Journal of Thermal Biology 140 (2026) 104541

Contents lists available at ScienceDirect

Journal of Thermal Biology

journal homepage: www.elsevier.com/locate/jtherbio

Responsiveness of the CALERA Research Sensor to heat-acclimation-induced reductions in rectal temperature

Timo van den Bogaard [a,*], Pooh Verheijen [a], Lisa Klous [b], Jan van Erp [b,c], Hein Daanen [a]

[a] Department of Human Movement Sciences, Faculty of Behavioural and Movement Sciences, Amsterdam Movement Sciences, Vrije Universiteit van Amsterdam, Amsterdam, De Boelelaan 1105, Amsterdam, 1081 HV, the Netherlands
[b] Netherlands Organization for Applied Scientific Research (TNO), Department of Human Performance, Unit Defence, Safety and Security, Kampweg 55, Soesterberg, 3769 DE, the Netherlands
[c] Department Human Media Interaction, Faculty Electrical Engineering, Mathematics and Computer Science, University of Twente, Drienerlolaan 5, Enschede, 7522 NB, the Netherlands

ARTICLE INFO

Keywords:
Calera research sensor
Responsiveness
Validation
Non-invasive body core temperature measurement
Heat acclimation

ABSTRACT

Military personnel and athletes increasingly use non-invasive methods to estimate body core temperature during heat acclimation. This study determined the responsiveness of the CALERA Research Sensor (CRS) to changes in resting and end-exercise body core temperature induced by heat acclimation as well as the criterion validity for determining the thermal dose, defined as time spent with a body core temperature ≥38.5 °C.

Fifteen participants completed a nine-day controlled-hyperthermia heat acclimation protocol. Body core temperature was continuously monitored prior to and during heat exposures using both the CRS (T_{crs}) and a rectal temperature (T_{re}) probe. This study assessed responsiveness and criterion validity using Bland-Altman analysis and intraclass correlation coefficients (ICC). Additionally, the analyses included a linear mixed model to further analyze responsiveness.

T_{crs} showed limited responsiveness to heat acclimation-induced changes in resting and end-exercise T_{re} with wide limits of agreement (±0.32 °C), proportional bias (indicating increasing underestimation with larger adaptations), and low ICC (≤0.18). Additionally, resting and end-exercise T_{re} decreased to a greater extent than T_{crs}, which showed a downward trend during end-exercise over the nine-day heat acclimation ($P < 0.001$). The CRS showed limited criterion validity for estimating thermal dose, with wide limits of agreement (±20 min), proportional bias (indicating increasing underestimation at higher thermal doses), and low ICC of 0.35.

These results indicate that the current version of the CRS underestimates both heat acclimation-induced changes in T_{re} and the thermal dose during controlled hyperthermia.

1. Introduction

Continuous monitoring of body core temperature (T_c) is important in military populations to maximize effective and safe performance in warm and cold environments (Buller et al., 2018; Dolson et al., 2022). Although there is no single proxy for T_c (Taylor et al., 2014), esophageal (T_{es}) and rectal temperature (T_{re}) are usually considered good indicators of T_c (ISO, 2004). Several disadvantages arise from the use of T_{es} and T_{re} in a military setting as they are intrusive and impractical during in-field activities (Priego-Quesada et al., 2025; Wartzek et al., 2011). Therefore, there is a need for more practical tools to assess T_c.

Non-invasive wearable sensors estimating T_c offer a small, easy-to-use, comfortable, and fast alternative for continuous estimation of T_c (Wartzek et al., 2011). Reliability of estimations of T_c using non-invasive wearable sensors has improved over the years and several wearable sensors show promising results, though further validation is required (Dolson et al., 2022; Notley et al., 2025; Priego-Quesada et al., 2025). The CALERA Research Sensor (CRS; CALERA Research Sensor, greenteg AG, Rümlang, Switzerland) estimates T_c based on skin temperature (T_{sk}), heart rate (HR), and heat flux using proprietary algorithms that are not publicly available. The CRS has received scientific interest, with recent studies focusing on its use to non-invasively monitor T_c (T_{crs}) in both laboratory settings (Daanen et al., 2023; Januario et al., 2024; Jolicoeur Desroches et al., 2023; Kaltsatou et al., 2024; Kubota et al., 2025; McLaughlin et al., 2025; Priego-Quesada et al., 2025; Richard et al., 2025; Verdel et al., 2021; Xu et al., 2025; Xu et al., 2024) and

* Corresponding author. Van der Boechorststraat 7, 1081 BT, Amsterdam, the Netherlands.
E-mail address: t.b.van.den.bogaard@vu.nl (T. van den Bogaard).

https://doi.org/10.1016/j.jtherbio.2026.104541
Received 4 February 2026; Received in revised form 7 July 2026; Accepted 14 July 2026
Available online 14 July 2026

applied settings (Ajčević et al., 2022; Etienne et al., 2023; Goods et al., 2023).

Despite variability in reported criterion validity,[1] the CRS demonstrates high test-retest reliability[2] as the reported mean bias between sessions is low (≤0.03 ± 0.40 °C) and the intraclass correlation coefficient (ICC) of 0.98 is high (Januario et al., 2024; Richard et al., 2025; Verdel et al., 2021). The high test-retest reliability theoretically makes the CRS suitable for monitoring changes in T_c over time, such as during heat acclimation (HA) to estimate HA status (Januario et al., 2024; Richard et al., 2025). During HA, both resting and mean exercise T_c reduce by approximately 0.2 °C and 0.2-0.4 °C, respectively (McDonald et al., 2025; Tyler et al., 2024). Moreover, HA generally lowers HR and T_{sk} (McDonald et al., 2025; Tyler et al., 2024), both of which are included in the CRS algorithm to estimate T_c. Therefore, successful adaptation may lead the CRS to generate lower T_c values as HA progresses. Both the high test-retest reliability and the HA-induced reductions in HR and T_{sk} theoretically strengthen the responsiveness of the CRS to detect reductions in T_c induced by HA. Within this study, we define responsiveness as the ability of an instrument to detect change over time in the construct to be measured (de Vet et al., 2011; Mokkink et al., 2010, 2021).

One scientific study used the CRS to quantify HA-induced adaptations in T_c, reporting a 0.4 °C lower end-exercise T_{crs} following a treadmill exercise bout after four weeks of HA (Xu et al., 2025). However, the authors acknowledge that no gold standard measure of T_c was included, making it impossible to quantify the true magnitude of change or evaluate the responsiveness of the CRS to HA-induced reductions in T_c. Furthermore, HA protocols are typically shorter than four weeks in military and athletic settings (Parsons et al., 2019; Periard et al., 2015). The responsiveness of the CRS during short- and medium-term HA protocols remains unexamined.

The controlled hyperthermia approach is a commonly used approach to HA (Ashworth et al., 2020; Esh et al., 2024; Parsons et al., 2019; Periard et al., 2015). Controlled hyperthermia involves elevating and maintaining T_{re}, the most commonly reported proxy of T_c in HA-research (Tyler et al., 2024), at and slightly above 38.5 °C for approximately 60 min (Periard et al., 2021). To determine the thermal dose during controlled hyperthermia HA, the CRS needs to accurately estimate the time spent with a $T_c \geq 38.5$ °C. One study evaluated if the CRS could be used to quantify the thermal dose during a four-day heat training camp for elite female hockey players (Goods et al., 2023). Their findings indicate that the CRS underestimated the thermal dose by 10 ± 25% (~12 ± 30 min) as the CRS typically underestimated gastrointestinal temperature. However, exercise with an unstable T_c, such as during hockey matches, reduces the criterion validity of the CRS to measure T_c due to the time it takes for heat to transfer from the body core to the skin (Xu et al., 2024). It remains to be investigated if conditions with a more stable T_c allow for a better estimation of the thermal dose.

The first aim of this study was to assess the responsiveness of the CRS to HA-induced reductions in resting and end-exercise T_{re} as an indicator of T_c. We hypothesized that the CRS detects HA-induced reductions in T_{re} as the CRS estimates T_c based on physiological variables that show a HA-induced reduction and the CRS has high test-retest reliability for estimating T_c. The second aim of this study was to report the criterion validity of the CRS to quantify the thermal dose as time spent with a $T_{re} \geq 38.5$ °C. We hypothesized that CRS underestimates the time spent with a $T_{re} \geq 38.5$ °C based on a previous study (Goods et al., 2023).

[1] Criterion validity is defined as how well the scores of a new measurement device agree with the gold standard (de Vet et al., 2011; Mokkink et al., 2010).

[2] Test-retest reliability is defined as the degree to which a measurement system is able to measure the same thing twice (de Vet et al., 2011; Mokkink et al., 2010).

2. Methods

2.1. Participants

Before the start of this study, potential participants filled in a screening questionnaire (Thompson, 2010) that aimed to assess potential medical risks. Potential participants were excluded if they had previously suffered from cardiovascular complications, had experienced heat stroke, had not engaged in weekly physical activity, or were older than 45 years for men or 55 years for women. Based on these criteria, all included participants were classified as low risk according to the American College of Sports Medicine guidelines (Thompson, 2010). Additionally, potential participants were excluded if they had spent seven days or more in a hot environment within the two months preceding the study to prevent heat acclimation effects.

This study commenced with twenty healthy, regularly active, and unacclimatized Caucasian volunteers. We included data from all participants who completed the entire protocol to answer the research question regarding the responsiveness (N = 15) and all participants who executed at least one session to answer the research question regarding the criterion validity for thermal dose (N = 20). Table 1 represents participant characteristics.

Prior to the study, participants were informed about the procedures and provided verbal and written consent. The ethics committee of the Faculty of Behavioral and Movement Sciences of the Vrije Universiteit Amsterdam approved all procedures (VCWE-2023-144R1), which were conform to the standards set out by the Declaration of Helsinki (2013).

2.2. Experimental design

Fig. 1 illustrates the study timeline. At least seven days (range: 7-31 days) before the HA protocol commenced, participants completed a preliminary visit to the laboratory for measurement of individual characteristics, including a maximal oxygen uptake (VO_{2max}) test. One day before the HA protocol commenced, participants performed a progressive heat stress test (HST) to assess thermoregulatory upper limits as elaborately described elsewhere (van den Bogaard et al., 2025). In short, the HST included cycling at ~1 W/kg external power output for ~120 min in a climatic chamber clamped at 38 °C dry-bulb temperature (T_{db}) with progressively increasing relative humidity (RH). On the following day, participants started a nine-day controlled hyperthermia HA protocol, which included another progressive HST on the day between the fifth and sixth day of HA (Fig. 1).

Before each HA session, participants were instructed to hydrate according to protocol (500 ml the night before and approximately 10 ml/kg body mass in the hours before) and to abstain from exercise and alcohol 24 h and from caffeine 12 h prior to the session. Upon arrival at

Table 1
Participant characteristics.

	Responsiveness group	Thermal dose group
N	15	20
Sex		
Male	6	9
Female	9	11
Age (years)	26.5 ± 5.9	26.8 ± 6.5
Height (cm)	174.0 ± 6.9	174.8 ± 7.8
Body mass (kg)	69.2 ± 9.2	69.0 ± 8.2
BSA (m^2)	1.8 ± 0.1	1.8 ± 0.1
BSA-to-mass ratio (cm^2/kg)	267.0 ± 16.1	267.2 ± 14.4
VO_{2max} (ml/kg/min)	46.3 ± 7.5	46.4 ± 7.4

Note. Characteristics of the participants included in both the criterion validity and responsiveness analysis. Values are average ± standard deviation. T_{re} = rectal temperature. N = number of participants. BSA = Body surface area. VO_{2max} = maximal oxygen uptake. No statistically significant differences ($P > 0.05$) were observed between the criterion validity and responsiveness groups.

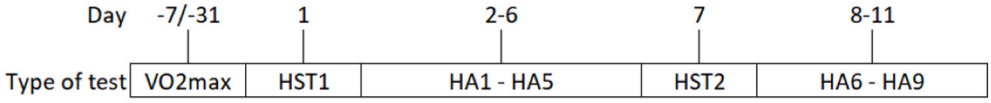

Fig. 1. Study timeline. HST = heat stress test. HA = heat acclimation. VO_{2max} = maximal oxygen uptake test.

the lab, participants provided a urine sample to verify euhydration, defined as urine-specific gravity <1.020 (PAL-10S, Atago Co. Ltd, Tokyo, Japan) (Kenefick and Cheuvront, 2012). If participants were hypohydrated, they were asked to drink 200 ml of water before they started the HA session. Participants performed the HA sessions in a climatic chamber set to warm-humid conditions (35 °C T_{db}, 65% RH, 31 °C wet-bulb globe temperature (WBGT)) with minimal airflow of 0.2 m/s. Each HA session consisted of three phases, all conducted in the warm-humid climate chamber: a constant resting phase, a constant cycling phase, and a controlled hyperthermia phase. During the constant resting phase, participants sat quietly for 15 min. During the constant cycling phase, participants cycled for 30 min on a cycle ergometer at an external workload of 1.75 W/kg. Immediately following the constant cycling phase, participants briefly exited the climatic chamber to have their body mass measured. During the controlled hyperthermia phase, participants adjusted cycling power output and resting intervals to accumulate 60 min with a T_{re} at and above 38.5 °C. Fig. 2 illustrates the three phases during each HA session.

2.3. Measurements

2.3.1. Preliminary visit

Participants visited the laboratory for measurement of body mass and height, and calculation of body-surface-area (BSA) using the formula proposed by DuBois and DuBois (Du Bois and Du Bois, 1916). Body mass was measured using a platform scale (SATEX SA-1 250, Weegtechniek Holland B.V., Zeewolde, The Netherlands) and height using a stadiometer (Seca 217, Seca, Hamburg, Germany). Next, participants performed a graded exercise test to volitional exhaustion on a cycle ergometer (Monark LC7TT, Monark Exercise AB, Vansbro, Sweden) to determine VO_{2max} in a temperate condition (~20 °C, ~30% RH). Cycling started at an external workload of 0 W for 3 min, after which the external workload increased linearly until participants reached volitional exhaustion. Workload increments were individualized a priori to elicit volitional exhaustion within approximately 8–12 min.

2.3.2. HA sessions

T_c was continuously estimated using the CRS device and a rectal probe (MSR, Seuzach, Switzerland; or Yellow Springs Instruments, Yellow Springs, OH, USA). The rectal probe, self-inserted ~12 cm past the anal sphincter, measured T_{re} at 4-sec intervals. The CRS device estimated T_{crs} continuously at 1-sec intervals. The CRS device was attached to a HR chest band monitor (Polar H10, Kempele, Finland), according to the manufacturer's guidelines, and estimated T_{crs} using the

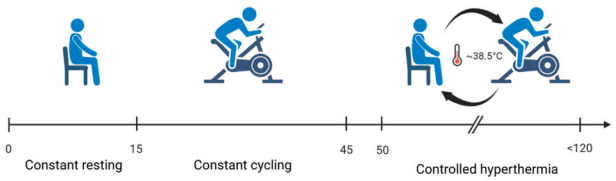

Fig. 2. Protocol for heat acclimation sessions. *The experiment was executed in a climatic chamber set to warm-humid conditions (*35 °C T_{db}, 65% RH*). Each session consisted of three phases: a constant resting phase were the participant sat quietly, a constant cycling phase were the participant cycled with an external power output of 1.*75 W/kg*, and a controlled hyperthermia phase were the participant altered cycling and resting intervals to accumulate* 60 min *with a rectal temperature above 38.5* °C. The sessions had a maximal duration of 120 min.

algorithm implemented on greenteg consumer wearables at the time of the study (2024; firmware version, core 0.8.4). The algorithm calculated T_{crs} based on heat flux and T_{sk} measured by the CRS and HR measured by the HR chest band monitor. Moreover, iButtons (DS192, Maxim Integrated Products, Inc., San Jose, CA, USA), placed on the neck, scapula, thigh, and shin, continuously measured T_{sk} at 1-min intervals. Weighted averages were used to calculate average T_{sk} (ISO, 2004).

2.4. Data processing

All data were synchronized, formatted, and analysed using R software (version 4.3.2, R Foundation for Statistical Computing, Vienna, Austria) in the RStudio environment (version 2023.12.1.402, RStudio, Inc., Boston, MA, USA).

After data extraction, 1-min averages were taken for all data. Subsequently, T_{re}, T_{crs}, T_{sk}, and HR responses were visualized for every participant. Two researchers (PV, TB) assessed the shape of the curves and discussed abnormalities through case-by-case review. T_{re} for the entire session was removed if outliers persisted for more than 10 min. If outliers lasted less than 10 min, data were interpolated. HR data points were excluded if resting values were below 40 bpm or if values were below 100 bpm after 10 min of cycling. Because T_{crs} depends on HR, corresponding T_{crs} data points were excluded if HR values were excluded.

HA adaptations were determined during both the constant resting and constant cycling phase by calculating the average of the final 5 min of each phase (Alkemade et al., 2021). The final 5 min were selected to minimize the T_{re} time lag due to poor rectal perfusion (Taylor et al., 2014) and to ensure a steady state to better estimate the true criterion validity of the CRS (Wartzek et al., 2011). To assess the responsiveness of the CRS and quantify adaptation, T_{re} and T_{crs} values from each HA session were compared with those recorded in the first session (or second session in case of missing values (N = 3)) to generate change scores. To quantify the thermal dose, the number of minutes at and above a T_{re} and T_{crs} of 38.5 °C per session was calculated over all phases of each HA session.

2.5. Statistical analysis

2.5.1. Responsiveness

To estimate the responsiveness of the CRS to HA-induced changes in T_c, this study employed two complementary approaches.

First, as proposed by de Vet et al. (2011), responsiveness was evaluated using a criterion-based approach to longitudinal validity by analyzing the agreement of the HA-induced adaptation of T_{crs} and T_{re} during rest and end-exercise. These adaptations were used to calculate the ICC and perform a Bland-Altman visualization. The Bland-Altman visualization shows systematic error (mean bias) and random error (limits of agreement) (Bland and Altman, 1985). For each participant i and session j, we defined the difference in adaptation as $d_{ij} = \Delta T_{CRS,ij} - \Delta T_{re,ij}$ and the mean adaptation as $m_{ij} = \left(\Delta T_{CRS,ij} + \Delta T_{re,ij}\right)/2$. To account for repeated measures within participants, d_{ij} was modelled using a random-intercept mixed model $d_{ij} = \beta_0 + b_i + e_{ij}$ with β_0 as mean bias, $b_i \sim N(0, \sigma_b^2)$ and $e_{ij} \sim N(0, \sigma_e^2)$. Population-level LoA for a single observation were computed as $\beta_0 \pm 1.96\sqrt{\sigma_b^2 + \sigma_e^2}$ using the estimated between-participant and residual variance components. Proportional bias was assessed by extending the model to $d_{ij} = \beta_0 + \beta_1 m_{ij} + b_i + e_{ij}$; when $\beta_1 \neq 0$, bias was presented as $\widehat{bias}(m) = \hat{\beta}_0 + \hat{\beta}_1 m$ with $LoA(m) =$

$\widehat{bias}(m) \pm 1.96\sqrt{\sigma_b^2 + \sigma_e^2}$. Confidence intervals for the mean bias and limits of agreement were estimated using parametric bootstrapping of the mixed effects Bland–Altman model and visualized as 95% confidence bands across the observed range of mean change scores. As no acceptability thresholds for the adaptation exist, mean bias and LoA were interpreted relative to the measured adaptation in T_{re}. Within this context, relatively large bias and wide LoA were considered indicative of limited responsiveness of the CRS to HA-induced changes in T_c. ICC estimates were based on a single-rater, absolute-agreement, 2-way mixed-effects model and calculated as the ratio of between-participant variance and total variance (de Vet et al., 2011; Koo and Li, 2016). This conservative ICC reflects ratio of between-participant variability to measurement error, with inference restricted to the specific devices used in this study. Total variance included both systematic differences between devices and residual variance. Variances were estimated using a mixed-effects model containing device as fixed effect and a random intercept for each participant. ICC's > 0.7 were interpreted as indicating acceptable agreement (de Vet et al., 2011).

Second, a linear mixed model was fitted to estimate the influence of HA over time across devices on the absolute values of T_c. The main aim of this analysis was to evaluate if changes over HA sessions, i.e. the slope, in T_c differed between devices. The model included: device (two levels: rectal probe, CRS) and HA session (modelled as a continuous variable (van Wely, 2024)) as independent variables, T_c as dependent variable, and a random intercept for each participant to account for baseline differences in T_c across individuals. The interaction term device x HA session indicated whether the slope of HA-induced changes in T_c differed between devices. The linear mixed model assumes a linear change in T_c over time whilst HA-induced changes in T_c are usually not linear (e.g. most adaptations happen in the first five days) (Periard et al., 2021). However, visual inspection of our data and linear mixed model showed that linear modelling of HA-induced changes in T_c was appropriate (appendix 1). To aid physiological interpretation, overall HA–induced changes in T_c were additionally quantified as the model-based estimated difference between the first and final HA session for each device, derived from the linear mixed model. Standardized within-subject effect sizes (Cohen's d_z) were calculated based on first-to-last session differences. Additionally, the SD of individual first-to-last session adaptations was calculated to describe inter-individual variability.

2.5.2. Criterion validity for thermal dose

To estimate the criterion validity to quantify the thermal dose, the amount of minutes with a T_{re} and $T_{crs} \geq 38.5$ °C were used to calculate the ICC and perform the Bland-Altman visualization. As acceptability thresholds for the amount of minutes with a $T_c \geq 38.5$ °C do not exist, agreement of thermal dose between the CRS and rectal probe was evaluated descriptively. Mean bias and LoA were interpreted relative to its practical implications for controlled hyperthermia HA, where time spent above T_c guides HA.

2.5.3. HA adaptation

To evaluate if our HA protocol also resulted in the expected adaptations in T_{sk} and HR, a paired *t*-test was executed using the first and last HA session. When the Shapiro-Wilk test indicated a violation of the assumption of normality, the Wilcoxon signed-rank test was reported.

3. Results

3.1. Responsiveness

A significant proportional bias during rest ($\beta_1 = -1.419$ (95% CI: −1.617 to −1.228), $P < 0.001$) and exercise ($\beta_1 = -0.396$ (95% CI: −0.607 to −0.178), $P < 0.001$) indicated that larger adaptations (more negative change scores) were associated with an increasing mean bias, reflecting increasing underestimation of the CRS. Consequently, regression-based mean bias and LoA are presented. Resting mean bias and LoA regressions were $\widehat{bias}(m) = 0.005 - 1.419 * m$ and $LoA(m) = \widehat{bias}(m) \pm 1.96 * \sqrt{0.011 + 0.015}$, respectively. Exercising mean bias and LoA regressions were $\widehat{bias}(m) = 0.005 - 0.396 * m$ and $LoA(m) = \widehat{bias}(m) \pm 1.96 * \sqrt{0.035 + 0.027}$, respectively. To aid interpretation, at the average T_c adaptation of −0.2 °C found in the literature (Tyler et al., 2024), mean bias and LoA of change scores during rest and exercise were 0.28 ± 0.32 °C and 0.08 ± 0.49 °C, respectively. Fig. 3 shows the Bland-Altman visualization for adaptations in rest and exercise.

The ICC was 0.134 (95% CI: 0.023 to 0.241) and 0.180 (95% CI: 0.024 to 0.364) for HA-induced adaptations during rest and exercise, respectively.

HA had a significant main effect on T_c at both rest and end-exercise ($P < 0.001$), while device effects were significant during exercise ($P < 0.001$) but not during rest ($P = 0.054$). There was a significant interaction between HA and measurement device during rest ($F(1, 245.05) = 23.09$, $P < 0.001$) and at end-exercise ($F(1, 247.03) = 6.17$, $P = 0.014$). Across HA sessions, reported as mean ± SE, T_{re} decreased by 0.04 ± 0.01 °C/day (total decrease: 0.33 ± 0.05 °C, $d_z = 1.06$) during rest and by 0.04 ± 0.01 °C/day (total decrease: 0.30 ± 0.05 °C, $d_z = 1.4$) at end-exercise. Across HA sessions, T_{crs} during end-exercise significantly decreased by 0.02 ± 0.01 °C/day (total decrease: 0.15 ± 0.05 °C; $P < 0.001$, $d_z = 0.8$) but not during rest (0.004 ± 0.006 °C/day, $P = 0.187$). The SDs in T_{re} adaptations were 0.27 °C during rest and 0.19 °C at end-exercise. Fig. 4 shows resting and at end-exercise T_{crs} and T_{re}.

3.2. Criterion validity for thermal dose

On average, the thermal dose, quantified as the time spent with a $T_c \geq 38.5$ °C, reported as mean ± SD, was 38 ± 23 min for the rectal probe and 19 ± 16 min for the CRS. A significant proportional bias ($\beta_1 = -0.403$ (95% CI: −0.572 to −0.234), $P < 0.001$) indicated that higher thermal dose was associated with larger mean bias, reflecting increasing underestimation of the CRS at higher thermal dose. Consequently, regression-based mean bias and LoA are presented. Mean bias and LoA regressions were $\widehat{bias}(m) = -8.117 - 0.403 * m$ and $LoA(m) = \widehat{bias}(m) \pm 1.96 * \sqrt{219.496 + 160.190}$, respectively. To aid interpretation, at the average thermal dose of 29 min measured across devices within this study, mean bias and LoA were −20 ± 19 min. Fig. 5 shows the Bland-Altman visualization for thermal dose.

The ICC was 0.350 (95% CI: 0.094 to 0.536) for thermal dose.

3.3. HA adaptations in HR and T_{sk}

The Shapiro-Wilk test showed that HR and T_{sk} data at rest and end-exercise were normally distributed ($P > 0.118$). The dependent *t*-test showed that HR, reported as mean ± SD, was significantly lower on HA9 compared to HA1 during rest and at end-exercise ($t(14) = 3.9$, $P = 0.002$; $t(14) = 2.6$, $P = 0.020$) by 6 ± 6 bpm and 6 ± 9 bpm, respectively. Also, T_{sk} was significantly lower on HA9 compared to HA1 during rest and at end-exercise ($t(14) = 2.8$, $P = 0.014$; $t(14) = 2.6$, $P = 0.022$) by 0.3 ± 0.4 °C and 0.3 ± 0.5 °C, respectively.

4. Discussion

4.1. Responsiveness

The first goal of this study was to assess the responsiveness, defined as the ability of an instrument to detect change over time in the construct to be measured, of the CRS to HA-induced reductions in resting and end-exercise T_c. We hypothesized that the CRS could detect HA-induced reductions in T_c due to the high test-retest reliability of the

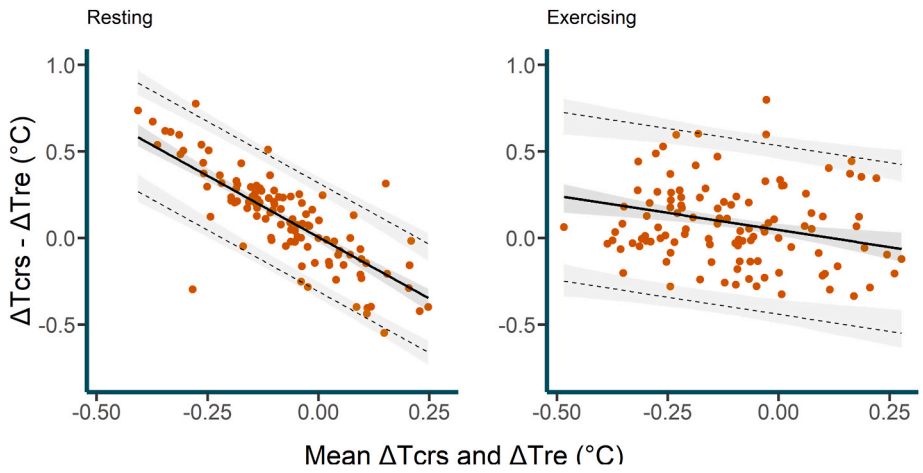

Fig. 3. Bland-Altman plot of heat acclimation induced changes in body core temperature. Adaptations in body core temperature over heat acclimation sessions were measured using a rectal probe (ΔT_{re}) and the CALERA Research Sensor (ΔT_{crs}) during rest (left panel) and end-exercise (right panel) in fifteen participants. The solid line indicates the mean bias; dashed line indicates the upper and lower limit of agreement. A positive value indicates that ΔT_{crs} was larger than ΔT_{re}. An increasingly positive difference on the y-axis with a decreasing mean on the x-axis indicates that ΔT_{crs} progressively overestimates ΔT_{re} during heat acclimation.

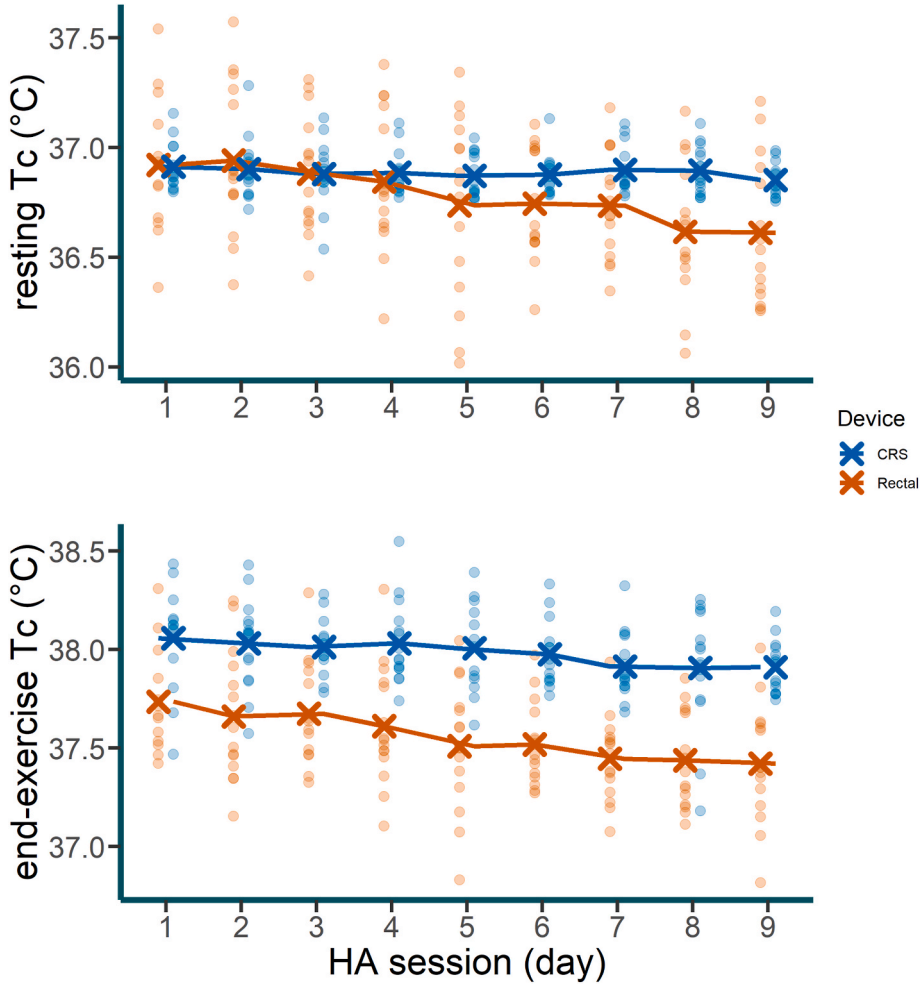

Fig. 4. Average body core temperature across heat acclimation sessions. *The average body core temperature (T_c) during rest (left panel) and end-exercise (right panel) across nine heat acclimation (HA) sessions for fifteen participants. T_c was measured using a rectal probe (T_{re}, orange) and a CALERA Research Sensor (T_{crs}, blue). Dots represent mean values of each participant and symbol X represents the mean value of that HA session.*

CRS as reported by Januario et al. (2024) and the HA-induced reductions in HR and T_{sk}. Contrary to our hypothesis, and despite the HA-induced reduction in HR and T_{sk} observed in this study, our findings show that the CRS did not detect HA-induced reductions in T_c similarly to a rectal probe. Agreement analyses showed a proportional bias, wide LoA, and low ICC of adaptations measured using the CRS compared to a rectal

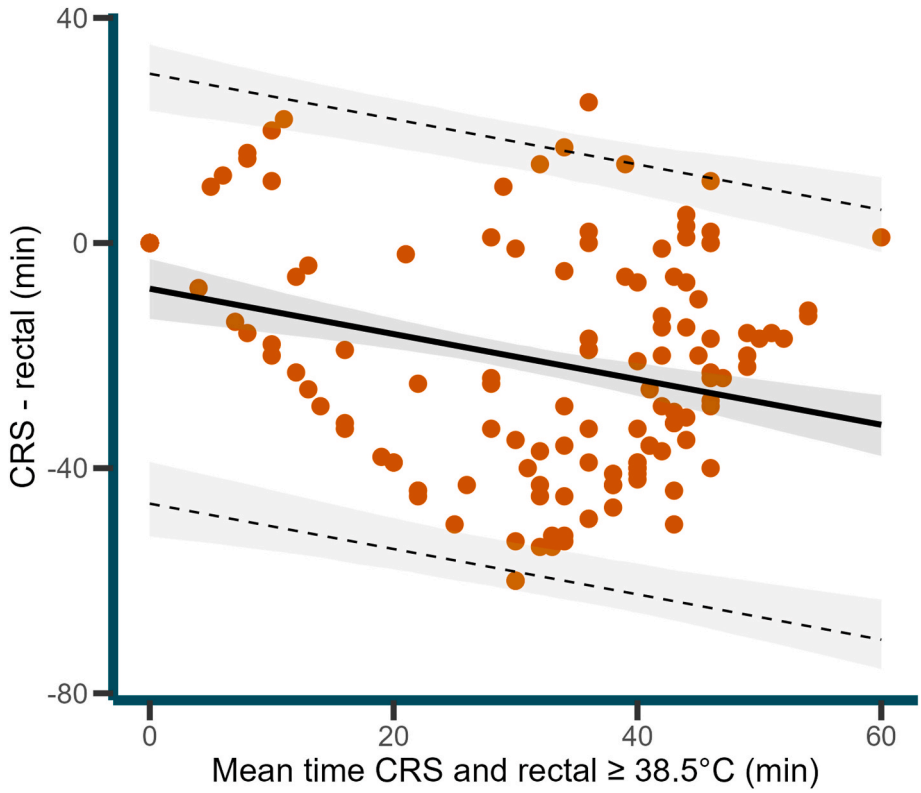

Fig. 5. Bland-Altman plot of the time spent with a body core temperature (T_c) at and above 38.5 °C during each heat acclimation session. The time spent with a T_c was measured using a rectal probe and the CALERA Research Sensor (CRS) for twenty participants. The solid line indicates the mean bias; dashed line indicates the upper and lower limit of agreement. A positive value indicates that CRS overestimated the time spent with a T_c at and above 38.5 °C. Dots outside the "V-shape" indicate sessions where $T_c \geq 38.5$ °C for longer than 60 min.

probe. Importantly, both mean bias and LoA are relatively large in magnitude compared to the measured adaptation in T_{re}. As a result, disagreement in change scores was large relative to inter-individual variability in adaptation magnitude. This is reflected by low ICC values and indicates limited relative agreement. Collectively, these findings indicate limited agreement of HA-induced changes in T_c during rest and exercise and therefore, within the criterion-based approach to longitudinal validity, suggest limited responsiveness of the CRS to HA-induced reductions in T_c during rest and at end-exercise.

The rate of change of HA-induced changes in T_c differs between the CRS and rectal measurements. Specifically, the CRS showed no HA-induced reduction in resting T_{crs}, whereas resting T_{re} decreased over the days. At end-exercise, CRS showed an HA-induced reduction in T_{crs} which was smaller than that in T_{re}. These observations indicate that HA status may affect the criterion validity of CRS for estimating absolute T_{re} because T_{re} decreases to a greater extent than T_{crs}. Supporting this interpretation, additional analyses (Appendix 2) reveal that mean bias increases and becomes more positive across HA sessions, reflecting greater overestimation over the course of acclimation. A plausible explanation for the observed increase in bias is that HA induces physiological adaptations that modify heat transfer from the core to the skin and from the skin to the environment, thereby altering the relationship between T_c and the variables used by the CRS to estimate T_c. Specifically, HA lowers the T_c threshold for sweating, increases whole-body sweat loss and skin wettedness, and redistributes blood flow (e.g. greater muscle and core perfusion and reduced skin perfusion at a given workload) (Lynch et al., 2025; Periard et al., 2015, 2016). Collectively, these adaptations enhance heat flux at a lower T_c, which may cause T_c measured at the rectum to decrease while the effect on T_c estimated from skin-based parameters remains more complex. This interpretation is speculative, as the CRS algorithms are not publicly available; therefore, no definitive conclusions can be drawn. Nevertheless, incorporating HA status into CRS algorithms could potentially improve criterion validity during HA, but this requires further investigation.

Despite the limited responsiveness in the current study, end-exercise T_{crs} showed a statistically significant HA-induced reduction of 0.02 °C/day and overall adaptation of 0.15 °C. The magnitude of overall adaptation was larger than the smallest detectable change in the CRS signal of 0.07 °C (derived from previous literature using intraclass correlation data (Januario et al., 2024)), indicating that the observed change reflects a detectable change in the CRS signal beyond measurement error. Additionally, the magnitude of this reduction is similar to the only available study using the CRS to determine HA-induced adaptations in end-exercise T_{crs} (Xu et al., 2025). Specifically, this study shows a reduction of 0.4 °C T_{crs} after twenty HA sessions during a four-week period. However, because no gold standard measure of T_c was included in that study, the true magnitude of T_c adaptation and therefore CRS's responsiveness cannot be directly compared between studies. Our results indicate that changes in end-exercise T_{crs} – although directionally consistent with T_{re} - are an underestimation of the HA-induced changes in end-exercise T_{re}. Therefore, when using CRS to quantify HA T_c changes, researchers should either include a gold standard measurement (e.g., T_{re}) for validation or explicitly acknowledge that CRS-derived changes represent a conservative estimate of true T_c adaptations. This raises the broader question of how HA status is quantified outside research settings. Different from the current study, greenteg reports a heat adaptation score based on multiple physiological parameters rather than T_c alone, acknowledging that HA-induced adaptations extend beyond changes in core temperature (Periard et al., 2021). By integrating T_{crs}, T_{sk} and HR into a single adaptation metric, the device aims to provide a more holistic assessment of HA status (Heat Adaptation Score, 2025). Although the simplicity of a single score may facilitate athlete uptake of HA practices (Esh et al., 2024), formal validation of this score against established physiological markers of heat adaptation

remains necessary.

Combined, these findings indicate that CRS shows limited responsiveness, as characterized by limited longitudinal validity of change scores and different adaptation patterns, to HA-induced reductions in T_{re} at rest and end-exercise. This limited responsiveness reflects an increasingly positive bias over the course of acclimation, resulting in reduced criterion validity of CRS estimates of T_{re} after multiple HA sessions.

4.2. Criterion validity for thermal dose

The second goal of this study was to assess the criterion validity for quantifying the thermal dose measured by the CRS compared to a rectal probe. We hypothesized that the CRS would underestimate the thermal dose, quantified as the time spent with a $T_{re} \geq 38.5$ °C. Aligning with our hypothesis, agreement analyses showed a proportional bias, wide LoA, and low ICC of thermal dose measured using the CRS compared to a rectal probe. Importantly, both mean bias and LoA are relatively large in magnitude compared to the thermal dose measured in T_{re}. These findings indicate low criterion validity of the CRS for quantifying the thermal dose during controlled hyperthermia HA, as evidenced by proportional bias and wide limits of agreement relative to the magnitude of the thermal dose measured in T_{re}.

Our findings are in accordance with the only study investigating the criterion validity of the thermal dose and that showed limited agreement with ICC of 0.36, a large mean bias of 10% deviation (~12 min), and wide LoA of ±59% (~30 min) compared to gastrointestinal temperature (Goods et al., 2023). Goods et al. (2023) argue that the low ICC and large mean bias might be caused by a reduction in CRS's T_c prediction accuracy at high T_c. This argument is supported by later studies showing that CRS's T_c prediction accuracy varies during exercise and deteriorates at higher T_c values (McLaughlin et al., 2025; Priego-Quesada et al., 2025). McLaughlin et al. (2025) suggests that the CRS shows reduced reliability and phase-dependent shifts in criterion validity, as indicated by wide LoA and changes in bias direction. Specifically, the shift in bias (i.e. larger under- or overestimation) occurs around a T_{re} of ~38.1-38.5 °C. This suggestion is supported by the Bland-Altman visualization in appendix 3, which shows increasingly negative mean bias and widening of the LoA for CRS at higher T_{re} values (>~38.2 °C). In other words, the CRS tends to underestimate T_c around the 38.5 °C threshold, therefore underestimating the thermal dose. To enhance CRS accuracy to determine the thermal dose, future algorithm development could incorporate additional data at higher T_c values during algorithm training (McLaughlin et al., 2025). Taken together, the results of our study and Goods et al. (2023) suggest that the current version of the CRS should not be used to quantify thermal dose during controlled hyperthermia HA.

Controlled hyperthermia HA protocols, where thermal impulse is defined by the time spent with $T_c \geq 38.5$ °C, were originally proposed to be the most effective HA protocols as they maintain a consistent thermal stimulus across sessions (Periard et al., 2015). However, a recent study indicates that controlled hyperthermia is not more effective than other methods (McDonald et al., 2025), and various other approaches that elevate T_c and T_{sk} and elicit profound sweating appear effective (Periard et al., 2021). In this context, the CORE sensor and associated app (i.e. consumer sports product by greenteg similar to CRS) quantify heat strain using T_{crs} and T_{sk} and determines thermal dose based on time spent at elevated heat strain (Heat Adaptation Score, 2025). Users are advised to train in Heat Zone 3, which can be reached with lower T_c than 38.5 °C when T_{sk} is elevated (>34 °C). However, whether prescribing HA based on these heat-strain zones elicits physiological adaptations comparable to established HA methods remains unknown and warrants future investigation.

4.3. Future directions

Collectively, our findings suggest that the current version of the CRS algorithm is unsuitable to measure HA-induced changes in T_{re} or determine the thermal dose during controlled hyperthermia HA. As HA status might influence the criterion validity of the CRS, HA status might be considered in the calculation of T_{crs}. HA status is not only determined by changes in T_c, but also by changes in HR, T_{sk}, sweat rate, and thermal sensation, among others (Periard et al., 2015; Tyler et al., 2024). Future research might focus on computing an individual's adaptation status using multiple variables such as estimated T_{crs}, the input parameters (T_{sk}, HR) of the CRS, and a thermal sensation score. Additionally, to improve the performance of the CRS to quantify the thermal dose, more data at high T_c values are required (McLaughlin et al., 2025; Priego-Quesada et al., 2025).

4.4. Limitations

Some aspects of the current study may limit the generalizability of the findings. First, T_{re} was chosen as the sole proxy of T_c as it is the most commonly reported measurement location in HA research (Tyler et al., 2024). However, T_c varies across measurement location and the time course of HA-induced changes might differ (Taylor et al., 2014; Tyler et al., 2024). For example, reductions in resting gastrointestinal temperature are slightly smaller and more variable compared to resting T_{re} and T_{es} (Tyler et al., 2024). Therefore, the findings on limited responsiveness cannot be generalized to other measurement locations of T_c. Second, the experiment was performed in a single environmental condition in a climatic chamber, which limits the generalizability to other environmental factors such as wind or solar radiation often encountered in a more practical setting, which might impact the validity of the CRS (Xu et al., 2025). Third, body fat of the healthy and regularly active participants in the current study was not measured. Higher body fat potentially increases the CRS bias (McLaughlin et al., 2025) and might therefore impact the bias found in the current study. Future studies should include the measurement of body fat.

5. Conclusion

Our results indicate that the current version of the CALERA research sensor underestimates HA-induced changes in rectal temperature (as an indicator of core temperature) and the thermal dose when quantified by the time spent with a rectal temperature at and above 38.5 °C. To improve the performance of the CALERA research sensor, it is recommended to include more data at high core temperature values and heat acclimation status in the training set for the estimated core temperature by the CALERA research sensor.

Data accessibility statement

Research data Most source data for this study is openly retrievable via https://doi.org/10.48338/VU01-PH2K50. Heat flux data cannot be shared due to a contractual agreement with the sensor manufacturer, which prevents disclosure to avoid reverse engineering of proprietary algorithms.

Declaration of generative AI and AI-assisted technologies in the manuscript preparation process

During the preparation of this work the authors used Microsoft Copilot in order to improve scientific English writing. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

Funding

This work was funded by TNO (the Netherlands Organisation of Applied Scientific Research) as part of research program V2316 of the Netherlands Ministry of Defence. GreenTeg AG provided the CALERA Research Sensors used in this study.

CRediT authorship contribution statement

Timo van den Bogaard: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Visualization, Writing – original draft, Writing – review & editing. **Pooh Verheijen:** Investigation, Writing – review & editing. **Lisa Klous:** Conceptualization, Funding acquisition, Methodology, Supervision, Writing – review & editing. **Jan van Erp:** Conceptualization, Methodology, Supervision, Writing – review & editing. **Hein Daanen:** Conceptualization, Funding acquisition, Project administration, Resources, Supervision, Writing – review & editing.

Declaration of competing interest

No competing interests.

Acknowledgements

The authors thank Hero Noort, Sanne de Baas, Jorge Dantas, Guido van Campen, and Viggo Spanjerberg for their contribution to the measurements, the participants for their physical efforts, Nienke Haakma for her assistance in the lab, and Marco Hoozemans for his help with the statistical procedures in this study.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jtherbio.2026.104541.

Data availability

I have shared the link to the data in the paper.

References

Ajčević, M., Buoite Stella, A., Furlanis, G., Caruso, P., Naccarato, M., Accardo, A., Manganotti, P., 2022. A novel non-invasive thermometer for continuous core body temperature: Comparison with tympanic temperature in an acute stroke clinical setting. Sensors 22 (13). https://doi.org/10.3390/s22134760.

Alkemade, P., Gerrett, N., Eijsvogels, T.M.H., Daanen, H.A.M., 2021. Individual characteristics associated with the magnitude of heat acclimation adaptations. Eur. J. Appl. Physiol. 121 (6), 1593–1606. https://doi.org/10.1007/s00421-021-04626-3.

Ashworth, E.T., Cotter, J.D., Kilding, A.E., 2020. Methods for improving thermal tolerance in military personnel prior to deployment. Milit. Med. Res. 7 (1), 58. https://doi.org/10.1186/s40779-020-00287-z.

Bland, J.M., Altman, D.G., 1985. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1 (8476), 307–310. https://doi.org/10.1016/S0140-6736(86)90837-8.

Buller, M.J., Welles, A.P., Friedl, K.E., 2018. Wearable physiological monitoring for human thermal-work strain optimization. J. Appl. Physiol. 124 (2), 432–441. https://doi.org/10.1152/japplphysiol.00353.2017, 1985.

Daanen, H.A.M., Kohlen, V., Teunissen, L.P.J., 2023. Heat flux systems for body core temperature assessment during exercise. J. Therm. Biol. 112, 103480. https://doi.org/10.1016/j.jtherbio.2023.103480.

de Vet, H.C.W., Terwee, C.B., Mokkink, L.B., Knol, D.L., 2011. Measurement in Medicine. A Practical Guide. Cambridge University Press.

Dolson, C.M., Harlow, E.R., Phelan, D.M., Gabbett, T.J., Gaal, B., McMellen, C., Geletka, B.J., Calcei, J.G., Voos, J.E., Seshadri, D.R., 2022. Wearable sensor technology to predict core body temperature: a systematic review. Sensors (Basel) 22 (19). https://doi.org/10.3390/s22197639.

Du Bois, D., Du Bois, F., 1916. A formula to estimate the approximate surface area if height and weight be known. Arch. Intern. Med. 17, 863–871.

Esh, C.J., Carter, S., Galan-Lopez, N., Garrandes, F., Bermon, S., Adami, P.E., Racinais, S., James, L., Stellingwerff, T., Adams, W.M., Chrismas, B., Stevens, C.J., Périard, J.D., Brade, C., Henderson, M.J., Taylor, L., 2024. A review of elite athlete evidence-based knowledge and preparation for competing in the heat. Journal of Science in Sport and Exercise 6 (3), 218–237. https://doi.org/10.1007/s42978-024-00283-y.

Etienne, S., Oliveras, R., Schiboni, G., Durrer, L., Rochat, F., Eib, P., Zahner, M., Osthoff, M., Bassetti, S., Eckstein, J., 2023. Free-living core body temperature monitoring using a wrist-worn sensor after COVID-19 booster vaccination: a pilot study. Biomed. Eng. Online 22 (1). https://doi.org/10.1186/s12938-023-01081-3.

Goods, P.S.R., Maloney, P., Miller, J., Jennings, D., Fahey-Gilmour, J., Peeling, P., Galna, B., 2023. Concurrent validity of the CORE wearable sensor with BodyCap temperature pill to assess core body temperature during an elite women's field hockey heat training camp. Eur. J. Sport Sci. 23 (8), 1509–1517. https://doi.org/10.1080/17461391.2023.2193953.

Heat adaptation score. Retrieved 16/04/2025 from. https://help.corebodytemp.com/en/articles/10447117-heat-adaptation-score?srsltid=AfmBOorMcaq0_NKZJiJkieLJ9tqC2m6lCGXIEgVuPlFLinx3-BmmPMDk.

ISO, 2004. ISO9886:2004. In: *Ergonomics - Evaluation of Thermal Strain by Physiological Measurements* Geneva.

Januario, W.M., Lessa, N.F., Schittine, A.J.O., Prata, E., Marins, J.C.B., Natali, A.J., Wanner, S.P., Primola-Gomes, T.N., 2024. Validity and reproducibility of the CALERA research sensor to estimate core temperature at different intensities of a cycling exercise in the heat. J. Therm. Biol. 123, 103907. https://doi.org/10.1016/j.jtherbio.2024.103907.

Jolicoeur Desroches, A., Naulleau, C., Deshayes, T.A., Pancrate, T., Goulet, E.D.B., 2023. CORE wearable sensor: comparison against gastrointestinal temperature during cold water ingestion and a 5 km running time-trial. J. Therm. Biol. 115, 103622. https://doi.org/10.1016/j.jtherbio.2023.103622.

Kaltsatou, A., Anifanti, M., Flouris, A.D., Xiromerisiou, G., Kouidi, E., 2024. Validity of the CALERA research sensor to assess body core temperature during maximum exercise in patients with heart failure. Sensors (Basel) 24 (3). https://doi.org/10.3390/s24030807.

Kenefick, R.W., Cheuvront, S.N., 2012. Hydration for recreational sport and physical activity. Nutr. Rev. 70 (Suppl. 2), S137–S142. https://doi.org/10.1111/j.1753-4887.2012.00523.x.

Koo, T.K., Li, M.Y., 2016. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J. Chiropract. Med. 15 (2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012.

Kubota, N., Okada, K., Yamanaka, Y., 2025. Circadian phase assessment of core body temperature using a wearable temperature sensor under the real world. Sleep Sci. https://doi.org/10.1055/s-0044-1800785.

Lynch, G.P., Mavros, Y., Jay, O., 2025. Quantifying the time course of changes in maximal skin wettedness with 7 days of heat acclimation. J. Appl. Physiol. 138 (1), 136–143. https://doi.org/10.1152/japplphysiol.00919.2023, 1985.

McDonald, P., Brown, H.A., Topham, T.H., Kelly, M.K., Jardine, W.T., Carr, A., Sawka, M. N., Woodward, A.P., Clark, B., Periard, J.D., 2025. Influence of exercise heat acclimation protocol characteristics on adaptation kinetics: a quantitative review with Bayesian meta-regressions. Compr. Physiol. 15 (3), e70017. https://doi.org/10.1002/cph4.70017.

McLaughlin, B.C., Aguilera, J.T., D'Lugos, A.C., 2025. Validity of the CORE wearable sensor during constant-load cycling exercise in the heat. J. Therm. Biol. 132, 104241. https://doi.org/10.1016/j.jtherbio.2025.104241.

Mokkink, L., Terwee, C., de Vet, H., 2021. Key concepts in clinical epidemiology: responsiveness, the longitudinal aspect of validity. J. Clin. Epidemiol. 140, 159–162. https://doi.org/10.1016/j.jclinepi.2021.06.002.

Mokkink, L.B., Terwee, C.B., Patrick, D.L., Alonso, J., Stratford, P.W., Knol, D.L., Bouter, L.M., de Vet, H.C., 2010. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. J. Clin. Epidemiol. 63 (7), 737–745. https://doi.org/10.1016/j.jclinepi.2010.02.006.

Notley, S.R., Meade, R.D., Looney, D.P., Chapman, C.L., Potter, A.W., Fogarty, A., Howlader, T., Main, L.C., Friedl, K.E., Kenny, G.P., 2025. Physiological monitoring for occupational heat stress management: recent advancements and remaining challenges. Appl. Physiol. Nutr. Metab. 1 (50), 1–14. https://doi.org/10.1139/apnm-2024-0395.

Parsons, I.T., Stacey, M.J., Woods, D.R., 2019. Heat adaptation in military personnel: mitigating risk, maximizing performance. Front. Physiol. 10, 1485. https://doi.org/10.3389/fphys.2019.01485.

Periard, J.D., Eijsvogels, T.M.H., Daanen, H.A.M., 2021. Exercise under heat stress: thermoregulation, hydration, performance implications, and mitigation strategies. Physiol. Rev. 101 (4), 1873–1979. https://doi.org/10.1152/physrev.00038.2020.

Periard, J.D., Racinais, S., Sawka, M.N., 2015. Adaptations and mechanisms of human heat acclimation: applications for competitive athletes and sports. Scand. J. Med. Sci. Sports 25 (Suppl. 1), 20–38. https://doi.org/10.1111/sms.12408.

Periard, J.D., Travers, G.J.S., Racinais, S., Sawka, M.N., 2016. Cardiovascular adaptations supporting human exercise-heat acclimation. Auton. Neurosci. 196, 52–62. https://doi.org/10.1016/j.autneu.2016.02.002.

Priego-Quesada, J.I., Gil-Calvo, M., Mundel, T., Hartmann, T., Marino, F.E., 2025a. Research directions in thermoregulation during exercise in extreme environments. J. Therm. Biol. 127, 104031. https://doi.org/10.1016/j.jtherbio.2024.104031.

Priego-Quesada, J.I., MacKay, N., Adejuwon, D.C., Keir, D.A., 2025b. Effect of aerobic fitness on the validity of the calera research sensor to estimate core temperature during exercise. J. Therm. Biol. 127, 104067. https://doi.org/10.1016/j.jtherbio.2025.104067.

Richard, N.A., Cheung, S.S., Claydon, V.E., Koehle, M.S., Cote, A.T., 2025. Accuracy and precision of the SlateSafety BandV2 and CORE devices in estimating resting and moderate hyperthermic exercise temperature in eumenorrheic females. Int. J. Sports Physiol. Perform. 20 (8), 1068–1078. https://doi.org/10.1123/ijspp.2025-0034.

Taylor, N.A., Tipton, M.J., Kenny, G.P., 2014. Considerations for the measurement of core, skin and mean body temperatures. J. Therm. Biol. 46, 72–101. https://doi.org/10.1016/j.jtherbio.2014.10.006.

Thompson, W.R., 2010. Acsm's Guidelines for Exercise Testing and Prescription, 8 ed. Lippincott Williams and Wilkins.

Tyler, C.J., Reeve, T., Sieh, N., Cheung, S.S., 2024. Effects of heat adaptation on physiology, perception, and exercise performance in the heat: an updated meta-analysis. Journal of Science in Sport and Exercise. https://doi.org/10.1007/s42978-023-00263-8.

van den Bogaard, T., Klous, L., Cottle, R., van Erp, J., Daanen, H., 2025. The effect of heat acclimation on critical environmental limits and rate of rectal temperature change. J. Appl. Physiol. https://doi.org/10.1152/japplphysiol.01004.2024.

van Wely, M., 2024. Mixed models for repetitive measurements: the basics. Fertil. Steril. 121 (6), 914–917. https://doi.org/10.1016/j.fertnstert.2024.04.009.

Verdel, N., Podlogar, T., Ciuha, U., Holmberg, H.-C., Debevec, T., Supej, M., 2021. Reliability and validity of the CORE sensor to assess core body temperature during cycling exercise. Sensors 21 (17). https://doi.org/10.3390/s21175932.

Wartzek, T., Muhlsteff, J., Imhoff, M., 2011. Temperature measurement. Biomed. Tech. 56 (5), 241–257. https://doi.org/10.1515/BMT.2011.108.

Xu, H., Wang, H., Lei, T.H., Wang, F., 2025. Validity of the wearable CORE temperature sensor during 8-hour indoor heat exposure with and without electric fan use. Build. Environ. 282. https://doi.org/10.1016/j.buildenv.2025.113288.

Xu, X., Wu, G., Lian, Z., Xu, H., 2024. Feasibility analysis of applying non-invasive core body temperature measurement in sleep research. Energy Build. 303. https://doi.org/10.1016/j.enbuild.2023.113827.

Xu, Y., Ye, C., Ma, S., Gao, B., 2025. Four-week heat acclimation lowers carbohydrate oxidation of trained runners during submaximal exercise in the heat. Front. Physiol. 16, 1581594. https://doi.org/10.3389/fphys.2025.1581594.