Assessment of alternative metrics in the application of infrared thermography to detect muscle damage in sports

IOP Publishing Physiol. Meas. 45 (2024) 095014 https://doi.org/10.1088/1361-6579/ad7ad3

Physiological Measurement

IPEM Institute of Physics and Engineering in Medicine

CrossMark

OPEN ACCESS

RECEIVED
16 July 2024

REVISED
25 August 2024

ACCEPTED FOR PUBLICATION
13 September 2024

PUBLISHED
24 September 2024

PAPER

Assessment of alternative metrics in the application of infrared thermography to detect muscle damage in sports

Lukas Verderber[1], Willian da Silva[2], Inmaculada Aparicio-Aparicio[3], Andresa M C Germano[1], Felipe P Carpes[4] and Jose Ignacio Priego-Quesada[3,5,*]

[1] Motor Control, Cognition and Neurophysiology, Institute of Human Movement Science & Health, Chemnitz University of Technology, Chemnitz, Germany
[2] Escuela de Kinesiología, Facultad de Ciencias, Pontificia Universidad Católica de Valparaíso, Valparaíso, Chile
[3] Research Group in Sports Biomechanics (GIBD), Department of Physical Education and Sports, University of Valencia, Valencia, Spain
[4] Applied Neuromechanics Research Group, Federal University of Pampa, Uruguaiana, Brazil
[5] Research Group in Medical Physics (GIFIME), Department of Physiology, University of Valencia, Valencia, Spain
[*] Author to whom any correspondence should be addressed.

E-mail: j.ignacio.priego@uv.es

Keywords: delayed muscle onset soreness, pain, skin temperature, thermal image

Abstract

Objective. The association between muscle damage and skin temperature is controversial. We hypothesize that including metrics that are more sensitive to individual responses by considering variability and regions representative of higher temperature could influence skin temperature outcomes. Here, the objective of the study was to determine whether using alternative metrics (TMAX, entropy, and pixelgraphy) leads to different results than mean, maximum, minimum, and standard deviation (SD) skin temperature when addressing muscle damage using infrared thermography. *Approach.* Thermal images from four previous investigations measuring skin temperature before and after muscle damage in the anterior thigh and the posterior lower leg were used. The TMAX, entropy, and pixelgraphy (percentage of pixels above 33 °C) metrics were applied. *Main results.* On 48 h after running a marathon or half-marathon, no differences were found in skin temperature when applying any metric. Mean, minimum, maximum, TMAX, and pixelgraphy were lower 48 h after than at basal condition following quadriceps muscle damage ($p < 0.05$). Maximum skin temperature and pixelgraphy were lower 48 h after than the basal condition following muscle damage to the triceps sural ($p < 0.05$). Overall, TMAX strongly correlated with mean ($r = 0.85$) and maximum temperatures ($r = 0.99$) and moderately with minimum ($r = 0.66$) and pixelgraphy parameter ($r = 0.64$). Entropy strongly correlates with SD ($r = 0.94$) and inversely moderately with minimum temperature ($r = -0.53$). The pixelgraphy moderately correlated with mean ($r = 0.68$), maximum ($r = 0.62$), minimum ($r = 0.58$), and TMAX ($r = 0.64$). *Significance.* Using alternative metrics does not change skin temperature outcomes following muscle damage of lower extremity muscle groups.

1. Introduction

Infrared thermography (IRT) is an imaging technique to detect infrared radiation emitted by a body surface (Hillen *et al* 2020). Being widely applied in medicine and industry (Jasti *et al* 2019), IRT has gained popularity in sports sciences due to its non-invasive, portable, cost-effective, and easy applicability (Priego Quesada 2017). Among the desired applications is the relationship between changes in skin temperature detected by IRT and physiological responses to exercise-induced muscle damage. However, results from investigations addressing this relationship are controversial (Stewart *et al* 2020, Da Silva *et al* 2021, Ferreira-Júnior *et al* 2021).

A recurrent question about these controversial outcomes is the methodological approaches conducted for data analysis. Standardized protocols for image capture are important, but technical aspects like selecting the metrics of the region of interest (ROI) could affect the results (Moreira *et al* 2017). A common metric for determining skin temperature is the mean temperature, which considers all pixels from a defined ROI (Formenti *et al* 2017). From the same ROI data, maximum and minimum temperatures can be obtained. However, both maximum and minimum temperatures can be heavily influenced by data collection settings since they refer to data from a single-pixel (Machado *et al* 2021). While mean skin temperature is more robust and reliable, maximum and minimum measures can relate more to physiological alterations resulting from changes in skin blood flow and inflammation (De Andrade Fernandes *et al* 2017, Formenti *et al* 2018, Da Silva *et al* 2022).

Including metrics more sensitive to individual responses by considering variability and clusters of pixels with higher temperatures in the ROI may influence skin temperature outcomes. In this regard, TMAX considers an area of 5×5 pixels around the five warmest pixels within an ROI to assess skin temperature changes after physical exertion (Formenti *et al* 2017). Entropy measures the average number of bits of an event with respect to the probability distribution for the random variable, and in IRT analysis represents the temperature distribution within a ROI that was correlated in a recent study with exercise workload, exercise duration, and pulmonary ventilation during exercise (Bogomilsky *et al* 2022). Additionally, the regions of high temperature can be identified using pixelgraphy, as this method examines the percentage of pixels in different temperature ranges assuming that higher percentages of pixels above 33 °C may be more related to muscle damage (Rodrigues Júnior *et al* 2021, Alburquerque Santana *et al* 2022). Although all these metrics seem relevant and have a physiological meaning, it remains unclear whether they provide different results compared to the central tendency and dispersion variables often employed in studies about muscle damage and skin temperature using IRT.

The objective of this study was to determine whether the use of TMAX, entropy, and pixelgraphy leads to different results compared to the conventional metrics when assessing skin temperature responses using IRT following muscle damage. Considering the results of previous studies (Alburquerque Santana *et al* 2022, Bogomilsky *et al* 2022), we hypothesized that results could differ using these metrics that consider regional variations of temperature and variability compared with descriptive statistics commonly employed. The results of this work could be a guide for researchers, as well as for thermography analysis software, helping them determine the most appropriate metrics for their studies.

2. Material and methods

2.1. Experimental design

IRT data from previously published research were reanalysed using alternative metrics. The results were compared with the conventional metrics. We also investigated the relationship between the conventional and alternative metrics.

2.2. Database

The database used includes data from four studies (Da Silva *et al* 2018, 2021, Pérez-Guarner *et al* 2019, Priego-Quesada *et al* 2020). These four studies were conducted by two collaborating laboratories, with some researchers who are authors of this work and therefore with access to the data. Two ROIs were considered in the anterior thigh and the posterior lower leg, according to the muscle damage protocols delivered to the quadriceps and triceps surae muscle groups (see table 1 for more details). In the case of the Half-marathon, 5 participants were removed from the analysis as we were unable to process the raw data for these participants for problems in their image files. In the four studies, the authors confirmed ethical approval, compliance with the Declaration of Helsinki and the signing of a consent term for voluntary participation.

2.3. Protocols

The four studies induced muscle damage by different exercise protocols. Two studies (half-marathon and marathon) had participants completing an actual outdoor competition to assess the pre and different days post-competition differences (Pérez-Guarner *et al* 2019, Priego-Quesada *et al* 2020). In the two other studies (quadriceps and triceps surae) muscle damage was induced by a single exercise with measures before and 48 h after the protocol (Da Silva *et al* 2018, 2021). More details about the protocol can be found in the respective publications.

Table 1. Characteristics of the database used for the analysis.

References	Mentioned in the present study as	N participants	Profile of participants	Characteristics of participants	ROIs assessed in the present study	Measurement times assessed in the present study
(Pérez-Guarner *et al* 2019)	Half-marathon	11	Recreational runners	4 females and 7 males 40 ± 7 years old BMI 21.7 ± 1.2 kg m^{-2}	Anterior thigh Posterior leg	Before post24 h post48 h
(Priego-Quesada *et al* 2020)	Marathon	16	Recreational runners	3 females and 13 males 40 ± 7 years old BMI 23.2 ± 3.1 kg m^{-2}	Anterior thigh Posterior leg	Before post24 h post48 h
(Da Silva *et al* 2021)	Quadriceps	22	Untrained	12 females and 10 males 26 ± 5 years old BMI 23.5 ± 2.8 kg m^{-2}	Anterior thigh	Before post48 h
(Da Silva *et al* 2018)	Triceps surae	20	Untrained males	24 ± 5 years old BMI 24.8 ± 2.0 kg m^{-2}	Posterior leg	Before post48 h

Note: Body Mass Index (BMI).

The four studies followed the thermographic imaging in sports and exercise medicine checklist (Moreira *et al* 2017) having therefore similar protocols of image acquisition. Participants were instructed about avoiding smoking, drinking alcohol, caffeine, or other stimulant beverages, taking large meals, using ointments and cosmetics, sunbathing, joining physiotherapy treatments, and refraining from high-intensity physical activity before the assessments. The studies used the same camera model (E-60, sensor array size of 320 × 240; FLIR Systems, Wilsonville, USA) and performed their measurements in a moderate environment (23 °C–24 °C room temperature and 20%–60% relative humidity). The cameras were turned on at least 10 min before the assessments, participants performed a thermal adaptation to the room lasting 10 min, and IRT images were registered at 1–1.5 m far from the ROIs with the camera lens placed perpendicularly to the ROI.

2.4. Image analysis

The ROI were defined for our investigation (see figure 1) using a thermography software (Thermacam Researcher Pro 2.10 software, FLIR Systems, Wilsonville, USA) and an emissivity of 0.98 (Steketee 1973). All the analysis was performed by the same researcher. From these ROIs, a CSV file was created containing all the temperature values corresponding to all pixels of the ROI. Using a custom-made script in RStudio (version 2023.12.1), each matrix of pixel values was assessed obtaining the mean, the maximum, the minimum, the standard deviation (SD), the TMAX (Ludwig *et al* 2014, Formenti *et al* 2017), the entropy (Bogomilsky *et al* 2022), and for pixelgraphy analysis, we assessed the percentage of pixels in the ROI with higher values than 33 °C (Alburquerque Santana *et al* 2022).

TMAX consisted of considering the 5 hottest values with a minimum distance of five pixels from each other and averaging the 5 matrices of 5 × 5 values around these hottest values (Formenti *et al* 2017). Although we tried to follow the methodology outlined by Ludwig *et al* (2014) to determine TMAX, neither their paper nor other publications from the same authors provided details about the procedure in case one of the 5 hottest values is close to a ROI edge. In such cases, it is impossible to have a 5 × 5 matrix around it due to the presence of non-numerical data (NAs). Therefore, we only considered a 5 × 5 matrix when there were at least 10 valid values around. If the required 10 valid values were not found, another hottest region was identified, always satisfying the separation criteria from the other determined hottest values (Ludwig *et al* 2014).

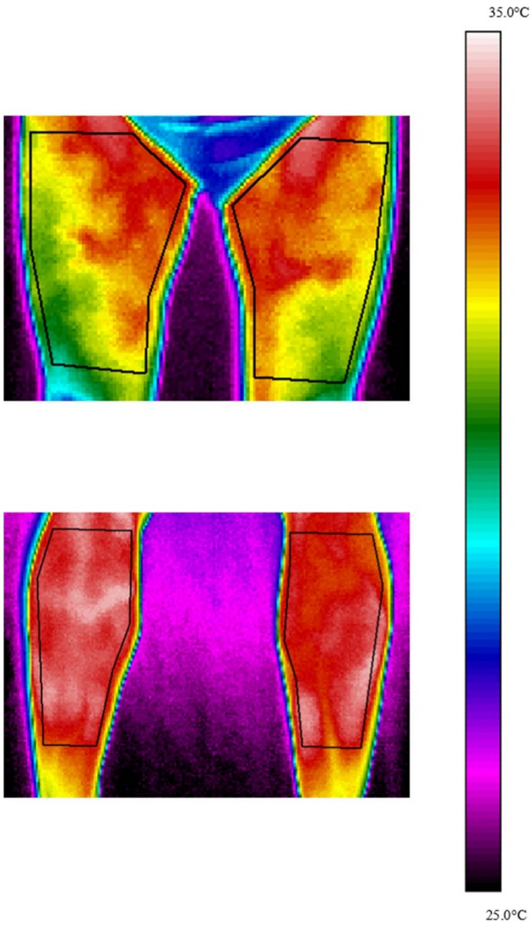

Figure 1. Regions of interest defined for the anterior thigh and posterior lower leg.

Entropy was computed according to the equation (1) (Bogomilsky *et al* 2022):

$$H(X) = -\sum_{i=1}^{n} p(x_i)\ \log_2 p(x_i) \tag{1}$$

where $p(x_1)$ is the probability of occurrence of each unique value.

2.5. Statistical analysis

Statistical analyses were performed using RStudio (version 2023.12.1) and considering an alpha set at 5%. For each metric, ROI, and study, the normality of data distribution was checked using the Shapiro–Wilk test before either repeated-measures ANOVA or Kruskal–Wallis tests were applied. Each of these tests was performed using 2 factors: measurement time (pre and post-muscle damage induction) and lower limb (right or left). When main effects or interactions were found, *post hoc* analyses were conducted with Bonferroni for ANOVAs, while Wilcoxon tests with Bonferroni correction were used for significant Kruskal–Wallis tests. Cohen effect sizes (ES*d*) were calculated for pairs of significant differences of parametrical data and classified as small (0.2–0.5), moderate (0.5–0.8), or large (>0.8) (Cohen 1988). For non-parametrical data, Rosenthal's *r* was calculated (ES*r*) and classified as small (0.1–0.3), moderate (0.3–0.5), or large (>0.5) (Rosenthal and Rubin 2003).

To determine whether the metrics explain different phenomena or are closely related, Pearson's correlation coefficient and principal component analysis (PCA) were performed with the data from the four studies together. Before applying these analyses, metrics were centered (subtracting the mean) and scaled (dividing by the SD). Significant correlations were classified as weak ($0.2 < |r| < 0.5$), moderate ($0.5 \leqslant |r| < 0.8$), or strong ($|r| \geqslant 0.8$) (O'Rourke *et al* 2005). For PCA, the *eigendecomposition* method was used on the correlation matrix.

Table 2. Results of the ANOVAs and Kruskal–Wallis tests for the measurement time factor for each study, region of interest and metric assessed.

		P-value of measurement time factor						
Study	Region of interest	Mean	Maximum	Minimum	SD	TMAX	Entropy	Pixelgraphy
Half-marathon	Anterior thigh	0.02 [A]	0.15	0.08 [A]	0.17	0.13	0.23 [A]	0.45
Half-marathon	Posterior leg	<0.01	<0.01	0.05	0.38 [A]	<0.01	0.35 [A]	<0.01
Marathon	Anterior thigh	0.68	0.85 [A]	0.63	0.55	0.80 [A]	0.46 [A]	0.75
Marathon	Posterior leg	0.65	0.86	0.18 [A]	0.22	0.97	0.32 [A]	0.68
Quadriceps	Anterior thigh	<0.001 [A]	0.04	<0.001 [A]	0.21	0.01	0.20 [A]	0.01
Triceps surae	Posterior leg	0.10 [A]	0.04	0.31 [A]	0.83	0.11 [A]	0.88 [A]	0.01

Note: [A]: *p*-value obtained with the ANOVA test, and if this letter is not present, it was obtained with the Kruskal–Wallis test.

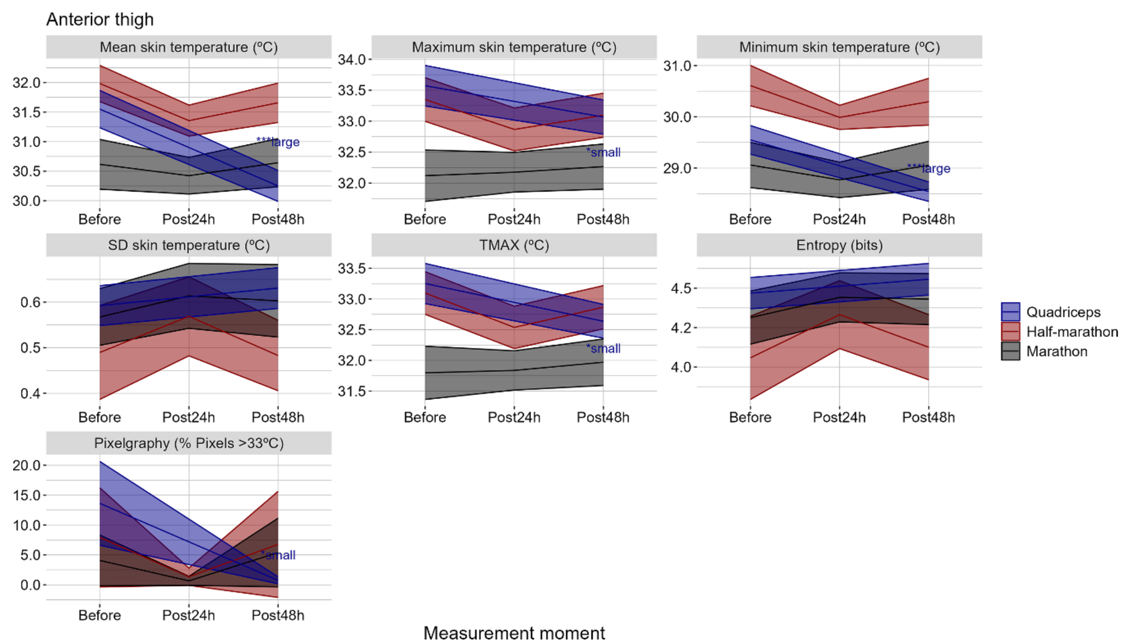

Figure 2. Mean (line) and 95% confidence interval (shadow) of the different parameters and measurement times of anterior thigh region of interest in three different studies. Significant differences between before and the other measurement times were indicated using symbols (* $p < 0.05$, *** $p < 0.001$) and the magnitude of the effect size in the color of the study when the differences were observed. Note: quadriceps study only has data in before and post48 h measurement time.

3. Results

Main effect of lower limb were not found ($p > 0.16$), except for the SD ($p = 0.02$) and entropy ($p = 0.04$) for the quadriceps study. No interactions were found between the lower limb and the measurement time factors in any of the studies ($p > 0.28$). Therefore, this lower limb factor was not considered in the following results.

Table 2 shows the ANOVAs and Kruskal–Wallis main effects for the measurement time factor. Figures 2 and 3 depicts the post hoc comparisons for the anterior thigh and posterior lower leg, respectively. Although the mean metric was significant for the ANOVA in the anterior thigh region in the half-marathon study, the *post hoc* did not reveal differences ($p > 0.82$). Similar results were observed for the posterior lower leg ($p > 0.97$ in the post hocs for mean, maximum, TMAX, and pixelgraphy). For the quadriceps study, most of the metrics were lower at Post48 than Pre: mean skin temperature ($p < 0.001$ and ESd = 1.4, 95%CI[0.9, 1.7 °C]), maximum ($p = 0.04$ and ESr = 0.2, 95%CI[0.0, 0.9 °C]), minimum ($p < 0.001$ and ESd = 1.3, 95%CI[0.7, 1.3 °C]), TMAX ($p = 0.01$ and ESr = 0.3, 95%CI[0.1, 1.0 °C]), and pixelgraphy ($p = 0.01$ and ESr = 0.3, 95%CI[0.0, 0.4%]). For the triceps surae study, maximum skin temperature ($p = 0.04$ and ESr = 0.2, 95%CI[0.7, 1.2 °C]) and pixelgraphy ($p < 0.01$ and ESr = 0.3, 95%CI[0.0, 7.6%]) were lower at Post48 than Pre.

Figure 4 shows the correlation matrix. Considering only the moderate or strong correlations, TMAX strongly correlated with mean and maximum, and moderately correlated with minimum and pixelgraphy. Entropy strongly correlated with SD, and inversely moderately correlated with minimum. Finally, pixelgraphy moderately correlated with mean, maximum, minimum and TMAX. PCA showed that only one component explained 91.9% of the variance, and a second component explained 6.1% of the variance (figure 5).

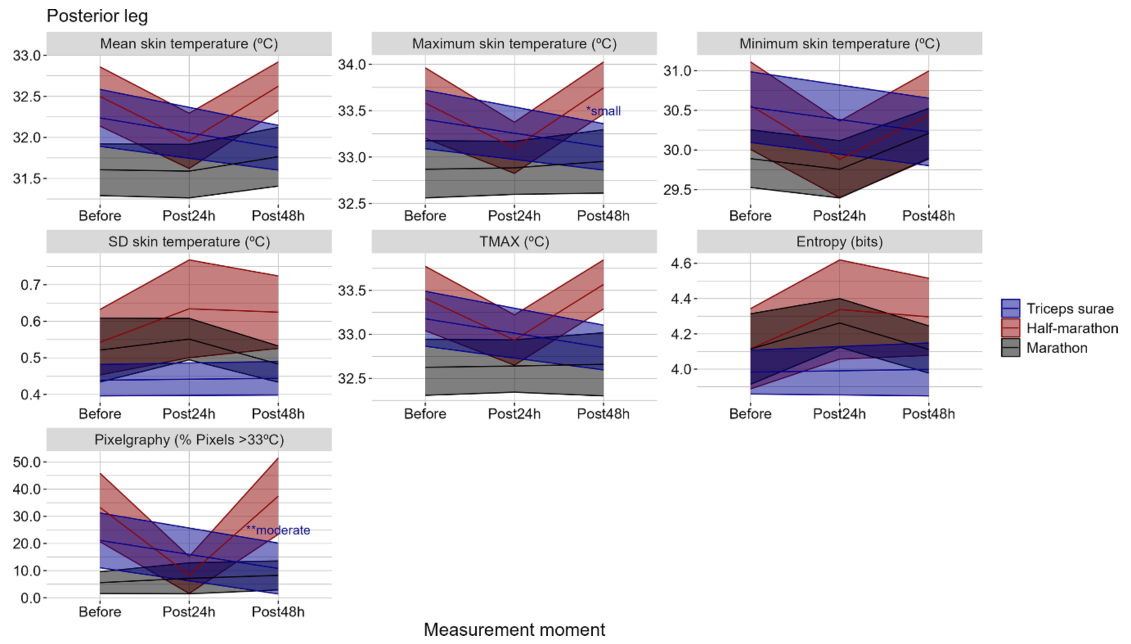

Figure 3. Mean (line) and 95% confidence interval (shadow) of the different parameters and measurement times of posterior leg region of interest in three different studies. Significant differences between before and the other measurement times were indicated using symbols (* $p < 0.05$, ** $p < 0.01$) and the magnitude of the effect size in the color of the study when the differences were observed. Note: triceps surae study only has data in before and post48 h measurement time.

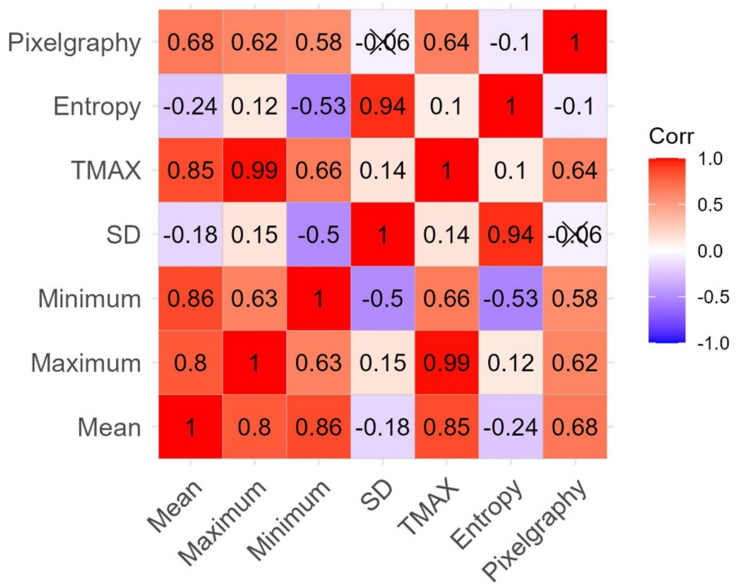

Figure 4. Correlation matrix of the thermal parameters assessed (Pearson's correlation). Color level is according to the correlation coefficient (Corr in the legend), which is also showed inside of each square. If the value is crossed out, it means that it is not significant.

4. Discussion

This study aimed to investigate whether alternative metrics of IRT (TMAX, entropy, and pixelgraphy) to quantify skin temperature following exercise-induced muscle damage led to different outcomes compared to the conventional descriptive and amplitude metrics commonly reported in previous studies. Comprehensive image analyses were conducted using IRT images from four studies (Da Silva *et al* 2018, 2021, Pérez-Guarner *et al* 2019, Priego-Quesada *et al* 2020). The selection of different metrics did not yield varying overall outcomes, and correlation and PCA analyses showed that some of these metrics were redundant and, therefore, should not be used together. Although one previous investigation compared mean, maximum and TMAX results (Formenti *et al* 2018), to the authors' knowledge, this can be considered the first study that assessed the utility of using TMAX, entropy, and pixelgraphy with conventional metrics.

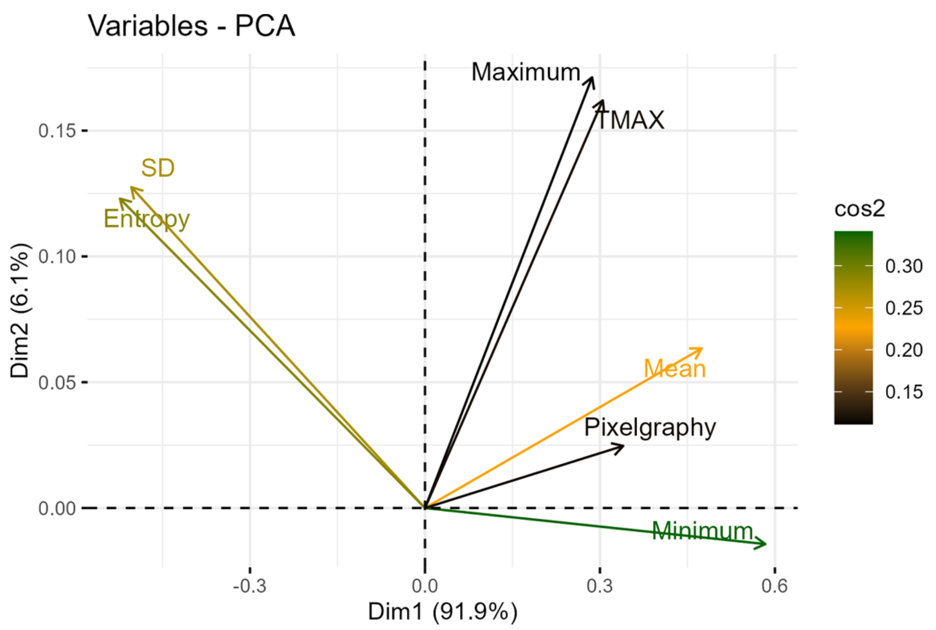

Figure 5. Biplot of the metrics concerning the principal components obtained.

Considering the alternative metrics assessed, although there was a difference in the measurement times evaluated in the quadriceps and triceps surae studies for TMAX and the pixelgraphy (percentage of pixels above 33 °C), these results did not provide any additional information to that obtained with the maximum temperature. These similar results using alternative or traditional metrics and also the strong correlations observed between the different metrics suggest that TMAX, entropy, and the pixelgraphy metric are interconnected with various temperature metrics, highlighting the complex interplay between different measures of skin temperature in the context of muscle damage. Implementing these metrics may allow a broader description of skin temperature responses following exercise aiming at inducing muscle damage. The re-analysis of thermal images recorded after a half-marathon race (Pérez-Guarner *et al* 2019), a marathon race (Priego-Quesada *et al* 2020), or muscle damage induction protocols of exercise targeting the quadriceps (Da Silva *et al* 2021) and triceps surae (Da Silva *et al* 2018) adding alternative metrics did not provide any additional information compared to the already described in those studies using the conventional metrics related to average and amplitude temperatures.

Our results may allow some recommendations for future studies addressing skin temperature and muscle damage. Correlation and PCA analyses showed that TMAX and maximum temperature are redundant and, therefore, should not be used together. Our results show that both metrics are highly correlated and therefore can be used interchangeably. However, we recommend considering the TMAX parameter instead of maximum temperature since it results from the measure of 25 pixels, resulting in a more robust metric less susceptible to the influence of single values (Formenti *et al* 2017). Similarly, entropy and SD seem to provide similar information. The SD may show low reproducibility as its values will be, in general, of low magnitude, and a small modification such as of one-tenth will lead to a large percentage of change (Machado *et al* 2021, Da Silva *et al* 2022). The entropy seemed more suitable for detecting differences between experimental conditions as it manifested differences in the image's texture and was suggested that could be more related to the higher presence of hot spots, and therefore of perforator vessels (Bogomilsky *et al* 2022). Although in our results, entropy values were not different between measurement times in any of the studies, this is in agreement with the possible explanation that there is not a higher skin blood flow 24 or 48 h post muscle damage induction. In previous studies, the presence of higher temperature was discussed as an indicator of physiological imbalances resulting from muscle damage (Stewart *et al* 2020) or even cumulative load during intense exercises (Machado *et al* 2023). In this regard, the pixelgraphy, considering temperature measures of pixels >33 °C could serve as another indicator of a body region with high temperature (Rodrigues Júnior *et al* 2021, Alburquerque Santana *et al* 2022). However, it is logical to think that increases in pixelgraphy parameter would result also in increases in TMAX. For this reason, and considering that correlation and PCA analyses showed that pixelgraphy combines information from conventional parameters (for example, mean, minimum, and maximal), our results did not support to use instead the other metrics.

The implementation of these alternative metrics can be challenging, as most of the current software available does not include these options. Standardized metrics that we consider worth including in software and codes to analyse IRT images include the mean temperature, mainly due to its robustness and repeatability (Machado *et al* 2021, Da Silva *et al* 2022). We also recommend the inclusion of TMAX, which is less susceptible to the influence of single pixel values than maximum temperature and relates to peripheral vasodilation and inflammation (Formenti *et al* 2018). While we have not investigated an alternative metric to the minimum temperature, we also recommend the inclusion of minimum temperature due to its theoretical relationship with peripheral vasoconstriction (Da Silva *et al* 2021). The calculation of a TMIN variable, under the same rationale as TMAX, is worth further investigation. Finally, the entropy is recommended, given its relationship with thermal distribution (Bogomilsky *et al* 2022). However, more research is necessary to validate these metrics with the physiological variables (e.g. skin blood flow, inflammation) since they are supposed to be related.

Regardless of the metric employed, the muscle damage resulting from running a half-marathon or marathon did not produce detectable changes in skin temperature. However, quadriceps damage decreased mean skin temperature, as well as reduced maximum, minimum, TMAX, and pixelgraphy metric. Similarly, damage to the triceps surae muscle reduced both the maximum skin temperature and pixelgraphy. These findings suggest that while general muscle damage from long-distance running does not impact skin temperature measurements significantly, localized muscle damage might result in discernible changes in specific temperature-related metrics.

The number of females in the 4 studies evaluated was limited (only one of the 4 studies has a considerable sample of women that would allow for statistical analysis), so it was decided not to assess the effect of sex on the results. However, future studies should analyse the effect of sex, since it has been observed that it can greatly influence the results (Da Silva *et al* 2021), because women usually present lower skin temperature due to their higher body fat proportion, lower body surface area, and lower peripheral blood flow (Barnes 2017, Jimenez-Perez *et al* 2020), and also their skin temperature variability is increased by the menstrual cycle (Stachenfeld *et al* 2000).

Some limitations were identified in our study. We decided to use metrics that were beginning to be used and had already been published with them instead of inventing new ones. We consider that a TMIN, like the TMAX approach but considering the minimum temperature, could be an option to explore in further studies. Another limitation we recognized was the combination of data from four different studies when performing the correlation and principal components analysis. Although the investigations were conducted with similar methodologies, the fact that they did not include the same participants may have affected the results. However, we believe that our results are robust enough to consider these limitations as of minimal influence and also is important to consider that higher data variability is positive for correlation and PCA. Finally, the four studies did not include a control group which is an important limitation in their study designs. However, this does not impact our objective, which is to assess if alternative metrics lead to different results than conventional metrics.

5. Conclusion

The use of alternative metrics is considered more sensitive to individual responses by considering variability and regions representative of higher temperature, which does not change skin temperature outcomes following muscle damage of lower extremity muscle groups. Mean skin temperature, the TMAX, the minimum, and the entropy can be used in combination to provide more information about an ROI.

Future studies should focus on validating the effectiveness of these alternative metrics in detecting changes in skin temperature and exploring their relationship with physiological phenomena like changes in skin blood flow, increase in the level of muscle damage markers, and presence of inflammation.

Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: https://data.mendeley.com/datasets/xfvrp7rpjy/1.

Conflict of interest

The authors declare they have no conflicts of interest regarding the content of this article.

Funding

The academic exchange of the first author was facilitated by the Erasmus+ program and funded by the European Union.

ORCID iDs

Lukas Verderber https://orcid.org/0009-0003-8961-6847
Willian da Silva https://orcid.org/0000-0002-6667-1646
Inmaculada Aparicio-Aparicio https://orcid.org/0000-0003-3784-8489
Andresa M C Germano https://orcid.org/0000-0002-3186-8588
Felipe P Carpes https://orcid.org/0000-0001-8923-4855
Jose Ignacio Priego-Quesada https://orcid.org/0000-0002-0375-1454

References

Alburquerque Santana P V, Alvarez P D, Felipe da Costa Sena A, Serpa T K, de Assis M G, Pimenta E M and Veneroso C E 2022 Relationship between infrared thermography and muscle damage markers in physically active men after plyometric exercise *J. Therm. Biol.* **104** 103187

Barnes J N 2017 Sex-specific factors regulating pressure and flow *Exp. Physiol.* **102** 1385–92

Bogomilsky S, Hoffer O, Shalmon G and Scheinowitz M 2022 Preliminary study of thermal density distribution and entropy analysis during cycling exercise stress test using infrared thermography *Sci. Rep.* **12** 14018

Cohen J 1988 *Statistical Power Analysis for the Behavioral Sciences* 2nd edn (Lawrence Erlbaum Associates, Publishers (https://doi.org/10.4324/9780203771587)

Da Silva W, Machado A, Kunzler M, Jimenez-Perez I, Gil-Calvo M, Priego-Quesada J I and Carpes F P 2022 Reproducibility of skin temperature analyses by novice and experienced evaluators using infrared thermography *J. Therm. Biol.* **110** 103345

Da Silva W, Machado Á S, Lemos A L, de Andrade C F, Priego-Quesada J I and Carpes F P 2021 Relationship between exercise-induced muscle soreness, pain thresholds, and skin temperature in men and women *J. Therm. Biol.* **100** 103051

Da Silva W, Machado Á S, Souza M A, Kunzler M R, Priego-Quesada J I and Carpes F P 2018 Can exercise-induced muscle damage be related to changes in skin temperature? *Physiol. Meas.* **39** 104007

De Andrade Fernandes A, Pimenta E M, Moreira D G, Sillero-Quintana M, Marins J C B, Morandi R F and Garcia E S 2017 Skin temperature changes of under-20 soccer players after two consecutive matches *Sport Sci. Health* **13** 635–43

Ferreira-Júnior J B, Suene C, Pinheiro M, Rezende V, Freitas E, Marins J and Costa C 2021 Is skin temperature associated with muscle recovery status following a single bout of leg press? *Physiol. Meas.* **42** 34002

Formenti D, Ludwig N, Rossi A, Trecroci A, Alberti G, Gargano M and Caumo A 2017 Skin temperature evaluation by infrared thermography: comparison of two image analysis methods during the nonsteady state induced by physical exercise *Infrared Phys. Technol.* **81** 32–40

Formenti D, Ludwig N, Rossi A, Trecroci A, Alberti G, Gargano M and Caumo A 2018 Is the maximum value in the region of interest a reliable indicator of skin temperature? *Infrared Phys. Technol.* **94** 299–304

Hillen B, Pfirrmann D, Nägele M and Perikles S 2020 Infrared thermography in exercise physiology: the dawning of exercise radiomics *Sports Med.* **50** 263–82

Jasti N, Bista S, Bhargav H, Sinha S, Gupta S, Chaturvedi S and Gangadhar B 2019 Medical applications of infrared thermography: a narrative review *J. Stem Cells* **14** 35–53

Jimenez-Perez I, Gil-Calvo M, Priego-Quesada J I, Aparicio I, Pérez-Soriano P and Ortiz de Anda R M C 2020 Effect of prefabricated thermoformable foot orthoses on plantar surface temperature after running: a gender comparison *J. Therm. Biol.* **91** 102612

Ludwig N, Formenti D, Gargano M and Alberti G 2014 Skin temperature evaluation by infrared thermography: comparison of image analysis methods *Infrared Phys. Technol.* **62** 1–6

Machado A, Priego-Quesada J I, Jimenez-Perez I, Gil-Calvo M, Carpes F P and Perez-Soriano P 2021 Influence of infrared camera model and evaluator reproducibility in the assessment of skin temperature responses to physical exercise *J. Therm. Biol.* **98** 102913

Machado Á S, Da Silva W, Priego-Quesada J I and Carpes F P 2023 Can infrared thermography serve as an alternative to assess cumulative fatigue in women? *J. Therm. Biol.* **115** 103612

Moreira D G, Costello J T, Brito C J, Adamczyk J G, Ammer K, Bach A J E and Sillero-Quintana M 2017 Thermographic imaging in sports and exercise medicine: a Delphi study and consensus statement on the measurement of human skin temperature *J. Therm. Biol.* **69** 155–62

O'Rourke N, Hatcher L and Stepanski E J 2005 *A Step-by-step Approach to Using SAS® for Univariate & Multivariate Statistics* 2nd edn 1. print (SAS Inst)

Pérez-Guarner A, Priego-Quesada J I, Oficial-Casado F, Cibrián Ortiz de Anda R M, Carpes F P and Palmer R S 2019 Association between physiological stress and skin temperature response after a half marathon *Physiol. Meas.* **40** 34009

Priego Quesada J I (Ed.) 2017 *Application of Infrared Thermography in Sports Science* (Springer International Publishing) (https://doi.org/10.1007/978-3-319-47410-6)

Priego-Quesada J I, Pérez-Guarner A, Gandia-Soriano A, Oficial-Casado F, Galindo C, Cibrián Ortiz de Anda R M and Salvador Palmer R 2020 Effect of a marathon on skin temperature response after a cold-stress test and its relationship with perceptive, performance, and oxidative-stress biomarkers *Int. J. Sports Physiol. Perform.* **15** 1467–75

Rodrigues Júnior J L, Duarte W, Falqueto H, Andrade A, Morandi R F, Albuquerque M R and Pimenta E M 2021 Correlation between strength and skin temperature asymmetries in the lower limbs of Brazilian elite soccer players before and after a competitive season *J. Therm. Biol.* **99** 102919

Rosenthal R and Rubin D B 2003 R equivalent: a simple effect size indicator *Psychol. Methods* **8** 492–6

Stachenfeld N S, Silva C and Keefe D L 2000 Estrogen modifies the temperature effects of progesterone *J. Appl. Physiol.* **88** 1643–9

Steketee J 1973 Spectral emissivity of skin and pericardium *Phys. Med. Biol.* **18** 686–94

Stewart I B, Moghadam P, Borg D N, Kung T, Sikka P and Minett G M 2020 Thermal infrared imaging can differentiate skin temperature changes associated with intense single leg exercise, but not with delayed onset of muscle soreness *J. Sports Sci. Med.* **19** 469–77