INTER AND INTRAEXAMINER RELIABILITY OF A NEW METHOD OF INFRARED THERMOGRAPHY ANALYSIS OF PATELLAR TENDON

Authors: Francisco J. Molina-Payá (1), José Ríos-Díaz (2), Jacinto J. Martínez-Payá (3)

Affiliation: 1) Doctoral Program in Health Sciences. Universidad Católica de Murcia, Murcia (Spain) ORCID: 0000-0002-2343-3613; 2) iNFIS Research Group. Centro de Ciencias de la Salud San Rafael. Universidad Antonio de Nebrija, Fundación de San Juan de Dios, Madrid (Spain) ORCID: 0000-0002-4786-3351; 3) Grupo de Investigación Discapacidad y Fisioterapia. Departamento de Fisioterapia, Facultad de Medicina. Universidad de Murcia, Murcia (Spain) ORCID: 0000-0002-9214-812X

Short title: Reliability of infrared thermography of tendon

Corresponding author:

Dr. José Ríos-Díaz jrios@nebrija.es
Centro de CC de la Salud San Rafael
Universidad Antonio de Nebrija
Pº de La Habana, 70 bis
28036 Madrid

ABSTRACT

Thermography is an imaging technique that records the infrared radiation of the skin with a potential to detect asymmetries in body temperature and relating them to pathologies or risk of injury. However, the location of landmarks and region of interest (ROI) can influence the recording process. In this study we describe the reliability of a method to select the ROI without modifying the original thermogram over 68 thermograms (17 participants) of patellar tendon with an infrared camera (softwareOptris 450PI). The width-height, X-Y axis and mean temperature of the ROI was measured by two examiners.All the lower limits of the Intraclass Correlation Coeficiente (CCI) were over 0.84 without biases. The skin temperature was similar for both intra-examiner ($F_{1,33}=0.488$; $p=0.490$) and inter-examiner ($F_{1,33}=0.011$; $p=0.917$). The mean differences were 0.006ºC (Limit of agreement (LoA): -0.10 to 0.10). The differences between examiners were 0.001 ºC (LoA: -0.13 to 0.13). All temperature differences were <0.25ºC. The method used has very good intra and inter-examiner reliability and reproducibility and the influence on the skin temperature is negligible. It will be of interest to extend the study to ascertain the reliability using different thermal cameras and software, which could increase the strength of the method.

KEYWORDS

Thermography; reliability; patellar tendon; infrared; interexaminer; intraexaminer

1. INTRODUCTION

Thermoregulation of the skin is a complex system that is mainly linked to cutaneous vascularization and to the sympathetic nervous system as regulator [1,2], maintaining, in healthy subjects, a thermal symmetry between both sides of the body [3–5].

Infrared thermography (IRT) is an imaging technique that captures infrared radiation in a fast, safe and low-cost way for both the diagnosis and monitoring of various pathologies associated with temperature changes, making it the technique of choice for many authors [6]. The extreme sensitivity of IRT permits the detection of temperature variations generated by metabolic changes in situations such as the regeneration of injured tissue [7], alterations of the sympathetic nervous system [8], inflammatory processes [9], infections [10], vascular disorders [11] or hormonal disorders [12].

Although some authors have reported good reproducibility [13], a comprehensive methodology and protocol is required to accurately determine temperature [14]. The study of factors that may influence thermal capture has allowed a protocol for obtaining thermographs to be normalized [15,16]. In this sense, the IRT of structures or areas determined on the body through the selection of regions of interest (ROI) can be considered as the most reliable method in static situations, whereas other methods are needed when the exploration is carried out in dynamic processes, involving the use averages of areas around hot spots within the area explored [17].

Unlike the protocol for IRT, there is no clear consensus for the determination of ROI [18]. Thus, some authors delimit the ROI by means of markers on the skin [9,19–21], with the disadvantage that manipulation of the area during positioning can alter the temperature of the ROI [22]. Other authors delimit the ROI on already captured thermographs [23,24], although accurately interpreting anatomical references in this way is difficult [25]. Additionally, software is being developed that allows the automatic positioning of ROI [26]. Recently, a protocol has been developed by consensus to improve the acquisition of thermographic images [27]. However, the process of placing reference markers on the skin may introduce bias into the recorded temperature.

We propose an alternative method to select the ROI, while avoiding any modification of the original thermogram. The aim of this study was to assess the reliability and reproducibility of this IRT method to select ROIs by means of markers on the skin without manipulation of the region, which might alter the interpretation of the thermogram.

2. MATERIAL AND METHODS

Study subjects

A total of 17 participants (34 bilateral images) (7 women and 10 men) between 18 and 62 years old (mean: 32.2 years; S.D:10.9 years) were voluntarily recruited from a private Physical Therapy Centre (Murcia, Spain) in July and August 2018. All participants were informed and signed the informed consent document. The study was approved by the Ethical Committee of the Catholic University of Murcia (CE111803). As it is a study that compares the size, location and average temperature of the thermograms, the intrinsic influence factors of the subjects were not considered.

Sample size determination

Sample size was calculated by the data proposed by Walter et al.[28]. For an expected intraclass correlation coefficient of 0.9 and a minimum acceptable value of 0.7, and two measurements per image, then the minimum required sample size is 19 participants with alpha=0.05 and beta=0.2.

IRT imaging

All images were recorded following the same standard protocol. The participant was acclimatized in an isolated room at a mean temperature of 22.7 ± 0.5ºC and a relative humidity of 43.4 ± 4% for 20 min without clothing on the lower limbs.

The participant was seated on a hydraulic stretcher with his/her feet on a step-in order to isolate contact with the ground. Then a point was marked bilaterally on the dorsal face of the metatarsophalangeal joint of the second toe and a vertical line on the Achilles tendon that was moved to the heel [29]. These two references were used to position the feet on two parallel lines (25 cm apart) drawn on the step. The knees were positioned at 90 degrees of flexion which allows a right angle of radiation while the skin that covers the patellar tendon a greater part of the patellar tendon [30], than in the standing position [4,9]. In addition, the sitting position improves perfusion and skin circulation by reducing the circulatory collapse of the area caused by gravity while standing, and also the influence on skin temperature [31].

The IRT images were recorded with an OPTRIS PI 450 IRT camera coupled to Optris PI Connect Software (Germany). The IRT camera has a Noise Equivalent Temperature Difference <40 mK with 38º x 29º FOV, a wide range of temperature from −20°C to +100°C, spectrum range of 7.5–13 µm, focal plane array sensor size of 382 × 288 pixels, emissivity set at 0.98 and a measurement uncertainty of ± 2% of the overall temperature reading. The size of the capture frame was 55.4 × 40.63 cm (1.5 mm/px).

The most mid-cranial part of the patella was marked with metallic ink using a 0.7 mm marker so that it could be easily observed on the thermogram (reference points) and placed outside the ROI. The camera was positioned 80 cm from the knees and aligned on the three axes by means of a self-leveling laser with respect to the knee and step reference points [32] (fig.1a,b). In this position a first thermogram (T1) was recorded without manipulation or modification (raw thermogram) (fig1c).

In order to delimit the area of skin corresponding to the patellar tendon, metallic adhesives (fig2a) were fixed on the skin with the following references: (1) midpoint: proximal-distal midpoint of the patella; (2) lateral point: 1 cm lateral to the most lateral point of the patella [33], p.3) medial point: over the most medial point of the patella; and (4) distal point: 5 cm distal to the external femoro-tibial articular interline [30]. Contralateral knee ROI was symmetrically adjusted[34].

A second thermogram (T2) was saved and ROI boxes were adjusted to the metal adhesives (fig.2b, c). The size and position of ROI boxes on the pixel matrix (X and Y axis) of the image provided by the software was calculated. On the second thermogram (T2), the pixel distance between the ROI boxes and the reference points was calculated (fig. 2c). The ROI boxes from the second thermogram (T2) to the first (T1) (fig. 2dwere transferred considering the size of the boxes and their distance (pixels) to the reference points (fig. 1c, 2b). Thermogram 1 (T1) was re-recorded with the transferred of the ROI boxes of thermogram 2 (T2) and the average temperature, position and size data were extracted (fig 2d).

The same first thermogram (T1) was used for both examiners as reference image. The reference points and ROI (T2) were located by the two examiners independently and blinded (interexaminer analysis) at two different times (1week) (intraexaminer analysis). The images were coded, and the order of analysis was random. The flowchart with the analysis phases is shown in figure 3.

The main outcome was mean temperature of ROI (ºC) but the box size (width and height in pixels) and the position on the X and Y axis were also analyzed.

Statistical Analysis

Descriptive statistics were calculated for ROI dimension variables: width, height, position on x-axis and y-axis, and mean temperature of the boxes in both knees corresponding to the different thermographs obtained by both observers. Although the sample size (n=34) allows the assumption of normality, it was checked with the coefficients of asymmetry, kurtosis, the Q-Q normality plots and the Kolmogorov-Smirnov test. Parametric tests were applied to all variables. Descriptive statistics were used (mean, standard deviation, range and quartiles) to summarize the data for each examiner's assessment.

The intraclass correlation coefficient (ICC) was calculated in total agreement with a two-factor alpha model and mixed effects ($ICC_{2,1}$) for each of the variables of interest [35,36]. This coefficient offers values of between 0 and 1, where 0 would be a lack of agreement and 1 would be total agreement. Although the interpretation of these cut-off points is, to a certain extent, arbitrary, very good (ICC>0.8), good (ICC=0.61-0.80), moderate (0.41 to 0.6), low (0.21 to 0.4) and poor (<0.21) reproducibility will be considered [37].

Measurement precision [38,39] was evaluated using standard error of measurement (SEm) [SEm=SD•√(1-ICC)] and its relative value with respect to the average of all measurements and the smallest real difference (SRD). SRD is useful for determining whether a change in the parameter is due to a real change or lies within the limits of error of the measuring method [SRD=1.96•SEm•√2].

The limits of agreement (LOA) were calculated according to the method described by Bland and Altman [40] and the presence of summative or multiplicative biases with Passing-Bablock's linear regression method [41].

For a direct clinical interpretation, the graphical method proposed by Luiz et al. [42], based on the Kaplan-Meier estimate representing the probability of survival as a function of the degree of disagreement, was applied.

Statistical analysis was performed using IBM SPSS Statistics 19.0 (SPSS Inc. IBM Company, 2010) and the *jmv package* (version 0.9) [43] for R (version 3.5.0; 2018) . P-values of <0.05 were considered to indicate statistical significance.

3. RESULTS

A total of 68 thermograms were recorded from 17 participants. Table 1 shows the descriptive data. The width-box variable showed no statistical differences in the intra-examiner ($F_{1,33}$=1.47; p=0.233) or inter-examiner ($F_{1,33}$=0.454; p=0.505) analysis. The intra-examiner differences in height-box were not significant ($F_{1,33}$=0.815; p=0.373) but statistical differences in the means were found for inter-examiner data ($F_{1,33}$=4.40; p=0.044). The X-axis did not show intra-examiner ($F_{1,33}$=0.045; p=0.833) or inter-examiner ($F_{1,33}$≈0; p≈1.0) differences. The means of the Y-axis were similar intra-examiner ($F_{1,33}$=2.65; p=0.133) but not inter-examiner ($F_{1,33}$=42.3; p=0.001). Finally, the skin temperature was similar intra-examiner ($F_{1,33}$=0.488; p=0.490) and inter-examiner ($F_{1,33}$=0.011; p=0.917).

The reliability results are shown in Table 2 and figs. 4-5. Excellent reliability was observed both intra and inter-examiner in all parameters. No bias was found for the parameters analyzed. Only, as expected, in the X-axis parameter was a systematic bias detected, although it did not affect reproducibility.

For skin temperature, the difference between intra-examiner measurements was 0.006ºC (LOA:-0.10 to 0.10) and the differences between examiners were–0.001 ºC (LOA: -0.13 to 0.13).

Further information is provided in the supplementary material (Suppl 1 and Suppl 2).

4. DISCUSSION

In this study, we analyzed the intra- and inter-observer reliability of a method of ROI determination for thermographic analysis that does not modify the original thermogram positioning through the positioning of markers.

The lower limits of the ICC confidence interval showed very close agreement for all variables (>0.85, except for Y-axis inter-examiner) and specifically superiors 0.99 for the intra- and inter-examiner temperature values. This finding may be due to several factors that are decisive for which extreme precautions were taken, such as the choice of easily locatable anatomical references, prior consensus by the observers and the meticulous positioning of the camera and the participant, which left little room for subjectivity on the part of the technicians.

Also facilitating the high degree of agreement in the size variables of the ROI was the use of symmetrical ROI boxes for both knees, which is an important factor when analyzing thermal differences on both sides of the body [34]. It should also be mentioned that the boxes are rectangular, which facilitates their positioning in relation to skin markers, whose inner edge is also rectilinear (fig.2).

From a practical point of view, mean differences were between 0 and 2 pixels, with R^2>82%, with a maximum relative SEm of 2.5% and a maximum relative SRD of 8.7%. All the differences were less than 10 px (figs. 4 and 5.)

As regards the temperature, the reliability was even higher, with mean differences in the order of the tenth degree, both in the intra-examiner and inter-examiner analysis. The total differences were below 0.20ºC.

It is possible that the close agreement found would be lower if, rather than using the same thermogram to determine the size, position and mean temperature of the ROI, different primary thermograms were used for each measurement and examiner. In this study it was decided to use a single primary thermogram to minimize the influence of both extrinsic and intrinsic factors that may affect the uptake of infrared radiation and thus analyze only the factors that influence the creation and positioning of ROI.

Inter-observer reliability has also been investigated in other studies with good results. Selfe et al.[19] observed ICC values of between 0.82 to 0.97 in patients with knee pathologies. Similar results were obtained by Spalding et al. [44] in patients with arthritis of the hands, by Costa et al. in cervical and facial musculature [45] with very good agreement (ICC between 0.852 to 0.998). By contrast, Mustacchi et al.[46] found poor reliability in breast thermograms.

With regard to intra-examiner reliability, the good results of Varju et al. [47] (ICC= 0.899) in hand measurements and Costa et al [21] (ICC= 0.879 to 0.998) in measurements of cervical and facial musculature should be highlighted. On the other hand, Denoble et al. [9] obtained a much lower ICC (0.5 to 0.72) in the knee. Zaproudina et al. [13] suggested that the relatively poor intra and inter-observer results obtained for ICC is because the selection of the ROI is essentially based on a manual procedure and therefore the skill of the technician to select the ROI.

In order to eliminate this subjective factor, different software have been developed to automate the ROI selection process [46]. Fernandez et al. used a software with which they obtained very good results (inter-examiner ICC =0.989 and intra-examiner= 0.997) [26]. However, it should be borne in mind that ROI obtained automatically by software include more general location regions and do not allow measurements to be focused on specific structures because they cannot recognize determining anatomical references.

In recent years the technology and methodology of thermographic recordings have improved and different authors have analyzed the asymmetries in body temperature of different body regions to discriminate preventively areas susceptible to injury in football players [46] and in different pathologies [8,17,22] so that it could become an imaging technique for preventive use and control of the evolution of both injuries and pathologies.

Limitations

The most important limitation of this study is that we located the ROI over only one raw thermogram for each participant. However, as discussed above, the objective was to determine the reproducibility of the reference points to locate ROI, without the influence of other sources of variability.

The main strength of our study is the description of the method of anatomical localization of reference points without modifying skin temperature and the statistical analysis of the reliability from different perspectives and not only based on the ICC, which has several limitations [36,38–40].

5. CONCLUSION

The results obtained confirm that the method used has a very good reliability and reproducibility, while any influence on the detected temperature is negligible.

In this sense, it will be of interest to extend the study in order to ascertain the reliability between different thermal cameras and software, which could increase the strength of the method.

6. REFERENCES

[1] Charkoudian N. Skin blood flow in adult human thermoregulation: how it works, when it does not, and why. Mayo Clin Proc. 2003;78(5):603–612.

[2] Kellog D, Pérgola P. Skin responses to exercise and training. In: Garrett WE, Kirkendall DT, editors. Exerc and Sport Sci. London: Lippincott Williams & Wilkins; 2000. p. 239–250.

[3] Vardasca R. Symmetry of temperature distribution in the upper and lower extremities. Thermol Int. 2008;18:154–155.

[4] Selfe J, Whitaker J. A narrative literature review identifying the minimum clinically important difference for skin temperature asymmetry at the knee. Thermol Int. 2008;18:51–54.

[5] Vardasca R, Ring F, Plassmann P, et al. Thermal symmetry of the upper and lower extremities in healthy subjects. Thermol Int. 2012;22:53–60.

[6] Ring EFJ, Ammer K. Infrared thermal imaging in medicine. Physiol Meas. 2012;33:R33. [7] Mangine RE, Siqueland KA, Noyes FR. The use of thermography for the diagnosis and management of patellar tendinitis. J Orthop Sports Phys Ther. 1987;9:132–140.

[8] Schuhfried O, Vacariu G, Lang T, et al. Thermographic parameters in the diagnosis of secondary Raynaud's phenomenon. Arch Phys Med Rehabil. 2000;81:495–499.

[9] Denoble AE, Hall N, Pieper CF, et al. Patellar skin surface temperature by thermography reflects knee osteoarthritis severity. Clin Med Insights Arthritis Musculoskelet Disord. 2010;3:69–75.

[10] Romanò CL, D'Anchise R, Calamita M, et al. Value of digital telethermography for the diagnosis of septic knee prosthesis: a prospective cohort study. BMC Musculoskelet Disord. 2013;14:7.

[11] Huang C-L, Wu Y-W, Hwang C-L, et al. The application of infrared thermography in evaluation of patients at high risk for lower extremity peripheral arterial disease. J Vasc Surg. 2011;54: 1074–1080.

[12] Costa APC, Maia JM, Brioschi ML, et al. Thermography evaluation in patients with hypothyroidism and fibromyalgia by analyzing the temperatures of the palms of hands. In: Lhotska L, Sukupova L, Lacković I, et al. editors. World congr med phys biomed eng 2018. [Internet]. Singapore: Springer Singapore; 2019 [citado 2019 de octubre de 10]. p. 15–19. Available from: http://link.springer.com/10.1007/978-981-10-9035-6_3

[13] Zaproudina N, Varmavuo V, Airaksinen O, et al. Reproducibility of infrared thermography measurements in healthy individuals. Physiol Meas. 2008;29:515–524.

[14] Costello JT, Stewart IB. Use of thermal imaging in sports medicine research: A short report. Int Sportmed J. 2013;14:94–98.

[15] IACT. Standards and protocols in clinical thermography imaging. Thermology Guidelines Electronic Version Internet; 2002 [cited 2019 Apr 15]. Available from: http://www.iact-org.org/professionals/thermog-guidelines.html

[16] Ammer K. The Glamorgan protocol for recording and evaluation of thermal images of the human body. Thermol Int. 2008;18:125–129.

[17] Ludwig N, Formenti D, Gargano M, et al. Skin temperature evaluation by infrared thermography: comparison of image analysis methods. Infrared Phys Technol. 2014;62:1–6.

[18] Choi E, Lee P-B, Nahm FS. Interexaminer reliability of infrared thermography for the diagnosis of complex regional pain syndrome. Skin Res Technol Off J Int Soc Bioeng Skin ISBS Int Soc Digit Imaging Skin ISDIS Int Soc Skin Imaging ISSI. 2013;19:189–193.

[19] Selfe J, Hardaker N, Thewlis D, et al. An accurate and reliable method of thermal data analysis in thermal imaging of the anterior knee for use in cryotherapy research. Arch Phys Med Rehabil. 2006;87:1630–1635.

[20] Selfe J, Sutton C, Hardaker NJ, et al. Anterior knee pain and cold knees: a possible association in women. Knee. 2010;17:319–323.

[21] Costa ACS, Dibai Filho AV, Packer AC, et al. Intra and inter-rater reliability of infrared image analysis of masticatory and upper trapezius muscles in women with and without temporomandibular disorder. Braz J Phys Ther. 2013;17:24–31.

[22] Ferreira JJA, Mendonça LCS, Nunes LAO, et al. Exercise-associated thermographic changes in young and elderly subjects. Ann Biomed Eng. 2008;36:1420–1427.

[23] Varjú G, Pieper CF, Renner JB, et al. Assessment of hand osteoarthritis: correlation between thermographic and radiographic methods. Rheumatol Oxf Engl. 2004;43:915–919.

[24] George J, Bensafi A, Schmitt AM, et al. Validation of a non-contact technique for local skin temperature measurements. Skin Res Technol Off J Int Soc Bioeng Skin ISBS Int Soc Digit Imaging Skin ISDIS Int Soc Skin Imaging ISSI. 2008;14:381–384.

[25] Herry CL, Frize M. Quantitative assessment of pain-related thermal dysfunction through clinical digital infrared thermal imaging. Biomed Eng OnLine. 2004;3:19.

[26] Fernández-Cuevas I, Marins JC, Carmona PG, et al. Reliability and reproducibility of skin temperature of overweight subjects by an infrared thermography software designed for human beings. Thermol Int. 2012;22:130–137.

[27] Moreira DG, Costello JT, Brito CJ, et al. Thermographic imaging in sports and exercise medicine: A Delphi study and consensus statement on the measurement of human skin temperature. J Therm Biol. 2017;69:155–162.

[28] Walter SD, Eliasziw M, Donner A. Sample size and optimal designs for reliability studies. Stat Med. 1998;17:101–110.

[29] Ribeiro AP, Trombini-Souza F, Tessutti V, et al. Rearfoot alignment and medial longitudinal arch configurations of runners with symptoms and histories of plantar fasciitis. Clinics. 2011;66:1027–1033.

[30] Zooker C, Pandarinath R, Kraeutler MJ, et al. Clinical measurement of patellar tendon: accuracy and relationship to surgical tendon dimensions. Am J Orthop Belle Mead NJ. 2013;42:317–320.

[31] Ratovoson D, Jourdan F, Huon V. Influence of gravity on the skin thermal behavior: experimental study using dynamic infrared thermography. Skin Res Technol Off J Int Soc Bioeng Skin ISBS Int Soc Digit Imaging Skin ISDIS Int Soc Skin Imaging ISSI. 2013;19:e397–408.

[32] Tká M, Hudák R, Foffová P, et al. An importance of camera subject distance and angle in musculoskeletal applications of medical thermography. Acta Electrotech Inf. 2010;10:57–60.

[33] Veeramani R. Gender differences in the mediolateral placement of the patella and tibial tuberosity: a geometric analysis. Anat Int J Exp Clin Anat. 2010;4:45–50.

[34] Nahm FS. Infrared thermography in pain medicine. Korean J Pain. 2013;26:219–222.

[35] McGraw KO, Wong SP. Forming inferences about some intraclass correlation coefficients. Psychol Methods. 1996;1:30–46.

[36] Weir JP. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. J Strength Cond Res. 2005;19:231–240.

[37] Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychol Bull. 1979;86:420–428.

[38] Lexell JE, Downham DY. How to assess the reliability of measurements in rehabilitation. Am J Phys Med Rehabil. 2005;84:719–723.

[39] Atkinson G, Nevill AM. Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. Sports Med. 1998;26:217–238.

[40] Atkinson G, Nevill A. Measures of reliability in sports medicine and science: correspondence. Sports Med. 2000;30:375–381.

[41] Bablok W, Passing H. Application of statistical procedures in analytical instrument testing. J Autom Chem. 1985;7:74–79.

[42] Luiz RR, Costa AJL, Kale PL, et al. Assessment of agreement of a quantitative variable: a new graphical approach. J Clin Epidemiol. 2003;56:963–967.

[43] Love J, Dropmann D, Selker R. Jamovi project; 2018. [Internet]. [cited 2019 Dec 08] Available from: https://www.jamovi.org

[44] Spalding SJ, Kwoh CK, Boudreau R, et al. Three-dimensional and thermal surface imaging produces reliable measures of joint shape and temperature: a potential tool for quantifying arthritis. Arthritis Res Ther. 2008;10:R10.

[45] Mustacchi G, Milani S, Ciatto S, et al. Observer variation in mammary thermography: results of a teaching file test carried out in four different centers. Tumori. 1990;76:29–31.

[46] Plassmann PMP. CTHERM for standardised thermography. 9thEuropean Congress of Medical Thermology, Krakow (Poland); 2003.

7. TABLES

Table 1. Descriptive values for intra and inter-observer analysis.

Parameter	Mean (SD)	C.I. 95%	Minimum	Median (IQR)	Maximum
Width box A	34.1 (3.28)	33 to 35	26	34 (32 to 36)	40
Width box A2	33.9 (3.11)	33 to 35	26	35 (32 to 36)	40
Width box B	33.9 (3.64)	33 to 35	25	35 (31 to 36.3)	40
Height box A	56.6 (4.61)	55 to 58	47	56 (53.8 to 61)	63
Height box A2	56.3 (4.85)	55 to 50	47	57 (54 to 61)	61
Height box B	57.2 (4.45)	56 to 59	49	58 (54 to 61)	64
X-axis location A	192.5 (82.01)	164 to 221	106	193 (112 to 272)	279
X-axis location A2	192.4 (81.47)	164 to 221	107	194 (112 to 273)	278
X-axis location B	192.5 (80.99)	164 to 221	109	192 (112 to 272)	278
Y-axis location A	122.3 (7.99)	120 to 125	102	123 (119 to 124)	152
Y-axis location A2	121.7 (8.27)	119 to 125	98	122 (119 to 125)	147
Y-axis location B	120.3 (7.98)	118 to 123	98	121 (118 to 123)	147
Temperature A	32.9 (0.85)	32.6 to 33.2	31.2	32.9 (32.4 to 33.6)	34.6
Temperature A2	32.9 (0.85)	32.6 to 33.2	31.2	32.9 (32.4 to 33.6)	34.6
Temperature B	32.9 (0.84)	32.6 to 33.2	31.2	32.9 (32.5 to 33.6)	34.6

Box parameters units are pixels. Temperature in degrees Celsius. A: first evaluation for examiner A. A2: second evaluation for examiner A. B: evaluation for examiner B..

Table 2. Reliability for both intra and inter-observer analysis.

Parameter	ICC (95% CI)	Agreement	Mean difference (95% LOA)	% Change	SEm	SEm%	SRD (95% CI)	SRD%
Width box intra-observer	0.937 (0.878 to 0.968)	Excellent	-0.235 (-2.45 to 1.98)	-0.6	0.80	2.3%	2.21 (-1.5 to 2.92)	6.5%
Width box inter-observer	0.904 (0.818 to 0.951)	Excellent	-0.176 (-3.17 to 2.82)	-0.5	1.07	3.1%	2.96 (-1.72 to 4.19)	8.7%
Height box intra-observer	0.92 (0.846 to 0.959)	Excellent	-0.294 (-4.02 to 3.43)	-0.5	1.33	2.4%	3.68 (-2.80 to 4.56)	6.5%
Height box inter-observer	0.92 (0.846 to 0.959)	Excellent	0.588 (-2.62 to 3.79)	1.1	1.41	2.5%	3.91 (-2.50 to 5.32)	6.9%
X-axis location intra-observer	1 (1 to 1)	Excellent	-0.059 (-3.22 to 3.10)	0.1	0.00	0.00%	0 (-)	0.0%
X-axis location inter-observer	1 (0.999 to 1)	Excellent	0.00 (-4.97 to 4.97)	0.3	0.00	0.00%	0 (-)	0.0%
Y-axis location intra-observer	0.968 (0.937 to 0.984)	Excellent	-0.559 (-4.48 to 3.37)	-0.5	1.44	1.2%	4 (-2.56 to 5.45)	3.3%
Y-axis location inter-observer	0.946 (0.544 to 0.984)	Excellent	-2.0 (-5.51 to 1.51)	-1.6	1.43	1.2%	3.96 (-1.78 to 6.14)	3.3%
Temperature intra-observer	0.998 (0.997 to 0.999)	Excellent	0.0059 (-0.090 to 0.10)	0.0	-	-	-	-
Temperature inter-observer	0.997 (0.994 to 0.999)	Excellent	0.0012 (-0.127 to 0.129)	0.0	-	-	-	-

Box parameters units are pixels. Temperature in degrees Celsius. ICC: intraclass correlation coefficient. LOA: limits of agreement. %Change: mean difference respect total average of measures. SEm: standard error of measurement. SRD: smallest real difference. SEM and CRD for temperature are not shown because it is an interval variable.

8. FIGURES

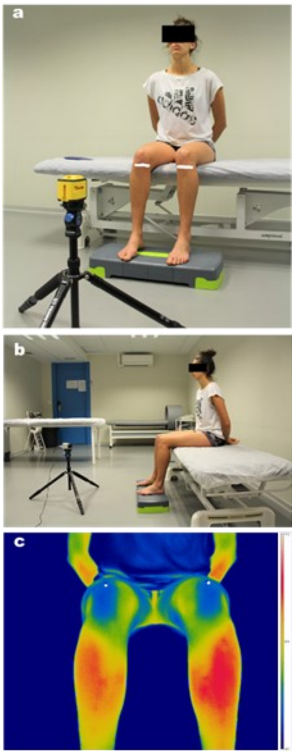
Fig. 1. Participant and camera positioning (a,b) and the first thermogram with reference points(c)

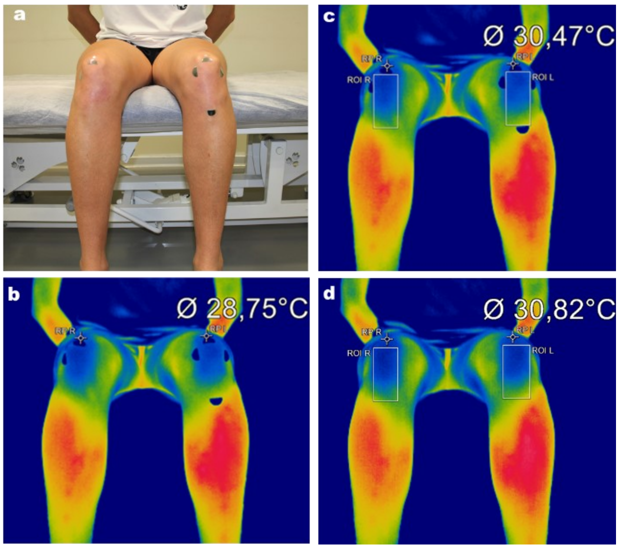
Fig. 2. Metallic markers (a), thermogram with reference points (b), ROI boxes on the second thermogram (c) and superimposed over first thermogram (d)

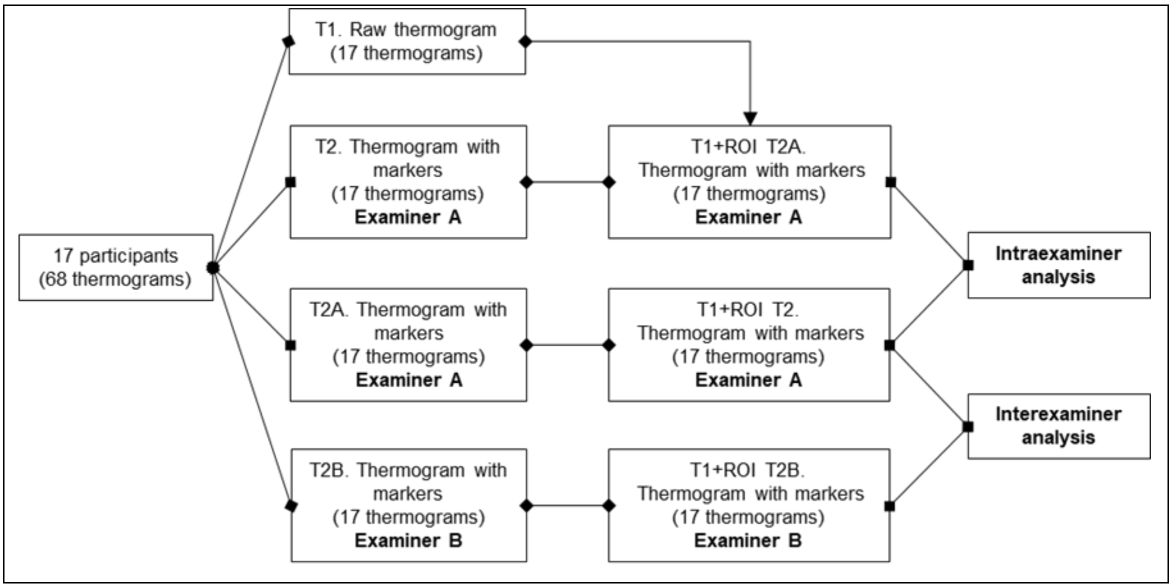
Fig.3. Flowchart of analysis method.

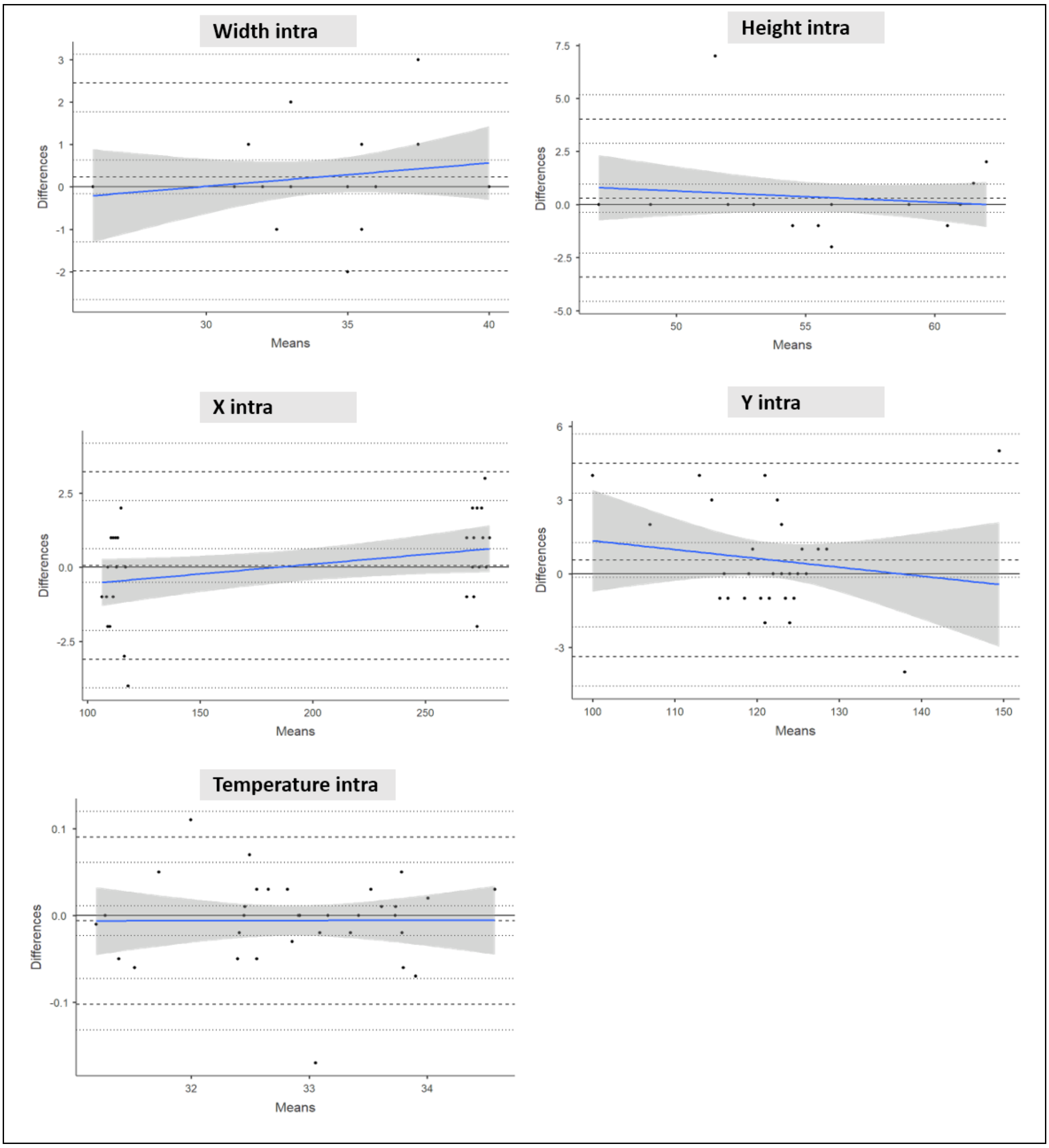
Fig. 4. Intra-examiner reliability Bland–Altman's plots. Each pair of measurements is represented by a point determined by the mean value (x-axis) and by the difference between them (y-axis). Dotted lines represent the limits of agreement with the confidence intervals. Dark zone and blue line represent the proportional bias line with confidence intervals. The Width and Height are in millimetres (mm), X and Y are coordinates in pixels and Temperature is in Celsius grades (ºC).

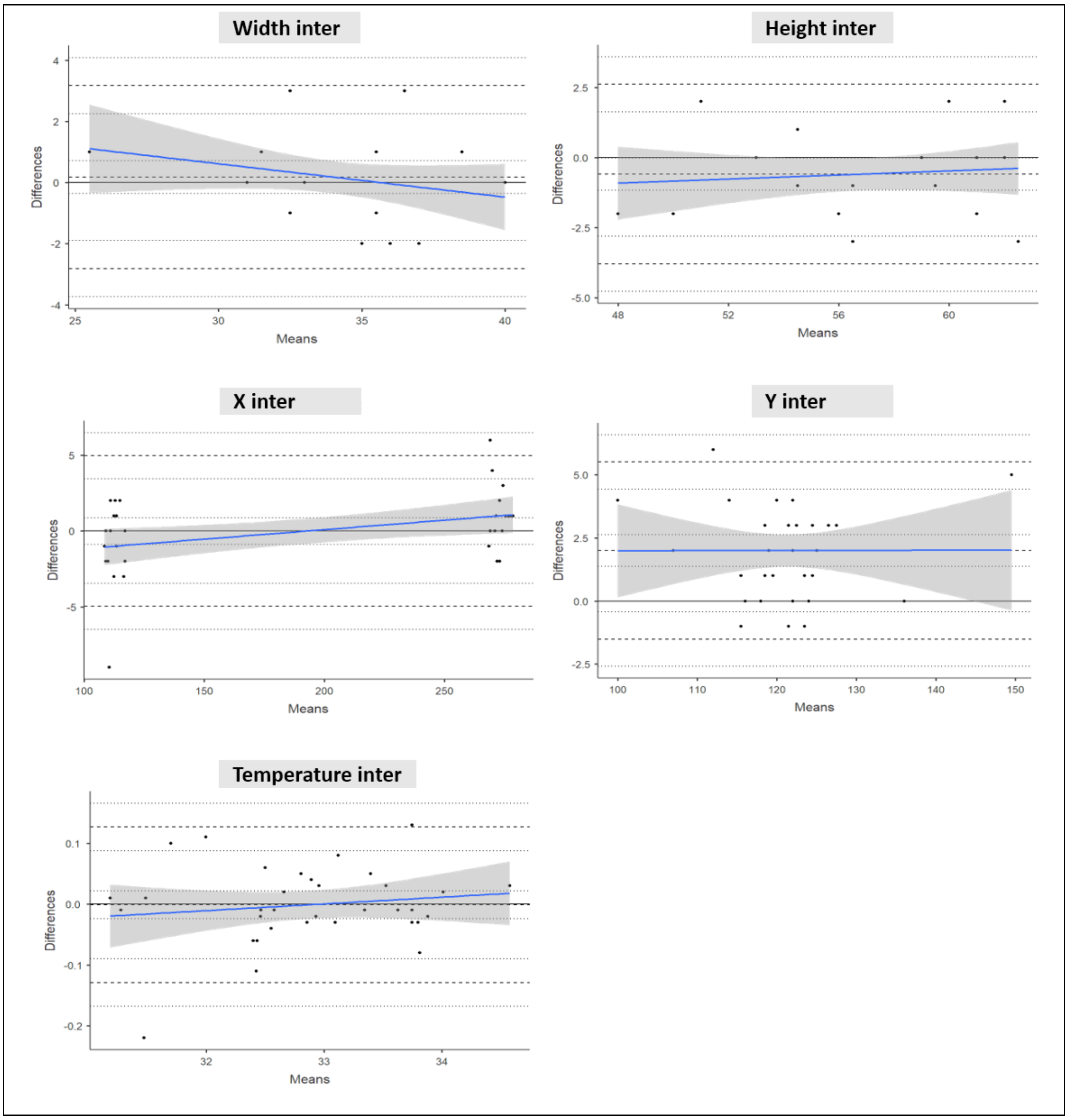
Fig. 5. Inter-examiner reliability Bland–Altman's plots. Each pair of measurements is represented by a point determined by the mean value (x-axis) and by the difference between them (y-axis). Dotted lines represent the limits of agreement with the confidence intervals. Dark zone and blue line represent the proportional bias line with confidence intervals. The Width and Height are in millimetres (mm), X and Y are coordinates in pixels and Temperature is in Celsius grades (ºC


a
b
c


a
c
Ø 30,47°C
ROI R
ROI L
b
Ø 28,75°C
d
Ø 30,82°C
ROI R
ROI L


17 participants
(68 thermograms)
T1. Raw thermogram
(17 thermograms)
T2. Thermogram with markers
(17 thermograms)
Examiner A
T2A. Thermogram with markers
(17 thermograms)
Examiner A
T2B. Thermogram with markers
(17 thermograms)
Examiner B
T1+ROI T2A.
Thermogram with markers
(17 thermograms)
Examiner A
T1+ROI T2.
Thermogram with markers
(17 thermograms)
Examiner A
T1+ROI T2B.
Thermogram with markers
(17 thermograms)
Examiner B
Intraexaminer analysis
Interexaminer analysis


Width intra
Differences
Means
Height intra
Differences
Means
X intra
Differences
Means
Y intra
Differences
Means
Temperature intra
Differences
Means



Width inter
Differences
Means
Height inter
Differences
Means
X inter
Differences
Means
Y inter
Differences
Means
Temperature inter
Differences
Means