

Comparison of Smartphone-Based 3D Body Measurements with Traditional Anthropometry and 4D Stereophotogrammetry: Reliability, Validity, and Accuracy in Healthy Adults

Sofia Scataglini^{1,*}, Femke Van Orden¹, Emma Van Den Hoek¹, Steven Truijen¹

4D4ALL, Department of Rehabilitation Sciences and Physiotherapy, Center for Health and Technology (CHaT), Faculty of Medicine and Health Sciences, MOVANT, University of Antwerp, 2000 Antwerpen, Belgium

[*sofia.scataglini@uantwerpen.be](mailto:sofia.scataglini@uantwerpen.be)

Abstract:

Background:

Accurate body measurements are central to clinical assessment, requiring high precision, validity and reliability. However, their routine use is often limited by the time-intensive nature of conventional anthropometric techniques.

Objective:

This study evaluated the reliability, validity and accuracy of smartphone-based body measurement applications compared with 4D stereophotogrammetry, using traditional anthropometry as the reference standard.

Methods:

In this cross-sectional study, 35 adults (≥ 18 years) were assessed using three approaches: 4D stereophotogrammetry, two smartphone applications (3D Measure Me and MeThreeSixty), and standard anthropometric methods (stadiometer and measuring tape). Validity and reliability were examined using intraclass correlation coefficients (ICC) and statistical testing (e.g. paired T-test), while accuracy was evaluated using root mean square error (RMSE) and Bland–Altman analyses.

Results and discussion:

4D stereophotogrammetry and 3D Measure Me demonstrated acceptable to excellent agreement with tape-based measurements, whereas MeThreeSixty showed consistently lower validity and accuracy. ICC values were largely significant for 4D stereophotogrammetry and 3D Measure Me, with the former showing the highest agreement across most variables. Measurements such as bust height and waist height did not reach statistical significance across methods. Reliability was generally excellent (ICC = 0.90–1.00), although right waist height (ICC = 0.022) and belly height (ICC = 0.066) showed poor reliability. Accuracy analyses indicated lower RMSE for extremity measurements and higher error for trunk variables. 4D stereophotogrammetry showed more stable performance (RMSE: 6.98–120.29 mm) than 3D Measure Me (14.88–197.67 mm), while MeThreeSixty exhibited increasing error with larger body segments. Most differences fell within acceptable limits of agreement.

Conclusion:

4D stereophotogrammetry and 3D Measure Me represent valid and reliable alternatives to conventional anthropometric methods in clinical settings.

Author Contributions:

Conceptualization: S.S., F.V.O., E.V.D.H., M.S., S.T.; Data curation, formal analysis, investigation, methodology, and software: S.S., F.V.O., E.V.D.H., M.S.; Project administration, supervision, and resources: S.S., S.T.; Writing – original draft: S.S., F.V.O., E.V.D.H., M.S.; Writing – review and editing: S.S., S.T. All authors have read and agreed to the published version of the manuscript.

Funding:

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors

Data availability:

All titles and abstracts resulting from the database search were directly imported into Rayyan for screening and management. The datasets generated and analyzed during the current study are available from the corresponding author upon reasonable request. Due to licensing restrictions from the database providers, raw data (i.e., full records retrieved from proprietary databases) cannot be publicly shared.

Ethics approval and consent to participate

This study was approved by the Ethics Committee of Antwerp University Hospital and the University of Antwerp (BUN: B3002024000113; Project ID: 6723). Final favorable ethical approval was granted on 30 September 2024. All procedures involving human participants were performed in accordance with the ethical standards of the institutional research committee and with the Declaration of Helsinki and its later amendments. Written informed consent was obtained from all participants prior to inclusion in the study.

Trial registration

The study was retrospectively registered at ClinicalTrials.gov on 8 May 2026 under registration number NCT07589998 (Protocol ID: ANTHRO-SMART-4D-2026).

Data availability

The datasets generated and/or analyzed during the current study are available from the corresponding author on reasonable request and subject to institutional and ethical approval. De-identified individual participant data, the data dictionary, statistical code, study protocol, and statistical analysis plan are available upon reasonable request.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

1 Introduction

Anthropometry, a subfield of metrology, is the discipline concerned with the measurement of the human body (1). Body measurements are essential for quantifying differences within and between populations (2). Advances in three-dimensional (3D) and four-dimensional (4D) scanning technologies have enabled detailed, marker less, and time-efficient assessments of human morphology (3). These measurements are of great interest to dietitians, nutritionists, health professionals, and sports scientists because of their clinical utility (4).

Traditionally, simple anthropometric measurements tools are flexible measuring tapes and calipers (5). These methods are indirect, because the body composition is not measured directly (6). Traditionally, anthropometric measurements are obtained using calipers or two-dimensional photography **Error! Reference source not found.**. These manual techniques are performed by trained personnel but remain subject to errors arising from equipment malfunction and human inaccuracy (8). Manual anthropometric techniques require standardized anatomical landmarks, consistent measurement methods and instruments, and well-trained examiners to ensure accurate and reliable data (9). Direct measurements are considered the most precise for assessing body composition and are typically obtained using imaging techniques such as dual-energy X-ray absorptiometry (DXA), magnetic resonance imaging (MRI), or computed tomography (CT), which allow internal body evaluation (10). However, these methods expose patients to radiation and are not cost-effective (11).

In recent years, practitioners have increasingly adopted smartphone applications as a cost-effective and space-efficient alternative for collecting essential anthropometric variables that would otherwise require bulkier, more expensive, and time-consuming equipment (12). A virtual twin of the scanned subject or object can be created, enabling researchers to revisit the 3D scan and extract additional contextual information (13). Avatars also enhance user engagement and play a crucial role in motor skill (re)acquisition by facilitating self-correction through avatar observation, thereby improving clinical rehabilitation outcomes (14).

When an additional dimension is introduced, 4D systems incorporate the temporal dimension along with the three spatial axes (x, y, z) (15). The study by Meletani et al. (2024) confirmed the validity and reliability of 4D digital human modeling and its potential for measuring biomechanical characteristics.

Building on last year's systematic review (16), this study compares traditional anthropometric measurements, two smartphone-based 3D scanning applications, and the 4D Stereophotogrammetry system, with a focus on assessing accuracy, validity, and reliability.

2 Methods

The method section was written based on the consort 2010 checklist. The study was retrospectively registered (8 May 2026) at ClinicalTrials.gov under registration number NCT07589998 (Protocol ID: ANTHRO-SMART-4D-2026). The trial record can be accessed via:

<https://clinicaltrials.gov/study/NCT07589998>. The date of registration is reported in the ClinicalTrials.gov registry entry. The trial protocol and statistical analysis plan are available from the corresponding author upon reasonable request. De-identified individual participant data, including the data dictionary, statistical code, and other relevant study materials, are available from the corresponding author upon reasonable request and subject to institutional and ethical approval.

2.1 Trial design

This study used a cross-sectional design, where all 35 participants underwent three types of measurement: 4D Stereophotogrammetry (scan), smartphone (scan) and traditional measurement. These different outcomes were taken at one point in time. Due to the large number of participants and the available time, the group was divided into subgroups, that were measured on different days.

2.2 Participants

For this study, 35 healthy participants were selected, comprising seventeen men and eighteen women. Participants were required to meet predefined inclusion criteria (table 1). Subjects were recruited through flyers and emails. To reach a broader audience, the flyers were also spread via social media. Interested individuals could contact the research team and received detailed information about the study purpose and design. Those who agreed to participate, were asked to sign an informed consent (addendum 1) 'Insert Table 1 here'

Table 1: inclusion-exclusion

	inclusion	exclusion
P	Human adults (³ 18 years)	Children; animals
I	iPhone; Samsung; smartphone; 4D Stereophotogrammetry	X-ray; MRI; CT
C	Traditional anthropometrics	Surgery
O	Whole body; shape; avatar	Dental; geographical mapping; non-digital human; body parts; skeleton body

2.3 Interventions

Participants will undergo a 3D full-body scan using two smartphone-based scanning applications (Iphone 17 ProMax), 3DMeasureMe (17) and MeThreeSixty (18) , a traditional anthropometric measurement and a reference scan in the 4D4ALL laboratory.

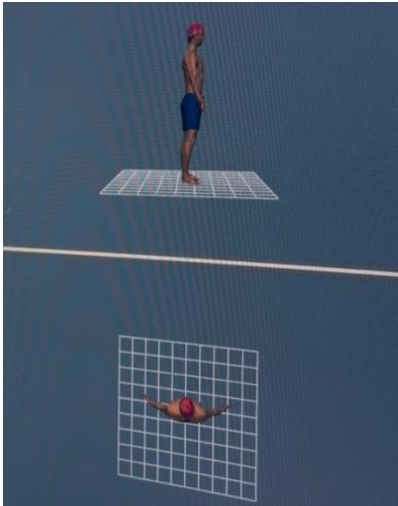
The two mobile applications employ different scanning procedures. The 3DMeasureMe app captures four photographs, one anterior, two lateral and one posterior, in two anatomical planes (frontal and sagittal). In contrast, the MeThreeSixty app captures only two photographs in the same planes: one from the anterior view and one from the lateral view.

Before we start with the scan, the participants must wear a swimming cap and tight-fitting clothing. Afterwards, they are instructed to stand in an A-pose during the scanning procedure. For 3DMeasureMe, the hands are positioned flat and facing the screen, whereas for MeThreeSixty, the hands are in pronation. Each scan takes approximately one minute to complete and generates a digital avatar for subsequent analysis. In addition to the difference in body position and the number of photographs taken, there is also a difference in the type of camera used. With the 3DMeasureMe app, users can choose to perform the scan themselves using the front camera, or have another person take the scan using the back camera. In contrast, the MeThreeSixty app does not offer this option and always uses the front camera.



Furthermore, traditional anthropometric measurements will be performed using two instruments, namely a stadiometer, weight scale and a measuring tape. The tape measurement will be taken one time by two different researchers for the inter-rater reliability of the traditional measurement. The estimated duration of this procedure is 20 to 30 minutes.

Finally, participants will undergo a reference scan in the 4D Stereophotogrammetry. During this procedure, they will stand in an A-pose at the center of the cameras' field of view. The scan will generate a full-body avatar composed of 14 modules. Each module includes two infrared cameras to



capture surface geometry, a projector, and an RGB camera to record skin and clothing texture. A central processor connected to all 14 modules transmits the collected data to a computer for processing.

Participants will wear identical swimming caps and minimal clothing. The clothing worn during the scan will be tight-fitting, dark-colored, and without socks. The scanning procedure will take approximately 10 seconds and will be performed once.

2.4 Outcomes

2.4.1 Primary outcomes

All of the primary measures are anthropometric and were obtained in the laboratory using 4D Stereophotogrammetry, smartphone apps and conventional methods. These measurements were separated into two categories: length and circumference.

1. Length measures: waist hip; waist to knee height; inside leg height; knee height; outer arm in; foot length and height
2. Circumference measures: neck girth; hip girth; belly girth; bust girth; under bust girth; back arm pit cont.; forearm girth; wrist girth; thigh girth; knee girth; ankle girth and upper arm girth

After the comma, both measurements were given in centimetres (cm) with one decimals. All the primary measurements are shown in Table 2, with the length metrics shown in light blue and the circumference represented in dark blue

As explained in the section intervention, a stadiometer and a measuring tape were used in traditional measures. Two apps were used to collect the smartphone scans and 14 modules were used to create an advanced digital body for the 4D Stereophotogrammetry. These outcomes were all taken in one sitting, which indicates that participants were invited in the laboratory and the four instruments were measured in the same day. Table 4 provides descriptions of all outcomes, including the primary and secondary, along with the precise reference points. 'Insert Table 2 here'

Table 2: Primary outcomes

<u>Traditional measures</u>	<u>4D Stereophotogrammetry</u>	<u>3D Measure Me</u>	<u>MeThreeSixty</u>
Waist ht	Waist ht	Waist to floor	
Waist to knee ht	Waist to knee ht	Waist to knee	
Knee ht	Knee ht	Knee height	
Outer arm in	Outer arm in	Arm length	
Foot length		Foot length	
height	height	height	height
Neck girth	Neck girth	Neck	
Hip girth	Hip girth	Hips	
Belly girth	Belly girth	Abdomen	Stomach
Bust girth	Bust girth	Chest / Bust	Chest
Under bust girth	Under bust girth	Under bust	
Back arm pit cont	Back arm pit cont	Back width	
Forearm girth	Forearm girth	Forearm	Forearm
Wrist girth	Wrist girth	Wrist	

Thigh girth	Thigh girth	Upper thigh	Thigh
Knee girth	Knee girth	Knee	
Ankle girth	Ankle girth	Ankle minimum or maximum	
Upper arm girth	Upper arm girth	Biceps	biceps

2.4.2 Secondary outcomes

These measurements do not form the basis of the power calculation in a physical practice. They are rather offered as supplementary information. These outcomes are also separated into length and circumference. They are displayed in Table 3 with the same colour coding as the primary and likewise measured in centimetres. 'Insert Table 3 here'

Table 3: Secondary outcomes

<u>Traditional measured</u>	<u>4D Stereophogrammetry</u>	<u>3D Measure Me</u>	<u>MeThreeSixty</u>
Belly height		Abdomen height	
Bust height		Bust height	
armspan		armspan	
Head girth	Head girth	Head circumference	
Bust width	Front armpit cont	Bust width	

‘Insert Table 4 here’

Table 4: Reference points

<u>Traditional measures</u>	<u>Reference point</u>
Waist ht	Length waist to floor (SIAS)
Waist to knee ht	Vertical length from waist to middle of knee cap
Knee ht	Length from the top of the knee to the bottom of the foot
Outer arm in	Arm length from bone at top edge of shoulder to elbow to wrist just below wrist bone
Foot length	Length from back of heel to tip of longest toe
height	Height from top of head to floor
Bust height	Distance from base of neck down to the centre of the fullest part of bust
Belly height	Vertical distance from waist to abdomen
armspan	Length from one end of a person's arms (measured at the middle fingertips) to the other when arms raised parallel to the ground at shoulder height at 90° angle
Neck girth	Circumference around neck over Adams apple following natural curve
Hip girth	Circumference at widest point of bum from side view
Belly girth	Circumference at belly button
Bust girth	Circumference around chest beneath armpits and in line with nipples
Under bust girth	Circumference just under bust
Back arm pit cont	Distance armpit to armpit straight across the back without going into the armpit
Forearm girth	Circumference at largest point on forearm
Wrist girth	Circumference just below wrist bone
Thigh girth	Circumference measurement taken 1cm below the gluteal line or fold
Knee girth	Circumference around knee cap at the widest point
Ankle girth	Circumference at the narrowest point of the ankle
Upper arm girth	Bicep circumference at upper arm mid-point
Head girth	Circumference from slightly above eyebrows above the ears and over the occipital prominence and the back of the head (with a bathing cap)
Bust width	Length from the left armpit to the right armpit across the front in a straight line over the upper part and not through the centre

2.4.3 Explanatory outcome

In order to comprehend why findings could vary between smartphone scans, traditional measures and the 4D Stereophotogrammetry, extra measurements were gathered. The researcher evaluated how thigh clothes were during the scanning process. There are two types of clothing: tight and loose. Movement during scanning was also rated: no movement, slight movement or significant movement. Finally, the participant skin colour is classified as light, medium or dark.

2.5 Sample size

To ensure a statistically valid and representative sample size, minimizing the influence of random variation and increasing the generalizability of the results, a balanced sample of 17 men and 17 women was targeted for this study. The required sample size was calculated using G*Power (19), based on an expected correlation coefficient of $r = 0.65$, a significance level of $\alpha = 0.05$ and a statistical power of $\beta = 0.80$. According to these parameters, the calculation resulted in a minimum required total number of 32 subjects. To limit the risk of dropouts or measurement errors, it was decided to recruit a larger group consisting of 35 subjects, evenly divided between men and women. This balanced composition increases the precision of the correlation analyses and provides a sufficient basis for evaluating the consistency and reliability between the different measurement methods. If possible, additional participants will be recruited to further strengthen the statistical power and reduce the confidence intervals of the results.

2.6 Randomization

Participants were randomly assigned to one of three possible sequences of measurement methods, following a balanced Latin-square design (20).

The four measurement methods included the traditional anthropometric assessment, the 4D Stereophotogrammetry measurement and two mobile applications (MeThreeSixty and 3DMeasureMe).

This design ensured that each measurement method appeared equally often in each order position, thereby minimizing potential order or carry-over effects. The numbers are linked to the participants, while the letters are linked to the researchers. This can be found in Table 5 and 6.

‘Insert Table 5 here’

Table 5: Randomisation 1

PARTICIPANT	A	B	C
1	Traditional	4D Stereophotogrammetry + traditional 2	Smartphones
2	4D Stereophotogrammetry + traditional 2	Smartphones	Traditional
3	Smartphones	Traditional	4D Stereophotogrammetry + traditional 2

‘Insert Table 6 here’

Table 6: Randomisation 2

PARTICIPANT	LETTER
1	A
2	B
3	C
4	D

In addition, the allocation of researchers was randomised so that each of the three researchers assessed a comparable number of participants, ensuring balance and reducing investigator bias. Randomisation was performed using a computer-generated list and the random seed was recorded to ensure reproducibility.

2.7 Blinding

Given that the study uses a cross-sectional design, blinding is not applicable. This means that we have no intervention, the data is collected at a single point in time and existing characteristics such as height and circumference are measured. However, each person will be linked to an ID number for anonymity, reliable and consistent data processing.

2.8 Statistical methods

The statistical analysis will focus on assessing the validity, accuracy and reliability of 3D body scans through smartphones compared with traditional anthropometric methods and the 4D Stereophotogrammetry. All results are at ratio/interval level and an Excel file will be used to import and do statistical analyses in SPSS 31, except for the RMSE and the CCC that will be determined in Excel. Table 8 till 12 provides an interpretation of the various findings.

2.8.1 Reliability

For reliability, we assessed the inter-rater reliability of the measurements. To evaluate the degree of agreement between the two researchers, the ICC was calculated. This statistical measure quantifies how consistently the two raters performed the measurements and indicates the extent to which the obtained values are reproducible across different observers.

2.8.2 Validity

To assess validity, traditional anthropometric measurements (stadiometer, weight scale, and tape measurements) were compared with the corresponding measurements obtained from the 4D Stereophotogrammetry laboratory, 3D Measure Me, and MeThreeSixty systems. The golden standard tape measurement was obtained by the researcher that only focussed on the tape measurement and not the person who also did the LAB. Intraclass correlation coefficients (ICCs) were calculated using

SPSS to evaluate agreement between methods. For each ICC, the 95% confidence interval (minimum and maximum values) and statistical significance were reported.

2.8.3 Accuracy

Lastly, root mean square error (RMSE) will determine accuracy, as it quantifies the average magnitude of the error between two measurements. For a graphical overview of the agreement and identify the systematic bias between the two measurement methods, a Bland-Altman plot was created. This is depicted in the section discussion and provides the visibility of the heterogeneity and distribution.

The formula for RMSE:
$$RMSE = \sqrt{\frac{\sum_{i=1}^N \|y(i) - \hat{y}(i)\|^2}{N}},$$

‘Insert Table 7 here’

Table 7: Interpretation of validity, reliability and accuracy

Validity	Reliability	Accuracy
P-value	ICC	RMSE
Not significant: $P > 0.05$	Poor: $0.0 < ICC < 0.50$	A lower RMSE shows a good prediction between reality and estimation
Significant: $P < 0.05$:	Moderate: $0.50 \leq ICC < 0.75$	
	Good: $0.75 \leq ICC < 0.90$	
	Excellent: $0.90 < ICC < 1$	
	Perfect: 1	
ICC		
Poor: $0.0 < ICC < 0.50$		
Moderate: $0.50 \leq ICC < 0.75$		
Good: $0.75 \leq ICC < 0.90$		
Excellent: $0.90 < ICC < 1$		
Perfect: 1		
Pearson:		
Poor: $r < 0.40$		
Modest: $0.40 \leq r < 0.75$		
Excellent: $r \geq 0.75$		

3 Results

The most important results should be summarized in the text and presented in tables or figures. This section should include only a report of the results, without interpretation or positioning.

2.8. Reliability

The measurements obtained by the two tape measures showed statistically significant reliability ($p < 0.001$) for all variables, except for right waist height and belly height, for which no significant reliability was observed. The ICC values further support these findings, as the same two variables (belly height and right waist height) that were not statistically significant also demonstrated poor reliability. The remaining variables ranged from moderate reliability, observed for left knee girth and right forearm girth, to excellent reliability. Six variables showed good reliability. All others were classified as excellent, except for height, which demonstrated perfect agreement. Overall, the majority of the measurements exhibited excellent reliability. 'Insert Table 8 here'

Table 8: Comparison between tape measurement and tape measurements

Tape measurements – tape measurements	ICC	Min.	Max.	Sig.		ICC	Min.	Max.	Sig.
Waist height (Le)	0.999	0.997	0.999	<.001	Forearm girth (Ri)	0.723	0.520	0.850	<.001
Waist height (Ri)	0.022	-0.309	0.349	0.449	Wrist girth (Le)	0.860	0.742	0.927	<.001
Waist to knee (Le)	0.991	0.982	0.995	<.001	Wrist girth (Ri)	0.884	0.783	0.939	<.001
Waist to knee (Ri)	0.995	0.989	0.997	<.001	Thigh girth (Le)	0.992	0.985	0.996	<.001
Knee height (Le)	0.995	0.991	0.998	<.001	Re thigh girth	0.892	0.792	0.945	<.001
Knee height (Ri)	0.994	0.988	0.997	<.001	Knee girth (Le)	0.664	0.430	0.814	<.001
Outer arm in (Le)	0.994	0.988	0.997	<.001	Knee girth (Ri)	0.947	0.898	0.973	<.001
Outer arm in (Ri)	0.995	0.989	0.989	<.001	Ankle girth (Le)	0.854	0.729	0.923	<.001
Foot length (Le)	0.978	0.957	0.989	<.001	Ankle girth (Re)	0.889	0.781	0.944	<.001
Foot length (Ri)	0.976	0.953	0.988	<.001	Upper arm girth (Le)	0.867	0.753	0.931	<.001
Height	1.000	1.000	1.000	<.001	Upper arm girth (Ri)	0.979	0.960	0.989	<.001
Neck girth	0.991	0.982	0.995	<.001	Belly height	0.066	-0.270	0.387	0.353
Hip girth	0.997	0.993	0.998	<.001	Bust height (Le)	0.963	0.925	0.982	<.001
Belly girth	0.988	0.976	0.994	<.001	Bust height (Ri)	0.953	0.909	0.976	<.001
Bust girth	0.999	0.998	0.999	<.001	Armspan	0.998	0.997	0.999	<.001
Under bust girth	0.999	0.998	1.000	<.001	Head girth	0.966	0.934	0.983	<.001
Back arm pit	0.972	0.945	0.986	<.001	Bust width	0.937	0.880	0.968	<.001
Li forearm girth	0.975	0.940	0.988	<.001					

2.9. Validity

The comparison between tape measurements and the MeThreeSixty app (Table 9) showed statistically significant agreement for chest, left thigh, left biceps and right biceps, whereas no significant agreement was observed for stomach, forearm (left and right), right thigh, and waist. ICC values were generally moderate for the significant measurement sites and low to moderate for the non-significant sites.

Overall, reliability appears acceptable for chest, thigh, and biceps measurements, while agreement for the remaining anatomical sites is limited and should be interpreted with caution.

Overall, the agreement between tape measurements and 3D Measure Me (Table 10) was statistically significant for most anatomical sites. ICC values ranged from low to excellent, with particularly high reliability observed for height, upper arm girths, forearm girths, neck girth, and under-bust girth. Moderate agreement was found for several limb lengths and girths, including thigh and ankle measurements. Lower reliability was observed for hip girth, bust girth, knee girth, and certain height-derived measures. No significant agreement was found for right waist height, hip girth, bilateral bust height, and knee girths. Overall, the findings indicate generally acceptable reliability between tape measurements and 3D Measure Me, although agreement varies depending on the anatomical site.

Agreement between tape measurements and the 4D Stereophotogrammetry (Table 11) was statistically significant for the majority of anatomical sites. ICC values ranged from low to excellent, with particularly high validity observed for height, upper and lower limb girths, under-bust girth, and several arm measurements. Moderate agreement was found for hip, bust, wrist, and knee measurements. Lower reliability was observed for upper thigh girth, belly girth, neck girth, and certain height-derived measures. No significant agreement was found for right waist height and bilateral bust height.

Overall, the results indicate generally good to excellent reliability between tape measurements and the 4D Stereophotogrammetry, with a few site-specific limitations. 'Insert Table 9 here',

Table 9: Comparison between MeThreeSixty and tape measurements

MeThreeSixty – tape measurements	ICC	Min.	Max.	Sig.
Stomach	0.107	-0.087	0.345	0.084
Chest	0.358	0.042	0.612	0.014
Forearm (Le)	0.256	-0.062	0.533	0.058
Forearm (Ri)	0.135	-0.212	0.448	0.221
Thigh (Le)	0.346	0.014	0.608	0.021
Thigh (Ri)	0.053	-0.287	0.378	0.382
Bicep (Le)	0.352	0.034	0.608	0.016
Bicep (Ri)	0.339	0.004	0.603	0.024
Waist	0.228	-0.105	0.517	0.089

‘Insert Table 10 here’, ‘Insert Table 11 here’

Table 10: Comparison between 3D Measure Me and tape measurements

3D Measure Me – tape measurements	ICC	Min.	Max.	Sig.
Waist height (Le)	0.810	0.657	0.899	<.001
Waist height (Ri)	0.025	-0.306	0.352	0.442
Waist to knee (Le)	0.543	-0.049	0.809	<.001
Waist to knee (Ri)	0.216	-0.068	0.487	0.041
Knee height (Le)	0.621	-0.087	0.877	<.001
Knee height (Ri)	0.644	-0.073	0.879	<.001
Outer arm in (Le)	0.667	-0.059	0.887	<.001
Outer arm in (Ri)	0.641	-0.068	-0.068	<.001
Foot length (Le)	0.537	0.124	0.766	<.001
Foot length (Ri)	0.553	0.093	0.786	<.001
Height	0.923	0.829	0.965	<.001
Neck girth	0.798	0.370	0.919	<.001
Hip girth	0.007	-0.287	0.316	0.483
Belly girth	0.298	-0.104	0.634	<.001
Bust girth	0.083	-0.225	0.388	0.303
Under bust girth	0.762	0.200	0.925	<.001
Back arm pit	0.554	0.282	0.746	<.001
Fore arm girth (Le)	0.807	0.646	0.899	<.001
Forearm girth (Ri)	0.667	0.435	0.817	<.001
Wrist girth (Le)	0.525	0.029	0.777	<.001
Wrist girth (Ri)	0.367	-0.101	0.713	<.001
Thigh girth (Le)	0.435	-0.014	0.711	<.001
Thigh girth Ri)	0.438	0.076	0.686	<.001
Knee girth (Le)	0.221	-0.125	0.516	0.103
Knee girth (Ri)	0.281	-0.061	0.561	0.052
Ankle girth (Le)	0.461	0.029	0.722	<.001
Ankle girth (Ri)	0.427	0.029	0.690	<.001
Upper arm girth (Le)	0.839	0.706	0.915	<.001
Upper arm girth (Ri)	0.820	0.671	0.905	<.001
Bust height (Le)	0.155	-0.137	0.487	0.15
Bust height (Ri)	0.171	-0.155	0.516	0.160
Armspan	0.429	0.126	0.662	0,004
Head girth	0.275	-0.098	0.620	<.001
Bust width	0.318	-0.007	0.592	0.016

Table 11: Comparison between 4D Stereophotogrammetry and tape measurements

4D Stereophotogrammetry – tape measurements	ICC	Min.	Max.	Sig.
Waist height (Le)	0.671	-0.37	0.885	<0.001
Waist height (Ri)	0.028	-0.307	0.356	0.436
Waist to knee height (Le)	0.436	-0.072	0.638	<.001
Waist to knee height (Ri)	0.268	-0.069	-0.069	<.001
Knee height (Le)	0.554	-0.084	0.85	<.001
Knee height (Ri)	0.489	-0.071	0.82	<.001
Outer arm in (Le)	0.824	0.312	0.937	<.001
Outer arm in (Ri)	0.767	0.303	0.906	<.001
Height	0.992	0.981	0.996	<.001
Neck girth	0.365	0.06	0.613	0.009
Hip girth	0.688	-0.061	0.899	<.001
Belly girth	0.269	-0.097	0.585	0.001
Bust girth	0.649	0.345	0.818	<.001
Under bust girth	0.882	0.779	0.779	<.001
Back arm pit	0.330	-0.078	0.635	<.001
Forearm girth (Le)	0.953	0.910	0.976	<.001
Forearm girth (Ri)	0.685	0.462	0.827	<.001
Wrist girth (Le)	0.632	0.35	0.802	<.001
Wrist girth (Ri)	0.599	0.336	0.775	<.001
Li upper thigh girth	0.240	-0.053	0.511	0.047
Upper thigh girth (Ri)	0.422	0.012	0.690	<.001
Knee girth (Le)	0.722	0.501	0.852	<.001
Knee girth (Ri)	0.851	0.726	0.922	<.001
Ankle girth (Le)	0.467	-0.101	0.785	<.001
Ankle girth (Ri)	0.518	-0.088	0.807	<.001
Upper arm girth (Le)	0.931	0.765	0.973	<.001
Upper arm girth (Ri)	0.955	0.903	0.978	<.001
Bust height (Le)	0.001	-0.001	0.005	0.228
Bust height (Ri)	0.001	-0.001	0.005	0.284
Head girth	0.519	0.159	0.742	<.001
Front arm pit	0.473	0.003	0.740	<.001

2.10. Accuracy

The accuracy was assessed using the root mean square error (RMSE), which was calculated in Excel. Overall the RMSE for 4D4ALL compared to traditional tape measurements (Table 13) were good for most measurements. However, for torso measurements greater values were found: as example belly girth (120.287 mm). The same trend is noticeable for 3D Measure Me contrasted to tape (Table 12). The extremities achieve a good RMSE, even better than the 4D4ALL that showed more variation. The MeThreeSixty app (Table 14) has structurally higher errors. Especially in the torso (stomach: 110.113 mm) and larger segments such as the right thigh (67.801 mm). 'Insert Table 12 here' 'Insert Table 13 here' 'Insert Table 14 here'

Table 13: Comparison between 4D4ALL and tape measurements

4D Stereophotogrammetry – tape measurements	RMSE (mm)
Waist height	62.237
Waist to knee height	68.715
Knee height (Le)	40.729
Knee height (Ri)	44.936
Outer arm in (Le)	24.249
Outer arm in (Ri)	27.950
Height	12.821
Neck girth	45.800
Hip girth	52.263
Belly girth	120.287
Bust girth	64.185
Under bust girth	40.875
Back arm pit	43.817
Forearm girth (Le)	6.978
Forearm girth (Ri)	20.424
Wrist girth (Le)	9.473
Wrist girth (Ri)	9.337
Upper thigh girth (Le)	78.079
Upper thigh girth (Ri)	64.758
Knee girth (Le)	14.138
Knee girth (Ri)	10.962
Ankle girth (Le)	20.198
Ankle girth (Ri)	17.768
Upper arm girth (Le)	11.248
Upper arm girth (Ri)	9.335
Head girth	19.553

Table 14: Comparison between Me Three Sixty and tape measurements

MeThreeSixty – tape measurements	RMSE (mm)
Stomach	110.113
Chest	56.912
Forearm (Le)	21.130
Forearm (Ri)	24.582
Thigh (Le)	39.975
Thigh (Ri)	67.801
Bicep (Le)	25.702
Bicep (Ri)	27.182
Waist	68.661

Table 12: Comparison between 3D Measure Me and tape measurements

3D Measure Me – tape measurements	RMSE (mm)
Waist height (Le)	43.170
Waist height (Ri)	41.306
Waist to knee (Le)	35.941
Waist to knee (Ri)	71.304
Knee height (Le)	38.619
Knee height (Ri)	34.457
Outer arm in (Le)	35.813
Outer arm in (Ri)	35.715
Foot length (Le)	19.022
Foot length (Re)	19.785
Height	42.112
Neck girth	21.185
Hip girth	197.665
Belly girth	117.461
Bust girth	166.310
Back arm pit	25.241
Forearm girth (Le)	14.884
Forearm girth (Ri)	20.869
Wrist girth (Le)	15.889
Wrist girth (Ri)	19.094
Thigh girth (Le)	62.429
Thigh girth (Ri)	64.387
Knee girth (Le)	60.3222
Knee girth (Ri)	53.416
Ankle girth (Le)	20.012
Ankle girth (Ri)	20.612
Upper arm girth (Le)	16.512
Upper arm girth (Ri)	18.262
Armspan	57.458
Head girth	35.964
Bust width	54.658

4 Discussion

This study evaluated the validity, reliability, and accuracy of smartphone-based 3D scanning applications and the 4D stereophotogrammetry system in comparison with traditional anthropometric measurements. Overall, the findings indicate that both the 4D stereophotogrammetry and the 3D Measure Me application demonstrate acceptable to excellent agreement with tape measurements across many anthropometric variables, while MeThreeSixty consistently showed lower agreement and reduced accuracy.

In terms of validity, intraclass correlation coefficients (ICC) were calculated, with p-values being predominantly significant for both the 4D stereophotogrammetry and the 3D Measure Me application. The highest level of consistency with tape measurements was observed for the 4D stereophotogrammetry, with ICC values ranging from good to excellent across most variables. The 3D Measure Me application also demonstrated substantial agreement with traditional measurements; however, this agreement was generally less consistent and more variable depending on the anatomical region. MeThreeSixty, on the other hand, showed lower validity across multiple outcomes, reflected in predominantly low ICC values.

Notably, only a few variables consistently failed to reach statistical significance across methods. Measurements such as left and right bust height and right waist height were not significant in both the 4D stereophotogrammetry and 3D Measure Me, suggesting that these parameters may be more difficult to capture in the female subjects accurately regardless of the technology used. This may be related to challenges in identifying anatomical landmarks or variations in posture during scanning.

Upon examining reliability, expressed as ICC values, the results indicate that the majority of measurements fall within a range of 0.90 to 1.00, corresponding to excellent reliability. This suggests that, for most parameters, the measurements can be considered highly consistent across raters. However, notable exceptions were observed for right waist height (0.022) and belly height (0.066), both demonstrating poor reliability and indicating substantial variability between measurements. As previously discussed in relation to validity, right waist height also performed poorly in terms of accuracy, further highlighting the difficulty of consistently capturing this parameter.

These two parameters were the only ones associated with non-significant p-values, suggesting a lack of consistent agreement beyond chance. In contrast, all other measurements showed statistically

significant p-values, supporting the presence of systematic agreement between raters. In addition, height yielded a perfect ICC value of 1.00, indicating extremely high consistency across measurements. Furthermore, slightly lower ICC values were observed for right forearm girth (0.723) and left knee girth (0.664). Although these values fall below the threshold for excellent reliability, they still indicate moderate to good reliability and suggest that these measurements are reasonably consistent, although with greater variability compared to the other parameters.

Overall, the findings suggest that the tape measurement reference method demonstrates generally good inter-rater reliability. Nevertheless, the variability observed in specific parameters indicates that certain anatomical measurements may be more challenging to assess consistently, and these should therefore be interpreted with appropriate caution.

Regarding accuracy, expressed as the root mean square error (RMSE), both the 4D stereophotogrammetry and 3D Measure Me demonstrated relatively low error values for measurements of the extremities, whereas larger errors were consistently observed for trunk-related outcomes. The 4D stereophotogrammetry showed a more stable performance across variables, with RMSE values ranging from 6.978 mm to 120.287 mm. In comparison, 3D Measure Me exhibited greater variability, with RMSE values ranging from 14.884 mm to 197.665 mm, indicating less consistent accuracy across different measurement sites.

For MeThreeSixty, a clear pattern of increased error was observed for larger body segments. While smaller measurements such as left and right biceps showed relatively low RMSE values (25.702 mm and 27.182 mm, respectively), substantially higher errors were found for trunk-related measures, such as the stomach (RMSE = 110.113 mm). This suggests that the accuracy of MeThreeSixty decreases as the size and shape of the measured body region increase.

For the best-performing method, the various Bland-Altman plots indicate how closely the measured values match the reference values. In all outcomes, the left arm girth consistently scored best, with the lowest RMSE. There are no discernible errors in the means (middle line) for the various plots, which are -0.743 (mm) for 4D stereophotogrammetry (Figure 1), -5.914 (mm) for MeThreeSixty (Figure 3) and -4.743 (mm) for 3DMeasureMe (Figure 2). The 4D stereophotogrammetries confidence intervals are between -1.454 and 1.305. This illustrates the distribution of individual difference around the mean. These are shown as the bottom and top horizontal lines. The intervals for MeThreeSixty are -4.625 and 3.442. 3DMeasureMe agreement limits range from -3.279 to 2.330. The majority of variations are within this range, indicating that the approach is useful in real-world situations. . 'Insert Figure 1 here' 'Insert Figure 2 here' 'Insert Figure 3 here'

Figure 1: Tape measurement - 4D stereophotogrammetry: forearm girth (Le)

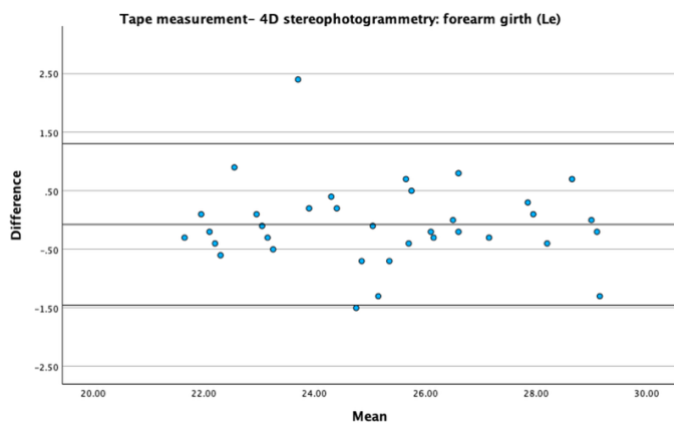


Figure 2: 3D Measure Me- Tape measurement: forearm girth (Le)

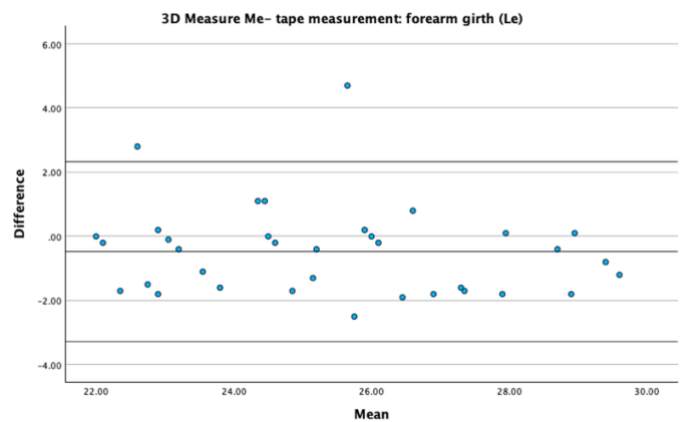
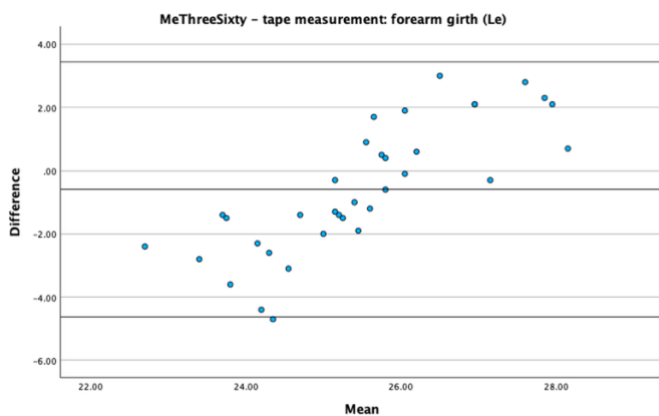


Figure 3: MeThreeSixty- tape measurement: forearm girth (Le)



We also provided a paired T-test for the measurements of the left forearm girth. Men can conclude that 3D Measure Me, 4D stereophotogrammetry and MeThreesixty app generally scores negative compared to tape. This can be concluded out of the negative mean. However, none show a statistically significant difference ($p > 0.05$). This means that there is no convincing evidence for a systematic deviation between traditional and 3D Measure Me ($p = 0.0582$), 4D stereophotogrammetry ($p = 0.537$) and MeThreesixty app ($p = 0.982$). Based on this value, it can be concluded that the scanners do not consistently score higher or lower for this value, but the p-value between traditional and 3D Measure Me is a borderline. When examining trunk outcomes, there was a significant difference ($p < 0.01$) between the smartphone apps and traditional measurements, with a large mean (e.g., waist circumference > 90 mm). However, this is not surprising since the values here were more deviant, as mentioned earlier.

Regarding the differences between the apps, MeThreeSixty used the iPhone's front camera and identifies two features: front and side view. However, the detection was not always easy with women. The lines often did not correspond with their arms. Possibly due to hyperextension in the elbows. The 3D Measure Me takes use of the back camera and takes four features: front, both sides and back view. Furthermore, compared to the MeThreeSixty app, the 3D Measure me offers a much greater number of results and measurements.

In order to ensure the most accurate measurements, the participants wore as little clothing as possible. Men frequently wore shorts and a bare torso, while women typically wore a sports bra and a fitting short. However, the men's shorts were generally a little looser.

Also, the participants skin tones were noted. One participant had dark skin, four had light tans and the rest had white skin.

Lastly, the presence of movement during the process was examined. One participant showed signs of slight movement. Nevertheless, this had no detrimental effect on the image materials quality in the 4D stereophotogrammetry.

5 Conclusion

Overall, the statistical analyses indicate that the measurements obtained from the different methods showed a generally good level of agreement and consistency. The 4D stereophotogrammetry demonstrated good to excellent validity and reliability, followed by the 3D Measure Me application, which also yielded favourable results. In contrast, MeThreeSixty consistently showed lower performance across outcomes.

Considering the clinical applicability of smartphone-based applications, the findings suggest that these tools may serve as a practical and accessible alternative for anthropometric measurements. However, important differences between applications must be taken into account. Based on the present findings, the use of 3D Measure Me appears to be preferable over MeThreeSixty, given its superior performance in terms of usability, validity, reliability, and accuracy. Furthermore, the use of such applications is particularly recommended for measurements of the extremities, where higher levels of agreement were observed. In contrast, caution is warranted when interpreting trunk-related measurements due to the systematic errors identified.

Several limitations of this study should be acknowledged. Although the sample size was determined a priori using a correlation-based calculation and can be considered sufficient, the relatively small

sample of 37 participants may limit generalizability. In addition, variability in clothing, posture, and individual body characteristics may have influenced the measurements. While participants were instructed to wear tight-fitting clothing, this was not fully standardized, resulting in some variation. Furthermore, factors such as posture and individual characteristics, vary between participants and are difficult to fully control within the measurement process. Finally, potential inaccuracies in the identification of anatomical landmarks may have contributed to measurement error.

Future research should focus on improving the detection of anatomical landmarks to enhance measurement accuracy, particularly for trunk-related outcomes. In addition, further development of algorithms that are robust across different body types, sexes, and postural variations is necessary. Moreover, larger and more heterogeneous samples are necessary to improve the robustness and generalizability of findings. Finally, the implementation of a more standardized measurement protocol, including uniform clothing and fixed camera positioning, is necessary to further enhance the reliability and validity of these technologies.

6 References

- (1) Bartol, K., Bojanić, D., Petković, T., Pribanić, T., Department of Electronic Systems and Information Processing, Faculty of Electrical Engineering and Computing, University of Zagreb, 10000 Zagreb, Croatia, Croatian Science Foundation, & Kristijan Bartol. (2021). A review of body measurement using 3D scanning. In Mohamad Forouzanfar (Ed.), *IEEE Access* (Vol. 9, p. 67281) [Journal-article]. <https://doi.org/10.1109/ACCESS.2021.3076595>
- (2) Bougourd, J., Treleaven, P., & University College London. (2010). UK National Sizing Survey - SizeUK. In *International Conference on 3D Body Scanning Technologies*, *International Conference on 3D Body Scanning Technologies*. https://proc.3dbody.tech/papers/2010/10327_32bougourd.pdf
- (3) Hodgdon, J., Martine Brandsma, Daanen, H., & Hervig, L. (2005). Prediction of body surface area based on 3-Dimensional laser scans. In T.R.U.E. Research Foundation & U.S. Army Medical Research and Materiel Command, T.R.U.E. Research Foundation. T.R.U.E. Research Foundation.
- (4) Rumbo-Rodríguez, L., Sánchez-SanSegundo, M., Ferrer-Cascales, R., García-D'Urso, N., Hurtado-Sánchez, J. A., & Zaragoza-Martí, A. (2021). Comparison of body scanner and manual anthropometric measurements of body shape: A Systematic review. *International Journal of Environmental Research and Public Health*, 18(12), 6213. <https://doi.org/10.3390/ijerph18126213>
- (5) Heymsfield, S. B., Bourgeois, B., Ng, B. K., Sommer, M. J., Li, X., & Shepherd, J. A. (2018). Digital anthropometry: a critical review. *European Journal Of Clinical Nutrition*, 72(5), 680–687. <https://doi.org/10.1038/s41430-018-0145-7>
- (6) Rumbo-Rodríguez, L., Sánchez-SanSegundo, M., Ferrer-Cascales, R., García-D'Urso, N., Hurtado-Sánchez, J. A., & Zaragoza-Martí, A. (2021). Comparison of Body Scanner and Manual Anthropometric Measurements of Body Shape: A Systematic Review. *International Journal Of Environmental Research And Public Health*, 18(12), 6213. <https://doi.org/10.3390/ijerph18126213>
- (7) Reliability and agreement of 3D anthropometric measurements in facial palsy patients using a Low-Cost 4D imaging system. (2020, August 1). *IEEE Journals & Magazine | IEEE Xplore*. <https://ieeexplore.ieee.org/document/9133557>
- (8) Bougma, K., Mei, Z., Palmieri, M., Onyango, D., Liu, J., Mesarina, K., Akelo, V., Mwando, R., Zhou, Y., Meng, Y., & Jefferds, M. E. (2022). Accuracy of a handheld 3D imaging system for child anthropometric measurements in population-based household surveys and

- surveillance platforms: an effectiveness validation study in Guatemala, Kenya, and China. *American Journal of Clinical Nutrition*, 116(1), 97–110. <https://doi.org/10.1093/ajcn/nqac064>
- (9) Susato, S. (2011). Development and Application of Portable Manual Non-Contact-Type Anthropometric Instruments for Measuring Human Anatomical Longitudinal Parameters. *Journal Of PHYSIOLOGICAL ANTHROPOLOGY*, 30(2), 55–67. <https://doi.org/10.2114/jpa2.30.55>
- (10) Lopez, M. G., Matthes, K. L., Sob, C., Bender, N., & Staub, K. (2023). Associations between 3D surface scanner derived anthropometric measurements and body composition in a cross-sectional study. *European Journal of Clinical Nutrition*, 77(10), 972–981. <https://doi.org/10.1038/s41430-023-01309-4>
- (11) Aydin, A. C., Karakol, P., Sulhan, A., Erk, H., & Bozkurt, M. (2024). Comparison of anthropometric and cephalometric measurements obtained by stereophotogrammetry and 3D computed tomography of the nose before septorhinoplasty. *Aesthetic Plastic Surgery*, 48(21), 4314–4322. <https://doi.org/10.1007/s00266-024-04097-9>
- (12) Graybeal, A. J., Brandner, C. F., Compton, A. T., Swafford, S. H., Henderson, A., Aultman, R., Vallecillo-Bustos, A., & Stavres, J. (2023). Smartphone derived anthropometrics: Agreement between a commercially available smartphone application and its parent application intended for use at point-of-care. *Clinical Nutrition ESPEN*, 59, 107–112. <https://doi.org/10.1016/j.clnesp.2023.11.021>
- (13) Hobbs Murphy, K. H. M. (2024, 16 april). Intra-rater and inter-rater reliability of 3D facial measurements. ScienceDirect. <https://www.sciencedirect.com/science/article/pii/S0003687023002569?via%3Dihub>
- (14) Liu, L. Y., Sangani, S., Patterson, K. K., Fung, J., & Lamontagne, A. (2020). Real-Time Avatar-Based feedback to enhance the symmetry of spatiotemporal parameters after stroke: Instantaneous effects of different avatar views. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 28(4), 878–887. <https://doi.org/10.1109/tnsre.2020.2979830>
- (15) Scataglini, S., Dens, A., Laermans, E., Naalden, A., & Truijen, S. (2025). Exploring and analysing 3D and 4D camera based markerless technologies for human body anthropometrics in healthy individuals: a systematic review. *International Journal of Medical Informatics*, 205, 106084. <https://doi.org/10.1016/j.ijmedinf.2025.106084>
- (16) Scataglini, S., Van Orden, F., Van Den Hoek, E., Stessel, M., & Truijen, S. (2025). Smartphone and tablet-based 3D scanner for anthropometric assessments in adults: a systematic review of

- reliability, validity, and accuracy. BioMedical Engineering OnLine, 25(1), 3. <https://doi.org/10.1186/s12938-025-01480-8>
- (17) 3D Measure Me. (2025b, oktober 11). 3D Measure Me Home Page - 3D Measure Me. 3D Measure Me - Get Your Body Measurements With The Touch Of A Button Using Your Smartphone. <https://3dmeasureme.com/>
- (18) MeThreeSixty | 3D Body Visualizer. (n.d.). MeThreeSixty. <https://www.methreesixty.com/>
- (19) G*Power. (2025, June 12). <https://www.psychologie.hhu.de/arbeitsgruppen/allgemeine-psychologie-und-arbeitspsychologie/gpower>
- (20) A balanced Latin square design allows each participant to complete all measurement conditions while controlling for order and carry-over effects across treatments (Dean & Voss, 1999; Cotton, 1993; Jones & Kenward, 1989).
- (21) Susato, S. (2011). Development and Application of Portable Manual Non-Contact-Type Anthropometric Instruments for Measuring Human Anatomical Longitudinal Parameters. Journal Of PHYSIOLOGICAL ANTHROPOLOGY, 30(2), 55–67. <https://doi.org/10.2114/jpa2.30.55>