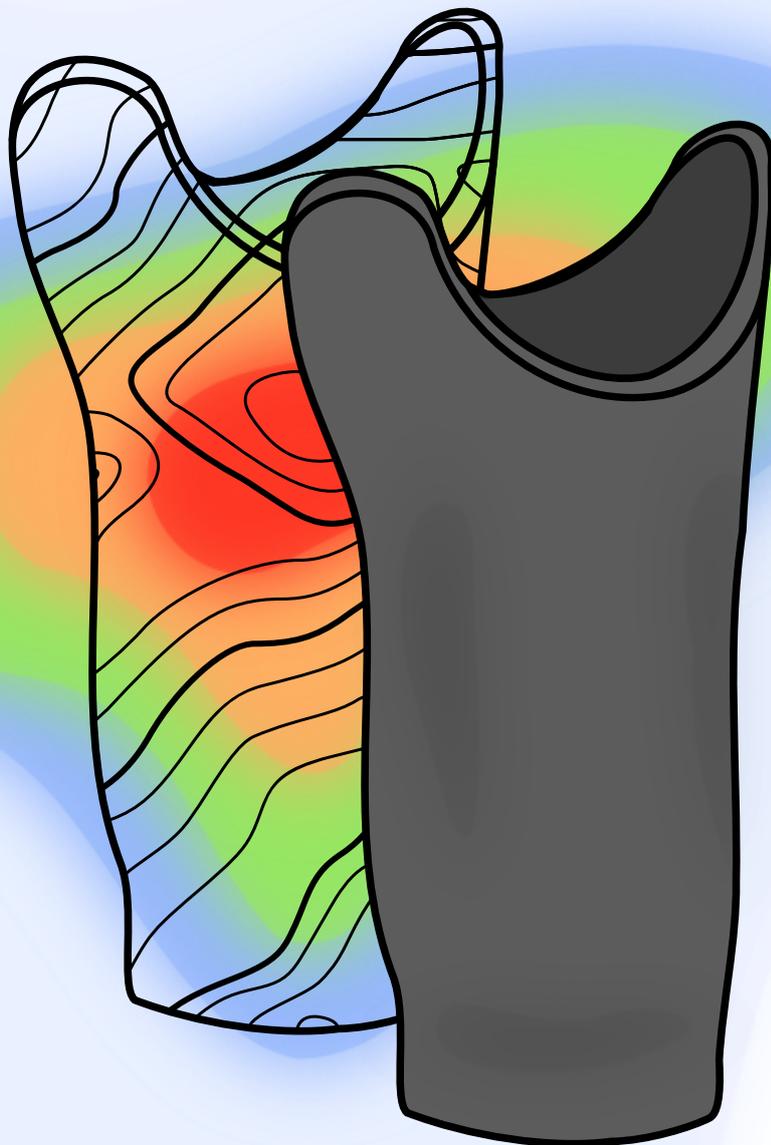


INFRARED THERMOGRAPHY OF THE RESIDUAL LIMB AS A NOVEL METHOD TO MEASURE SOCKET FIT

a proof-of-concept study



Pien van Gastel



Universiteit
Leiden

TUDelft Delft
University of
Technology

Erasmus
ERASMUS UNIVERSITEIT ROTTERDAM

INFRARED THERMOGRAPHY OF THE RESIDUAL LIMB AS A NOVEL METHOD TO MEASURE SOCKET FIT

a-proof-of-concept study

Pien, van Gastel

Student number : 5389879

15 August 2024

Thesis in partial fulfilment of the requirements for the joint degree of Master of Science in

Technical Medicine

Leiden University ; Delft University of Technology ; Erasmus University Rotterdam

Master thesis project (TM30004 ; 35 ECTS)

Dept. of Biomechanical Engineering, TUDELFT

6 Nov 2023 –15 Aug 2024

Feasibility study (15 ECTS): 20 May -26 Jul 2024

Supervisors:

Dr. Ruud Leijendekkers

Dr. ir. Gerwin Smit

Drs. Merel van der Stelt

Thesis committee members:

Dr. ir. Gerwin Smit, TU Delft (chair)

Dr. Ruud Leijendekkers , Radboudumc

Prof. dr. ir. Jaap Harlaar, Erasmus University Medical Center

Prof. dr. Ruud Selles, Erasmus University Medical Center

An electronic version of this thesis is available at <http://repository.tudelft.nl/>.



Universiteit
Leiden

TUDelft
Delft
University of
Technology

Erasmus
ERASMUS UNIVERSITEIT ROTTERDAM

Preface

I am very pleased to present my thesis for the conclusion of the Master's in Technical Medicine. This thesis represents not just the end of this Master's journey, but also the completion of a decade of education, including a Bachelor's in Medicine and a pre-master in Technical Medicine. Reflecting on these past ten years fills me with happiness, pride, countless learning opportunities, and significant personal growth.

This graduation project has seamlessly combined my passion for people, medicine, technology, and creativity. I very much enjoyed the physiotherapy sessions with my colleagues, patients, and especially with the other half of our dynamic duo, Pim. These sessions taught me invaluable lessons about patient care, communication, and dedication, which will remain with me throughout my career.

As a final chapter of my studies, I had the opportunity to travel to Sierra Leone, where I continued to work on the 3D Prosthetics project as an extension of this research. This experience allowed me to explore my enduring interest in tropical medicine within the context of medical technology. It was an incredibly enriching period, filled with stories worthy of an entire book.

I want to express my heartfelt gratitude to my RadboudUMC supervisors, Merel and Ruud, and my TU Delft supervisor, Gerwin, for their guidance, advice, and honest feedback over the past year. Merel, your passion for the Sierra Leone project is inspiring and has motivated me to contribute to the fullest extent. Thank you for your support. Ruud, your trust in my ability to manage my own patients was invaluable, and our insightful discussions sparked many new solutions (and questions), which I deeply appreciate. Gerwin, you ensured I stayed aligned with academic requirements amid the evolving project, sharing your expertise in the prosthetics field. I also want to thank all my colleagues at the 3D Lab and at the physiotherapy for their support. Specifically, Pim, Iris, and Celena, thank you for making every day at the 3D Lab enjoyable and for all the fun we had during our many coffee breaks and lunch runs.

Now, let's continue with the next step!

Amsterdam, August 2024

Abstract

Background: Currently, there is no accurate method to objectively assess transtibial prosthetic socket fit. This study aims to evaluate the use of infrared thermography (IRT) to analyse temperature changes in response to pressure points, and their relation to pain and comfort assessments. This study also explores the reliability and agreement parameters of IRT.

Methods: A within-subject study was conducted on seven participants wearing a vacuum or pin suspension socket. Each participant was examined with two sockets: their prescribed socket (unmodified) and their prescribed socket modified with added pressure pads inside the prosthetic socket at the fibular head and distal tibia end. Thermal images were obtained after 10 minutes of walking with the unmodified socket (TW1), after a subsequent 15-minute resting period (TR1), after 10 minutes of walking with the modified socket (TW2), and after a second resting period (TR2). Difference maps showing temperature changes relative to the baseline (TR1) were created. Subjective assessments included localised pain and Socket Comfort Scores. Regions of interest (ROI) for temperature analysis were selected at the fibular head, distal tibia end, popliteal fossa, and gastrocnemius belly. Reliability and agreement were assessed using the Intraclass Correlation Coefficient (ICC), standard error of measurement (SEM), SEM%, smallest detectable change (SDC), SDC%, and Bland-Altman plots.

Results: Visual inspection and ROI temperature values revealed hotspots in 86% (6/7) of participants at the fibular head post-walking with the modified socket, contrasting with the unmodified socket in the majority of cases. Pain corresponded with the presence of hotspots at this location. Hotspots at the distal tibia end post-walking with the modified socket were observed in 43% (3/7) of participants, but without apparent difference from the unmodified socket. Only two-thirds of these coincided with reported pain. Reliability of IRT was very good to excellent across the residual limb and ROIs ($ICC \geq 0.85$), except for fair reliability at the fibular head ROI ($ICC 0.55$). For TR1 and TR2 test-retest agreement, SEM ranged from 0.3 °C (1.0%) to 0.7 °C (2.1%), with SDC from 0.9 °C (2.8%) to 2.0 °C (5.9%). On the difference maps of unmodified and modified sockets, SEM ranged from 0.6 to 0.8 °C, and SDC from 1.7 to 2.1 °C.

Conclusion: Based on the observed thermal response to pressure in the majority of cases, IRT demonstrated potential in identifying peak pressure points and pain through increased temperature measurements at the fibular head, but had limited capability in detecting these at the distal tibia end. This limitation is likely due to differences in circulation, socket suspension effects, and uncertainty about the exerted pressure. Future research should simplify the protocol to enhance clinical applicability, enabling IRT to complement subjective socket fit assessments and improve prosthetic socket designs and patient outcomes.

Contents

Preface	ii
Abstract	iii
Nomenclature	vi
1 Introduction	1
1.1 Clinical background	1
1.2 Measurement of socket fit	2
1.3 Infrared thermography	2
1.4 Objectives	2
2 Methods	4
2.1 Participants	4
2.2 Study design	4
2.3 Measurement instruments and setup	4
2.3.1 Infrared camera system	4
2.3.2 Thermographic markers	5
2.3.3 Measurement setup	5
2.3.4 Subjective assessment tools	6
2.4 Experimental procedure	6
2.4.1 Data collection moments	6
2.4.2 Procedure steps	6
2.5 Data analysis	7
2.5.1 Image registration	7
2.5.2 Residual limb temperatures	7
2.5.3 Relation with subjective assessments	8
2.5.4 Reliability and agreement parameters	8
2.5.5 Software	8
3 Results	9
3.1 Participant characteristics	9
3.2 Experimental conditions and registration accuracy	9
3.3 Residual limb temperatures	9
3.3.1 Temperature at different measurement moments	9
3.3.2 Temperature changes at pressure pads	9
3.3.3 Temperature changes for different suspension systems	10
3.3.4 Temperature distribution for unmodified and modified sockets	10
3.4 Subjective assessments related to temperature changes	12
3.5 Reliability and agreement parameters	12
4 Discussion	14
4.1 Relationship between thermal patterns, pressure, and pain	14
4.2 Relationship between thermal patterns and socket comfort	15
4.3 Thermal patterns of TSB-vacuum and SWB-pin sockets	15
4.4 Reliability and agreement parameters	15
4.5 Limitations	15
4.6 Clinical implications and future perspective	16
5 Conclusion	18
References	19
A Evaluation of measurement systems	22
A.1 List of requirements	22

A.2	Measurement systems	23
A.2.1	Pressure sensors	23
A.2.2	Pressure mapping film	23
A.2.3	Infrared thermography (IRT)	23
A.2.4	Thermal sensors	24
A.3	Scoring	24
B	Participant data	25
B.1	Data collected at the start of the procedure	25
B.1.1	Amputation side, reason, and time of amputation	25
B.1.2	Residual limb	25
B.1.3	Prosthetic socket	25
B.1.4	Mobility level	25
B.1.5	Skin type	25
B.1.6	Medication use	26
B.1.7	Use of alcohol and smoking habits	26
B.1.8	Height and weight	26
B.1.9	Body temperature	26
B.2	Data collected during the procedure	26
B.2.1	Decubitus classification	26
B.2.2	Socket Comfort Score (SCS) and localisation	26
C	Data analysis code	27
C.1	Automatic marker detection	27
C.2	Image registration	27
C.3	Difference maps	29
C.4	ROI analysis	29
D	Reproducibility calculations	31

Nomenclature

Abbreviations

Abbreviation	Definition
2D	2-dimensional
3D	3-dimensional
ICC	Intraclass Correlation Coefficient
IRT	Infrared thermography
PROM(s)	Patient-reported outcome measure(s)
RMSE	Root mean square error
ROI(s)	Region(s) of interest
SCS	Socket Comfort Score
SDC	Smallest detectable change
SEM	Standard error of measurement
SWB	Specific weight-bearing
TR	Thermal imaging after a 15-minute resting period
TR1	First set of post-resting thermal images
TR2	Second set of post-resting thermal images
TSB	Total surface-bearing
TW	Thermal imaging post-walking
TW1	Thermal images post-walking with the unmodified prosthetic socket
TW2	Thermal images post-walking with the modified prosthetic socket

Symbols

Symbol	Definition	Unit
ΔT	Temperature changes between post-walking and post-resting conditions	degrees Celsius ($^{\circ}\text{C}$)

1

Introduction

1.1. Clinical background

In the Netherlands, 57% of major lower limb amputations (i.e., above the ankle) are performed at the transtibial level.¹ Transitioning to a well-fitting prosthesis can significantly restore mobility, allowing individuals to regain independence and resume daily activities.² A transtibial prosthesis comprises a prosthetic socket, a connecting tube, and a foot (Figure 1.1). The prosthetic socket, acting as the integral interface connecting the human body to the rest of the prosthesis, must fit closely around the residual limb to minimize interface movement and enhance stability. Prosthetic socket designs consider the weight-bearing capacity of the residual limb, including pressure-tolerant areas, such as the patellar tendon, medial and lateral flares of the tibia and the gastrocnemius muscle, and pressure-sensitive areas, such as the distal tibia and fibular head³ (Figure 1.2). Common approaches for prosthetic sockets include specific weight-bearing (SWB) with pin suspension and total surface-bearing (TSB) with vacuum suspension, each tailored to distribute weight accordingly (Figure 1.1). In SWB sockets, the residual limb experiences both normal and shear forces. Conversely, TSB sockets utilise a vacuum seal that minimises residual limb movement within the socket⁴, primarily resulting in normal pressure and reducing shear stress. Inadequate socket fit can introduce excessive forces at the residual limb-socket interface, leading to pain, skin damage, and altered gait^{2,5,6}, which diminishes quality of life². Moreover, frequent prosthetic socket modifications and refitting are time-consuming and labour-intensive, leading to significant financial consequences for healthcare systems.⁷

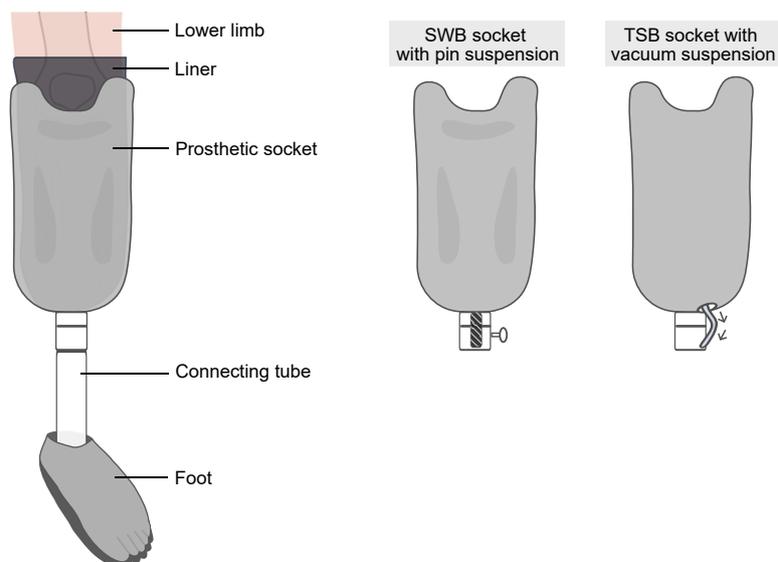


Figure 1.1: Left: Schematic representation of a transtibial prosthesis, including a liner, prosthetic socket, connecting tube, and foot. Right: Specific weight-bearing (SWB) socket with pin suspension and total surface-bearing (TSB) socket with vacuum suspension.

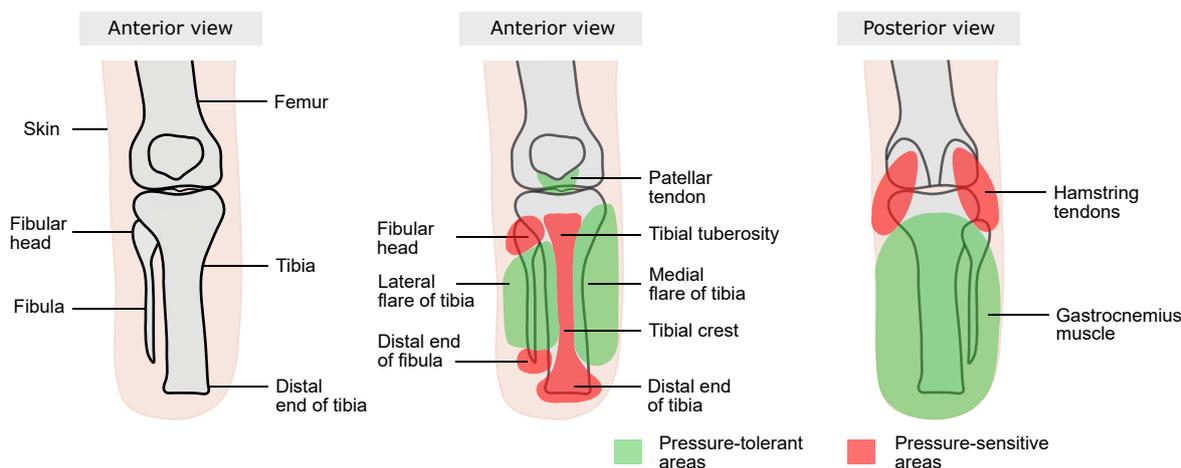


Figure 1.2: Schematic representation of a right-sided residual limb after a transtibial amputation indicating areas of pressure tolerance and sensitivity. Left: Anatomical landmarks are identified on the anterior view. Centre: Anterior view showing pressure-tolerant (green) and pressure-sensitive (red) areas relevant to prosthetic socket fitting. Right: Posterior view of the residual limb with the same mapping.

1.2. Measurement of socket fit

The fit of a prosthetic socket is a multi-dimensional construct, assessed through the distribution of forces at the residual limb-socket interface, stability, wearer comfort, and perspiration management. The current assessment is primarily subjective, in which prosthetic socket fit is determined by patients using Patient Reported Outcomes (PROMs), questionnaires, or scores.^{8,9} Although PROMs provide insight into the patient's perception and can be easily applied in clinical care, they have inherent limitations, such as the risk of biased responses or lack of detail.¹⁰ Specifically in patients with lower extremity amputations, decreased sensation in the residual limb impairs the reliable use of self-reported pain and comfort of the prosthetic socket. Moreover, evaluation of prosthetic socket fit by a prosthetist is strongly dependent on the prosthetist's expertise. These limitations necessitate an objective measurement system. Current objective methods predominantly involve pressure measurements at the interface, but these can be inaccurate, lack durability, and interfere with the socket fit.^{11,12} A comparison of different objective measurement technologies can be found in Appendix A.

1.3. Infrared thermography

Infrared thermography (IRT) is a promising alternative for assessing socket fit. It requires no modifications to the socket and does not disturb the residual limb-socket interface, offering distinct advantages over traditional pressure measurement systems. IRT detects heat emissions in the infrared spectrum and translates these into thermal images that visually map skin temperature patterns. Its applications in medicine are broad, aiding in the management of diabetic feet¹³, peripheral arterial disease¹⁴, and wound care^{15,16,17}. Moreover, studies show that pressure and shear forces on the skin cause reduced blood flow¹⁸, and subsequently induce reactive hyperaemia, i.e., a temporary post-occlusion increase in blood flow¹⁹. This, in turn, leads to a rise in skin temperature.²⁰ This reactive hyperaemia can be measured using IRT.²¹ The application of thermal imaging in evaluating prosthetic sockets has been explored in several studies, which indicate the potential of thermography in identifying loading zones on the residual limb.^{22,23,24,25,26} However, these studies lack a targeted analysis of the precise correspondence between thermographic data and specific areas of pressure and shear forces on the skin. Consequently, there is a need for further research to clarify the value of thermal imaging in this context. Such research is essential for evaluating prosthetic socket designs and optimising the fit of the prosthetic socket.

1.4. Objectives

This study evaluates the use of IRT to objectively assess socket fit in transtibial prostheses. The objectives include 1) analysing temperature changes in response to pressure and shear forces, 2) examining these changes in relation to subjective assessments of fit and pain, and 3) determining the reliability and agreement parameters of the IRT measurements. It is hypothesised that pressure applications will result in detectable

hot spots, and that these hotspots will align with the current reference standard of subjective experiences of fit and pain. Moreover, it is expected that different suspension systems result in different thermal patterns on the residual limb. Reliability is anticipated to be acceptable, with an Intraclass Correlation Coefficient (ICC) exceeding 0.7, similar to IRT studies in other medical applications in the lower limb^{27,28}. Based on limited IRT reliability data from studies in breast cancer patients²⁹ and on healthy lower limbs³⁰, the standard error of measurement (SEM) and smallest detectable change (SDC) are expected to be approximately 0.5 °C and 1.0 °C, respectively. However, data on these parameters remains scarce.

2

Methods

This proof-of-concept study with within-subject design was conducted in collaboration between the 3D lab and the Department of Rehabilitation at Radboud University Medical Centre. Ethical approval was obtained from the Radboud University Medical Centre's Research Ethics Committee (ID: 2024-17056).

2.1. Participants

Participants were recruited from the Radboud University Medical Centre and Papenburg Orthopedie. Informed consent was obtained from all participants before initiation of the study procedures. Inclusion criteria included: 1) adults with transtibial amputations using pin- or vacuum-suspension socket prostheses, and 2) the ability to walk for a minimum of 10 minutes with their prosthesis. Exclusion criteria were residual limb neuropathy, active skin issues on the residual limb, or cognitive impairments affecting communication.

2.2. Study design

Participants were evaluated under two conditions: 1) using their as-prescribed (referred to here as unmodified) socket, and 2) using a modified socket. If the unmodified socket was normally worn with a sock, this was considered the unmodified state. The modified socket consisted of the participant's unmodified socket, with the same number of socks, into which drop-shaped, polyethylene foam (hardness Shore A 35) pads were added to introduce pressure (Figure 2.1). A pad measuring 7.0 × 5.0 cm with a thickness of 6.0 mm in the middle was placed at the distal tibia end, and a pad of 6.0 × 4.0 cm with the same middle thickness was added to the fibular head region inside the prosthetic socket. The thickness of the pressure pads is comparable to the additional space that prosthetists typically add at the distal tibia end and fibular head to the prosthetic socket. In cases where a participant could not adequately don (i.e., put on) the prosthetic socket with the initial pad size at the distal tibia end, a smaller pad identical in size to the one used at the fibular head was used as substitute. Inserting these pads into a pin-suspension socket was assumed to introduce both normal and shear pressures, while in a vacuum-suspension socket, the pads were presumed to generate predominantly normal pressure.

Participants were aware of the order of the unmodified and modified socket. However, they were blinded to the positioning of the pads during donning and doffing (i.e., removing) of the modified socket.

2.3. Measurement instruments and setup

2.3.1. Infrared camera system

Thermal images were captured using a FLIR E75 infrared camera (FLIR Systems, USA), with specifications detailed in Table 2.1. Ambient temperature and humidity were monitored using a thermo-hygrometer (Model 625, B&K Precision, USA) during measurements. The human skin emissivity coefficient was set at 0.98.^{31,32} Reflective temperature was set to ambient temperature.³³ These measurement parameters were set prior to each measurement moment. Spatial calibration was performed using a cooled, known-dimension object positioned at one metre distance as a reference. Data were extracted using FLIR software (FLIR ResearchIR Max, version 4.40.12.38), and exported as MATLAB files. Concurrent visual images were also captured by the FLIR camera.

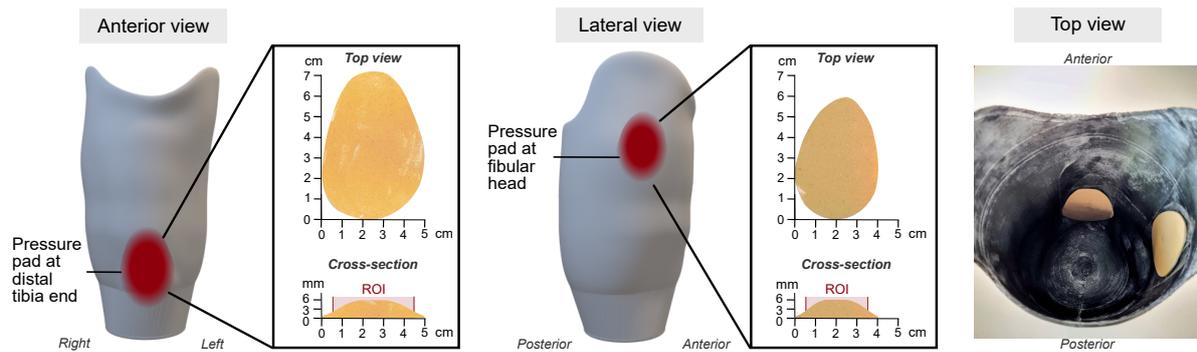


Figure 2.1: Anterior, lateral, and top views of a right-sided prosthetic socket, showing the placement of temporary pressure pads (red) at the fibular head and distal tibia end. These pads simulate pressure points for the study. The dimensions of the pads are displayed together with an image of the pad, in which the ROIs used for data analysis are highlighted in red. Abbreviations: ROI = region of interest, mm = millimetre, cm = centimetre.

Table 2.1: Specifications of FLIR E75 infrared camera system. Abbreviations: C = Celsius, μm = micrometre.

Temperature range ($^{\circ}\text{C}$)	Image resolution (pixels)	Accuracy (%)	Spectral range (μm)	Thermal sensitivity ($^{\circ}\text{C}$) at 25 $^{\circ}\text{C}$
-20 to 120	240 x 320	2	7.5 to 13	<0.03

2.3.2. Thermographic markers

3D printed thermographic markers (Figure 2.2) were used to facilitate the registration of thermal images at different measurement moments. The hollow design permits ambient air to fill the interior of the thermographic marker, creating a temperature contrast with the residual limb such that it is visible in thermal images.²³ The thermographic markers were secured to the residual limb using skin tape and positioned such that at least three markers were visible from each camera position.

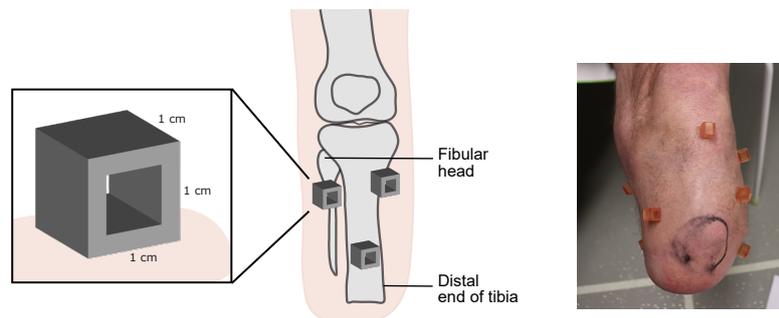


Figure 2.2: Left: Illustration of the residual limb with custom 1x1x1 cm thermographic markers. The hollow markers were 3D printed and positioned to aid in thermal image registration. Right: Photograph of a residual limb with thermographic markers. The distal tibia end is outlined with a permanent marker. Abbreviations: cm = centimetre.

2.3.3. Measurement setup

Measurements were conducted in a room where temperature was maintained by keeping doors closed and avoiding external heating sources, ensuring minimal interference. The room was free of windows and mirrors to prevent reflections. During thermal imaging, participants were instructed to stand with the aid of a walking device or to sit on a chair, depending on their mobility, without their prosthesis. The residual limb was positioned above a marked spot on the floor, at which two lasers were used to align the residual limb in frontal and sagittal directions. To systematically capture thermal images around the residual limb, a setup consisting of an octagon with each vertex positioned one metre from the centre was used (Figure 2.3). Primary thermal images were captured from the 1st, 3rd, 5th, and 7th vertex, providing anterior, lateral, posterior, and medial views of the residual limb. Additional images at the 2nd, 4th, 6th, and 8th vertex were collected as backups for the primary images and potential extended analysis in the future. This two-phase approach ensured complete coverage of the limb within a concise time frame, limiting the variations due to cooling of the residual limb.

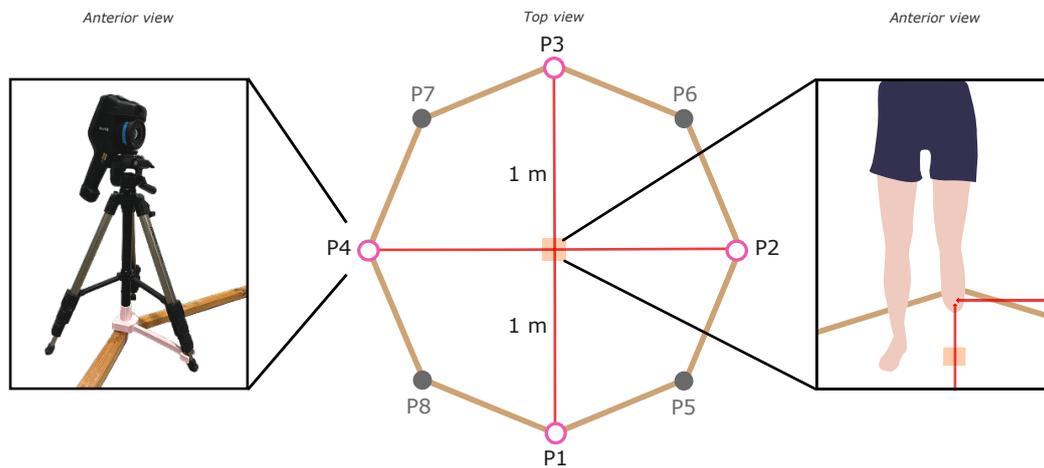


Figure 2.3: Thermal imaging setup. The left panel shows the infrared camera positioned on a tripod fixed to a vertex of the octagonal framework (brown). A custom-designed and 3D printed connector (light pink) was used to quickly lock the tripod on each of the 3D printed vertices and to allow repeatable camera positioning. The middle panel illustrates a top view of the octagonal setup with 3D printed vertices spaced one metre from the centre (orange). Primary (pink) and secondary (grey) camera positions are indicated. Lasers (red) align the residual limb in the frontal and sagittal direction. The right panel shows the positioning of the residual limb above the centre (orange), aligned with lasers (red). Abbreviations: P = infrared camera position, m = metre.

2.3.4. Subjective assessment tools

Subjective assessment consisted of the Socket Comfort Score (SCS), a validated tool for evaluating the overall comfort of the prosthetic socket.³⁴ Comfort levels were rated on a scale from 0 (least comfortable) to 10 (most comfortable). Additionally, discomfort and pain in specific regions of the residual limb were assessed using a similar 0 (no pain) to 10 (maximum pain) scale. These areas were systematically recorded on a schematic diagram of the residual limb. For the detailed assessment, refer to Appendix B.

2.4. Experimental procedure

2.4.1. Data collection moments

Residual limb skin temperature was recorded at four moments: 1) post-walking with the unmodified socket (TW1), 2) following a 15-minute resting period (TR1), 3) post-walking with the modified socket (TW2), and 4) following a second 15-minute resting period (TR2). These temperature measurements were used to create differential thermal maps, which visualise the temperature changes between post-walking and post-resting conditions (ΔT). Additionally, each subject serves as their own control such that individual physiological differences can be accounted for.

2.4.2. Procedure steps

At the start of the procedure, participants underwent assessment of their residual limb and prosthetic socket, and participant information was obtained (see Appendix B). Body temperature was measured in the participant's right ear using an infrared thermometer (Dr. Original, OTC Medical, Netherlands) to screen for fever. The bony prominences of the distal tibia and fibular head were palpated and delineated on the residual limb (Figure 2.2), and thermographic marker locations were indicated on the residual limb.

Participants walked for 10 minutes at a self-selected speed with their unmodified socket, during which videos were captured to document gait patterns. Following this, the prosthetic socket and liner were removed, and photographs of the limb were taken to document any skin changes. Subsequently, thermographic markers were secured to the residual limb. Thermal imaging was performed by the same researcher after each walking and resting period. Primary thermal images were acquired within one minute. During the resting periods, the same researcher evaluated the SCS and areas of discomfort or pain on the residual limb.

2.5. Data analysis

2.5.1. Image registration

To facilitate the comparison of temperature differences between the unmodified and modified sockets, thermal images post-walking (TW) and post-resting (TR) were aligned. This alignment was necessary to subtract TR from TW to highlight areas of thermal change. Due to observed variability in thermal patterns between TR1 and TR2 during visual inspection, TR1 was selected as the baseline for alignment. Therefore, TR1 was used as the moving image, meaning that it was adjusted to align with both TW1 and TW2. TW1 and TW2 were treated as the fixed images and, therefore, remained unaltered. Given the absence of distinctive features on the residual limb suitable for registration and the expected variability in thermal patterns between TW and TR, feature-based image registration using the thermographic markers was employed. To quantify registration accuracy, Euclidean distance and root mean square error (RMSE) between positions of the thermographic markers in TW and their corresponding positions in the aligned TR were calculated. The Euclidean distance measures the straight-line distance between the markers, while RMSE provides the square root of the average of the squared differences between the positions of the markers. The detailed methodology of the image registration process is visualised in Figure 2.4 and elaborated in Appendix C.

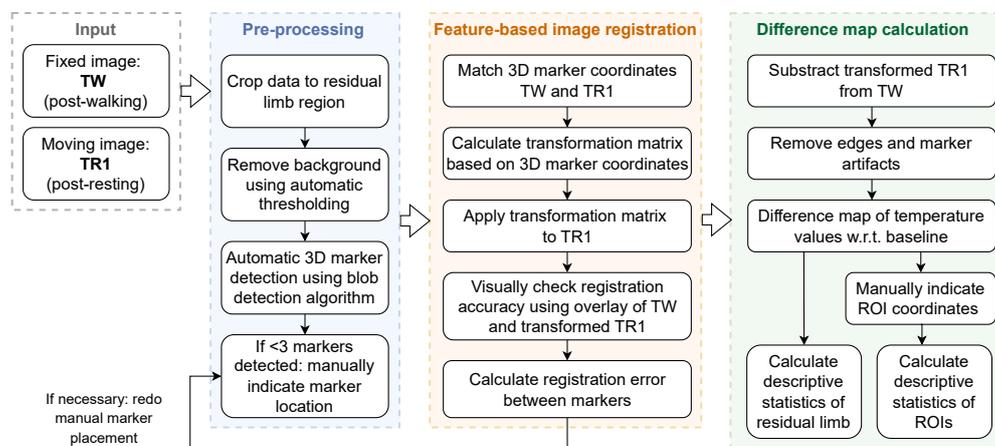


Figure 2.4: Flowchart of the data analysis framework for thermal images. Input for the analysis was thermal data post-walking (TW) and post-resting (TR1). TW consisted of the first (TW1) or second (TW2) set of post-walking images. The thermal images were pre-processed to isolate the residual limb and identify markers. Feature-based registration aligned the images. Finally, difference maps were calculated, and statistics were extracted for the residual limb and specific ROIs to assess thermal changes. Abbreviations: TW = thermal imaging post-walking, TR = thermal imaging after a 15-minute resting period, 3D = 3-dimensional, w.r.t. = with respect to, ROI = region of interest.

2.5.2. Residual limb temperatures

Temperatures of the residual limb at different measurement moments were measured and analysed using descriptive statistics. To investigate the effects of normal and shear pressures on residual limb temperatures, ΔT for the unmodified and modified sockets was calculated (Figure 2.4). The values for ΔT were depicted on thermal difference maps and summarised using descriptive statistics. The analysis focused on the entire residual limb and two specific regions of interest (ROIs): the distal tibia end and the fibular head. These ROIs were delineated as circles on the thermal images, sized to match the diameter of the pressure pads. Given the minimal pressure exerted by the pad's thin edges, the circles were set at 4 cm in diameter for the distal tibia end and 3 cm for the fibular head (Figure 2.1). The positioning of the ROIs in the thermal images was adjusted through a side-by-side comparison with corresponding visual images on which bony prominences were delineated. Hotspots, predefined as localised areas exhibiting higher temperature values than surrounding regions, were assessed in the context of de SDC.

To assess the impact of the unmodified and modified sockets on all views of the residual limb, data from all four views were filtered using a consistent cut-off value of 8 degrees Celsius ($^{\circ}\text{C}$), identified through data inspection, to indicate outliers. Values exceeding this threshold were removed, and the remaining data were aggregated to generate histograms and calculate measures of skewness and kurtosis. Furthermore, to assess the impact of different suspension systems on thermal patterns, visual comparisons were made between the

difference maps of SWB-pin suspension sockets and TSB-vacuum suspension sockets. This analysis was conducted to compare the effects of combined normal and shear pressures versus predominantly normal pressure on residual limb temperatures.

2.5.3. Relation with subjective assessments

To explore the relationship between ΔT and subjective experiences, comfort and pain scores were graphically compared with mean and maximum ΔT values at the locations where pain was reported.

2.5.4. Reliability and agreement parameters

To assess reproducibility under test-retest conditions, reliability, and agreement parameters were calculated using resting states TR1 and TR2. Intervention effects from pressure pads at TW1 and TW2 on the below-knee region necessitated the assessment of difference map reliability using a ROI above the knee. This same ROI was consequently also evaluated in both TR1 and TR2.

Reliability was quantified using the ICC, calculated through a two-way mixed effects model (ICC3.1 agreement) with 95% confidence intervals (CI). The ICC values were interpreted according to Byrt's guidelines: poor (0.01–0.20), slight (0.21–0.40), fair (0.41–0.60), good (0.61–0.80), very good (0.81–0.92), and excellent (0.93–1.00).³⁵ Agreement was assessed using the SEM and SDC, both expressed in °C, and SEM% and SDC%, expressed as percentages of the average test and retest values for the corresponding region. Due to the average temperature values of the difference map expected to be close to zero, calculating SEM% and SDC% was not feasible for this data, as these metrics become unreliable with near-zero denominators. Detailed formulas are provided in Appendix D.

Bland-Altman plots were constructed to determine if there was systematic bias in the measurements and if variability in temperature measurements related to the temperature of the residual limb was present. The Bland-Altman plots visualise the relationship between temperature measurements of the residual limb at TR1. The visualised 95% limits of agreement (95% LoA) were determined as the mean difference ± 1.96 SD of the difference.³⁶

2.5.5. Software

Data pre-processing, image registration, difference map calculations, and ROI measurements, along with the construction of Bland-Altman plots, were performed using Python. All additional statistical analyses were conducted with IBM SPSS Statistics version 29.0.1.0. Due to the small sample size, p-values were not considered in the analysis.

3

Results

3.1. Participant characteristics

The study included seven participants with a mean age of 46.5 years (range 28.0-67.0). Four participants used TSB sockets with a vacuum suspension system (TSB-vacuum sockets), while three participants used SWB sockets with a pin suspension system (SWB-pin sockets). Body temperature at the start of the session was 36.7 ± 0.1 °C (range 36.5-36.8). None of the participants exhibited a fever during the session. Detailed participant characteristics are presented in Table 3.1.

3.2. Experimental conditions and registration accuracy

Room temperature during the experiments averaged 20.6 ± 1.2 °C (range 19.0-22.0). Image registration of TR1 to TW1 and TW2 resulted in a mean Euclidean distance of $2.8 \times 10^{-2} \pm 5.0 \times 10^{-2}$ mm (range 1.0×10^{-3} - 3.9×10^{-1}) and a mean RMSE of $1.1 \times 10^{-2} \pm 2.0 \times 10^{-2}$ mm (range 3.0×10^{-4} - 1.4×10^{-1}).

3.3. Residual limb temperatures

3.3.1. Temperature at different measurement moments

After resting, the mean skin temperature of the residual limb varied between participants, with a mean temperature of 31.6 ± 1.2 °C (range 29.2-33.8). Walking with the unmodified socket increased this by 0.7 °C to an average temperature of 32.3 ± 1.2 °C (range 29.8-33.9). After walking with the modified socket, the average temperature returned to 31.6 ± 1.1 °C (range 29.0-32.9), similar to the post-resting value.

3.3.2. Temperature changes at pressure pads

Figure 3.1A displays thermal difference maps for unmodified and modified sockets across three participants, selected as representative examples due to their clear illustration of the thermal patterns, showing the lateral view of the fibular head where pressure pads are located. These pads were associated with hotspots in 86%

Table 3.1: Participant and prosthesis characteristics.

PID	Sex	Age (y)	BMI (kg/m ²)	Skin type	Amp side	Amp reason	Time since amp (y)	RL length (cm)	Tibia length (cm)	Socket + suspension type	Liner type	K-level
1	M	56	30.6	2	Right	Oncological	0.3	28.0	26.0	SWB-pin	Gel	K3
2	M	32	26.0	3	Left	Traumatic	0.5	18.0	18.0	TSB-vacuum	Gel	K4
3	M	67	33.6	2	Left	Infectious	7.0	15.0	10.0	TSB-vacuum	Gel	K2
4	M	58	23.4	2	Left	Infectious	12.3	17.0	13.0	TSB-vacuum	Gel	K3
5	F	28	19.7	2	Right	Traumatic	3.3	15.0	13.0	TSB-vacuum	Gel	K3
6	M	38	25.8	2	Left	Traumatic	0.4	15.0	12.0	SWB-pin	Gel	K4
7	M	28	23.5	2	Left	Traumatic	12.0	13.5	11.5	SWB-pin	Gel	K4

Tibia length is measured from the patella apex to the distal end of the tibia. Abbreviations: PID = participant identifier, M = male, F = female, BMI = body mass index, kg/m = kilogram/metre, Amp = amputation, y = years, RL = residual limb, cm = centimetre, SWB = specific weight-bearing, TSB = total surface-bearing.

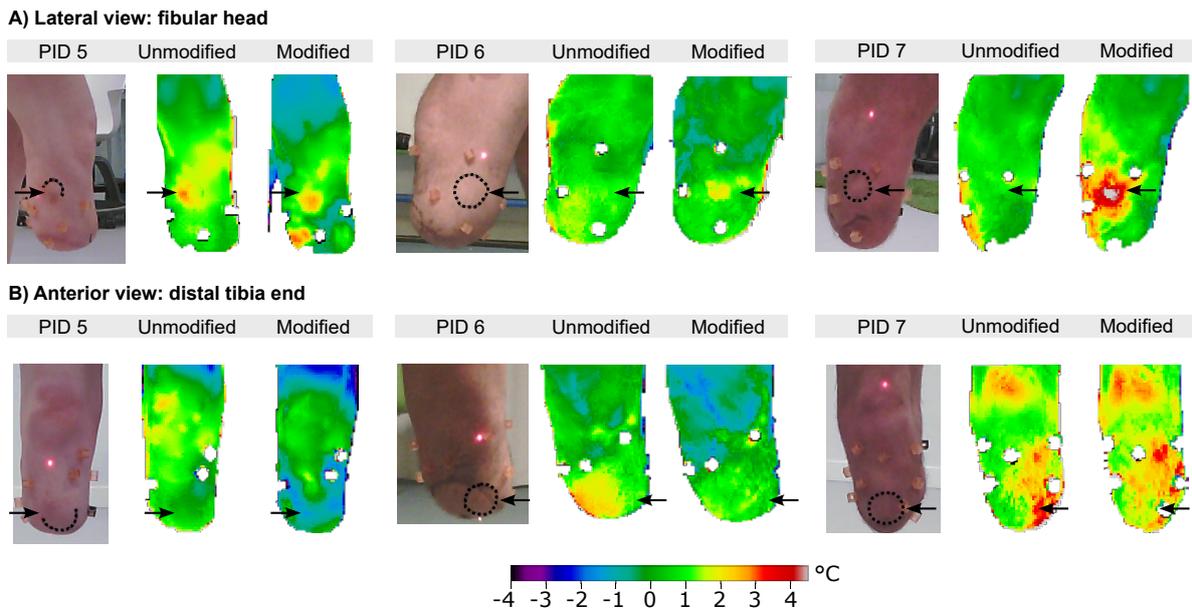


Figure 3.1: Thermal difference maps and concurrent visual images captured by the FLIR camera, displaying the thermal response of the residual limb relative to baseline for three representative participants with unmodified and modified sockets. The lateral views highlighting the fibular head region, and the anterior views focusing on the distal tibia end, are displayed in panel A and B, respectively. Bony prominence of the fibular head and distal tibia end were marked on the residual limb during the procedure and traced here with a dotted line, and indicated with arrows. Abbreviations: PID = participant, C = Celsius.

(6/7) of participants. In the remaining participant, participant 1 (14%, 1/7), a hotspot was observed slightly distal to the pressure pad. Additionally, participant 5 reported pain at the fibular head in both prosthetic sockets, indicating that pressure was already present here in the unmodified socket. Notably, the temperature increase at the fibular head of participants 3 and 4 is also present in the unmodified socket, but appears less localised compared to the modified socket (Figure 3.2). All participants that reported pain at this location (participants 2, 5, 6, and 7) also showed a hotspot here. Figure 3.1B presents thermal difference maps from the anterior view, highlighting pressure pad locations at the distal tibia end. Hotspots at these locations were less consistent, appearing in only 43% (3/7) of participants (participant 3, 4 and 7). Notably, participant 3 and 4 already exhibited hotspots in the unmodified socket at this location (Figure 3.2). Two-third of the cases with hotspots were associated with the presence of pain.

3.3.3. Temperature changes for different suspension systems

Figure 3.2 shows the difference maps for all seven participants across all views, grouped by suspension type. The SWB-pin socket and TSB-vacuum socket do not exhibit distinct thermal patterns corresponding to the load-bearing areas of the prosthetic socket type on the difference map. With the modified TSB-vacuum socket, the addition of pressure pads leads to slightly more localised temperature changes, such as observed in participant 2. In contrast, the the modified SWB-pin socket shows a more diffused temperature response following the introduction of pressure pads, such as seen in participants 1, suggesting potential shear stress extending beyond the pressure pad areas.

3.3.4. Temperature distribution for unmodified and modified sockets

Figure 3.3 shows the histogram analysis of ΔT across all views of the residual limbs combined. The largest peak represents the mode of the temperature values. Hotspots are reflected by additional peaks, most clearly visible in participants 2, 6, and 7. Participants 1, 3, 4, 5, and 6 exhibited an increase in skewness from the unmodified to the modified socket, indicating a shift toward a more right-tailed distribution with more pixels showing large positive ΔT values. Participant 2 remained stable. Participant 7 showed a decrease in skewness, suggesting a left-tailed shift with more pixels having large negative ΔT values. Participants 2, 3, 4, and 7 experienced a decrease in kurtosis, indicating a more uniform data spread in the modified socket. Participants 1 and 6 remained stable, while participant 5 showed increased kurtosis, suggesting more outliers in the modified socket. The histogram analysis did not reveal a clear relationship between skewness, kurtosis, and the thermal images in Figure 3.2.

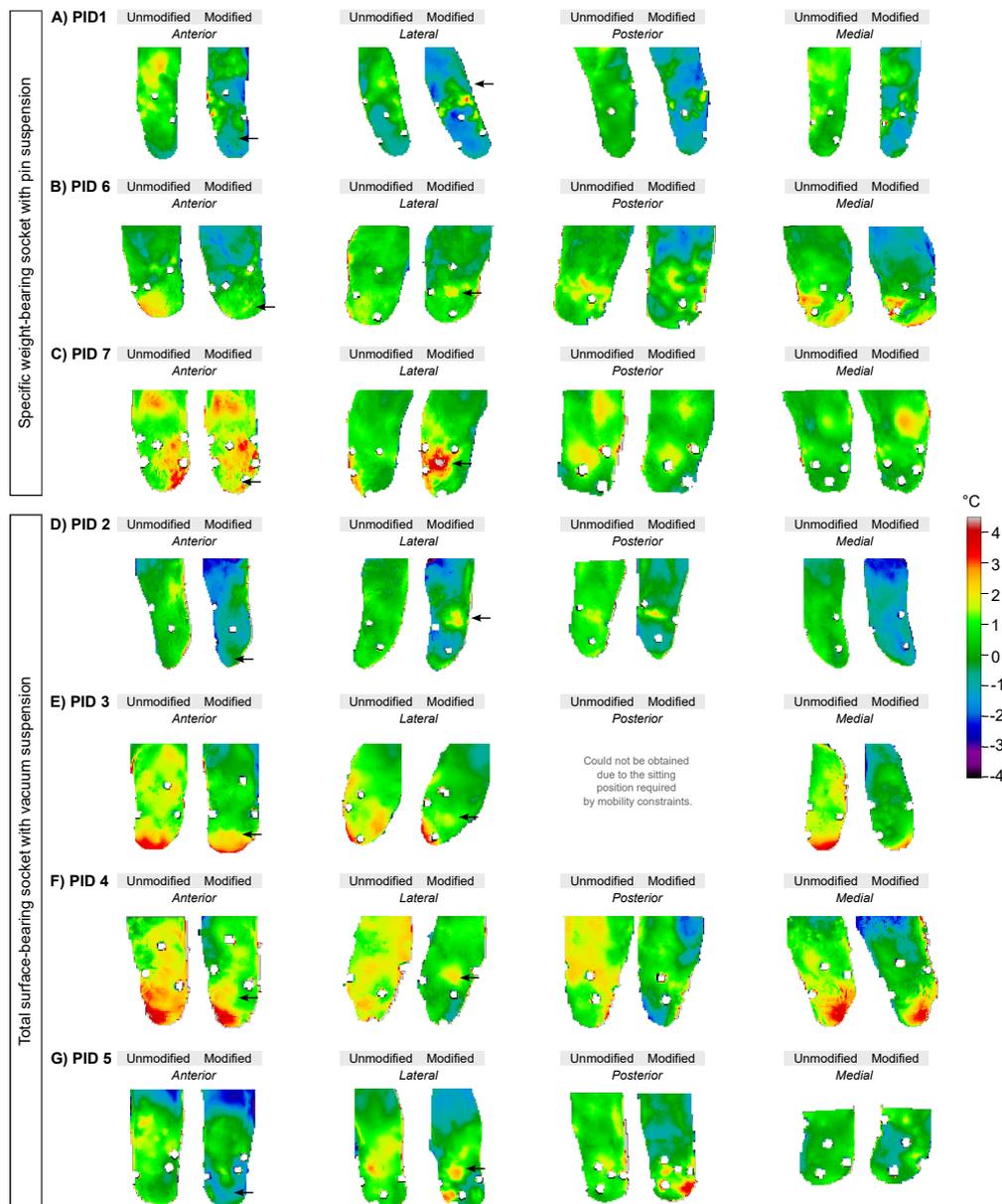


Figure 3.2: Difference maps for the unmodified and modified sockets of all participants. Hotspots appear as temperature increases of 2.0 to 4.5 °C. Arrows indicate the positioning of the pressure pads. Abbreviation: PID = participant, C = Celsius.

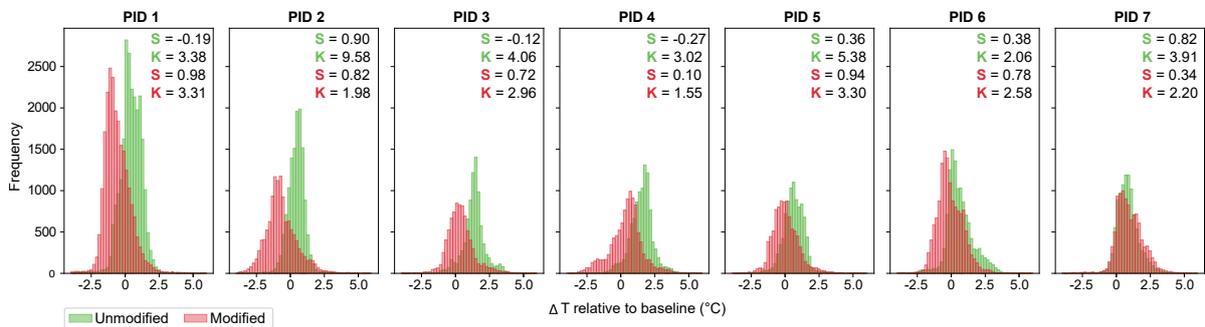


Figure 3.3: Histograms displaying the frequency of temperature change values relative to baseline temperatures (ΔT) for residual limbs across seven participants with both unmodified (green) and modified (red) prosthetic sockets. Posterior images could not be captured for participant 3. The histograms highlight the skewness (S) and kurtosis (K) values for each prosthetic socket. Abbreviations: PID = participant, S = skewness, K = kurtosis.

3.4. Subjective assessments related to temperature changes

In Figure 3.4A, the SCS for unmodified sockets averaged 7.5 ± 1.2 (range 6.0-9.0), reflecting baseline comfort levels. Participants self-reported the unmodified socket fit as too small (14%, 1/7), good (71%, 5/7), or too big (14%, 1/7). Modified sockets showed a reduced average SCS of 4.6 ± 2.1 (range 2.0-8.0). Figure 3.4B details the pain scores at specific anatomical sites. In unmodified sockets, pain was reported at the fibular head by 14% (1/7) of participants, at the distal tibia end by 43% (3/7), and at the popliteal fossa by 14% (1/7). There was a notable increase in pain at the fibular head in 57% (4/7) of participants and at the distal tibia end in 71% (5/7) of participants using modified sockets. Lesser increases were observed at the gastrocnemius belly (14%, 1/7) and the popliteal fossa (29%, 2/7).

Figure 3.4C illustrates ΔT at various ROIs. Participants 2, 6, and 7, who experienced pain at the fibular head only with the modified socket, showed a noticeable increase in ΔT at this ROI. Conversely, participant 5, who reported pain with both socket types, exhibited similar ΔT values at this ROI for both sockets. Participants without pain generally exhibited stable or reduced ΔT at this ROI. At the gastrocnemius belly ROI, one participant reported pain with the modified socket, corresponding to an increase in ΔT compared to the unmodified socket, while those without pain showed a decrease in ΔT . At the distal tibia end and popliteal fossa ROIs, there was no association between reported pain and ΔT , with ΔT generally lower in the modified sockets compared to the unmodified sockets.

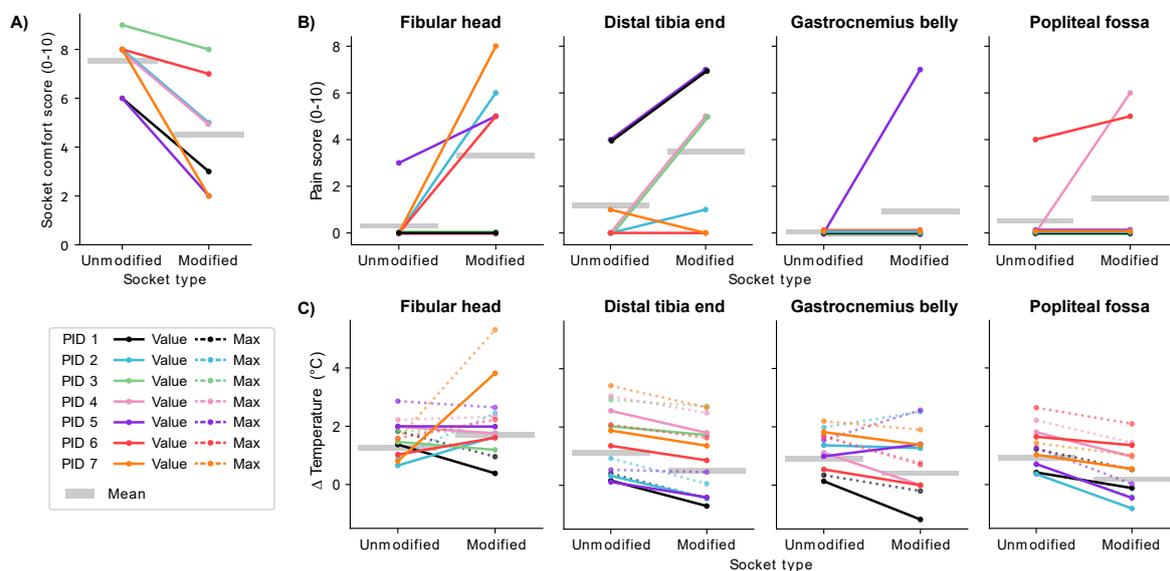


Figure 3.4: Subjective assessments and associated temperature changes for the unmodified and modified prosthetic sockets. Panel A illustrates the Socket Comfort Score reported by seven participants (PID 1-7) on a scale of 0-10, where higher scores indicate greater comfort. Panel B and C depict changes in pain and temperature changes relative to baseline temperatures, respectively, at four anatomical points (fibular head, distal tibia end, gastrocnemius belly, popliteal fossa). In these panels, lines connect the mean scores (solid) and maximum scores (dashed) for each participant. Mean scores or temperature values are indicated in each panel (grey) for unmodified and modified prosthetic sockets. Posterior images could not be captured for participant (PID) 3.

3.5. Reliability and agreement parameters

Table 3.2 indicates that test-retest reliability was very good to excellent ($ICC \geq 0.83$) across all measurements, with the exception of the ROI at the fibular head, demonstrating fair reliability ($ICC 0.55$, $95\% CI -1.62-0.92$). For test-retest with TR1 and TR2, SEM ranged from $0.3^{\circ}C$ (SEM% 1.0%) to $0.7^{\circ}C$ (SEM% 2.1%), and SDC ranged from $0.9^{\circ}C$ (SDC% 2.8%) to $2.0^{\circ}C$ (SDC% 5.9%). For test-retest on the difference maps of the unmodified and modified prosthetic socket, SEM ranged from 0.6 to $0.8^{\circ}C$, and SDC ranged from 1.7 to $2.1^{\circ}C$.

The Bland-Altman plots (Figure 3.5) reveal a systematic bias ranging from $0.3^{\circ}C$ to $0.4^{\circ}C$, and a spread of variability across the measured temperatures of the residual limb.

Table 3.2: Test-retest values, and reliability and agreement metrics for thermal images across different tested areas and views.

Tested area	View	Test: TR1 (°C) mean±SD	Retest: TR2 (°C) mean±SD	Difference test-retest (°C) mean±SD	ICC (CI 95%)	SEM (°C)	SEM %	SDC (°C)	SDC %
Total RL	Anterior	31.5 ± 1.4	31.2 ± 1.0	0.5 ± 0.2	0.95 (0.62-0.99)	0.4	1.3	1.2	3.7
	Lateral	31.8 ± 1.0	31.5 ± 0.8	0.4 ± 0.1	0.97 (0.81-0.99)	0.3	1.0	0.9	2.8
	Posterior	31.6 ± 1.0	31.2 ± 0.9	0.4 ± 0.2	0.98 (0.86-1.00)	0.4	1.1	1.0	3.1
	Medial	31.5 ± 1.4	31.2 ± 1.1	0.5 ± 0.3	0.97 (0.83-1.00)	0.4	1.2	1.1	3.4
ROI distal tibia end	Anterior	31.8 ± 1.9	31.7 ± 1.6	0.4 ± 0.3	0.98 (0.88-1.00)	0.4	1.1	1.0	3.1
ROI fibular head	Lateral	32.8 ± 0.8	32.9 ± 1.0	0.7 ± 0.7	0.55 (-1.62-0.92)	0.7	2.1	2.0	5.9
ROI upper leg	Anterior	31.3 ± 1.5	31.0 ± 1.1	0.6 ± 0.4	0.94 (0.65-0.99)	0.5	1.6	1.4	4.5
	Lateral	31.8 ± 0.9	31.4 ± 0.7	0.4 ± 0.3	0.95 (0.73-0.99)	0.4	1.2	1.0	3.3
	Posterior	32.3 ± 0.7	32.1 ± 0.9	0.5 ± 0.4	0.83 (-0.19-0.98)	0.5	1.5	1.3	4.2
	Medial	31.7 ± 1.5	31.1 ± 1.2	0.7 ± 0.5	0.95 (0.66-0.99)	0.6	1.9	1.7	5.4

Tested area	View	Test: Diff unmodified (°C) mean±SD	Retest: Diff modified (°C) mean±SD	Difference test-retest (°C) mean±SD	ICC (CI 95%)	SEM (°C)	SEM %	SDC (°C)	SDC %
ROI upper leg	Anterior	0.1 ± 1.0	-0.8 ± 1.1	0.9 ± 0.5	0.94 (0.63-0.99)	0.8	-	2.1	-
	Lateral	0.1 ± 0.5	-0.7 ± 0.7	0.8 ± 0.4	0.90 (0.40-0.98)	0.6	-	1.7	-
	Posterior	0.1 ± 1.6	-0.8 ± 1.3	0.9 ± 0.5	0.96 (0.74-1.00)	0.8	-	2.1	-
	Medial	0.2 ± 0.8	-0.7 ± 0.8	0.9 ± 0.6	0.83 (-0.22-0.98)	0.8	-	2.1	-

Abbreviations: RL = residual limb, ROI = region of interest, TR1 = first set of thermal images post-resting, TR2 = second set of thermal images post-resting, Diff = difference map, C = Celsius, SD = standard deviation, ICC = intraclass correlation coefficient, CI = confidence interval, SEM = standard error of measurement, SDC = smallest detectable change.

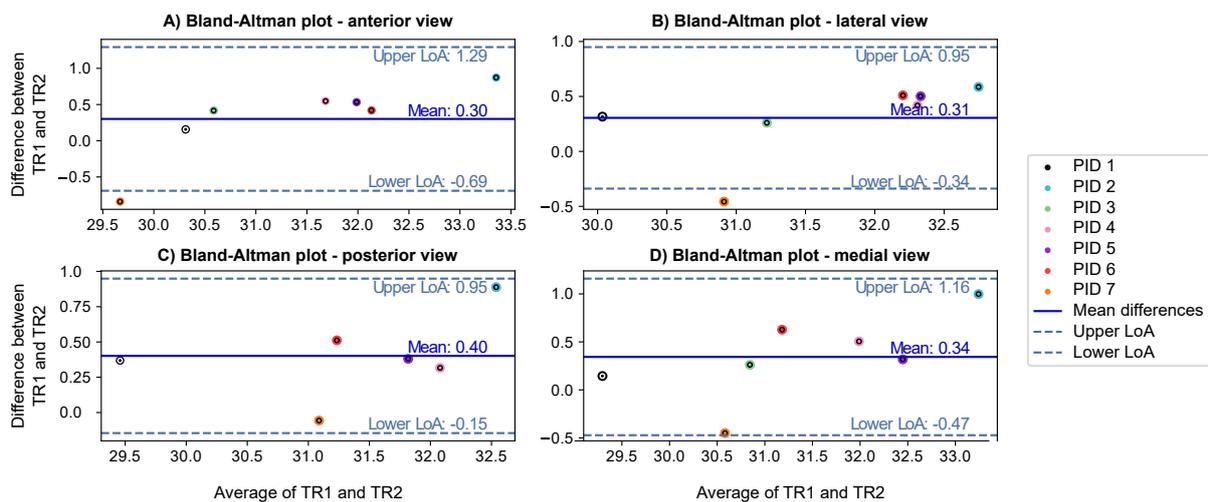


Figure 3.5: Bland-Altman plots for test-retest agreement of thermal images TR1 and TR2 of the residual limb. Panels A, B, C, and D show the measurements in the anterior, lateral, posterior, and medial views, respectively. The solid line represents the mean difference (systematic bias) and the dashed lines represent the 95% limits of agreement (mean difference \pm 1.96 SD of the difference). Posterior images could not be captured for participant 3. Abbreviations: TR1 = first set of thermal images post-resting, TR2 = second set of thermal images post-resting, LoA = limits of agreement, PID = participant.

4

Discussion

Current methods for evaluating transtibial prosthetic socket fit are largely subjective, lacking an objective measurement system. This study therefore assessed the viability of IRT in providing an objective method to evaluate prosthetic socket fit. The findings of this study show that IRT can effectively capture thermal images of the entire residual limb in the current setup and generate differential thermal maps that isolate the prosthetic socket's impact from the baseline resting state. Moreover, temperature variations between unmodified sockets and modified sockets that are perceived as less comfortable are observed. Notably, the introduction of pressure pads at the fibular head coincided with the emergence of hotspots and reported pain among the majority of participants. In contrast, similar interventions at the distal tibia end showed a less consistent association with hotspots and pain, suggesting a more complex interaction or less contact at this site.

4.1. Relationship between thermal patterns, pressure, and pain

At the fibular head, the correspondence between the presence of pressure and localised hotspots in 86% (6/7) of cases supports the hypothesis that IRT is sensitive to localised pressure on the residual limb at this location. The remaining participant (14%) showed a hotspot distal to the pressure pad, likely because their residual limb was positioned higher in the prosthetic socket due to the pressure pads, resulting in a more distal hotspot. In all participants who reported pain here, hotspots were observed, demonstrating the potential of IRT to identify painful spots at the fibular head. Hotspots appeared as ΔT values of 2.0 to 4.5 °C (Figure 3.2). These ΔT values were equal to or exceeded the largest SDC measured at the difference map (2.1 °C). Therefore, these hotspots are considered to be valid changes in temperature, beyond the intrinsic measurement errors.

The observation of a lesser incidence of hotspots at the distal tibia end and lesser coincidence with localised pain, as compared to the fibular head, opens discussions on several possible underlying reasons. First, circulatory differences at the distal residual limb, which typically shows lower temperatures compared to more proximal areas^{26,37}, may contribute to its varied thermal response. Second, predominantly pin, but also vacuum, suspension systems create negative pressure during the swing phase of walking, potentially leading to distal oedema and venous stasis.³⁸ This might add to the difference in thermal response, due to reactive hyperaemia, at the distal tibia end compared to the fibular head. Third, the movement of the distal residual bone within the prosthetic socket resembles a pseudo-joint³⁹, potentially causing the pressure pad to shift the distal tibia inward. Fourth, the study's methodology may have influenced these findings. Prosthetists typically provide extra space in the prosthetic socket at the distal tibia, especially when pain complaints are present, but the dimensions of this added space were not measured in this study. Observations during the study also indicated that some TSB-sockets unexpectedly included additional space at the distal tibia, resembling a tibial-relief TSB-socket. The pressure pads may not have exerted enough pressure at these locations to produce a noticeable thermal response. Lastly, several participants noted that their residual limb was positioned higher in the modified socket, potentially preventing the distal tibia end from adequately contacting the pressure pads. This could be addressed by using a transparent prosthetic socket to visually confirm the contact points between the residual limb and the pressure pad. This positioning issue might be related to the placement of the pressure pad at the fibular head. For example, in transfemoral prosthetic sockets, the addition of silicone pads at the proximal end can hold the residual limb tissue higher within the prosthetic socket, thus reducing pressure on the distal end of the residual limb.⁴⁰

Visual inspection of the difference maps was more effective for distinguishing hotspots than examining the absolute ΔT values in the ROIs. This is because the ROI values did not adequately reflect the temperature relative to the surrounding areas, making it difficult to determine the presence of hotspots. Since the difference map values were close to zero, using a ratio relative to the surrounding temperature was not effective.

4.2. Relationship between thermal patterns and socket comfort

In the majority of cases, visual examination of thermal images after wearing the more uncomfortable, modified sockets showed a more heterogeneous temperature distribution, contrasting with the homogeneous distribution seen in the more comfortable, unmodified sockets. The study sought to validate these observations through histogram analyses, however, the histograms did not clearly depict the heterogeneous patterns visible in the images. This discrepancy likely arises because the hotspots, being relatively small in pixel count compared to the entire residual limb, do not prominently show as peaks in the histograms. Additionally, despite filtering outliers, the small registration errors skewed the ΔT values around the markers, further complicating histogram interpretation. No alternative analysis method proved more effective in quantifying these thermal patterns, but this would be valuable, especially with a larger sample size.

4.3. Thermal patterns of TSB-vacuum and SWB-pin sockets

Pressure pads at the fibular head appeared to create slightly more localised hotspots in TSB-vacuum sockets, compared to slightly more spread-out hotspots in SWB-pin sockets, possibly due to additional shear pressure in the latter. However, the analysis of thermal patterns between the unmodified sockets did not reveal distinct variations related to weight distribution specific to each type of prosthetic socket. It is plausible that the gel liners worn by all participants mitigated some of the effects of the pressure pads on the residual limb. Additionally, the pressure pads may have created a snugger fit in both types of sockets, reducing limb movement and thereby limiting the generation of shear pressure. A longer walking period might be beneficial for observing the thermal effects on the residual limb associated with different socket types. Nonetheless, the current study lacked sufficient power to effectively compare these two subgroups.

4.4. Reliability and agreement parameters

The test-retest reliability analysis revealed that the ICC ranged from very good to excellent for all measurements, with the exception of the ROI at the fibular head. The lower ICC observed in this region may be attributed to the insufficient 15-minute washout period after TW2, during which the effects of the pressure pads, most pronounced in the fibular head, likely persisted into TR2. Consequently, despite a high ICC for mean residual limb temperature between TR1 and TR2, TR1 was selected as the baseline for calculation of the difference maps. Although a 15-minute resting period has been used before in another study²², these findings suggest this may not be sufficient when additional heat is generated due to peak pressures.

The ROI chosen proximal to the knee proved suboptimal for reliability analysis. This study was not equipped to reliably assess regions above the knee joints due to the placement of image registration markers solely on the below-knee residual limb. The knee, being a joint, introduces potential registration errors not present in the fixed bone of the below-knee residual limb. Despite these limitations, no alternative ROIs were available to quantify the reliability of the difference maps.

It is important to acknowledge that the small sample size in this study, which falls below the recommended threshold according to COSMIN guidelines⁴¹, compromises the validity of the ICC, SEM, and SDC values. Additionally, this limited sample size could account for the observed deviations in SEM and SDC from the initial hypotheses. Given the limitations imposed by the small sample size, this study primarily focused on a method to interpret the findings. Future research should include larger sample sizes to validate these reliability and agreement values more definitively.

4.5. Limitations

This study has several limitations that warrant discussion. Firstly, the impact of socket shape on the position of the residual limb within the prosthetic socket, and hence how the load is transmitted on the residual limb, introduces variability in where normal and shear pressures are exerted. This makes it challenging to predict the precise locations of thermal hotspots relative to the placement of pressure pads within the prosthetic

socket. Although some participants showed temperature changes near the pressure pad areas, these could not be definitively attributed to the pressure pads. Additionally, participants reported feeling positioned higher in the modified socket, adding further uncertainty regarding whether the distal tibia end was in adequate contact with the pressure pads. Future research should use transparent prosthetic sockets with markers on the liner to visually verify the residual limb's position and contact with the pressure pads.

Secondly, in this study, participants used their own prosthetic sockets, which exhibited varying levels of fit quality at the time of the study. Some participants experienced their prosthetic socket to be too large or too small, had low scores on the SCS, reported pain, or had not recently had their alignment checked by a prosthetist. This variability raises concerns about whether the hotspots observed in the unmodified sockets are indicative of a poorly- or a well-fitted socket, or are influenced by misalignment²⁴. Moreover, it questions the precise impact of pressure pads in the modified sockets. Future studies should consider designing a new, precisely fitted and aligned prosthetic socket to serve as a controlled baseline for comparisons. Modified sockets could also be newly produced with an intentionally poor fit to introduce peak pressure points without the need for pressure pads. This approach would simultaneously address the differences in hardness between the pressure pads and the prosthetic socket, which may influence the magnitude of normal and shear pressures.

Thirdly, the study relied on the subjective assessment of comfort and pain, known for its inherent limitations. For instance, the SDC of the SCS is relatively high (2.7 points)⁴², meaning that changes in scores below this threshold may be indistinguishable from measurement error. Moreover, participants showed difficulty precisely localising pain on the residual limb, and there were instances where high SCSs were reported despite the presence of pain post-walking. These factors compromise the reliability of subjective evaluations and complicate its association with hotspots. An improved approach would be to incorporate the assessments of experienced prosthetists who can professionally evaluate the fit and comfort of the prosthetic socket, adding a layer of external evaluation to the study. Alternatively, a combination of IRT and pressure sensors could be used to determine if these two methods yield similar results.

Fourthly, the practical feasibility of the current protocol is low. The protocol is too complex for use in clinical care, primarily because it was designed for a proof-of-concept study. To make this technique more usable in clinical settings, simplification of the protocol is necessary. This could involve developing a markerless registration method, although this presents its own set of challenges due to the lack of anatomical landmarks, to limit the steps and streamline the process. Alternatively, simplifying the image acquisition by using an automatically rotating platform for the participant could enhance practicality. Addressing these complexities is essential for translating this technique into a more widely applicable clinical tool.

Fifthly, the accuracy of the IRT camera ideally should be verified. However, this verification requires black box calibration, which is a costly method that was not feasible within the resource constraints of this project. Future studies are recommended to include this calibration to ensure the precision of thermal measurements.

Lastly, the long-term consequences of the observed hotspots, including whether they could lead to blisters or other skin injuries, remain unexplored. Extended monitoring of prosthetic sockets, especially those that are poorly fitted, along with repeated thermal imaging, would provide valuable insights into the potential for IRT as an early warning sign for skin problems. This approach could also enable preventative measures in both socket design and patient care.

4.6. Clinical implications and future perspective

Further development of IRT could lead to a measurement technique capable of not only distinguishing between well-fitted and poorly-fitted prosthetic sockets based on thermal hotspots, but also of identifying early warning signs of potential skin injury during the prosthetic socket fitting and rehabilitation process. Such a technique would be invaluable for less experienced prosthetists and supplementing subjective evaluations. Moreover, given its non-invasive, rapidly deployable, and easily interpretable nature, without the need for wires or extensive system requirements, this technique could be suitable for use by prosthetists in low-resource settings. However, this suitability is contingent upon streamlining the protocol.

To build on these findings, subsequent studies could investigate the combination of 2D thermal images with 3D scans of the residual limb to enhance interpretability. 3D models reveal more anatomical landmarks than 2D thermal images, thus providing a clearer relationship between locations on the limb and locations in the prosthetic socket that require adjustments. While one previous study successfully implemented this integration for a single participant's residual limb²³, and others have done so for diabetic feet^{15,16,17}, the technique often requires multiple thermal images from different angles to achieve accurate 2D-3D mapping. This can be challenging when rapid image acquisition is necessary to prevent the residual limb from cooling

after pressure application, a concern not as prevalent with diabetic feet. This also impacts the overall feasibility of the study. This study attempted to use a structure from motion algorithm⁴³ to estimate the 3D structure from the 2D thermal images and relate this to 3D scans, but the current setup did not allow the necessary amount of images. UV-mapping⁴⁴ was also considered, however, it proved inaccurate as the 2D images became skewed. Future studies could explore other options and design experiments specifically tailored to facilitate this advanced imaging integration.

5

Conclusion

The need for an objective system to assess transtibial prosthetic socket fit and the limitations of existing methods highlight the importance of exploring alternative measurement systems. This study evaluated the use of IRT to analyse temperature changes at the residual limb in response to normal and shear pressure, and their relation to subjective assessments. In addition, the reliability and agreement parameters of IRT measurements were examined. IRT successfully generated detailed thermal images of the residual limb, post-walking with unmodified and modified sockets, and post-resting, revealing temperature increases in response to pressure. It demonstrated promise in identifying peak pressure points, particularly at the fibular head, where localised pain was reported. However, IRT showed limited capability in detecting added pressure at the distal tibia end, likely due to differences in circulation and socket suspension effects, and uncertainty about the exerted pressure. Despite these challenges, the overall agreement of IRT measurements was very good to excellent. These findings highlight the potential of IRT as a valuable tool in prosthetic socket evaluation, complementing subjective assessments with objective analysis. Future research should focus on refining this technology, including the use of newly designed transparent sockets and long-term follow-up, to overcome current limitations and enhance clinical applicability. With further development, IRT could become an essential component in the objective assessment of prosthetic socket fit, improving prosthetic design and patient outcomes.

References

1. Frölke JPM, Rommers GM, Boer AW de, Groenveld TD, and Leijendekkers R. Epidemiology of Limb Amputations and Prosthetic Use During COVID-19 Pandemic in the Netherlands. *Archives of Physical Medicine and Rehabilitation* 2023
2. Dillingham TR, Pezzin LE, Mackenzie EJ, and Burgess AR. Use and Satisfaction with Prosthetic Devices Among Persons with Trauma-Related Amputations A Long-Term Outcome Study. *American Journal of Physical Medicine Rehabilitation* 2001; 80(8):563–71
3. Radcliffe C and Foort J. Biomechanics of Below Knee Prosthetics. *The Patellar Tendon Bearing Below Knee Prosthesis*. University of California, Berkeley, 1961 :26–52
4. Safari R and Meier MR. Systematic review of effects of current transtibial prosthetic socket designs—Part 2: Quantitative outcomes. *Journal of Rehabilitation Research and Development* 2015; 52(5):509–26
5. Safari MR and Meier MR. Systematic review of effects of current transtibial prosthetic socket designs—Part 1: Qualitative outcomes. *Journal of Rehabilitation Research and Development* 2015; 52(5):491–508
6. Sanders JE, Youngblood RT, Hafner BJ, Cagle JC, McLean JB, Redd CB, Dietrich CR, Ciol MA, and Allyn KJ. Effects of socket size on metrics of socket fit in trans-tibial prosthesis users. *Medical Engineering and Physics* 2017; 44:32–43
7. Frossard L. Trends and opportunities in health economic evaluations of prosthetic care innovations. *Canadian Prosthetics and Orthotics Journal* 2021; 4(2)
8. Won N, Paul A, Garibaldi M, Baumgartner R, Kaufman K, Reider L, et al. Scoping review to evaluate existing measurement parameters and clinical outcomes of transtibial prosthetic alignment and socket fit. *Prosthet Orthot Int* 2022; 46(2):95–107
9. Gastel P van. Comparisons of (cost-)effectiveness of manual, hybrid, and digital shape capture and shape design techniques for transtibial and transfemoral prosthetic sockets: a scoping review. Unpublished 2024
10. Kingsley C and Patel S. Patient-reported outcome measures and patient-reported experience measures. *BJA Educ* 2017; 17(4):137–44
11. Al-Fakih EA, Osman NAA, and Adikan FRM. Techniques for interface stress measurements within prosthetic sockets of transtibial amputees: A review of the past 50 years of research. *Sensors (Switzerland)* 2016 Jul; 16(7)
12. Young PR, Hebert JS, Marasco PD, Carey JP, and Schofield JS. Advances in the measurement of prosthetic socket interface mechanics: a review of technology, techniques, and a 20-year update. *Expert Review of Medical Devices* 2023; 20(9):729–39
13. Hernandez-Contreras D, Peregrina-Barreto H, Rangel-Magdaleno J, and Renero-Carrillo F. Plantar thermogram database for the study of diabetic foot complications. *IEEE Access* 2019; 7:161296–307
14. Zenunaj G, Lamberti N, Manfredini F, Traina L, Acciari P, Bisogno F, Scian S, Serra R, Abatangelo G, and Gasbarro V. Infrared thermography as a diagnostic tool for the assessment of patients with symptomatic peripheral arterial disease undergoing infrapopliteal endovascular revascularisations. *Diagnostics* 2021; 11(9)
15. Gutierrez E, Castañeda B, Treuillet S, and Lucas Y. Combined thermal and color 3D model for wound evaluation from handheld devices. *SPIE-Intl Soc Optical Eng* 2021
16. Gutierrez E, Castañeda B, Treuillet S, and Hernandez I. Multimodal and multiview wound monitoring with mobile devices. *Photonics* 2021; 8(10)
17. Gutierrez E, Castañeda B, Mendoza H, Hernandez I, and Treuillet S. Smartphone-based 3D thermal monitoring of chronic wounds in a clinical setting. *Biomedical Signal Processing and Control* 2023
18. Zhang M, Turner-Smith A, and Roberts V. The reaction of skin and soft tissue to shear forces applied externally to the skin surface. *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine* 1994; 208:217–22
19. Wang-Evers M, Casper M, Glahn J, Luo T, Doyle A, and Karasik D. Assessing the impact of aging and blood pressure on dermal microvasculature by reactive hyperemia optical coherence tomography angiography. *Sci Rep* 2021; 11
20. Goller H, Lewis W, and McLaughlin R. Thermographic studies of human skin subjected to localized pressure. *Am J Roentgenol Radium Ther Nucl Med* 1971; 113(4):749–54
21. Tang Y, Xu F, Lei P, Li G, and Tan Z. Spectral analysis of laser speckle contrast imaging and infrared thermography to assess skin microvascular reactive hyperemia. *Skin Research and Technology* 2023; 29(4):e13308
22. Cutti A, Perego P, Fusca M, Sacchetti R, and Andreoni G. Assessment of lower limb prosthesis through wearable sensors and thermography. *Sensors (Switzerland)* 2014; 14(3):5041–55
23. Cutti A, Morosato F, Gentile C, Gariboldi F, Hamoui G, and Santi M. A Workflow for Studying the Stump–Socket Interface in Persons with Transtibial Amputation through 3D Thermographic Mapping. *Sensors* 2023; 23(11)
24. Cárdenas A, Uribe J, Font-Llagunes J, Hernández A, and Plata J. The effect of prosthetic alignment on the stump temperature and ground reaction forces during gait in transfemoral amputees. *Gait Posture* 2022; 95:76–83

25. Vieira I, Silva Honório J da, Santos PB dos, Lorenzetti Branco L, Lorenzetti Branco H, and Luz CT da. Thermographic Evaluation, Residual Limb Skin Sensitivity, and Adaptation to the Prosthesis of Individuals with Lower-Limb Amputation with Prosthesis Provided by the Universal Health System. 2023
26. Mendes E, Silva A, Correia R, Crisóstomo C, Vaz F, and Gabriel J. Thermography as an Alternative Tool to Determine Pressure Distribution on the Stump of Transfemoral Amputees. *Thermology International* 2012; 22(3):99–100
27. Silva N, Castro H, Carvalho L, Chaves E, Ruela L, and Lunes D. Reliability of infrared thermography images in the analysis of the plantar surface temperature in diabetes mellitus. *Journal of chiropractic medicine* 2018; 17:30–5
28. Hildebrandt C and Raschner C. An Intra-Examiner Reliability Study of Knee Temperature Patterns With Medical Infrared Thermal Imaging. *Thermology international* 2009; 19(3)
29. Jesus Guirro R de, Oliveira Lima Leite Vaz M, Neves L das, Dibai-Filho A, Carrara H, and Oliveira Guirro E de. Accuracy and reliability of infrared thermography in assessment of the breasts of women affected by cancer. *Journal of medical systems* 2017; 41:1–6
30. Calvo-Lobo C, San-Antolín M, García-García D, Becerro-de-Bengoa-Vallejo R, Losa-Iglesias ME, Cosín-Matamoras J, Casado-Hernández I, Martínez-Jiménez EM, Mazoteras-Pardo V, and Rodríguez-Sanz D. Intra-and inter-session reliability and repeatability of an infrared thermography device designed for materials to measure skin temperature of the triceps surae muscle tissue of athletes. *PeerJ* 2023; 11:e15011
31. Steketee J. Spectral emissivity of skin and pericardium. *Physics in Medicine Biology* 1973; 18(5):686–94
32. Ramirez-GarcíaLuna J, Bartlett R, Arriaga-Caballero J, Fraser R, and Saiko G. Infrared thermography in wound care, surgery, and sports medicine: a review. *Frontiers in physiology* 2022; 3(13)
33. Teledyne FLIR. How Does Emissivity Affect Thermal Imaging? 2021. Available from: <https://www.flir.eu/discover/professional-tools/how-does-emissivity-affect-thermal-imaging/> [Accessed on: 2023 Dec 1]
34. Hanspal RS, Fisher K, and Nieveen R. Prosthetic socket fit comfort score. *Disability and Rehabilitation* 2003; 25(22):1278–80
35. Byrt T. How good is that agreement? *Epidemiology* 1996; 7:561
36. Bland J and Altman D. Measuring agreement in method comparison studies. *Stat Methods Med Res* 1999; 8(2):135–60
37. Peery JT, Klute GK, Blevins JJ, and Ledoux WR. A three-dimensional finite element model of the transibial residual limb and prosthetic socket to predict skin temperatures. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 2006; 14:336–43
38. Beil TL and Street GM. Comparison of interface pressures with pin and suction suspension systems. *Journal of Rehabilitation Research & Development* 2004; 41
39. Jonkergouw N, Prins MR, Donse D, Wurff P van der, Dieën JH van, Buis A, and Houdijk H. Application of ultrasound to monitor in vivo residual bone movement within transtibial prosthetic sockets. *Scientific Reports* 2024; 14:9725
40. Karamousadakis M, Porichis A, Ottikkutti S, Chen D, and Vartholomeos P. A sensor-based decision support system for transfemoral socket rectification. *Sensors* 2021; 21:3743
41. Mokkin LB, Prinsen CA, Patric DL, Alonso J, Bouter LM, Vet HC de, and Terwee CB. COSMIN Study Design checklist for Patient-reported outcome measurement instruments. 2019. Available from: https://www.cosmin.nl/wp-content/uploads/COSMIN-study-designing-checklist_final.pdf [Accessed on: 2024 Aug 1]
42. Hafner BJ, Morgan SJ, Askew RL, and Salem R. Psychometric evaluation of self-report outcome measures for prosthetic applications. *Journal of rehabilitation research and development* 2016; 53:797
43. Abreu de Souza M, Alka Cordeiro DC, Oliveira Jd, Oliveira MFAd, and Bonafini BL. 3D multi-modality medical imaging: Combining anatomical and infrared thermal images for 3d reconstruction. *Sensors* 2023; 23:1610
44. Domina T, Kinnicutt P, and MacGillivray M. Thermal pattern variations analyzed using 2D/3D mapping techniques among females. *Journal of Textile and Apparel, Technology and Management* 2011; 7
45. Zeiler W. Basisboek Ontwerpen: van methodisch ontwerpen tot integraal ontwerpen. Noordhoff, 2014
46. Hollinger A and Wanderley MM. Evaluation of commercial force-sensing resistors. *Proceedings of the International Conference on New Interfaces for Musical Expression, Paris, France* 2006 :4–8
47. Swanson EC, Weathersby EJ, Cagle JC, and Sanders JE. Evaluation of force sensing resistors for the measurement of interface pressures in lower limb prosthetics. *Journal of biomechanical engineering* 2019; 141:101009
48. Buis A and Convery P. Calibration problems encountered while monitoring stump/socket interface pressures with force sensing resistors: techniques adopted to minimise inaccuracies. *Prosthetics and orthotics international* 1997; 21:179–82
49. Swanson EC, McLean JB, Allyn KJ, Redd CB, and Sanders JE. Instrumented socket inserts for sensing interaction at the limb-socket interface. *Medical engineering & physics* 2018; 51:111–8
50. Tekscan, Inc. Force Sensors. Available from: <https://www.tekscan.com/force-sensors> [Accessed on: 2023 Dec 1]
51. Tekscan, Inc. F-Socket System. Available from: <https://www.tekscan.com/products-solutions/systems/f-socket-system> [Accessed on: 2023 Dec 1]
52. Dumbleton T, Buis AW, McFadyen A, McHugh BF, McKay G, Murray KD, and Sexton S. Dynamic interface pressure distributions of two transtibial prosthetic socket concepts. *Journal of Rehabilitation Research & Development* 2009; 46
53. Goh JCH, Lee PVS, and Chong SY. Comparative study between patellar-tendon-bearing and pressure cast prosthetic sockets. *Journal of Rehabilitation Research & Development* 2004; 41
54. Tabor J, Agcayazi T, Fleming A, Thompson B, Kapoor A, Liu M, Lee MY, Huang H, Bozkurt A, and Ghosh TK. Textile-based pressure sensors for monitoring prosthetic-socket interfaces. *IEEE sensors journal* 2021; 21:9413–22

55. Tekscan, Inc. Fujifilm Prescale Film. Available from: <https://www.tekscan.com/products-solutions/pressure-sensing-film/fujifilm-prescale-film> [Accessed on: 2023 Dec 1]
56. Ghoseiri K, Zheng YP, Hing LLT, Safari MR, and Leung AK. The prototype of a thermoregulatory system for measurement and control of temperature inside prosthetic socket. *Prosthetics and orthotics international* 2016; 40:751–5
57. Majdič N, Vidmar G, and Burger H. Establishing K-levels and prescribing transtibial prostheses using six-minute walk test and one-leg standing test on prosthesis: A retrospective audit. *International Journal of Rehabilitation Research* 2020; 43(3):266–71
58. NHG-werkgroep: Wiersma T. NHG-Standaard Decubitus. 2015
59. Koo TK and Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of chiropractic medicine* 2016; 15:155–63
60. De Vet H, Terwee C, Mokkink L, and Knol D. Chapter 5: Reliability. *Measurement in Medicine*. Cambridge University Press, 2011 :96–149
61. De Vet H, Terwee C, Mokkink L, and Knol D. Chapter 8: Interpretability. *Measurement in Medicine*. Cambridge University Press, 2011 :227–74

A

Evaluation of measurement systems

This Appendix provides an evaluation of various measurement systems for objective transtibial prosthetic socket fit assessment. First, a list of requirements for the measurement system was defined. Second, measurement systems from both the literature, including the preceding scoping review, and commercial sources were identified. Third, these measurement systems were compared against the predefined requirements and scored based on insights from literature, expert opinions, and conversations with manufacturers, to select the most suitable system for this study.

A.1. List of requirements

Following a standardised methodology⁴⁵, the requirements for the measurement system are outlined in Table A.1. These requirements were categorised as fixed (i.e., essential criteria), flexible (i.e., adjustable based on conditions), and wishes (i.e., desirable, but not mandatory). The weighting system for flexible requirements ranged from 1 to 3. The value assigned to the measurement system ranged from 1 to 4 and was multiplied by this weight, influencing the system's overall score. The maximum possible score for each measurement system is 48 points.

Table A.1: List of system requirements for socket fit measurement

Category and requirement	Priority	Weight
<i>Functionality</i>		
Can assess the construct of socket fit	Fixed	-
Measures both normal and shear pressure at the residual limb-prosthetic socket interface	Fixed	-
No prosthetic socket modifications necessary to use the measurement system	Flexible	2
Provides coverage for the complete surface of the residual limb	Flexible	3
<i>Durability</i>		
Maintains operational integrity and performance throughout the warranty period	Flexible	2
<i>Validity</i>		
Previously used in assessments of prosthetic socket fit for patients	Fixed	-
Scientifically validated for assessment of socket fit in transtibial prostheses	Flexible	3
<i>User characteristics</i>		
User-friendly, requiring no advanced training for operation	Fixed	-
Ensures accurate measurements with minimal calibration steps	Flexible	1
<i>Patient interaction</i>		
Non-invasive method	Fixed	-
Safe for all patient demographics, minimizing risk to skin integrity	Fixed	-
Comfortable for patient during use with minimal interference with the residual limb	Flexible	1
<i>Availability and financial aspect</i>		
Commercially available system	Fixed	-
Affordability, without compromising key functionalities	Wish	-

A.2. Measurement systems

The following relevant measurement systems with potential for application in this study were identified.

A.2.1. Pressure sensors

Force Sensitive Resistor (FSR)

FSRs are sensors whose resistance changes when force, pressure, or mechanical stress is applied. These systems need to be configured in a custom-made setup on the residual limb. FSRs capture force at a single point (unlike pressure, which is force across a surface) and measure only normal pressure, not shear pressure. Although they can be placed inside the prosthetic socket, their thickness can interfere with the residual limb-socket interface. Drift error values ranged from 4% to 33.2%, accuracy error from 2.8% to 92%, and hysteresis from 0.02% to 41.88%.^{12,11,46} When used on curved surfaces, as is common in prosthetic socket assessments, FSRs show even greater measurement errors (e.g., a 32% decrease in accuracy on curved surfaces for FlexiForce sensors). They are also negatively affected by shear pressure.^{47,12} Calibration is challenging due to the complex pressure-voltage relationship, which varies between studies and sensors.^{48,47} Despite their low cost and commercial availability, numerous sensors would be required to adequately cover the residual limb. FSRs have been employed in various research applications for prosthetics.^{49,12} For example, FlexiForce sensors from Tekscan are commonly used.⁵⁰

FSR systems

Commercially available systems consisting of FSRs, specifically designed to quantify prosthetic socket-residual limb interface pressures, are available. The most commonly used system in prosthetic research is the Tekscan F-Socket system.^{51,52} Similar inaccuracies and difficulties as with singular FSRs are present. Additionally, manufacturer discussions revealed that the F-Socket system must be cut to fit a specific residual limb shape, rendering it unusable for other shapes. The system may also be too short for long residual limbs. The purchase of multiple pressure pads is necessary to cover the entire limb. Furthermore, the F-socket is prone to breakage, requiring the replacement of the entire pad, making it a costly solution.

Force transducer (FTs)

FTs, such as elastic element load-cells, measure force by detecting changes in electrical resistance on a deformable element. Due to their large profile, these sensors cannot be placed directly between the prosthetic socket and the residual limb. Instead, they must be installed through holes made in the prosthetic socket. This installation method can cause the sensor to protrude from the socket wall, potentially leading to inaccurate measurements. FTs offer advantages over FSR and printed circuit technologies, with lower drift error (0.35-1.35%), accuracy error (0.25-0.71%), and hysteresis (0.5-8%). Moreover, they can measure both normal and shear forces. However, like FSRs, they measure force at a single point, and installing multiple sensors can alter the properties of the prosthetic socket interface.¹² These sensors are used in studies assessing prosthetic sockets.⁵³

Alternative sensors

Alternative sensors, such as capacitive, fibre optic, fluid-filled, textile-based sensors, and electronic skin, offer promising benefits over conventional pressure sensors. They typically exhibit higher accuracy, greater sensitivity, lower hysteresis and drift, and the capability to measure both normal and shear pressure. However, these technologies are largely in the experimental stage, with limited patient application data.^{12,54} Initiating new research with these technologies necessitates extensive groundwork and specialised expertise.

A.2.2. Pressure mapping film

Pressure mapping film, produced by Fujifilm, is a pressure-sensitive film that reveals pressure distribution through a coloured impression after normal or shear pressure is applied. Its precision is relatively low (10%), and it has not been used in prosthetic socket evaluation. Additionally, the film is single-use and expensive. To quantitatively interpret the measurements, a separate and costly system (Prescale Digital Analysis System) is required.⁵⁵ Samples of this film were obtained to evaluate its functionality in a mock setup. However, even the act of donning a prosthetic socket or touching the film substantially coloured the film.

A.2.3. Infrared thermography (IRT)

IRT using a thermal camera system is a novel method for assessing socket fit, capable of mapping hotspots on the residual limb, making it applicable for prosthetic socket evaluation.^{22,23,24,26} It offers high thermal

sensitivity and the ability to provide a complete image of the stump.²³ Its durability is advantageous since it does not rely on separate sensors that could break. However, limited research has been conducted on the direct relationship between peak pressures on the residual limb and hotspots, and the reliability of IRT measurements. Although it has a relatively high initial cost, discussions with manufacturers and the successful borrowing of an advanced FLIR camera system for preliminary studies support its potential use.

A.2.4. Thermal sensors

Thermal sensors that measure temperature at specific points are sometimes used in prosthetic socket research.^{22,56,37} However, they do not quantify the overall construct of socket fit, but rather a subset of socket comfort.

A.3. Scoring

Table A.2 provides the evaluation of various measurement systems based on predefined requirements. The table indicates that while most systems can quantify socket fit and are user-friendly, only a few, like FTs and IRT, can capture both normal and shear forces. Durability and minimal calibration are strengths for some systems but not all. Proven use for socket fit and scientific validation varied. Affordability and commercial availability are limiting factors for some systems, with only FTs and IRT satisfying all fixed requirements. IRT showed the highest percentage of the maximum score for flexible requirements at 77%, hence indicating a strong potential for practical application.

Table A.2: Evaluation of measurement systems according to predefined requirements. Compliance with fixed criteria and wishes is indicated as yes (Y) or no (N), while compliance with flexible criteria is assessed using weighted scores (total score in brackets). The total percentage of the maximum score represents the proportion of the total possible score achieved for the flexible requirements. Abbreviations: FSR = force sensitive resistors, FT = force transducer, AS = alternative sensors, PMF = pressure mapping film, TS = thermal sensor, IRT = infrared thermography, Y = yes, N = no.

Requirement	FSR	FSR system	FTs	AS	PMF	TS	IRT
Assess socket fit	Y	Y	Y	Y	Y	N	Y
Normal and shear forces	N	N	Y	Y	Y	N	Y
No socket modifications	2 × 4 (8)	2 × 4 (8)	2 × 1 (2)	2 × 4 (8)	2 × 4 (8)	2 × 4 (8)	2 × 4 (8)
Residual limb coverage	3 × 1 (3)	3 × 3 (9)	3 × 1 (3)	3 × 4 (12)	3 × 2 (6)	3 × 1 (3)	3 × 4 (12)
Durability	2 × 3 (6)	2 × 1 (2)	2 × 3 (6)	2 × 3 (6)	2 × 1 (2)	2 × 3 (6)	2 × 4 (8)
Proven use for socket fit	Y	Y	Y	Y	N	N	Y
Scientifically validated	3 × 2 (6)	3 × 2 (6)	3 × 2 (6)	3 × 1 (3)	3 × 1 (3)	3 × 2 (6)	3 × 1 (3)
User-friendly	Y	Y	Y	Y	Y	Y	Y
Minimal calibration	1 × 1 (1)	1 × 1 (1)	1 × 1 (1)	1 × 1 (1)	1 × 4 (4)	1 × 2 (2)	1 × 2 (2)
Non-invasive	Y	Y	Y	Y	Y	Y	Y
Safe for patient use	Y	Y	Y	Y	Y	Y	Y
Comfortable for patient	1 × 3 (3)	1 × 2 (2)	1 × 3 (3)	1 × 4 (4)	1 × 2 (2)	1 × 3 (3)	1 × 4 (4)
Affordability	Y	N	Y	?	N	Y	N
Commercially available	Y	Y	Y	N	Y	Y	Y
Satisfied all fixed requirements?	N	N	Y	N	N	N	Y
Percentage of maximum score (flexible requirements)	44%	58%	44%	71%	52%	58%	77%

B

Participant data

This Appendix describes the participant data collection at the start and during the experimental procedure.

B.1. Data collected at the start of the procedure

B.1.1. Amputation side, reason, and time of amputation

The researcher documents the side of amputation, along with the reason (vascular, traumatic, infectious, oncological, congenital, other) and the time (month and year) of amputation.

B.1.2. Residual limb

The researcher measures the length of the residual limb, including the length of the residual limb (apex patella to distal residual limb end) and length of the tibial bone (apex patella to distal tibia end), and assesses the residual limb characteristics (shape, firmness, boniness, and skin integrity).

B.1.3. Prosthetic socket

The type of prosthetic socket, suspension, and liner are documented, along with whether the participant is wearing socks around the residual limb and, if so, how many. The researcher asks the participant to describe the fit of the unmodified socket as 'too small,' 'good,' or 'too large.'

B.1.4. Mobility level

The researcher inquires about the participant's current mobility and applies the K-level classification⁵⁷ (levels K1 through K4) based on the information provided. The categories are:

- K1: Indoor ambulator; can manage walking distances on a level surface at a steady pace with the aid of a prosthesis.
- K2: Limited community ambulator; can handle limited walking distances and low environmental barriers (such as curbs, steps, and uneven floors) with a prosthesis.
- K3: Unlimited community ambulator; can move freely with a prosthesis on various types of terrain, walk at different speeds, and handle most environmental barriers, and can engage in therapeutic, occupational, or recreational activities without overloading the prosthesis.
- K4: High community ambulator with extraordinary demands; can move without restrictions with a prosthesis, including enduring higher load, stress, or energy levels, walking duration and distance are unlimited, the prosthesis can accommodate the typical needs of a child, active adult, or athlete.

B.1.5. Skin type

The researcher determines the Fitzpatrick classification of skin type by visual assessment of the skin and inquires about sensitivity to sunlight. The Fitzpatrick classification includes:

- Skin type 1: Burns easily, never tans with sun exposure.
- Skin type 2: Burns easily, tans minimally with sun exposure.
- Skin type 3: Rarely burns, tans gradually with sun exposure.
- Skin type 4: Rarely burns, tans easily with sun exposure.

- Skin type 5: Rarely burns, tans profusely with sun exposure.
- Skin type 6: Never burns, tans profusely with sun exposure.

B.1.6. Medication use

The researcher records the participant's medication use.

B.1.7. Use of alcohol and smoking habits

The researcher inquires about the participant's alcohol consumption and smoking habits (frequency and amount).

B.1.8. Height and weight

The researcher measures the participant's height and weight, including clothing and prosthesis, and weight of the prosthesis. If the participant has recently measured their height and weight, these values are used.

B.1.9. Body temperature

The researcher measures the participant's body temperature in degrees Celsius using an ear thermometer.

B.2. Data collected during the procedure

B.2.1. Decubitus classification

The researcher determines the decubitus classification⁵⁸ through visual inspection and evaluation of the skin by pressing observed redness on the skin. The categories are:

- No skin abnormality.
- Category 1: Non-blanchable redness of intact skin.
- Category 2: Excoriation or blister.
- Category 3: Full-thickness skin loss, but bone, tendons, and muscles are not exposed.
- Category 4: Full-thickness tissue loss; bone, tendon, or muscle visible.

B.2.2. Socket Comfort Score (SCS) and localisation

The researcher conducts a comfort assessment by asking the following questions twice to each participant, once after walking with the unmodified socket and once after walking with the modified socket:

- Comfort assessment using the SCS: "On a scale from 0 to 10, where 0 represents the most uncomfortable feeling you can imagine while wearing a prosthetic socket, and 10 represents the most comfortable feeling, how would you rate the comfort of the prosthetic socket you have just walked with?"
- Localisation of uncomfortable areas: "Did you experience any specific spots of discomfort or pain during walking with the prosthetic socket? If yes, please point out these spots on your residual limb. For each indicated spot, how would you rate the level of pain or discomfort on the same scale from 0 to 10, where 0 represents severe pain or discomfort and 10 represents no pain or discomfort?" The researcher draws the indicated spots on a schematic representation of the residual limb (Figure B.1).

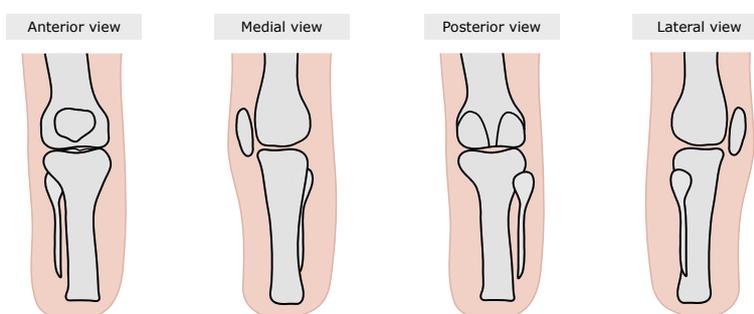


Figure B.1: Anterior, medial, posterior and lateral views of a right lower leg residual limb, with the residual limb skin in light pink and the bones in grey. These visualisations serve as a guide for the researcher to mark the indicated locations of pain or discomfort. For a left lower leg residual limb, a mirrored view is used.

C

Data analysis code

This Appendix includes essential Python code segments used for analysing thermal data.

C.1. Automatic marker detection

The following section presents the Python function developed to automatically detect markers on the binary mask of the cropped residual limb. Figure 2.2 illustrates an example of the markers as detected by the script.

```
1 def detect_markers(binary_mask):
2     """
3     Detects markers on a binary image of a cropped residual limb using blob detection.
4     Parameters:
5     - binary_mask (numpy.ndarray): A binary image on which the markers can be detected.
6     Returns:
7     - im_with_keypoints (numpy.ndarray): The binary image overlaid with detected keypoints,
8       represented as blue circles.
9     - keypoints (list of cv2.KeyPoint): A list of keypoints detected in the binary mask. Each
10      keypoint represents a detected marker.
11     """
12
13     if binary_mask.max() == 1:
14         binary_mask = (binary_mask * 255).astype(np.uint8)
15
16     params = cv2.SimpleBlobDetector_Params()
17     params.filterByColor = True
18     params.blobColor = 0
19     params.minThreshold = 0
20     params.maxThreshold = 100
21     params.filterByArea = True
22     params.minArea = 10
23     params.maxArea = 30
24     params.filterByCircularity = True
25     params.minCircularity = 0.5
26     params.filterByConvexity = True
27     params.minConvexity = 0.2
28     params.filterByInertia = True
29     params.minInertiaRatio = 0.2
30
31     detector = cv2.SimpleBlobDetector_create(params)
32     keypoints = detector.detect(binary_mask)
33     im_with_keypoints = cv2.drawKeypoints(binary_mask, keypoints, np.array([]), (0, 0, 255),
34     cv2.DRAW_MATCHES_FLAGS_DRAW_RICH_KEYPOINTS)
35
36     return im_with_keypoints, keypoints
```

C.2. Image registration

The following section describes the Python function for feature-based image registration of TR1 to TW1 and TW2. Feature-based registration aligns images by identifying and matching relevant features across

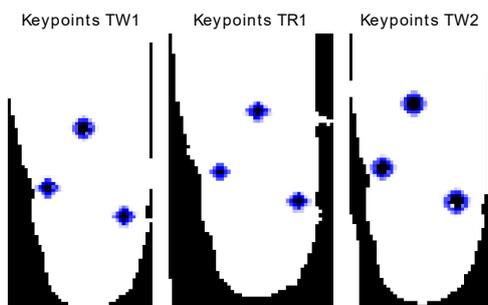


Figure C.1: Binary masks of the cropped residual limb of one participant with detected markers (blue). Abbreviations: TW1 = thermal images post-walking with unmodified socket, TR1 = first set of post-resting thermal images, TW2 = thermal images post-walking with modified socket.

different images. Figure 2.2 shows examples of the fixed, moving, and transformed images as a result of this registration process.

```

1 def match_and_transform_points(points_TR, points_TW, cropped_foreground_TR,
2   cropped_foreground_TW, distance_threshold=3):
3   """
4   Matches points between two sets using Euclidean distance and computes an affine
5   transformation to align TR with TW on these points.
6   Parameters:
7   - points_TR (numpy.ndarray): Array of points from the reference image (TR).
8   - points_TW (numpy.ndarray): Array of points from the target image (TW).
9   - cropped_foreground_TR (numpy.ndarray): Cropped foreground of the reference image (TR).
10  - cropped_foreground_TW (numpy.ndarray): Cropped foreground of the target image (TW).
11  - distance_threshold (float, optional): Maximum allowed distance between matched points.
12  Defaults to 3.
13  Returns:
14  - transformation_matrix (numpy.ndarray or None): The 2x3 affine transformation matrix if
15  successful; None otherwise.
16  - matched_points_TR (numpy.ndarray): Array of points from points_TR that were
17  successfully matched and transformed.
18  - matched_points_TW (numpy.ndarray): Array of points from points_TW that correspond to
19  matched_points_TR.
20  - transformed_image (numpy.ndarray or None): The transformed version of
21  cropped_foreground_TR aligned to cropped_foreground_TW; None if transformation fails.
22  """
23
24  distances = cdist(points_TR, points_TW)
25  min_distances_indices = np.argmin(distances, axis=1)
26  min_distances = distances[np.arange(len(points_TR)), min_distances_indices]
27
28  valid_matches = min_distances < distance_threshold
29  matched_points_TR = points_TR[valid_matches]
30  matched_points_TW = points_TW[min_distances_indices][valid_matches]
31
32  if matched_points_TR.shape != matched_points_TW.shape:
33      print("Shapes of matched point arrays do not match.")
34      transformed_image = None
35
36  transformation_matrix, status = cv2.estimateAffinePartial2D(matched_points_TR,
37  matched_points_TW)
38
39  if transformation_matrix is not None:
40      print("Transformation matrix:", transformation_matrix)
41      height, width = cropped_foreground_TW.shape[:2]
42      transformed_image = cv2.warpAffine(cropped_foreground_TR, transformation_matrix, (
43  width, height))
44  else:
45      print("Failed to estimate affine transformation. Check input points.")
46      transformed_image = None
47
48  return transformation_matrix, matched_points_TR, matched_points_TW, transformed_image

```

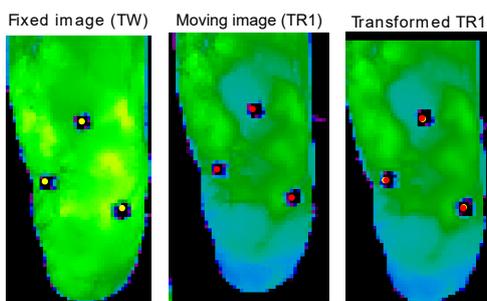


Figure C.2: Cropped residual limb images of one participant with fixed image markers (yellow), moving image markers (red), and transformed image showing both fixed (yellow) and transformed (red) markers. Background values are excluded. Abbreviations: TW1 = thermal images post-walking with unmodified socket, TR1 = first set of post-resting thermal images, TW2 = thermal images post-walking with modified socket.

C.3. Difference maps

The following section details the function developed to calculate difference maps of the residual limb, focusing on areas without background thermal values. The process involves subtracting the transformed moving image TR1 from the fixed images TW1 and TW2, respectively. This subtraction highlights thermal differences with respect to the post-resting state.

```

1 def create_diff_map_without_background(cropped_foreground_TW, transformed_image,
2   cropped_binary_mask_TW, moment_TW, vmin_value_diff, vmax_value_diff):
3   """
4   Creates a differential map by subtracting the transformed TR (cropped and only foreground
5   ) from the TW (cropped and only foreground), excluding non-overlapping areas.
6   Parameters:
7   - cropped_foreground_TW (numpy.ndarray): Cropped foreground image of TW.
8   - transformed_image (numpy.ndarray): Transformed version of the moving image (TR1)
9     aligned to the fixed image (TW).
10  - cropped_binary_mask_TW (numpy.ndarray): Binary mask of TW.
11  Returns:
12  - diff_map_corrected (numpy.ndarray): The corrected difference map where differences are
13    only calculated for overlapping areas of interest, excluding background and refined
14    through erosion to reduce alignment errors.
15  """
16  cropped_foreground_TW = cropped_foreground_TW.astype(np.float32)
17  transformed_image = transformed_image.astype(np.float32)
18
19  valid_foreground_mask = (cropped_foreground_TW != 0)
20  valid_transformed_mask = (transformed_image != 0)
21  overlap_mask = valid_foreground_mask & valid_transformed_mask
22
23  diff_map = np.full_like(cropped_foreground_TW, np.nan, dtype=np.float32)
24  diff_map[overlap_mask] = cropped_foreground_TW[overlap_mask] - transformed_image[
25    overlap_mask]
26
27  kernel_size = 2
28  kernel = np.ones((kernel_size, kernel_size), np.uint8)
29  eroded_mask = cv2.erode(cropped_binary_mask_TW.astype(np.uint8), kernel, iterations=1)
30  eroded_mask_overlap = cv2.erode(overlap_mask.astype(np.uint8), kernel, iterations=1)
31  eroded_mask = eroded_mask.astype(bool)
32  eroded_mask_overlap = eroded_mask_overlap.astype(bool)
33
34  final_mask = eroded_mask & eroded_mask_overlap
35  diff_map_corrected = np.where(final_mask, diff_map, np.nan)
36
37  return diff_map_corrected

```

C.4. ROI analysis

The following section describes the Python function developed to place circular regions of interest (ROIs) on the difference maps at manually selected coordinates. Figure 2.2 displays examples of these difference maps annotated with ROIs.

```

1 def analyse_circular_roi(diff_map_nan_background, center_coordinates, circle_radius):
2     """
3     Analyses a circular region of interest (ROI) within a difference map. This function
4     calculates various statistical measures for the pixel values inside the defined
5     circle.
6     Parameters:
7     - diff_map_nan_background (numpy.ndarray): The difference map where non-interest areas (
8       background) are set to NaN.
9     - center_coordinates (tuple): A tuple (x, y) representing the center of the circular ROI
10    within the differential map.
11    - circle_radius (int): The radius of the circle defining the ROI in pixels.
12
13    Returns:
14    - mean_values_roi (float): The mean value of the pixel intensities within the ROI,
15      providing an average temperature difference with respect to baseline.
16    - sd_values_roi (float): The standard deviation of the pixel intensities within the ROI.
17    - min_values_roi (float): The minimum pixel intensity within the ROI, highlighting the
18      lowest temperature difference with respect to baseline.
19    - max_values_roi (float): The maximum pixel intensity within the ROI, highlighting the
20      highest temperature difference with respect to baseline.
21    - median_values_roi (float): The median pixel intensity within the ROI.
22    - circle_mask (numpy.ndarray): A boolean array where True values indicate pixels within
23      the specified circle, used for further analysis.
24    """
25    x, y = np.meshgrid(np.arange(diff_map_nan_background.shape[1]), np.arange(
26      diff_map_nan_background.shape[0]))
27
28    distance_from_center = np.sqrt((x - center_coordinates[0])**2 + (y - center_coordinates
29      [1])**2)
30
31    circle_mask = distance_from_center <= circle_radius
32    values_in_roi = diff_map_nan_background[circle_mask]
33
34    mean_values_roi = np.nanmean(values_in_roi)
35    sd_values_roi = np.nanstd(values_in_roi)
36    min_values_roi = np.nanmin(values_in_roi)
37    max_values_roi = np.nanmax(values_in_roi)
38    median_values_roi = np.nanmedian(values_in_roi)
39
40    return mean_values_roi, sd_values_roi, min_values_roi, max_values_roi, median_values_roi,
41      circle_mask

```

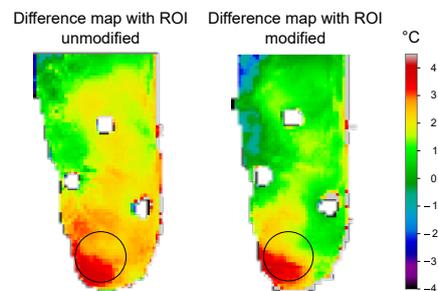


Figure C.3: Difference maps for the unmodified and modified socket of one participant. The black circle indicates the positioning of the ROI in both images. Abbreviation: ROI = region of interest, C = Celsius.

D

Reproducibility calculations

This Appendix describes the calculations for the reliability and agreement parameters.

The two-way mixed-effects model ($ICC_{3.1 \text{ agreement}}$), accounting for both random and systematic variations, was calculated in IBM SPSS Statistics. The ICC was based on Formula D.1⁵⁹:

$$ICC_{3.1 \text{ agreement}} = \frac{MS_R - MS_E}{MS_R + (k - 1)MS_E}, \quad (\text{D.1})$$

where MS_R is the mean square for rows, MS_E is the mean square for error, and k is the number of measurements.

The standard error of measurement (SEM) was calculated using Formula D.2⁶⁰:

$$SEM_{\text{agreement}} = \sqrt{\sigma_O^2 + \sigma_{\text{residual}}^2}, \quad (\text{D.2})$$

where σ_O^2 is the variance attributed to systematic differences between measurement occasions, $\sigma_{\text{residual}}^2$ is the residual or error variance (including random error). The VARCOMP analysis in IBM SPSS Statistics was used to determine variances.⁶⁰

Subsequently, the smallest detectable change (SDC), indicating the minimal detectable difference that can be considered above the threshold of measurement error, was derived using Formula D.3⁶¹:

$$SDC = 1.96 \times \sqrt{2} \times SEM, \quad (\text{D.3})$$

where SEM is the standard error of measurement.