades, the survival rate among these burn patients has increased significantly due to major improvements in burn care [3]. Because more people nowadays survive a severe burn injury, there are more survivors with extensive scars. Therefore, scar formation has become an important outcome of burn care [4]. Scar formation may lead to aesthetic, functional, social, and psychological problems, which consequently may influence patients' health-related Quality of Life [5,6].

Besides subjective scar assessment tools, such as Vancouver Scar Scale (VSS) [7,8] and 'Patient and Observer Scar Assessment Scale' (POSAS) [9–11], objective scar measuring instruments are of interest in daily practice [4,12]. Objective scar quantity assessment, through scar surface area or scar volume measurement, enables quantification of scar formation over time, comparison of outcome measures, and evaluation of scar treatment efficacy [13]. The surface area measurement could potentially be used to assess hypopigmented, hyper-pigmented, or scar contraction. Scar volume measurement could potentially be used to evaluate pathologic scarring, such as hypertrophic or atrophic scars and keloids.

The most basic method of surface measurement is calculating length times width while using a ruler. However, digital innovation has led to scar surface area measurement by three-dimensional (3D) imaging. The advantage of 3D imaging is that it also takes body curvature into account. Therefore, 3D scar surface area measurement is more valid and reliable than measurements with 2D techniques [14–17]. Devices range from stereo photogrammetry such as Vectra H1 (Canfield Scientific, Fairfield, NJ, USA) [18,19] and 3D Lifeviz™ (Quantificare S.A., Sophia Antipolis, France) [15] to (laser-assisted) three-dimensional imaging systems such as Artec Eva (Artec Group, Luxembourg, Luxembourg) [14]. These systems are the most advanced products in their product range. However, very advanced devices often go along with high costs ranging from € 8.000 – 40.000 [20]. Moreover, these devices are time-consuming since the analysis of the 3D image is not performed directly on the cameras themselves. The images are mostly analyzed afterward on a computer with additional software [14,19]. These high costs and time-consuming procedures make it less feasible in a clinical setting.

To overcome these limitations, 3D imaging techniques develop towards other depth sensors. These depth sensors are low-cost (range € 500–800) and low definition or high definition image quality devices, which are attached to or integrated into mobile devices. They are therefore easily applicable in a clinical setting [21–23]. Analysis of 3D images is performed through real-time mobile applications and can be performed within a one-minute procedure [24,25]. In this clinimetric assessment, we evaluate the validity and reliability of a low definition depth sensor system on scar surface area measurement, to evaluate its potential as an objective scar measurement instrument [26].

2. Patients and Methods

2.1. *Patients and volunteers*

The Medical Ethics Review Committee (MERC) of Amsterdam University Medical Centers (Amsterdam UMC), location VUmc classified this study as not subject to the Medical Research Involving Human Subjects Act (WMO). Therefore, no further assessment by the MERC was needed. The principles outlined in the Declaration of Helsinki were followed. Oral informed consent was obtained from 20 volunteers and 36 patients which were 18 years of age or older. Staff members volunteered for the validity measurements. For reliability assessment, patients were included from the outpatient burn scar clinic of the Red Cross Hospital (RKZ), Beverwijk, and the scar outpatient clinic of the Department of Plastic, Reconstructive, and Hand surgery of the Amsterdam UMC, location VUmc in Amsterdam. All types of scars were eligible. Scars were classified into three categories: hypertrophic scars, keloids, or normotrophic scars. Hypertrophic scars were defined as scars with a prolonged period of collagen accumulation up to 6–12 months, resulting in a scar that is elevated and thickened, and is very stiff. Erythema will be prolonged and the scar is often intensely pruritic or even painful [27]. Keloids are defined as scars, which unlike hypertrophic scars behave more like tumors in that growth can occur even years after the original injury and extend far

Table 1 – Patient and scar characteristics. Patient and scar characteristics. Patients were recruited from the burn scar clinic at the Red Cross Hospital and the outpatient clinic of the Plastic, Reconstructive, and Hand surgery of the Amsterdam UMC, location VUmc.

Characteristics	Value (N)	Percentage (%)
No. of scars		
Amsterdam UMC, VUmc	23	33.3
Red Cross Hospital	46	66.7
No. of patients		
Amsterdam UMC, VUmc	10	27.8
Red Cross Hospital	26	72.2
Gender		
Male	11	30.6
Female	25	69.4
Age (years)		
Mean (SD)	39 (12.9)	
Range	18–59	
Ethnicity		
Caucasian	24	66.7
Asian/Mediterranean	10	27.8
African-American	2	5.6
Causes of scarring		
Burns	36	52.2
Trauma/Surgery	33	47.8
Scar location		
Head and neck	6	8.7
Trunk	18	26.1
Upper extremity	26	37.7
Lower extremity	19	27.5
Scar pathology		
Normotrophic	31	44.9
Keloid	20	29.0
Hypertrophy	18	26.1

beyond the confines of the original scar [27]. Patient- and scar characteristics are listed in Table 1.

2.2. *Three-dimensional image capturing*

3D image capturing was performed with the Occipital© (San Francisco, CA, USA) Structure Sensor, a depth sensor (Fig. 1). The device was attached to an Apple© (Cupertino, CA, USA) iPad Air 2. The sensor casts thousands of invisible infrared dots onto an object and calculates the distance between the sensor and the infrared dot on the object to create a 3D image irrespective of curvature. The depth sensor is aligned with the build-in iOS camera thereby creating a colored live 3D image in the software application (3DU HealthCare). Thereafter, the software uses this 3D image to calculate the requested surface area of the object. The Structure Sensor has a VGA resolution of 640 × 480 pixels, its precision ranges from 0.5 mm to 3.0 mm on respectively 40–350 cm distance and it has a field of view of 58 × 45 degrees (horizontal x vertical) [26]. For the fact sheet of the Occipital© (San Francisco, CA, USA) Structure Sensor see appendix 1.

2.3. *Real-time mobile application measurement*

Real-time image capturing and calculations on the 3D image was performed using the 3DUniversum (Amsterdam, The Netherlands) (3DU) Healthcare iOS application [28]. The

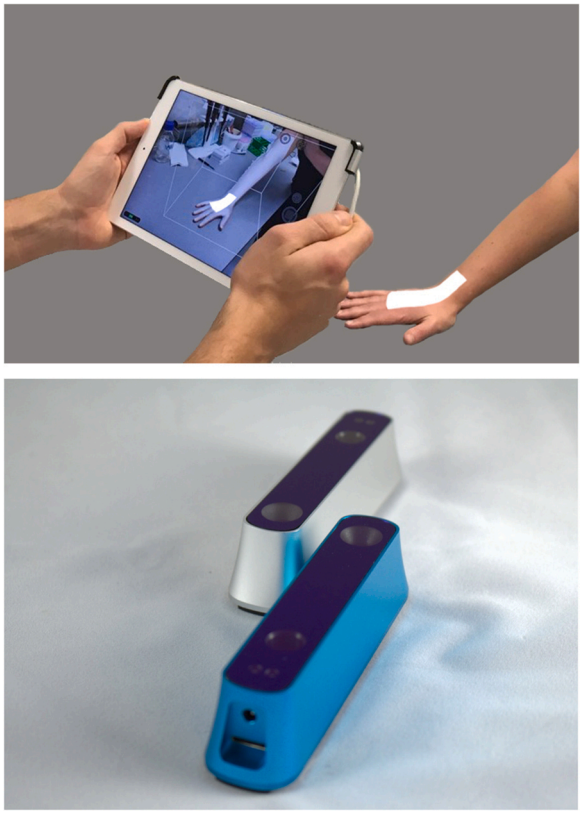

Fig. 1 – 3D depth sensor mounted on a mobile device for instant 3D image capturing.

application is suitable on both iOS and Android platforms and analyzes images from Structure Sensor, Google Tango, and Intel RealSense depth sensors. The technology allows users to capture images, view and analyze surface area by the circumference and depth of the object. Additional software measurement data are color and volume measurement. Borders were drawn manually using the iPad touchscreen. To improve manual alignment object enlargement was performed on the device screen. Image capturing and analyzing of objects are performed in real-time and within a one-minute standardized procedure [28] (Fig. 1).

2.4. *Image capturing procedure*

To achieve standardized three-dimensional images, the image capturing procedure was performed perpendicular to the sticker or scar. The device is slowly pulled back until it meets the minimum threshold distance of 60 cm, thereby automatically starting the 3D image capturing process. Each object was captured in one keyframe. One keyframe is defined as an image capturing procedure in which the entire object is visible and captured within one 3D image. Overlapping keyframes from different angles, particularly used in circular scarring patterns, were not taken into account to diminish artifacts. Despite of the keyframe setting, the depth sensor always takes body curvature into account. Image capture was performed in a pre-arranged room with consistent lighting conditions. Patients had to remain

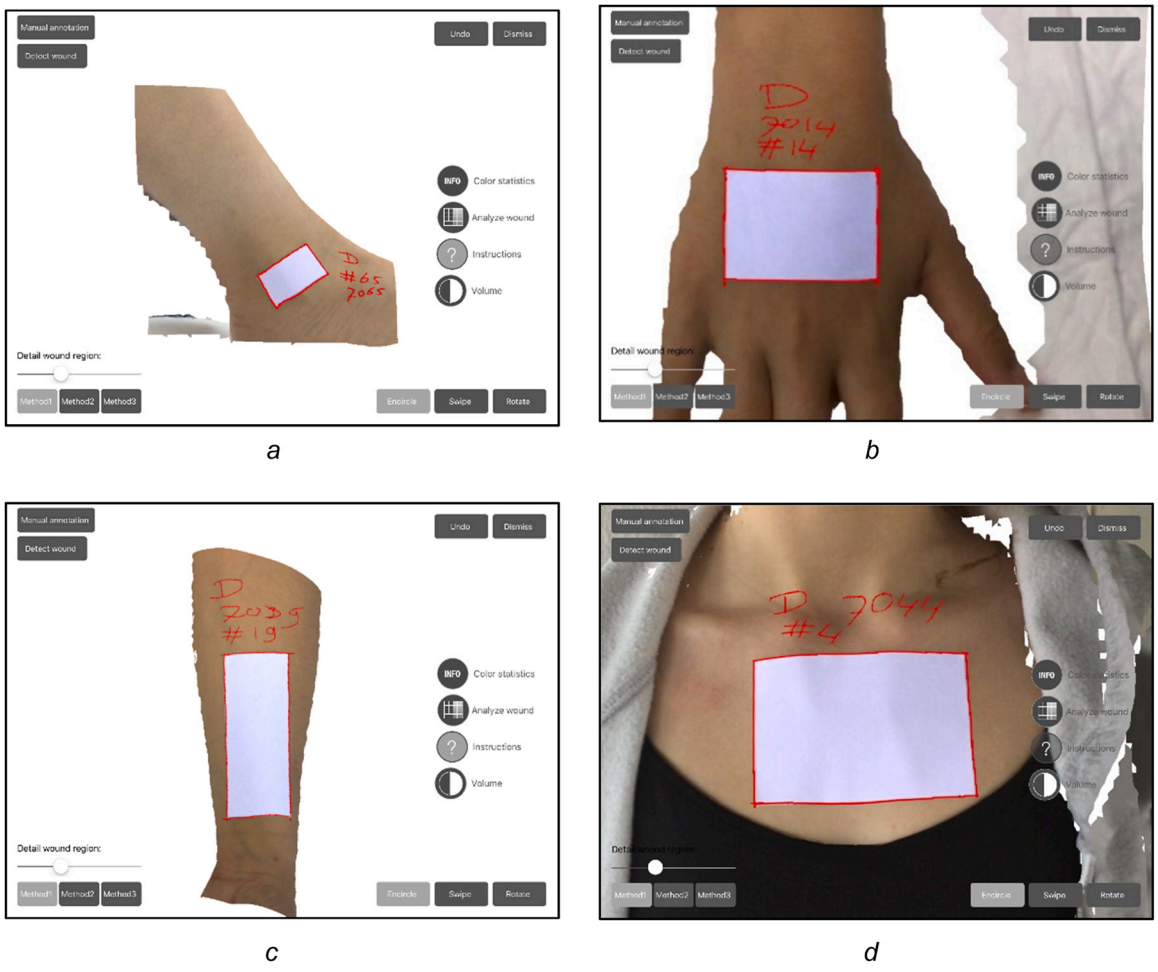

Fig. 2 – Manual sticker annotation on: a) ankle, b) dorsum of the hand, c) anterior forearm and, d) chest. Scanning was performed with the use of 3D technology, a depth sensor on a mobile device, and 3 analysing software. Scanning procedure was performed in a single-observer setting, perpendicular to the object and performed in one 3D keyframe (total object visible with the eye). Validity – Sticker measurement on ifferent body parts.

motionless for a few seconds to diminish image-capturing artifacts. Observers received instructions from colleagues with experience in 3D imaging of scars. Moreover, observers had access to the online device manuals of the Occipital© (San Francisco, CA, USA) Structure Sensor and, had support of 3DUniversum (Amsterdam, The Netherlands). All users were familiar with the use of the Apple© (Cupertino, CA, USA) iPad Air 2.

2.5. *Validity procedure*

Stickers with predefined dimensions were used to act as a gold standard measurement instrument[29]. Twenty volunteers permitted to apply four stickers of various sizes on their ankle (8.01 cm^2), hand (16.33 cm^2), forearm (44.42 cm^2), and chest (77.70 cm^2) to represent increasing scar surface area. A sample of volunteers provides measurement variability through difference in body size and curvature. Stickers are imaged once within a single-observer setting (D.R.) resulting in twenty 3D images per sticker and a total of eighty images. These eighty images were analyzed by both observer A (D.R.) and observer B (M.D.) independently using the device touchscreen. The validity image capturing procedure and flowchart are shown in Figs. 2 and 3.

2.6. *Reliability procedure*

Thirty-six patients with hypertrophic scars, keloids and normotrophic scars who visited the outpatient clinic of the RKZ and Amsterdam UMC, location VUmc were considered eligible and gave oral informed consent. Following the predefined image capturing procedure, scars were imaged by two observers (D.R. and M.D.) separate from each other. Both 3D images were analyzed by both observers independently. This provided four surface area measurements per scar. The reliability image capturing procedure is shown in Fig. 4 and the flowchart for the reliability study in Fig. 5.

2.7. *Statistical analysis*

Data were analyzed using SPSS 24.0 (SPSS Inc. Chicago, USA). Descriptive statistics were performed on patient- and scar

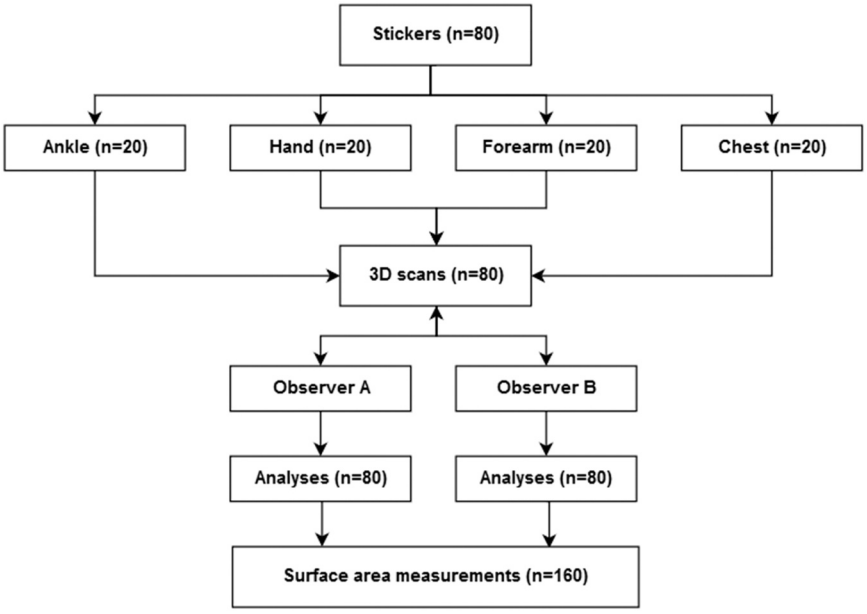

Fig. 3 – Validity - Flowchart of the scanning procedure. flowchart of the scanning procedure for validity assessment of the camera and 3D software.

characteristics. Validity is expressed as mean sticker surface area measurements (cm^2) and mean differences between absolute sticker size and the mean sticker surface area measurements (cm^2). Boxplots illustrate the variance of measurement error per sticker for each observer independently. In addition, coefficients of variation (CV) were calculated. CV displays the percentage of one standard deviation (SD) in correspondence to the mean. A lower coefficient of variation represents a more accurate measurement instrument.

$$CV_{sticker1} = (SD_{sticker1} / Mean_{sticker1}) * 100\%$$

Reliability is displayed through mean scar surface area measurements (cm^2), mean measurement difference inter- and intra-observer scar surface area measurements (cm^2), intra-observer and inter-observer coefficient of variation (CV),

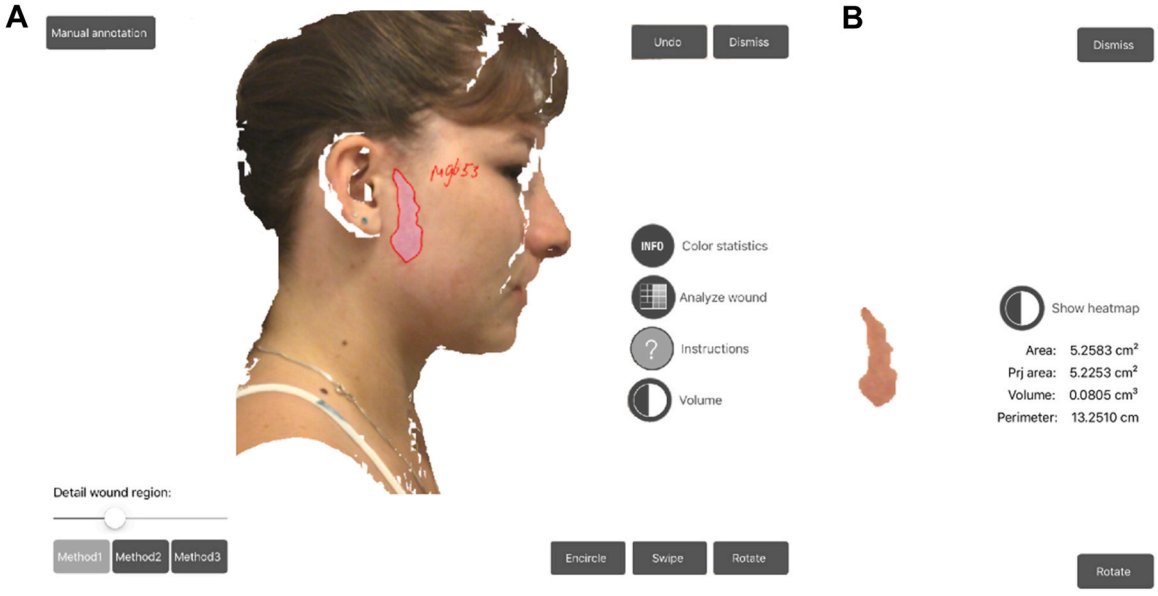

Fig. 4 – Reliability – Scar surface area measurement output of 3D program A) Manual annotation of a burn scar on the right cheek. Red line defines the manual boundaries for surface area measurement. B) A cut-out of the annotated scar area and presented 3D software output. The software provides data on scar surface area, volume and perimeter based on annotation and curvature.

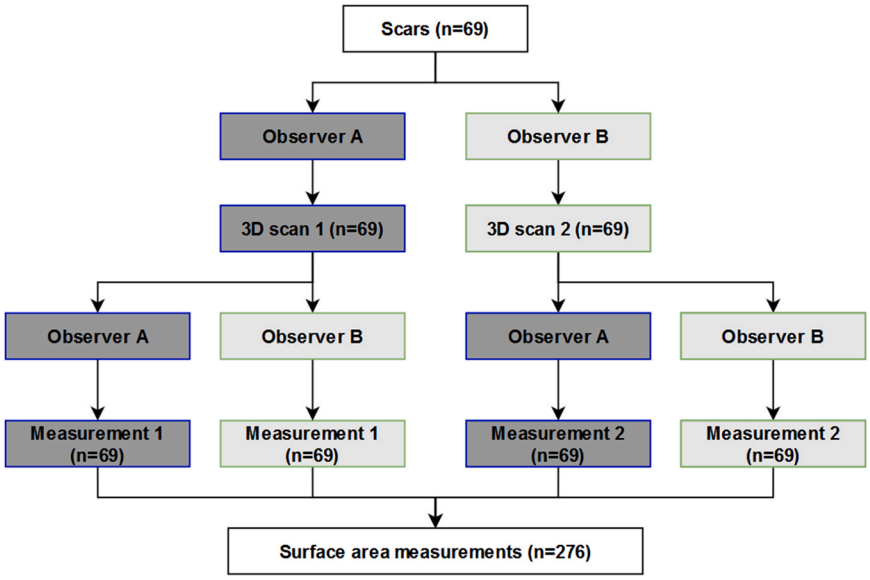

Fig. 5 – Reliability – Flowchart of scanning procedure for scar surface area assessment: flowchart of the scanning procedure for assessment of scar surface area. Both observers made a 3D scan of each scar. A total of 69 scars on 36 patients were included. Intraclass Correlation Coefficients are extracted both within as between observers.

limits of agreement (LoA), and Intraclass Correlation Coefficients (ICC). To account for an increase in observer measurement difference (cm^2) alongside the increase in scar surface area (cm^2), surface area measurements were 10 log-transformed to approximate a normal distribution. Transformed data were used to calculate standard error of measurement (SEM). SEM was transformed back to calculate CVs and LoA[15,28]. CV is formulated as SEM * 100%. A Bland-and-Altman plot was made in which LoA[25] were calculated as a function of the mean scar area[15]. The LoA give an indication of the absolute agreement between observers based on 95% percent of surface area measurement differences between observers. ICC is a ratio of variance components of a random-effects model. ICC expresses the correlation of scar measurements within (ICC_{intra}) and between (ICC_{inter}) observers and ranges from 0.00 to 1.00. A value of 0.00 implies no correlation, whereas an ICC value of 1.00 indicates an excellent correlation. All ICCs were calculated on 10 log-transformed data.

In this study, five correlations were calculated based on the reliability flowchart (Fig. 5).

- ICC_{intra}A: interobserver correlation of observer A based on both 3D images: measurement of 3D image 1 by observer A and 3D image 2 by observer A (Blue).
- ICC_{intra}B: interobserver correlation of observer B based on both 3D images: measurement of 3D image 1 by observer B and 3D image 2 by observer B (Green).
- ICC_{inter}1: interobserver correlation based on independent surface area ratings on the same 3D image: measurement of 3D image 1 by observer A and observer B
- ICC_{inter}2: interobserver correlation based on independent surface area ratings on the same 3D image: measurement of 3D image 2 by observer A and observer B
- ICC_{inter}1–2: interobserver correlation based on independent surface area ratings on different 3D images: measurement of 3D image 1 by observer B and 3D image 2 by observer A.

$$ICC_{inter} = \sigma^2 scar / (\sigma^2 scar + \sigma^2 observer + \sigma^2 error)$$

Results are presented as an analysis of complete data, accompanied by an optimistic analysis in which data are used which could be obtained after training of observers. In this analysis, outliers expected to be due to clinical interpretation differences in the image capturing, protocol are excluded. We examined outliers with more than 20% interobserver measurement differences.

3. Results

3.1. *Patients and scar characteristics*

A total of 69 scars on 36 patients were included. All patients from the Amsterdam UMC presented with keloids, which were caused by trauma or surgery. At the Red Cross Hospital, mainly normotrophic scars (18/26 patients) and hypertrophic scars (11/26 patients) due to burns were included. Patient and scar characteristics are shown in Table 1.

3.2. *Validity*

Sticker surface area measurement showed small measurement errors. The CV for observer A was 0.62%, 1.53%, 1.67%, and 1.06% for ankle, hand, forearm, and chest measurements respectively. CV for observer B was 0.75%, 0.92%, 1.19% and 1.02% in the same sticker order. Results are shown in Table 2.

Table 2 – Results for validity measurement using stickers with predefined dimensions. Mean surface area displays the mean measurement of twenty stickers per body part on twenty healthy volunteers. Mean difference area displays the difference between true sticker size and the mean surface area. Coefficient of Variation (CV) displays the percentage of one standard deviation (SD) in correspondence to the mean measurement per sticker. A = observer A. B = observer B. CV = (SDsticker / Meansticker) * 100%).

		Ankle (n = 20)	Hand (n = 20)	Forearm (n = 20)	Chest (n = 20)
Sticker size	cm^2	8.01	16.33	44.42	77.70
Mean surface area (A)	cm^2 (SD)	7.97 (0,05)	16.59 (0.25)	44.98 (0.74)	77.34 (0.82)
Mean difference area (A)	cm^2 (SD)	0.04 (0.05)	-0.26 (0.25)	-0.56 (0.74)	0.36 (0.82)
Mean surface area (B)	cm^2 (SD)	8.01 (0.06)	16.44 (0.15)	45.62 (0.53)	77.30 (0.79)
Mean difference area (B)	cm^2 (SD)	-0.01 (0.06)	-0.11 (0.15)	-0.20 (0.53)	0.40 (0.79)
$CV_{observerA}$	(%)	0.62	1.53	1.67	1.06
$CV_{observerB}$	(%)	0.75	0.92	1.19	1.02

The boxplots in Fig. 6A and B show the distribution of sticker measurements for observers A and B separately. These figures illustrate an increase in measurement difference according to the increase in object size. Observers A and B show similar patterns due to the use of a standard image capturing protocol.

3.3. *Reliability*

Mean scar surface area ranged from 0.83 cm^2 to 155.59 cm^2. Table 3 shows that the mean scar surface area measurement was 13.83 cm^2 (SD 23.06) for observer A and 13.59 cm^2 (SD 23.31) for observer B. Intra-observer CV was 6.23% for observer A and 2.06% for observer B, indicating a difference in measuring accuracy. Inter-observer CV on the same scar were 8.11% (image 1) and 8.66% (image 2), indicating a discrepancy in clinical interpretation between observers. In addition, inter-observer CV on different images was 8.77%. This indicates that no significant additional measurement error is due to the image capturing device and image capturing procedure.

Seven outliers differing more than 20% between inter-observer measurements (20.39 – 45.78%) were examined. Four of these outliers were caused by interobserver differences in clinical interpretation of scar borders. Three outliers were based on the unknown intra-observer differences on the same 3D image. Adjusted analyses, excluding the seven outliers, resulted in an intra-observer measurement error of 4.07% (observer A) and 1.88% (observer B). Corresponding with an improved adjusted interobserver CV improved from 8.77% to 5.59%.

Limits of agreements were calculated for the interobserver variation of measurement A2 versus B1: LoA = 0 ± 0.24 x mean surface area (Fig. 7a). This implies that 95% of all scar surface area measurements in general, image capturing error, and interobserver variability will range 24% above or below the actual scar surface area. Adjusted limits of agreement improved: LoA = 0 ± 0.15 x mean surface area (Fig. 7b). Since a significant part of the measurements concerned small scars, an additional Bland and Altman plot was performed on scar surface areas smaller than 5 cm^2. LoA for

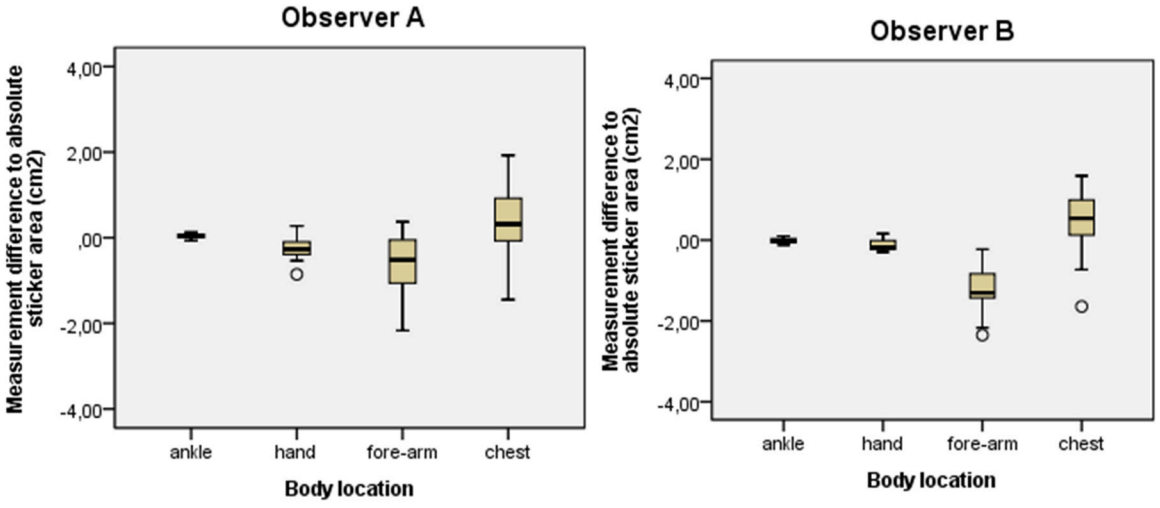

Fig. 6 – Validity – Boxplot of measurement difference to absolute sticker surface area A. boxplots showing variance between measured sticker areas made by observer A and observer B compared to absolute sticker area (cm2). Body location on the x-axis represent four sticker locations and corresponding sizes; ankle (8.01 cm2), hand (16.33 cm2), forearm (44.42 cm2) and chest (77.70 cm2). Y-axis displays the measurement difference to absolute sticker surface area (cm^2). Each location is measured once per observer on twenty volunteers. The boxplots illustrate an increase of measurement difference according the increase in object size. Observer A and B show similar patters due to the use of a standard scanning protocol.

Table 3 – Results for reliability. Reliability measurement on 69 scars within a multi-observer setting. Scar surface area displays the mean (SD) of measurement per observer or between observers. Standard error of Measurement (SEM) expresses the error margin around the measurement. Due to skewed distribution, data were first 10 log-transformed to calculate SEMs, then SEMs were transformed back and presented in this table.

		Intra-observer A1 and A2	Intra-observer B1 and B2	Interobserver A1 and B1	Interobserver A2 and B2	Interobserver A2 and B1
Scar surface area (mean, SD)	cm2	13.83 (23.06)	13.59 (23.31)	13.76 (23.42)	13.67 (22.95)	13.71 (23.10)
Scar surface area.cor (mean, SD)	cm2	14.39 (23.44)	14.12 (23.71)	14.30 (23.81)	14.21 (23.33)	14.25 (23.48)
Measurement difference (mean, SD)	cm2	0.11 (1.05)	0.08 (0.72)	0.25 (0.92)	0.22 (1.21)	0.15 (1.50)
Measurement difference.cor (mean, SD)	cm2	0.13 (1.10)	0.10 (0.76)	0.21 (0.89)	0.18 (1.22)	0.08 (1.53)
SEM	cm2	0.06	0.02	0.08	0.09	0.09
SEMcor.	cm2	0.05	0.02	0.06	0.05	0.06
CV	%	6.23	2.06	8.11	8.66	8.77
CVcor.	%	5.49	1.91	6.46	5.23	5.59
ICC		0.99	1.00	0.99	0.99	0.99
ICCcor.		0.99	1.00	0.99	0.99	0.99

Coefficient of Variation (CV) displays the percentage of Standard error of Measurement (SEM) in correspondence to the measurement mean. CV = SEM * 100%. Intraclass Correlation Coefficients display the agreement between observers (correlations 1 and 2) or between observers (correlations 3, 4, and 5). Due to skewed distribution, data were 10 log-transformed to calculate ICCs. ICC = σ^2scar / (σ^2scar + σ^2observer + σ^2error).cor represents adjusted analyses with the outcomes corrected for seven outliers based on intra-observer measurement errors and interobserver measurement differences based on clinical interpretation of more than 20% compared to the other surface area measurements.

scars < 5 cm^2 slightly increased: LoA = 0 ± 0.27 x mean surface area indicating higher measurement error on smaller scars (Fig. 8a). Adjusted LoA were 0 ± 0.16 x mean surface area (Fig. 8b). ICCs were high and ranged from 0.99 to 1.00 for both intra-observer and interobserver reliability.

4. Discussion

In this study, we showed that scar surface area measurement could be valid and reliably performed by using three-dimensional technology on mobile devices. The technology provides an affordable and feasible system that can be used quite easily in scar care worldwide.

In absence of a clinical gold standard, validity was expressed as measurement error (cm^2) and coefficient of variation (%) around stickers with predefined dimensions. These sticker dimensions and locations represented increasing scar surface area and differences in body curvature. The latter could interfere with calculations on the 3D image. Overall mean measurement differences and corresponding CV were found to be very low and thereby illustrate the accuracy of this 3D depth sensor in combination with the used software on scar surface area measurement. In addition, the tablet enlarging ability of the 3D image might have led to the more precise delineation of sticker borders and therefore more accurate measurements.

Reliability was assessed in a clinical two-observer setting. Data was presented as mean surface area (SD) and mean difference (SD). Intra- and interobserver reliability was expressed as coefficient of variation, limits of agreement, and, Intraclass Correlation Coefficients. For scars, the coefficient of variation between observers ranged from 8.11% to 8.77%, which was higher than found in the validity assessment of stickers with clear boundaries (0.62–1.67%). Border interpretation is more challenging in scars than with the use of white stickers. This is especially the case in normotrophic scars since these colors merge with non-traumatized skin tones, whereas keloids or hypertrophic scarring show a clearer transition. Software upgrades, device resolution, standardized lighting conditions, and the zoom function might decrease these measurement errors. Increased measurement error is to be expected in untrained personnel and the lack of a clear standard operating procedure.

In this study, data are presented both as an analysis of complete data and after the exclusion of seven outliers which could be avoided by appropriate training of the observers. The training should include a set of challenging scars, with a two-folded aim: to have some practice with image capturing and analysis of scars and to make clear agreements on clinical interpretation of scar borders. Before implementing it in daily practice, we recommend health care providers train personnel with such training to prevent these kinds of measurement errors.

Both intra-observer and interobserver ICC were 0.99 or higher, indicating excellent reliability. ICCs with a minimum of 0.7 are considered acceptable for research purposes and ICCs of at least 0.9 are required in a clinical setting [29]. The high ICC, however, can be explained by the wide variation of scar sizes in this study, ranging from 0.83 cm^2 – 155.59 cm^2.

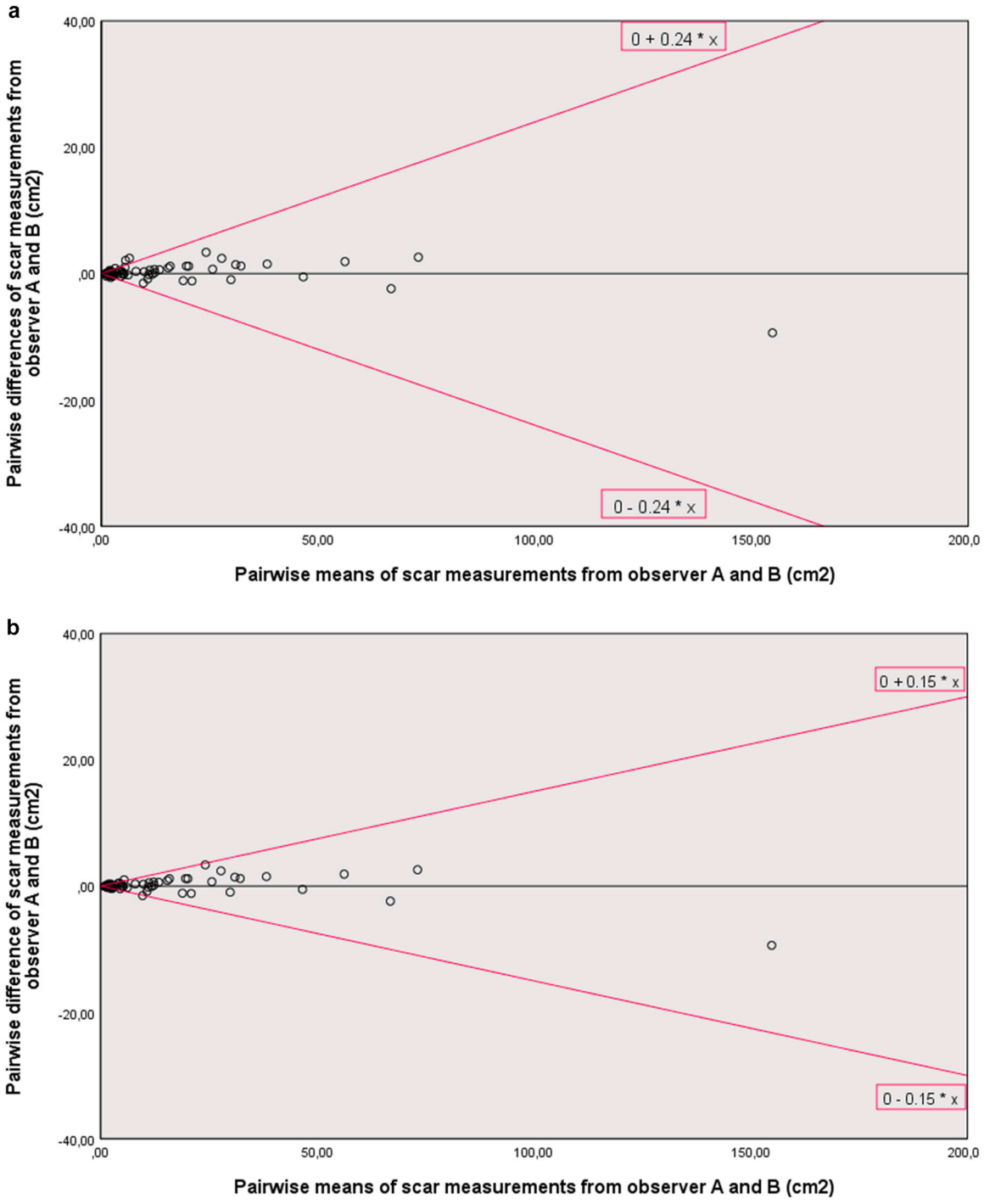

Fig. 7 – a. Reliability – Bland and Altman Plot Interobserver Agreement – Overall Limits of Agreement. a. A Bland and Altman plot presenting the interobserver agreement between two observers (measurement A2 versus B1). Due to expected skewed distribution of scar surface area measurements, the data was natural log-transformed to approximate a normal distribution and obtain a regression coefficient for limits of agreement. The limits of agreement indicate the absolute agreement between observers based on 95% percent of surface area measurement differences between observers. The interobserver agreement is based on the mean surface areas of scan 2 analyzed by observer A and scan 1 analysed by observer B (see Fig. 5). b. Reliability – Bland and Altman Plot Interobserver Agreement – Overall Adjusted Limits of Agreement. b. A Bland and Altman plot adjusted for outliers, presenting an interobserver agreement between two observers (measurement A2 versus B1). Due to expected skewed distribution of scar surface area measurements, the data was natural log-transformed to approximate a normal distribution and obtain a regression coefficient for limits of agreement. The limits of agreement indicate the absolute agreement between observers based on 95% percent of surface area measurement differences between observers. The interobserver agreement is based on the mean surface areas of scan 2 analyzed by observer A and scan 1 analysed by observer B (see Fig. 5).

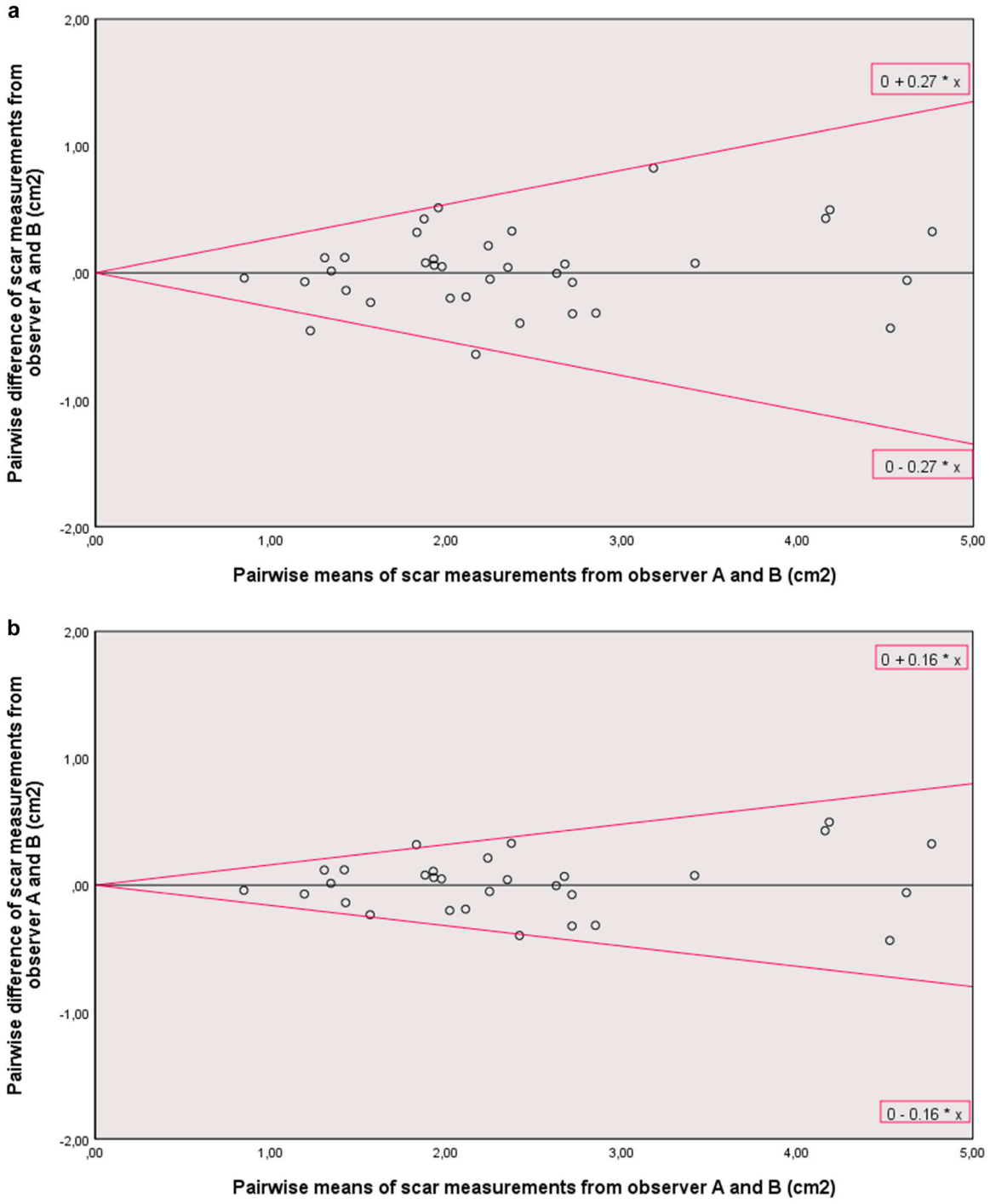

Fig. 8 – a. Reliability – Bland and Altman Plot Interobserver Agreement – Limits of Agreement for Small Scars (< 5 cm^2)a. A Bland and Altman plot presenting an interobserver agreement between two observers (measurement A2 versus B1). The interobserver agreement is based on the mean surface areas of scan 2 analyzed by observer A and scan 1 analysed by observer B (Fig. 5) on scar surface areas less than 5 cm^2. 8b. Reliability – Bland and Altman Plot Interobserver Agreement – Adjusted Limits of Agreement for Small Scars (< 5 cm^2) b. A Bland and Altman plot adjusted for outliers, presenting an interobserver agreement between two observers (measurement A2 versus B1). The interobserver agreement is based on the mean surface areas of scan 2 analyzed by observer A and scan 1 analysed by observer B (Fig. 5) on scar surface areas less than 5 cm^2.

The influence of device, software, and observer measurement error is small compared to the variation in scar size and therefore has little influence on the correlation.

Limits of Agreement (LoA) project a more comprehensive measurement outcome since it takes the wide range of object sizes into account. Since the LoA showed that the

measurement error increased with increasing scar size, data were log-transformed to approximate a normal distribution. LoA can be extrapolated toward a clinical setting and used in comparison to other instruments.

The adjusted 3D depth sensor LoA = 0 ± 0.15 x mean surface area. This implies that 95% of all scar surface area measurements in a clinical setting, adjusted for scar size, image capturing error and interobserver variability will range 15% above or below the scar surface area. LoA of scars < 5 cm^2 slightly increased to LoA = 0 ± 0.16 x mean surface area. These LoA found in this study are considered appropriate for clinical use compared with other 3D cameras [14,15]. A study by Rashaan et al., 2016 [14] assessed wound surface areas with the Artec 3D camera and Stekelenburg et al., 2013 [15] performed scar surface area measurements with stereo photogrammetry. The mean scar surface area of this study was 13.71 cm^2, whereas a mean surface area of 20.58 cm^2 and 59.65 cm^2 were found in the two compared studies respectively. Limits of Agreement of the Artec MHT™ 3D camera was LoA = 0 ± 0.17 x mean surface area [14] whereas stereo photogrammetry showed a LoA = 0 ± 0.19 x mean surface area [15]. In general, this mobile 3D technology with a depth sensor is not inferior to counterparts such as the Artec MHT™ 3D camera or stereo photogrammetry regarding scar surface area measurement. Since mobile devices and 3D analyzing software provide object zooming function it provides more accurate and reliable measurements of the 3D image. Alongside less time-consuming data analysis and lower costs, this demonstrates the strengths of these 3D depth sensors and its mobile software applications in comparison to 3D cameras mentioned above.

This mobile 3D technology does have limitations. The difficulty of capturing wounds or scars on highly curved body parts is a well-known problem within 3D imaging techniques [14–16,30–33]. Body curvature remains a bigger challenge than with 3D techniques mentioned above, due to a lower resolution and thus fewer data points. High curvature and circular objects require multiple keyframes, which, to date, show too much overlap and therefore higher measurement error. Especially, scars on noses, ears, and fingers could not be imaged. This is unfavorable, since a substantial part of the patients that visit the outpatient clinics of the Red Cross Hospital have scars on their face and hands.

Ongoing this 3D software innovation is expected to surpass these limitations as they have already shown its accuracy and reliability within one keyframe. Automatic annotation of object borders based on color segmentation is currently in development. It might account for more reliable scar assessment by surpassing clinical interpretation and increasing usability in a clinical setting. A High Definition depth sensor that is currently on the market [34] might be more capable to measure the volume of a scar, which is an important clinical outcome parameter. Moreover, it might measure the extent of circular scar patterns or pigmentation changes within a scar more accurately. Alongside the integration of 3D depth sensors in current smartphones, the next innovations in healthcare applications are on the horizon.

5. Conclusion

3D technology by using affordable and feasible 3D depth sensors mounted on mobile devices provides instant object data in a clinical setting and thereby surpasses expensive and time-consuming 3D cameras. With the ongoing software innovation and integration of depth sensors in mobile devices, the implementation of this 3D technology for scar surface area measurements can be considered worldwide.

Funding

This research project was funded by the Dutch Burns Foundation, project number 16.02.

Disclosure

Authors S. Karaoglu and T. Gevers are founders of 3DUniversum (Amsterdam, The Netherlands) (3DU) and supervised the development of the 3DU Healthcare iOS application.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests, Authors S. Karaoglu and T. Gevers are founders of 3DUniversum (Amsterdam, The Netherlands) (3DU) and supervised the development of 3DU Healthcare iOS application. 3DU had no interference with the study data.

REFERENCES

[1] Haagsma JA, et al. The global burden of injury: incidence, mortality, disability-adjusted life years and time trends from the Global Burden of Disease study 2013. Inj Prev 2016;22(1):3–18.
[2] WHO, Global Health Estimates: deaths, disability (YLD) by cause, age and sex, 2000–2012. 2016.
[3] Brusselaers N, et al. Outcome and changes over time in survival following severe burns from 1985 to 2004. Intensive Care Med 2005;31(12):1648–53.
[4] Lee KC, et al. A systematic review of objective burn scar measurements. Burns Trauma 2016;4:14.
[5] Van Loey NE, Van Son MJ. Psychopathology and psychological problems in patients with burn scars: epidemiology and management. Am J Clin Dermatol 2003;4(4):245–72.
[6] Brown BC, et al. The hidden cost of skin scars: quality of life after skin scarring. J Plast Reconstr Aesthet Surg 2008;61(9):1049–58.
[7] Baryza MJ, Baryza GA. The vancouver scar scale: an administration tool and its interrater reliability. J Burn Care Rehabil 1995;16(5):535–8.
[8] Sullivan T, et al. Rating the burn scar. J Burn Care Rehabil 1990;11(3):256–60.
[9] Draaijers LJ, et al. The patient and observer scar assessment scale: a reliable and feasible tool for scar evaluation. Plast Reconstr Surg 2004;113(7):1960–5. discussion 1966-7.
[10] van der Wal MB, et al. The modified patient and observer scar assessment scale: a novel approach to defining

pathologic and nonpathologic scarring? Plast Reconstr Surg 2012;129(1):172e–4e.
[11] The Patient and Observer Scar Assessment Scale (POSAS). Available from: https://www.posas.org/.
[12] Lindeboom JA, et al. Interpretation and precision of the observer scar assessment scale improved by a revised scoring. J Clin Epidemiol 2008;61(12):1289–95.
[13] Perry DM, McGrouther DA, Bayat A. Current tools for noninvasive objective assessment of skin scars. Plast Reconstr Surg 2010;126(3):912–23.
[14] Rashaan ZM, et al. Three-dimensional imaging: a novel, valid, and reliable technique for measuring wound surface area. Ski Res Technol 2016;22(4):443–50.
[15] Stekelenburg CM, et al. Three-dimensional digital stereophotogrammetry: a reliable and valid technique for measuring scar surface area. Plast Reconstr Surg 2013;132(1):204–11.
[16] Stekelenburg CM, et al. In a clinimetric analysis, 3D stereophotogrammetry was found to be reliable and valid for measuring scar volume in clinical research. J Clin Epidemiol 2015;68(7):782–7.
[17] Darwin ES, et al. Comparison of 3-dimensional wound measurement with laser-assisted and hand measurements: a retrospective chart review. Wound Manag Prev 2019;65(1):36–41.
[18] Tanaka Y. Long-term three-dimensional volumetric assessment of skin tightening using a sharply tapered non-insulated microneedle radiofrequency applicator with novel fractionated pulse mode in asians. Lasers Surg Med 2015;47(8):626–33.
[19] Urbanova P, Hejna P, Jurda M. Testing photogrammetry-based techniques for three-dimensional surface documentation in forensic pathology. Forensic Sci Int 2015;250:77–86.
[20] Verhulst, et al. Three-dimensional imaging of the face: a comparison between three different imaging modalities. Aesthetic Surg J 2018;15;38(6):579–85.
[21] Huang G, et al. Design and optimization of a novel three-dimensional force sensor with parallel structure. Sens (Basel) 2018(8):18.
[22] Kieser DC, Hammond C. Leading wound care technology: the ARANZ medical silhouette. Adv Ski Wound Care 2011;24(2):68–70.
[23] Jin J, et al. 3-D wound scanner: a novel, effective, reliable, and convenient tool for measuring scar area. Burns 2018;44(8):1930–9.
[24] Anghel EL, et al. The reliability of a novel mobile 3-dimensional wound measurement device. Wounds 2016;28(11):379–86.
[25] Bills JD, et al. Pilot study to evaluate a novel three-dimensional wound measurement device. Int Wound J 2016;13(6):1372–7.
[26] Occipital Inc. Technical specifications of Structure Sensor. 2016; Available from: https://support.structure.io/article/157-what-are-the-structure-sensors-technical-specifications.
[27] Téot L, et al. Textbook on Scar Management: State of the Art Management and Emerging Technologies. Springer; 2020. https://doi.org/10.1007/978-3-030-44766-3
[28] 3DUniversum, 3DUniversum: low-cost software solutions for real-time 3D mobile scanning and recognition. 2017. https://3duniversum.com/product/mobile-scanner/.
[29] de Vet HC, et al. When to use agreement versus reliability measures. J Clin Epidemiol 2006;59(10):1033–9.
[30] van Zuijlen PP, et al. Reliability and accuracy of techniques for surface area measurements of wounds and scars. Int J Low Extrem Wounds 2004;3(1):7–11.
[31] Euser AM, Dekker FW, le Cessie S. A practical approach to Bland-Altman plots and variation coefficients for log-transformed variables. J Clin Epidemiol 2008;61(10):978–82.
[32] Plassmann P, Jones TD. MAVIS: a non-invasive instrument to measure area and volume of wounds. Measurement of area and volume instrument system. Med Eng Phys 1998;20(5):332–8.
[33] Shetty R, et al. A novel and accurate technique of photographic wound measurement. Indian J Plast Surg 2012;45(2):425–9.
[34] Occipital.Inc, Structure Sensor (Mark II). 2019. https://occipital.com/2019/mark-ii-announcement.