Selecting Appropriate 3D Scanning Technologies for Prosthetic Socket Design and Transtibial Residual Limb Shape Characterisation

(Appropriate 3D Scanners for Socket Design & Shape Characterisation)

Alexander S. Dickinson[1*], Maggie K. Donovan-Hall[2], Sisary Kheng[3,4], Ky Bou[3,4], Auntouch Tech[3,4], Joshua W. Steer[1], Cheryl D. Metcalf[2], Peter R. Worsley[5]

1: Bioengineering Science Research Group, Faculty of Engineering and Physical Sciences, University of Southampton, UK

2: Active Living and Rehabilitation Group, Faculty of Environmental and Life Sciences, University of Southampton, UK

3: Exceed Worldwide, Phnom Penh, Cambodia

4: Department of Prosthetics and Orthotics, National Institute of Social Affairs, Cambodia

5: Skin Health Research Group, Faculty of Environmental and Life Sciences, University of Southampton, UK

* Corresponding Author:

Bioengineering Science Research Group, Mechanical Engineering Department,

Faculty of Engineering and Physical Sciences,

University of Southampton,

Highfield,

Southampton, SO17 1BJ

United Kingdom

Tel: +442380595394

Email: alex.dickinson@soton.ac.uk

Funding: the Engineering and Physical Sciences Research Council (EPSRC) / National Institute for Health Research (NIHR) Global Challenges Research Fund (grants EP/R014213/1 & EP/N02723X/1), and the Royal Academy of Engineering (RAEng grant RF/130).

ABSTRACT:

Introduction: Plaster casting and manual rectification represent the benchmark prosthetic socket design method. 3D technologies have increasing potential for prosthetic limb design and fabrication, especially for enhancing access to these services in lower and middle income countries (LMICs). However, the community has a responsibility to verify the efficacy of these new digital technologies. This study's objective was to assess the repeatability of plaster casting *in vivo*, specifically for clinically-relevant residuum shape and landmark capture, and to compare this with three clinically-used 3D scanners.

Materials and Methods: A comparative reliability assessment of casting and 3D scanning was conducted in eleven participants with established transtibial amputation. For each participant, two positive moulds were cast by a prosthetist and digitised using a white light 3D surface scanner. Between casts, each participant's residuum was scanned. The deviation between scan volumes, cross-sections and shapes was calculated.

Results: 95% of the clinically-relevant socket shape surface area had a deviation between manual casts <2.87mm (S.D. 0.44mm). The average deviation by surface area was 0.18mm (S.D. 1.72mm). The repeatability coefficient of casting was 46.1ml (3.47%) for volume, and 9.6mm (3.53%) for perimeters. For all clinically-meaningful measures, greater reliability was observed for the Omega scanner, and worse for the Sense and iSense scanners, although it was observed that the Sense scanner performance was comparable to casting (95^{th} percentile shape consistency).

Conclusions: This study provides a platform to appraise new clinical shape capture technologies in the context of best practice in manual plaster casting, and starts the conversation of which 3D scanning devices are most appropriate for different types of clinical use. The methods and benchmark results may support prosthetists in acquiring and applying their clinical experience, as part of their continuing professional development.

KEYWORDS:

amputee, CAD/CAM, consistency, prosthetic socket, plaster casting, rectification, shape error, transtibial, volume error

Introduction

A prosthetic limb user's functional outcome depends fundamentally upon a comfortable and robust human-prosthesis interface[1], which most commonly features a personalised socket. This is especially important after transtibial amputation where individuals may attempt to be more active than those with higher-level amputations, and where the socket-limb load transfer is particularly influenced by the underlying bony anatomy. A variety of transtibial socket design strategies exist, most notably patella-tendon bearing (PTB) and total surface bearing (TSB) approaches. According to PTB principles the residual limb is loaded proportionally to the load tolerance of the underlying soft tissue and bony areas. Despite studies on different aspects of transtibial sockets and residual limbs, there is a lack of knowledge to enable consistent manufacturing of a comfortable socket and optimal alignment without the need for iterative socket fittings[2].

Plaster casting and manual rectification is considered the benchmark shape capture and socket design method. It remains the technique whereby the majority of sockets are designed prior to conventional manufacturing routes[3], creating a standard against which new technologies should be measured. A Plaster of Paris (POP) wrap cast is manually applied over the residuum with the aim to capture a modified shape of the soft tissues. Prosthetists shape the POP during casting for the PTB socket using their hands, to create areas of load bearing around the tibial plateau. This shape is used to produce a positive mould, which is subsequently rectified according to similar design principles as listed above. These procedures can be highly individual and are based on the experience, skill, and preference of the individual prosthetist and their patient[4].

CAD/CAM technologies (Computer Aided Design / Manufacturing) are established in some communities for residuum shape capture, prosthetic socket design and fabrication, with claimed advantages including higher consistency and a perpetual digital design record.

Perceived disadvantages include a clinician learning curve and high capital equipment costs, although the development of lower cost 3D scanning devices has been proposed in an attempt to overcome this barrier[5,6]. Other low-cost devices have been proposed for socket reproduction, including smartphone-based photogrammetry[7]. However there is relatively little evidence in the scientific literature for the accuracy or reliability of these lower cost devices, either in absolute terms or in comparison to clinically meaningful benchmarks. Benchmark measures might be taken from research on manual plaster methods, the traditional and most frequently used approach. The consistency of plaster cast rectification has been investigated in terms of the location and depth of focal rectifications[4], and the influence of prior activity on the residuum's volume and shape[8]. Others have compared hands-off vs. hands-on casting methods in terms of the cast shape radius in a manikin model[9], and the cast shape volume and length[10]. However, more understanding is needed regarding the benchmark metrics relating to the reliability of both clinically-relevant shape metrics and volumetric parameters, against which to compare 3D scanning technologies.

This study's objective was to conduct an *in vivo* assessment of the repeatability of plaster casting specifically for residuum shape capture (i.e. pre-rectification), employing high accuracy and resolution CAD/CAM scanning and digitised shape analysis techniques[5,11,12] The motivation was to investigate a comprehensive set of clinically-relevant shape metrics, and provide benchmarking data to assess digital shape capture technologies. The work was approached from a global challenges research perspective, where people may seek lower cost technologies to improve P&O access in lower and middle income countries (LMICs), at the potential expense of accuracy and reliability. Therefore, the study was conducted with prosthetists in an ISPO-certified Cambodian P&O school and clinic.

MATERIALS AND METHODS

An assessment of the reliability of transtibial residual limb casting was conducted, for comparison to the reliability of 3D scanning with three different devices. Approval was granted by institutional (ERGO 25100) and national ethics boards (Cambodian National Ethics Committee for Health Research 073NECHR). Participants were recruited by convenience sampling from a single prosthetics centre. All participants' residual limbs were cast twice during one session, by one of two ISPO-certified prosthetists (authors AT and KB). Negative plaster casts were produced according to the prosthetists' normal practice when producing a patella tendon bearing, supracondylar suspended socket (Figure 1 A-F). The prosthetists then converted their negative casts into positive moulds and performed light surface abrasion using wire mesh (Figure 1 G). The positive mould shapes were digitised using a structured white light surface scanner (Go!SCAN (Creaform Inc., Lévis, Canada), which was previously shown to have a surface height accuracy of 0.2 mm ± 0.07 mm (mean ± standard deviation error) on a similar object[5]. Between casts, each participant's residual limb was scanned by two observers (ASD and PRW), to produce 3D .stl surface mesh files (Figure 1 H). The study used three scanners in a randomised order: the Creaform Go!SCAN device (equivalent to structured white light Omega scanner, Ohio WillowWood Company, USA), the 2nd generation Sense scanner and the iSense / Structure Sensor (3DSystems, USA).

Figure 1: Participants were cast seated, with a cellophane wrap on their residual limb (A, B). Indentations were marked by palpation either side of the patellar tendon (C) and at the supracondylar level (D). After doffing (E) the posterior shelf and flare for knee flexion was formed with additional plaster (F). Positive moulds were produced and lightly abraded (G). Between casts, the limb was scanned directly (H).

According to established methods[5,11,13], in the AmpScan open-source software package[14], pairs of scan files were compared. The shapes were aligned in 3D space, using both manual and automated approaches. Rigid registration was used to match the pairs of positive mould scans over each other, to assess the pairwise deviation. The shapes were sliced serially from the distal tip to the supracondylar ridge at 1% intervals, and the volume and cross-sectional profile dimensions were calculated using the enclosed cross-section area of each slice, its perimeter length, and the maximal widths in the coronal and sagittal planes.

Shape deviation was analysed further using a 'height' deviation, presenting the surface-surface normal deviation data following the visualisation standard set by Sanders and Severance[15]. Deviations were mapped between each cast and a direct scan of the participant's residual limb, and between the pairs of repeat manual casts and scans.

Quantitative analysis was conducted in several ways, all using MATLAB (MathWorks, USA), including:

1. Surface shape repeatability was characterised by calculating average and 95th area-percentile surface height deviation between aligned cast or scan pairs (i.e. 95% of the surface area deviated between scans by this value or less).
2. An Intraclass Correlation Coefficient (ICC) was calculated for reliability[16,17] for the eleven volume pairs, and for eleven pairs of perimeter measures from the mid-length of each shape. The ICC(1,1) equation was used for intra-rater repeatability of casting, and the ICC(3,1) equation for inter-rater reproducibility of scanning.
3. Bland Altman plots[18] were used to assess mean and within-subject differences between volume and perimeter measurements, indicating potential bias and changes in variance with measurement size, producing study population mean and standard deviation values for each. One volume measure for each participant was plotted, and 9 perimeters (at 10% intervals over the proximal 90%).
4. Where the Bland Altman plots showed no changes in bias or variance with measurement size, a repeatability coefficient (CR) was calculated as $\sqrt{2}$ x 1.96 x the standard deviation, to give the boundary within which a repeat measurement would lie with 95% probability[19].
5. Finally, to provide context to clinically relevant calliper and tape measurements, the pairwise mean absolute difference, root mean squared difference and Pearson correlation coefficients were calculated from the width and perimeter profiles along the shape lengths.

RESULTS

Eleven people with established (>2yrs), unilateral transtibial amputation were recruited and provided informed, written consent. All were male and had their amputation following traumatic injury resulting from landmine or road traffic accident; ten had unilateral amputation, and one bilateral.

The cast shapes represented a modification of the limb shape in several key regions (*Figure 2*A&B), and were similar for all participants. These included focal indentation either side of the patella tendon (ref. Figure 1C), medial and lateral supracondylar indentation (ref. Figure 1D), and relief posteriorly for the hamstring (ref. Figure 1F). Positive shape change was observed distally at the scar site and anteriorly over the tibia (i.e. the cast was larger than the limb in these areas).

Comparison of the repeat casts from the prosthetists revealed a high level of reliability (*Figure 2* left). Greatest surface height deviation between casts was observed in regions where the most substantial shape modifications were introduced during casting. The largest deviations were observed in the posterior block corresponding with the hamstring cut-out, introduced after the cast was removed from the residuum (ref. Figure 1F). The other notable region of deviation between cast shapes was on the distal posterior aspect associated with the calf muscles and scar site.

QUANTITATIVE ANALYSIS 1: SURFACE HEIGHT DEVIATION BETWEEN CASTS AND SCANS:

The cast pairs were compared quantitatively (Table 1). There was no apparent systematic error or bias between first and second casts (across participants, mean surface height error -0.18 mm, range from -0.70 to +0.55 mm). Consistency between cast pairs of local shape capture and modification is represented by the standard deviation in surface height over the surface area (across participants, mean 1.72 mm, range 1.07 to 2.16 mm). On average across the participants, 95% of the shape surface area had an absolute deviation between casts of <3.60 mm (S.D. 0.81 mm). In addition, 95% of the surface area most clinically relevant to socket-limb loading (patella tendon to distal tip, below the eventual socket brim) had a deviation between casts of <2.87 mm (S.D. 0.44 mm).

		Pairwise Cast Deviations /mm			**Pairwise Go!SCAN / Omega Scan Deviations /mm**		**Pairwise Sense Scan Deviations /mm**		**Pairwise iSense Scan Deviations /mm**	
Participant	**Mean Cast Volume (l)**	**Raw mean (s.d.)**	**Absolute 95th percentile**		**Raw mean (s.d.)**	**Absolute 95th percentile**	**Raw mean (s.d.)**	**Absolute 95th percentile**	**Raw mean (s.d.)**	**Absolute 95th percentile**
			Full Surface	**Patella Tendon to Distal Tip**						
1	1.299	0.55 (2.16)	4.05	3.21	-0.02 (0.39)	0.71	-0.57 (1.04)	2.14	-0.01 (2.34)	4.40
2	1.477	0.34 (2.15)	4.60	3.24	0.10 (0.54)	1.05	1.02 (1.03)	2.87	-	-
3	1.649	0.31 (1.52)	3.48	2.80	0.00 (0.69)	1.34	0.21 (1.21)	2.54	0.45 (2.02)	3.68
4	1.109	-0.18 (1.17)	2.75	2.34	0.07 (0.34)	0.74	-0.61 (1.11)	2.42	-1.76 (2.16)	5.22
5	1.296	-0.70 (1.94)	4.31	3.43	-0.20 (0.42)	0.94	-0.04 (0.84)	1.78	-0.30 (1.39)	2.86
6	1.010	-0.14 (1.95)	4.74	3.04	0.09 (0.49)	0.96	0.69 (1.44)	3.16	0.51 (1.53)	2.77
7	1.628	0.08 (1.54)	3.27	3.01	-0.16 (0.42)	0.86	-0.01 (1.51)	2.68	-0.74 (1.69)	3.56
8	1.299	-0.56 (1.95)	4.09	2.95	-0.15 (0.81)	1.42	-0.59 (1.07)	2.68	-0.52 (1.69)	3.93
9	0.807	-0.37 (1.66)	3.22	3.19	0.07 (0.52)	1.00	0.13 (1.26)	2.60	-0.17 (1.39)	2.95
10	1.383	-0.21 (1.07)	2.36	2.10	0.05 (0.42)	0.81	-0.98 (1.44)	3.14	0.11 (2.01)	3.86
11	1.154	0.01 (1.75)	2.70	2.28	-0.10 (1.51)	1.58	0.99 (1.13)	2.89	0.03 (1.01)	2.02
Mean (s.d.)	**1.283 (0.255)**	**-0.18 (1.72)**	**3.60 (0.81)**	**2.87 (0.44)**	**-0.02 (0.60)**	**1.04 (0.29)**	**0.02 (1.19)**	**2.63 (0.41)**	**-0.22 (1.57)**	**3.53 (0.92)**

Table 1: Reliability of residual limb cast and scan surface height measurement expressed as mean (s.d.) raw deviation and 95th percentile absolute deviation between cast or scan pairs, by area.

The scanned shapes represented a non-contact characterisation of the residual limb shape, i.e. without any soft tissue manipulation or pre-rectification landmarking. Comparison

of the repeat scans revealed differing reliability between devices (Figure 2 right), and less spatial trend in surface height deviation was observed between scans than between casts. The scan pairs were compared quantitatively (Table 1). The Omega scanner was more reliable than casting, the Sense scanner was similar and the iSense scanner less reliable. Systematic error or bias between first and second scans for any device was small compared to the scanner's corresponding consistency (standard deviation in surface height deviation over the surface area of 0.60 mm for Omega, 1.19 mm for Sense and 1.57 mm for iSense). On average across the participants, 95% of the surface area from patella tendon to distal tip (below the eventual socket brim) had a deviation between casts of <1.04 mm (S.D. 0.29 mm) for Omega, <2.63 mm (S.D. 0.41 mm) for Sense and <3.53 mm (S.D. 0.92 mm) for iSense.

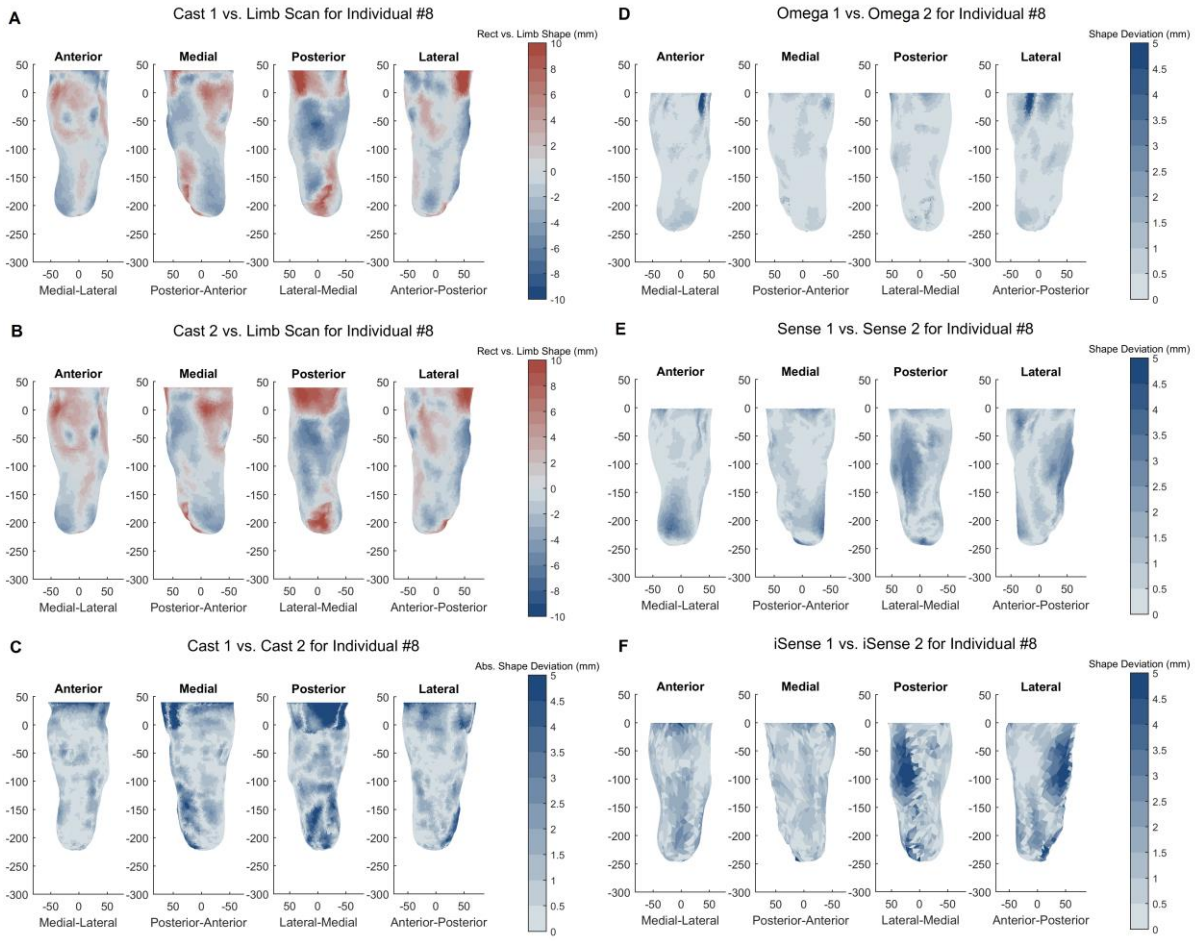

Figure 2: Scan surface deviation plots for one example participant: casts 1 and 2 vs. limb scan (A&B), and absolute deviation between cast 1 vs. cast 2 (C). Absolute deviations between scans are plotted for the Omega (D), Sense (E) and iSense (F) scanners.

213 *QUANTITATIVE ANALYSIS 2-4: VOLUME AND SERIAL SECTION PERIMETER*

214 *DEVIATION BETWEEN CASTS*

215 The intra-prosthetist residual limb casting repeatability (Table 2) was calculated from
216 raw cast volume data, and perimeters at 10% increments along the cast length, represented on
217 Bland-Altman plots (Figure 3, Figure 4). Casting and all scanners had very high ICC(1,1)
218 scores, above 0.977. No volume or perimeter bias was observed between first and second
219 casts or scans with any device. The standard deviation of pairwise volume differences of 16.6
220 ml equates to a repeatability coefficient of 46.1 ml or 3.47% (Figure 3 top left). The
221 repeatability coefficient for cast perimeters was 9.6 mm or 3.53% (Figure 4 top left). For both
222 measures, casting lay between the Omega and Sense scanners for repeatability, with the
223 iSense scanner less repeatable.

		Absolute /ml		Relative /%			
	Reliability Test (Cast 1 vs. Cast 2 or Scan 1 vs. Scan 2)	Pairwise Difference (mean ± SD)	Repeatability Coefficient	Pairwise Difference (mean ± SD)	Repeatability Coefficient	ICC	*t*-test p-Value
Volume	Cast	-1.1 ± 16.6	46.1	-0.17 ± 1.25	3.47	0.998*	<0.001
	Go!SCAN / Omega	-1.0 ± 8.8	24.3	-0.02 ± 0.65	1.81	0.999	<0.001
	Sense	27.8 ± 28.5	78.8	2.02 ± 1.89	5.23	0.994	<0.001
	iSense	21.0 ± 44.2	122.4	1.67 ± 3.23	8.94	0.980	<0.001
Perimeter	Cast	0.46 ± 3.46	9.60	0.15 ± 1.27	3.53	0.997*	<0.001
	Go!SCAN / Omega	0.08 ± 1.77	4.91	0.00 ± 0.61	1.69	0.999	<0.001
	Sense	4.04 ± 5.53	15.32	1.33 ± 1.81	5.03	0.994	<0.001
	iSense	-0.10 ± 7.24	20.07	0.09 ± 2.56	7.10	0.978	<0.001

224 Table 2: Intra-rater reliability statistics on volume and perimeter measures obtained from digital
225 measures from Omega scans of cast pairs (n=11), and direct limb scan pairs by Omega (n=11), Sense
226 (n=11) and iSense (n=10). * Cast-cast intra-rater reliability used equation ICC(1,1); scan-scan inter-
227 rater reliabilities used equation ICC(3,1).

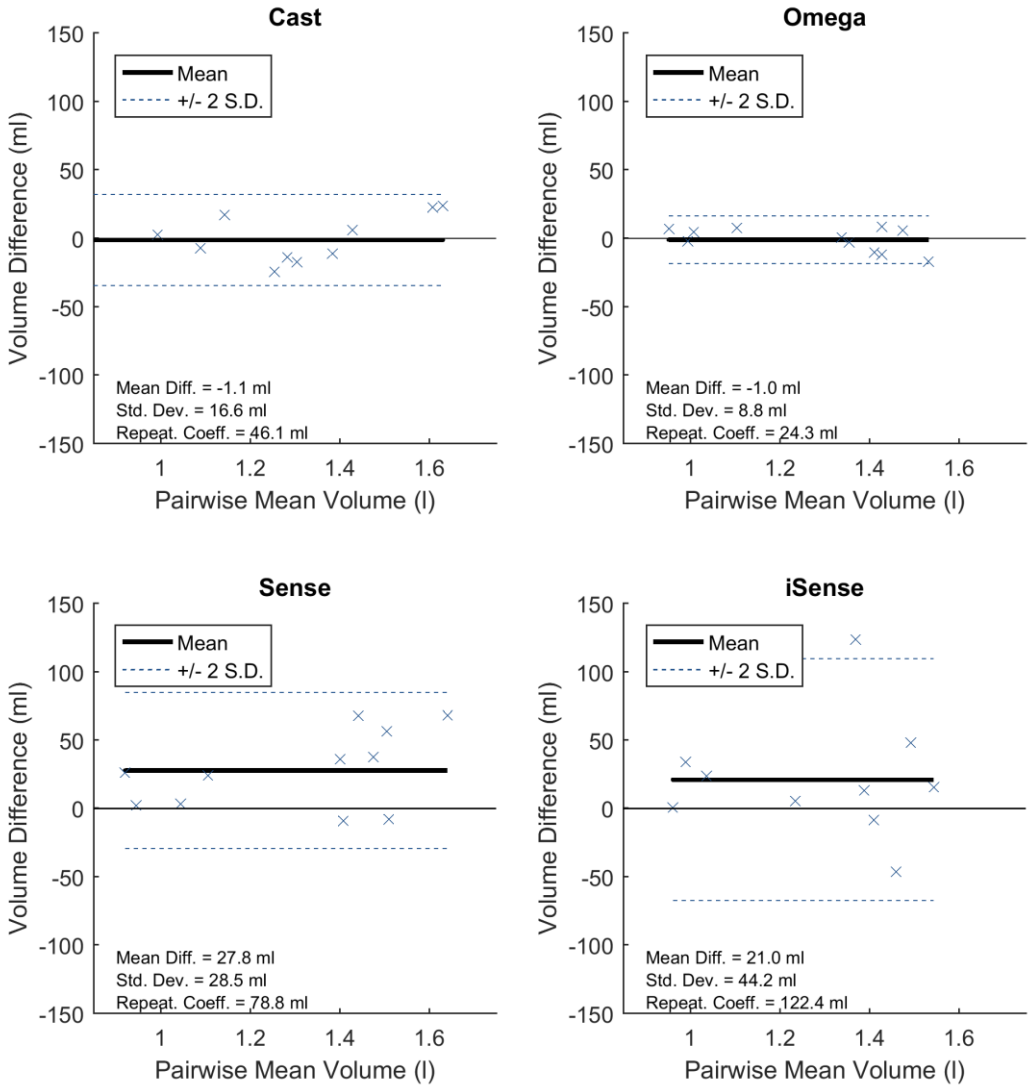

Figure 3: Bland-Altman plots of pairwise difference in volume measures for casting and the three scanners (one measure per limb).

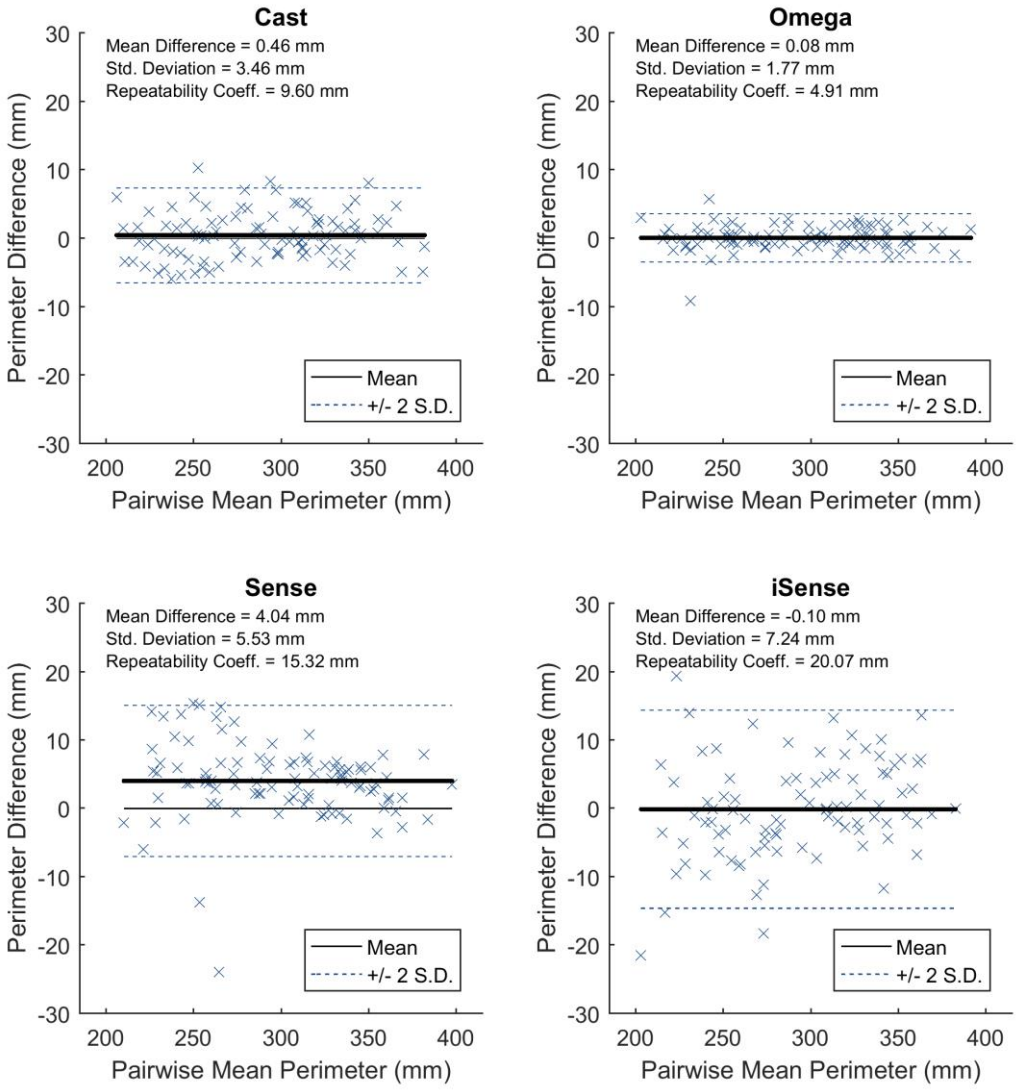

Figure 4: Bland-Altman plots of pairwise difference in perimeter measures for casting and the three scanners (9 measures per limb, at 10% intervals along length).

QUANTITATIVE ANALYSIS 5: WIDTH AND PERIMETER PROFILE DEVIATION BETWEEN CASTS

Width and perimeter profiles along the residuum length were used for further inspection of the shape deviation between casts and scans (Figure 5, Table 3). All cast 1 vs. cast 2 measures were highly correlated (r = 0.993-0.999), and all width measurements deviated by less than 3.25 mm. Slightly greater difference between measures was observed for the coronal (M-L) compared to sagittal width (A-P) measurements. This may be attributed

to the residuum having lower flexibility in the anterior-posterior direction than the medial-lateral direction, owing to the bony tibial crest. Conversely, no difference was observed between coronal and sagittal width reliability for the non-contact scan-based measurements. The Sense and iSense scanners had similar median and interquartile range deviations, both larger than casting, and the Omega scanner was again observed to be the most reliable tool.

		Absolute Difference /mm		Relative Difference /%		
	Measure (Cast 1 vs. Cast 2)	Root Mean Squared	Mean Absolute	Root Mean Squared	Mean Absolute	Pearson Correlation Coefficient
Sagittal Width (Anterior-Posterior)	Cast	1.355 (1.098 - 1.722)	1.055 (0.863 - 1.448)	1.634 (1.158 - 2.051)	1.266 (0.927 - 1.660)	0.997 (0.996 - 0.999)
	Go!SCAN / Omega	0.553 (0.391 - 0.729)	0.427 (0.310 - 0.521)	0.634 (0.383 - 0.682)	0.383 (0.327 - 0.613)	1.000 (0.999 - 1.000)
	Sense	2.311 (1.347 - 3.663)	2.005 (1.101 - 3.228)	2.802 (1.451 - 4.171)	2.159 (1.136 - 3.949)	0.999 (0.998 - 1.000)
	iSense	2.519 (1.663 - 3.580)	2.238 (1.428 - 3.197)	2.930 (1.701 - 4.559)	2.659 (1.416 - 4.337)	0.999 (0.999 - 0.999)
Coronal Width (Medial-Lateral)	Cast	2.123 (1.312 - 2.420)	1.630 (1.069 - 1.992)	2.101 (1.465 - 2.420)	1.643 (1.236 - 2.005)	0.994 (0.993 - 0.995)
	Go!SCAN / Omega	0.735 (0.646 - 0.829)	0.507 (0.468 - 0.677)	0.714 (0.635 - 0.811)	0.550 (0.458 - 0.631)	0.999 (0.998 - 1.000)
	Sense	2.042 (1.446 - 2.516)	1.635 (1.186 - 2.179)	2.337 (1.476 - 2.700)	1.779 (1.205 - 2.352)	0.998 (0.996 - 0.999)
	iSense	2.516 (2.009 - 3.048)	2.213 (1.798 - 2.655)	2.651 (2.112 - 3.070)	2.267 (1.898 - 2.594)	0.997 (0.995 - 0.998)
Perimeter	Cast	3.121 (2.673 - 3.589)	2.324 (1.948 - 2.472)	1.003 (0.885 - 1.116)	0.823 (0.728 - 0.898)	0.998 (0.997 - 0.999)
	Go!SCAN / Omega	1.477 (1.310 - 1.870)	0.786 (0.465 - 1.047)	0.534 (0.450 - 0.638)	0.410 (0.356 - 0.498)	0.999 (0.999 - 1.000)
	Sense	5.088 (4.213 - 8.023)	4.616 (3.733 - 6.377)	1.841 (1.551 - 2.974)	1.511 (1.201 - 2.260)	0.997 (0.998 - 0.999)
	iSense	5.043 (4.329 - 8.000)	3.404 (2.382 - 5.474)	1.684 (1.447 - 2.849)	1.350 (1.153 - 2.327)	0.996 (0.995 - 0.997)

Table 3: Intra-rater reliability statistics for clinically-relevant measures of width and cross-section perimeter profiles. Data were non-parametric, so presented as median (interquartile range).

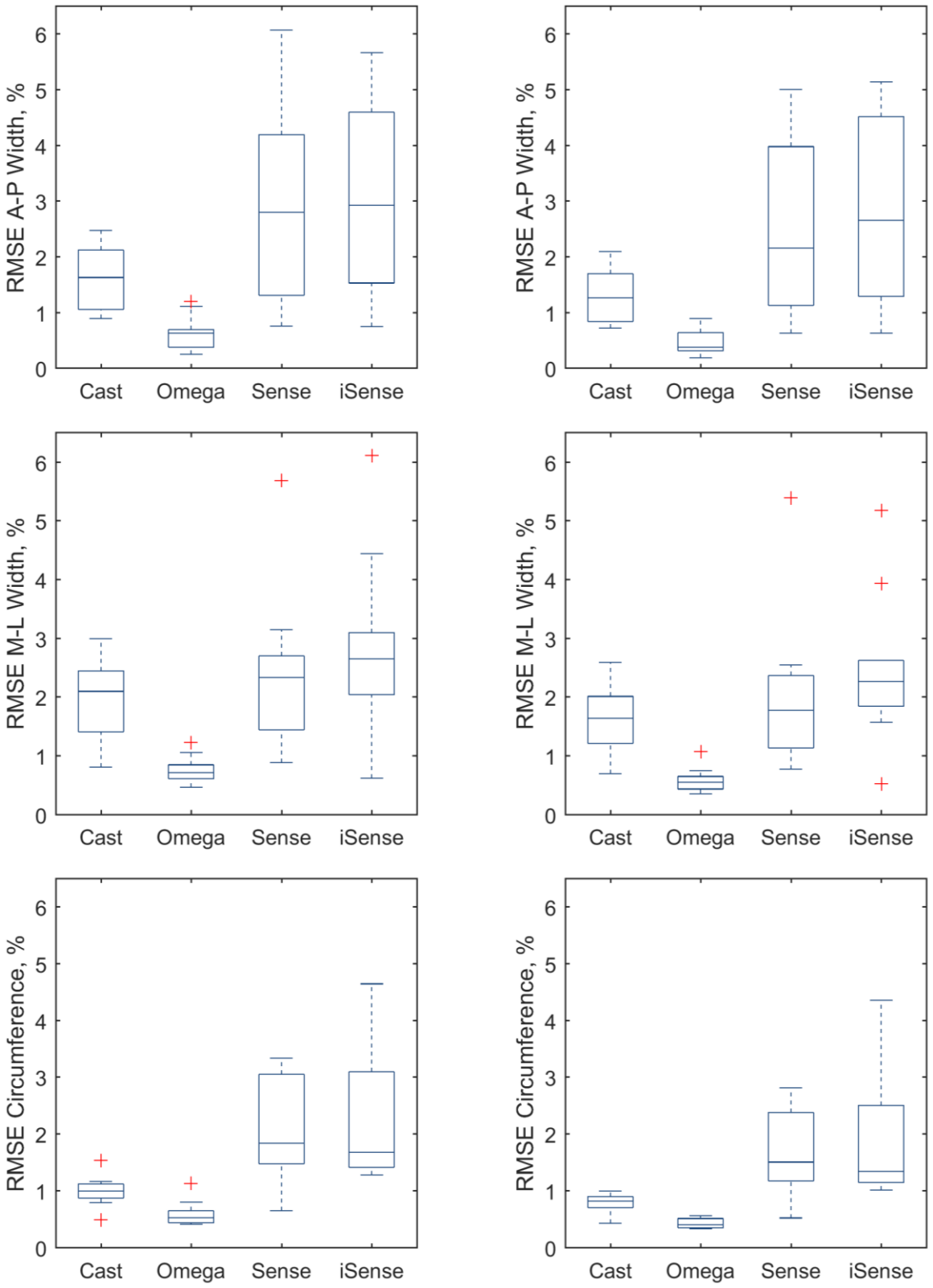

Figure 5: Intra-rater reliability of casting and scanning for clinically relevant measures of sagittal and coronal plane widths, and perimeters, expressed by box and whisker plots.

DISCUSSION

This study set out to conduct a new, high-resolution assessment of the repeatability of plaster casting, using gross volume and detailed shape metrics, specifically in the context of clinically relevant residuum shape capture. The intention was to provide stringent but appropriate benchmarking data for reliability of other technologies including 3D scanners and pressure casting. Exemplar measures were conducted for three 3D scanners in current clinical and research use, selected to cover a range from established, high accuracy devices to low cost consumer products. Participants with transtibial amputation were recruited to assess manual plaster cast and CAD/CAM methods in prosthetic device design and fabrication. In an effort to contribute most clearly to the clinical community the present study's results are presented in a consistent manner to the standards of data visualisation established by Sanders and Severance[15], and a consensus of the statistical methods presented by Seminati et al[13] and Kofman et al[20].

Considering consistency in both detailed shape measures and gross volume and perimeter measures, the high specification Omega scanner was more reliable than casting, and the Sense scanner was similar or less reliable. The iSense scanner was the least reliable in all tests. Considering clinically-comparable measures of residuum width and perimeter profiles, the Sense and iSense scanners had similar deviations, both larger than casting, and the Omega scanner was again observed to be the most reliable tool.

To provide clinical context, we can compare casting variability to clinically manageable volume changes. Lilja and Öberg[21] proposed that the volume change corresponding to donning one (+5%) and two stockings (+10%) is clinically significant, as a new socket is typically prescribed once two stockings are required. Sanders and colleagues reported that a limb volume change of ~6% (simulated by a uniform ±1.8 mm socket surface

offset)[22] may produce clinically detectable effects on gait, quality of fit, comfort and satisfaction measures. In the present study the repeatability coefficients of casting volume and perimeter were both ~3.5%, below these indicated limits for clinical significance. All these measurements' ICC scores also comfortably exceeded the 0.90 threshold for clinically-relevant reliability[23]. These values were also below that reported for volume change upon muscle activation during casting, with a mean deviation of +5.5% (range -4.2% to +14.2%)[24]. A recent study identified similar variability in volume and perimeter measures for transtibial sockets produced by casting and the Biosculptor CAD/CAM method, for a single individual[25], and the absolute volume reliability measures were slightly better than achieved using water displacement[26].

Considering the more detailed shape measures, the mean surface height deviation in serial castings was larger in 6 of the 11 participants than the +0.25 mm socket manufacturing bias (mean radial error) reported by Sanders et al[27] as clinically noticeable (Table 1). However, the present data includes the more variable residuum tip, and deviation was less than the thickness of a 1-ply sock[28] in all cases. The mean height error was smaller than the +0.25 mm bias for all 11 participants with the Omega scanner, but larger for 7/11 participants with the Sense and for 6/10 participants with the iSense. This study's high resolution shape deviation mapping extends prior casting reliability investigations which considered global metrics including volume, length and cross section area[9,10] to compare casting methods, and analysis of rectifications *subsequent* to the shape capture itself [4]. The most notable prior application of these shape deviation mapping techniques to casting was conducted by Sanders et al, addressing the specific question of the influence of the time delay between an activity protocol and casting[8].

Considering other clinical measurement tools, the width and perimeter measurement reliability for casting was in the same range or slightly lower than the calliper, tape measure

and anthropometer data presented by Geil[29]. Shape and volume reliability data were comparable to those obtained from existing CAD scanning technologies[5,13,20,30] and the lower cost scanners were comparable to smartphone photogrammetry for digitising sockets[7]. The standard deviation and 95th percentile surface height difference between casts, indicative of the greatest local variability in shape capture, were similar to the focal rectification consistency data reported by Convery et al[4].

The present study's main limitation is a relatively small convenience sample with restricted inclusion criteria, representing amputation due to trauma only (landmines and road traffic accidents), and as such may have limited generalisability. Indeed, it should be noted that the presented data are relevant to transtibial residual limb casting only, and may not be directly applied to other amputation levels or orthoses that have greater reliance upon bony prominences. A recent review[31] concluded that shape capture tools "require more consistent 'gold standards'" and highlighted a lack of CAD assessments on residual limbs, where most prior work has used models. This study offers *in vivo* benchmarks, and the participant cohort represents a stringent benchmarking test for new technologies, as several of the more established participants displayed slender, long residual limbs with clear bony prominences, and would be expected to produce more consistent contact-based shape measurements than individuals with more fleshy residua.

This study only included two prosthetists, although one was a recent graduate and the other had 18 years' clinical experience. Further, we did not address inter-prosthetist variability, as it is likely that different prosthetists would each produce differently shaped sockets that would achieve acceptable user comfort. The study also considers variability arising from plaster casting for residuum shape capture and landmarking for subsequent rectification, but not variability in the final rectification features. The presented data are intended for use in appraisal of other shape capture methods; consistency of subsequent

rectifications may be compared to the results presented by Convery et al[4]. As a final key limitation, it is acknowledged that the relationship between socket comfort and socket fit is not understood[31], and assessing user satisfaction with sockets produced by the different approaches was outside this study's remit.

In an LMIC clinical context, important discussion should consider the specific use of 3D scanning and other CAD/CAM technologies. There are risks associated with embedding these digital technologies in established manual plaster-based socket design and fabrication workflows, and with present CAD/CAM technologies the crucial sustainable maintenance, servicing and replaceability factors of Appropriate Technologies may not yet be met[32]. There are different use cases for these 3D technologies short of adopting a full CAD/CAM workflow, such as i) replication of a well-fitting but damaged or lost manual socket[7], ii) detailed residual limb volume and shape surveying, and iii) remote assessment for individuals who cannot easily attend P&O clinics. Regarding socket replication, the 'well-fitting' caveat is key, as the residuum's volume and shape are known to change over time as well as fluctuate, and it is uncommon for sockets to wear out. Anecdotally, some clinicians perform residuum shape capture by plaster casting and then proceed to digitise the cast by 3D scanning, prior to CAD rectification and CNC fabrication. Different levels of scanning accuracy and reliability will be necessary for these different cases, with greatest accuracy required for socket replication and cast digitisation. Lower specification devices may offer benefits instead for enhancing LMIC P&O service access for people living in remote communities, but not at the expense of providing an accurate, well-fitting socket. There is a moral imperative to ensure that only appropriate technology is deployed, which has been tested and validated for the particular P&O service in question, irrespective of geographical or financial constraints. Furthermore, it is essential that these digital methods and devices, high or low cost, are not seen as a replacement for clinician training. As a fundamental

principle of ethical research and development, all new technologies must be proven before clinical use, as we must prove the prosthetic devices themselves.

This study's results support the established body of evidence around plaster casting as the benchmark of prosthetic socket design, whereby expert prosthetists apply their skill and experience, and against which novel technologies should be measured. Three case-study devices in current clinical use were compared, and different use-cases proposed. Appraisal of new technologies should consider quantitative accuracy and reliability metrics alongside qualitative usability factors[20] and ultimately prosthesis user outcome measures[10,33] and acceptability scoring[34]. This study presents one element of the data to support such appraisal and selection of appropriate technologies. The shape capture and measurement method may also support prosthetists in measuring their plasterwork skills as part of their training and continuing professional development.

ACKNOWLEDGEMENTS

The authors are grateful to funders the Engineering and Physical Sciences Research Council (EPSRC) / National Institute for Health Research (NIHR) Global Challenges Research Fund (grants EP/R014213/1 & EP/N02723X/1), and the Royal Academy of Engineering (RAEng grant RF/130).

The authors thank Exceed Worldwide for facilitating, and Thearith Heang, Carson Harte and Sam Simpson of the Exceed Research Network (ERN) for providing critical review. We also thank the University of Southampton's Institute for Life Sciences / FortisNet interdisciplinary musculoskeletal research network for supporting our preliminary work.

We have no conflicts of interest relevant to this study.

Data availability Supporting data are openly available from the University of Southampton repository at https://doi.org/10.5258/SOTON /D0381 NOTE: TO BE ACTIVATED UPON PUBLICATION.

REFERENCES

1. Legro MW, Reiber G, Del Aguila MD, et al. Issues of importance reported by persons with lower limb amputations and prostheses. *J Rehabil Res Dev*. 1999;36(3):155-163.

2. Trower TA. Changes in lower extremity prosthetic practice. *Phys Med Rehab Clin North Am*. 2006;17:23-30.

3. ICRC. *Trans-Tibial Prosthesis Manufacturing Guidelines*. Geneva; 2006.

4. Convery PP, Buis AWP, Wilkie R, Sockalingam S, Blair A, McHugh B. Measurement of the consistency of patellar-tendon-bearing cast rectification. *Prosthet Orthot Int*. 2003;27(3):207-213. doi:10.1080/03093640308726683

5. Dickinson AS, Steer JW, Woods CJ, Worsley PR. Registering a methodology for imaging and analysis of residual-limb shape after transtibial amputation. *J Rehabil Res Dev*. 2016;53(2). doi:10.1682/JRRD.2014.10.0272

6. Armitage L, Kwah LK, Kark L. Reliability and validity of the iSense optical scanner for measuring volume of transtibial residual limb models. *Prosthet Orthot Int*. 2019;43(2):213-220. doi:10.1177/0309364618806038

7. Hernandez A, Lemaire E. A smartphone photogrammetry method for digitizing prosthetic socket interiors. *Prosthet Orthot Int*. 2017;41(2):210-214. doi:10.1177/0309364616664150

8. Sanders JE, Severance M, Swartzendruber D, Allyn K, Ciol M. Influence of prior activity on residual limb volume and shape measured using plaster casting: results from individuals with trans-tibial limb loss. *J Rehabil Res Dev*. 2013;50(7):1007-1016.

9. Buis AWP, Blair A, Convery P, Sockalingam S, McHugh B. Pilot study: Data-capturing consistency of two trans-tibial casting concepts, using a manikin stump model: A comparison between the hands-on PTB and hands-off ICECAST Compact® concepts. *Prosthet Orthot Int*. 2003;27(2):100-106. doi:10.1080/03093640308726665

10. Safari MR, Rowe P, McFadyen A, Buis A. Hands-off and Hands-on casting consistency of amputee below knee sockets using magnetic resonance imaging. *Sci World J*. 2013;2013. doi:10.1155/2013/486146

11. Zachariah SG, Sorenson E, Sanders JE. A method for aligning trans-tibial residual limb shapes so as to identify regions of shape change. *IEEE Trans Neural Syst Rehabil Eng*. 2005;13(4):551-557. doi:10.1109/TNSRE.2005.858459

12. Sanders JE, Rogers EL, Sorenson EA, Lee GS, Abrahamson DC. CAD/CAM transtibial prosthetic sockets from central fabrication facilities: How accurate are they? *J Rehabil Res Dev*. 2007;44(3):395-405. doi:10.1682/JRRD.2006.06.0069

13. Seminati E, Talamas DC, Young M, Twiste M, Dhokia V, Bilzon JLJ. Validity and reliability of a novel 3D scanner for assessment of the shape and volume of amputees' residual limb models. *PLoS One*. 2017;12(9):1-16. doi:10.1371/journal.pone.0184498

14. Steer JW, Stocks O, Worsley PR, Dickinson AS. AmpScan: Open-source scanning analysis for prosthetics and orthotics. In: *Prosthetics and Orthotics International*. Vol Suppl. 1. ; 2019:94. https://ampscan.readthedocs.io/en/latest/index.html.

15. Sanders JE, Severance MR. Assessment technique for computer-aided manufactured sockets. *J Rehabil Res Dev*. 2011;48(7):763-774. doi:10.1682/JRRD.2010.11.0213

16. Shrout PE, Fleiss JL. Intraclass Correlations: uses in assessing rater reliability. *Psychol Bull*. 1979;86(2):420-428. doi:10.1007/978-0-387-78665-0_5945

17. Rankin G, Stokes M. Reliability of assessment tools in rehabilitation: An illustration of appropriate statistical analyses. *Clin Rehabil*. 1998;12(3):187-199. doi:10.1191/026921598672178340

18. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986:307-310. doi:10.1016/S0140-6736(86)90837-8

19. Vaz S, Falkmer T, Passmore AE, Parsons R, Andreou P. The Case for Using the Repeatability Coefficient When Calculating Test-Retest Reliability. *PLoS One*. 2013;8(9):1-7. doi:10.1371/journal.pone.0073990

20. Kofman R, Beekman AM, Emmelot CH, Geertzen JHB, Dijkstra PU. Measurement properties and usability of non-contact scanners for measuring transtibial residual limb volume. *Prosthet Orthot Int*. 2018;42(3):280-287. doi:10.1177/0309364617736088

21. Lilja M, Oberg T. Proper time for definitive transtibial prosthetic fitting. *J Prosthetics Orthot*. 1997;9(2):90-95.

22. Sanders JE, Youngblood RT, Hafner BJ, et al. Effects of socket size on metrics of socket fit in trans-tibial prosthesis users. *Med Eng Phys*. 2017;44:32-43. doi:10.1016/j.medengphy.2017.03.003

23. Portney LG. *Foundations of Clinical Research: Applications to Practice*. Third Edit. Upper Saddle River NJ: Prentice Hall Inc.; 2007.

24. Lilja M, Johansson S, Öberg T. Relaxed versus activated stump muscles during casting for trans-tibial prostheses. *Prosthet Orthot Int*. 1999;23(1):13-20. doi:10.3109/03093649909071606

25. Mehmood W, Abd Razak NA, Lau MS, Chung TY, Gholizadeh H, Abu Osman NA. Comparative study of the circumferential and volumetric analysis between conventional casting and three-dimensional scanning methods for transtibial socket: A preliminary study. *Proc Inst Mech Eng Part H J Eng Med*. 2019;233(2):181-192. doi:10.1177/0954411918816124

26. Van Stuivenberg C, De Laat F, Meijer R, Van Kuijk A. Inter- and intra-observer reproducibility and validity of an indirect volume measurement in transtibial amputees. *Prosthet Orthot Int*. 2010;34(1):20-30. doi:10.3109/03093640902929285

27. Sanders JE, Severance MR, Allyn KJ. Computer-socket manufacturing error: How much before it is clinically apparent? *J Rehabil Res Dev*. 2012;49(4):567-582. doi:10.1682/JRRD.2011.05.0097

28. Cagle JC, D'Silva KJ, Hafner BJ, Harrison DS, Sanders JE. Amputee socks: Sock thickness changes with normal use. *Prosthet Orthot Int*. 2016;40(3):329-335. doi:10.1177/0309364614568412

29. Geil MD. Consistency and accuracy of measurement of lower-limb amputee anthropometrics. *J Rehabil Res Dev*. 2005;42(2):131-140.

458 doi:10.1682/JRRD.2004.05.0054

459 30. Sanders JE, Fatone S. *Residual Limb Volume Change: Systematic Review of*
460 *Measurement and Management*. Vol 48.; 2011. doi:10.1682/JRRD.2010.09.0189

461 31. Suyi Yang E, Aslani N, McGarry A. Influences and trends of various shape-capture
462 methods on outcomes in trans-tibial prosthetics: A systematic review. *Prosthet Orthot*
463 *Int*. 2019;43(5):540-555. doi:10.1177/0309364619865424

464 32. Dickinson A, Donovan-hall M, Kheng S, et al. Technologies to Enhance Quality and
465 Access to Prosthetics & Orthotics: the importance of a multidisciplinary, user-centred
466 approach. In: *Global Report on Assistive Technology (GReAT) Consultation*. Geneva:
467 World Health Organisation; 2019. doi:https://doi.org/10.5258/SOTON/P0014

468 33. Karakoç M, Batmaz I, Sariyildiz MA, Yazmalar L, Aydin A, Em S. Sockets
469 Manufactured by CAD/CAM Method Have Positive Effects on the Quality of Life of
470 Patients with Transtibial Amputation. *Am J Phys Med Rehabil*. 2017;96(8):578-581.
471 doi:10.1097/PHM.0000000000000689

472 34. Fuller M, Fraser J, Czyniewski S. Clinical Trial of Digitally Fabricated Ankle & Foot
473 Orthoses. In: *Prosthetics and Orthotics International*. Vol Suppl. 1. ; 2019:95.

474