Comparison of Rectal and Aural Core Body Temperature Thermometry in Hyperthermic, Exercising Individuals: A Meta-Analysis

Robert Huggins, MEd, ATC*; Neal Glaviano, MEd, ATC†; Naoki Negishi, MEd, ATC‡; Douglas J. Casa, PhD, ATC, FNATA, FACSM*; Jay Hertel, PhD, ATC, FNATA, FACSM§

*Korey Stringer Institute, Department of Kinesiology, University of Connecticut, Storrs; †Advanced Physical Therapy, Wolcott, CT; ‡University of Miami, Coral Gables, FL; §Kinesiology Program, University of Virginia, Charlottesville

Objective: To compare mean differences in core body temperature (Tcore) as assessed via rectal thermometry (Trec) and aural thermometry (Taur) in hyperthermic exercising individuals.

Data Sources: PubMed, Ovid MEDLINE, SPORTDiscus, CINAHL, and Cochrane Library in English from the earliest entry points to August 2009 using the search terms aural, core body temperature, core temperature, exercise, rectal temperature, thermistor, thermometer, thermometry, and tympanic.

Study Selection: Original research articles that met these criteria were included: (1) concurrent measurement of Trec and Taur in participants during exercise, (2) minimum mean temperature that reached 38°C by at least 1 technique during or after exercise, and (3) report of means, standard deviations, and sample sizes.

Data Extraction: Nine articles were included, and 3 independent reviewers scored these articles using the Physiotherapy Evidence Database (PEDro) scale (mean = 5.1 ± 0.4). Data were divided into time periods pre-exercise, during exercise (30 to 180 minutes), and postexercise, as well as Trec ranges <37.99°C, 38.00°C to 38.99°C, and >39.00°C. Means and standard deviations for both measurement techniques were provided at all time intervals reported. Meta-analysis was performed to determine pooled and weighted mean differences between Trec and Taur.

Data Synthesis: The Trec was conclusively higher than the Taur pre-exercise (mean difference [MD] = 0.27°C, 95% confidence interval [CI] = 0.15°C, 0.39°C), during exercise (MD = 0.96°C, 95% CI = 0.84°C, 1.08°C), and postexercise (MD = 0.71°C, 95% CI = 0.65°C, 0.78°C). As Trec measures increased, the magnitude of difference between the techniques also increased with an MD of 0.59°C (95% CI = 0.53°C, 0.65°C) when Trec was <38°C; 0.79°C (95% CI = 0.72°C, 0.86°C) when Trec was between 38.0°C and 38.99°C; and 1.72°C (95% CI = 1.54°C, 1.91°C) when Trec was >39.0°C.

Conclusions: The Trec was consistently greater than Taur when Tcore was measured in hyperthermic individuals before, during, and postexercise. As Tcore increased, Taur appeared to underestimate Tcore as determined by Trec. Clinicians should be aware of this critical difference in temperature magnitude between these measurement techniques when assessing Tcore in hyperthermic individuals during or postexercise.

Key Words: temperature assessment, tympanic membrane temperature, aural temperature, hyperthermia, exercise

Key Points

- Aural temperature underestimated core temperature as determined by rectal temperature.
- At the most extreme levels of hyperthermia, the relationship between rectal and aural temperatures is weakest.
- Recognizing this temperature difference is critical for clinicians assessing core temperature in hyperthermic individuals.

Since 1995, a total of 46 American football players have died from exertional heat stroke (EHS), 4 in 2010 alone.1 This number does not take into consideration many of the other activities that are potential settings for this condition, such as soccer, wrestling, and running activities (eg, cross-country, road races, marathons).2 The number may be even higher if we consider nonfatal cases of EHS, which often go unreported, and the misdiagnosis of many EHS-related deaths.3,4 In other cases, organ failure is listed as the cause of death when EHS is the actual cause or there is a misdiagnosis. Research3,4 supports an EHS incidence rate of 1 per 1000 participants who perform either athletic exercise or military activity in intense heat. Given this incidence of heat-related illnesses each year, certified athletic trainers and other sports medicine clinicians must be aware of the validity of different techniques to assess core body temperature (Tcore), so that this key vital sign is obtained accurately and subsequent medical decisions are appropriate.

To distinguish a nonfatal condition such as heat exhaustion from potentially fatal EHS, Tcore measurement is needed. An accurate Tcore assessment may help to rule in EHS5,6 over other emergency medical considerations, such as hyponatremia, cardiac emergency, diabetic emergency, head trauma, and exertional sickling, which can present with similar signs and symptoms. An athlete with a Tcore ≥40°C who exhibits signs and symptoms of central nervous system (CNS) dysfunction is most likely suffering from EHS and must be cooled immediately.5,6 Signs and symptoms of CNS dysfunction include but are not limited
to disorientation, altered consciousness, hypotension, tachycardia, increased respiration rate, vomiting, diarrhea, dehydration, coma, and convulsions.\textsuperscript{2,5} If cooling is not initiated promptly, physiologic complications may result: Of greatest concern are visceral organ failure and mortality. Therefore, a reliable and valid technique must be used to obtain accurate T\textsubscript{core} in exercising individuals. For proper diagnosis of EHS, T\textsubscript{core} must be greater than 40°C at the time of collapse and CNS dysfunction must be present.\textsuperscript{6} Although the patient may have a lucid initial period, both criteria must be met for proper diagnosis, making the temperature assessment even more critical.\textsuperscript{7} An accurate diagnosis is essential for appropriate treatment selection and duration. If the clinician is unaware of the starting core temperature, the treatment modality chosen may be less effective than necessary and the effective treatment time will be uncertain. Immediate cooling via ice-water immersion has shown a 100\% survival rate\textsuperscript{8,9}; however, before that treatment can be provided, the patient must first be properly diagnosed with EHS.\textsuperscript{7,10} The inherent problem is the observed inaccuracy of temperature devices used during exercise sessions in hyperthermic individuals.\textsuperscript{11,12}

To date, few reliable and valid methods for measuring T\textsubscript{core} in hyperthermic, exercising individuals have been demonstrated.\textsuperscript{11,12} Rectal, esophageal, and ingestible temperature assessment have the greatest validity in obtaining an individual’s true core temperature.\textsuperscript{11} Yet a plethora of other tools are used by health care professionals to assess T\textsubscript{core}, such as aural, axillary, forehead, oral, and temporal artery measurements, as well as others.\textsuperscript{11,12} These do not give accurate representations of actual T\textsubscript{core}, and few have been tested for validity in hyperthermic, exercising individuals.\textsuperscript{11-13}

Aural temperature (T\textsubscript{aur}) is commonly referred to as tympanic temperature, but this is a misnomer because the thermometer does not usually come into contact with the tympanic membrane. Devices that touch the tympanic membrane can cause extreme pain and damage the membrane (ie, perforation).\textsuperscript{14} Many aural devices actually only detect auditory canal temperature\textsuperscript{15} through infrared emission radiating from the tympanic membrane or auditory canal conduction via thermistors or thermocouples.\textsuperscript{14} Measurement of T\textsubscript{aur} is popular because it is quick, noninvasive, and simple, but its efficacy in actually measuring T\textsubscript{core} is debatable. In a clinical trial,\textsuperscript{15} T\textsubscript{aur} and T\textsubscript{re} in hospitalized patients undergoing surgical intervention showed a strong correlation. Other authors\textsuperscript{16-17} have examined whether T\textsubscript{aur} can predict T\textsubscript{core}, yet very few\textsuperscript{17-19} have demonstrated a strong correlation between T\textsubscript{aur} and T\textsubscript{re} in exercising individuals.\textsuperscript{16}

Rectal temperature assessment has encountered scrutiny and has a negative perception within sports medicine.\textsuperscript{13} It is viewed as an invasive and unnecessary technique. In a recent survey,\textsuperscript{13} certified athletic trainers were asked which T\textsubscript{core} assessment tool was most accurate and which tool they used. A total of 88\% (119/136) reported the rectal thermometer was the proper assessment tool, yet only 4\% (5/136) used T\textsubscript{re} for T\textsubscript{core} assessment.\textsuperscript{13} According to the position statements of both the American College of Sports Medicine\textsuperscript{10} and the National Athletic Trainers’ Association,\textsuperscript{7} it is vital to have an accurate T\textsubscript{core}, and use of a rectal thermometer is recommended. However, T\textsubscript{re} is not always used to diagnose an individual displaying a heat-related illness.

Our purpose was to perform a meta-analysis to compare T\textsubscript{aur} with T\textsubscript{re} core temperature measurement in hyperthermic individuals during exercise. We hypothesized that significant temperature differences between T\textsubscript{aur} and T\textsubscript{re} would occur when T\textsubscript{core} was assessed in hyperthermic, exercising individuals.

**METHODS**

**Data Sources**

We systematically reviewed the current literature to identify all studies that used T\textsubscript{re} and T\textsubscript{aur} to assess T\textsubscript{core} in hyperthermic individuals. PubMed, Ovid MEDLINE, SPORTDiscus, CINAHL, and Cochrane Library databases were searched from their earliest entry points to August 2009 using the search terms aural, core body temperature, core temperature, exercise, rectal, temperature, thermistor, thermometer, thermometry, and tympanic in various combinations. Previously identified articles, review articles, and reference lists of available studies were cross-referenced for possible articles that met the inclusion criteria. We limited our search to English-language research articles that involved human participants. A total of 2044 initial articles were identified via this search.

**Study Selection**

The specific inclusion criteria identified before data analysis included (1) hyperthermia \( \geq 38°C (100.4°F) \) during exercise, (2) simultaneous T\textsubscript{aur} and T\textsubscript{aur} mean and standard deviation measurements, and (3) original research studies with human participants. Studies in which hyperthermia was induced via passive methods were excluded to better reflect a traditional exercise setting. A minimum mean T\textsubscript{re} measurement of \( \geq 38°C (100.4°F) \) was chosen because athletes’ temperatures often exceed this value during exercise. Only measurements obtained via indwelling rectal thermistors were used to compare T\textsubscript{re} with T\textsubscript{aur} devices. Animal studies and case reports were not included in the data analysis.

The titles of 1981 studies were reviewed and excluded based on irrelevance. Abstracts of the remaining 63 identified studies were reviewed and 32 were excluded. Of the 31 remaining studies, 22 were removed because they either did not involve hyperthermic individuals or did not use T\textsubscript{re} or T\textsubscript{aur}. We manually cross-referenced the final 22 studies to look for other studies that should be included, but we did not find any additional relevant articles. This resulted in 9 articles for review\textsuperscript{2,4,11,12,14,16-19} (Figure 1). When articles lacked the necessary data for extraction, we obtained the data through personal communication with the authors.\textsuperscript{16-18}

**Quality Assessment**

All articles were rated using the Physiotherapy Evidence Database (PEDro) scale. The PEDro scale is a grading rubric that is widely used to determine the internal validity of clinical research studies. It consists of an 11-question, 10-point, yes or no questionnaire in which 1 point is awarded for each yes response and 0 points for each no.
response. Question 1 of the PEDro scale determines only if the study is eligible to be scored, whereas questions 2 through 11 are included in the point total. Questions 2 and 3 examine the extent of the random and concealed allocation of the participants, and questions 4 through 7 identify the level of blinding within the study design. Questions 8 through 11 grade the outcome measures reported and the statistical comparisons via point estimates and measures of variability.

Three authors independently reviewed the 9 articles with the PEDro scale. The PEDro scale has been commonly used in previous systematic reviews to assess the quality of included studies. If a discrepancy existed among the scores, the reviewers met to achieve a consensus and to ensure that one had not missed or misinterpreted an aspect of the study. Initial \( \kappa \) statistics revealed agreement of 1.000. Final PEDro scale scores for the 9 articles are shown in Table 1. Eight articles received a score of 5, and 1 article received a score of 6.

Except for 1 study that blinded the assessor of the aural reading to the rectal reading, none of the studies included blinded allocation of participants, assessors, or researchers and none randomly allocated participants to groups. A total of 5 studies were observational laboratory studies and 4 were observational field studies. All studies included in Table 1 obtained simultaneous \( T_{re} \) and \( T_{au} \) measurements. In 2 studies, \( T_{re} \) was compared with 9 other measurements, including \( T_{au} \). For these studies, only \( T_{au} \) and \( T_{re} \) data were extracted. One article contained data from 3 different studies, 2 of which (A and C) met our inclusion criteria. Study A could not be included because the standard deviation was not obtainable. Only Study C, an observational field study of individuals meeting heat stroke criteria in a medical tent at the end of a road race, was included.

To categorize the mean differences between \( T_{re} \) and \( T_{au} \) more clearly, the results from the included studies were placed into 1 of 3 groups based on the time of measurement: (1) pre-exercise, (2) during exercise, or (3) postexercise. Additionally, pooled weighted-mean averages were grouped based on \( T_{re} \) assessment ranges: (1) \(<38.0^\circ C\), (2) \(38.0^\circ C\) to \(38.99^\circ C\), or (3) \(>39^\circ C\).

**Data Extraction and Management**

The following information was extracted, when possible, from included studies: study design; number of participants; types of exercise; duration of exercise; \( T_{re} \) and \( T_{au} \) at pre-exercise, during exercise, and postexercise; environmental conditions; and manufacturers and models of the rectal and aural measurement devices. The primary outcome measure used by all studies was the comparison of \( T_{re} \) with \( T_{au} \).

In this analysis, we examined only temperature assessment via rectal thermistor and aural thermometer. None of the other variables included in the studies were analyzed. This allowed us to focus specifically on the differences between the temperature assessment methods in hyperthermic individuals during exercise.
**Table 1. Characteristics of the 9 Studies Included in Core Body Temperature Assessment Using Rectal and Aural Devices in Hyperthermic Exercising Individuals**

<table>
<thead>
<tr>
<th>Authors</th>
<th>Observational Design</th>
<th>Participants, N</th>
<th>Timing of Measurements</th>
<th>Types of Exercise and Settings</th>
<th>PEDro Scorea (Range, 1–10)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Armstrong et al14</td>
<td>Field study</td>
<td>5</td>
<td>Postexercise</td>
<td>11.5-km road race</td>
<td>5</td>
</tr>
<tr>
<td>Casa et al11</td>
<td>Field study</td>
<td>25</td>
<td>Pre-exercise; during exercise at 60, 120, 180 min; postexercise at 20, 40, 60 min</td>
<td>Field games: soccer, Ultimate Frisbee, etc</td>
<td>5</td>
</tr>
<tr>
<td>Coso et al14</td>
<td>Laboratory study</td>
<td>9</td>
<td>During exercise at 15 and 90 min</td>
<td>Cycle ergometer at 90 min @ 63% VO2max</td>
<td>5</td>
</tr>
<tr>
<td>Ganio et al12</td>
<td>Laboratory study</td>
<td>25</td>
<td>Pre-exercise; during exercise at 30, 60, 90 min; postexercise at 20, 40, 60 min</td>
<td>Treadmill at 4 m/s @ 10% grade</td>
<td>5</td>
</tr>
<tr>
<td>Hansen et al17</td>
<td>Laboratory study</td>
<td>8</td>
<td>During exercise (30°C, 40°C)</td>
<td>Treadmill at 5 km/min @ 6% grade</td>
<td>5</td>
</tr>
<tr>
<td>Hansen et al16</td>
<td>Field study</td>
<td>12</td>
<td>Postexercise</td>
<td>14-km road race</td>
<td>5</td>
</tr>
<tr>
<td>Hansen et al18</td>
<td>Laboratory study</td>
<td>7</td>
<td>During exercise at 50 min</td>
<td>Cycle ergometer at 50 min @ 180 W</td>
<td>5</td>
</tr>
<tr>
<td>Newsham et al2</td>
<td>Laboratory study</td>
<td>10</td>
<td>Pre-exercise; during exercise at 36.5 min, postexercise at 5, 10, 15, 20 min</td>
<td>Stair climber at self-selected intensity</td>
<td>5</td>
</tr>
<tr>
<td>Roth et al19</td>
<td>Field study</td>
<td>37</td>
<td>Postexercise</td>
<td>26.2-mile road race</td>
<td>6</td>
</tr>
</tbody>
</table>

Abbreviation: PEDro, Physiotherapy Evidence Database.

**Data Synthesis**

For each individual comparison, the mean difference (MD) and associated 95% confidence interval (CI) was calculated. Additionally, z scores or effect sizes (ESs) for each of the comparisons were calculated using the following equation: ES = (T<sub>re</sub> mean – T<sub>au</sub> mean)/(T<sub>au</sub> SD). Strength of the effect size was determined using the Cohen interpretation of effect size.20

Additionally, pooled weighted means from the data in each study were used to determine the pooled weighted MD pre-exercise, during exercise, and postexercise, as well as at T<sub>re</sub> ranges of <38.0°C, 38.0°C to 38.99°C, and >39°C. The MD (and 95% CI) between pooled weighted-mean T<sub>re</sub> and pooled weighted-mean T<sub>au</sub> was calculated for all comparisons, as were z scores and their CIs.

Combined, all included studies2,4,11,12,14,16–19 provided measurements for 138 physically active or highly trained participants. The selected studies used a variety of exercise types, intensities, durations, and environments. Five studies2,12,14,17,18 conducted exercise in a climate-controlled chamber, whereas 4 studies4,11,16,19 took place in outdoor settings. For the laboratory studies,2,12,14,17,18 researchers instructed participants to run or walk on a treadmill, use a stair climber, or ride on a cycle ergometer. For the 4 outdoor field studies,4,11,16,19 the participants ran on a field or road for different distances. The intensity of exercise was controlled in 4 laboratory studies,2,12,14,17,18 but it was inconsistent in the field studies.4,11,16,19 One laboratory study2 allowed self-selected intensity. Exercise duration ranged from 20 to 180 minutes (Table 2). All researchers used a digital thermometer except for one who used a glass mercury thermometer, and they used a digital aural thermometer in all studies to measure T<sub>core</sub>. Regardless of the specific exercise condition, 26 of the 28 measurements in all of the individual studies demonstrated a positive MD, with T<sub>re</sub> consistently being reported as higher than T<sub>au</sub> (Table 2).

**Baseline Measurements**

The authors of 5 studies reported baseline pre-exercise measurements,2,4,11,12,17 but the authors of 2 studies did not provide standard deviations.4,17 The other 3 studies2,11,12 showed positive MD values, with T<sub>re</sub> being consistently higher than T<sub>au</sub> at baseline (Table 2). Casa et al11 reported the highest MD at baseline in their field study (0.41°C, 95% CI = 0.25°C, 0.57°C); Newsham et al2 reported the lowest MD (0.04°C, 95% CI = −0.28°C, 0.36°C). Variability of the baseline measures between T<sub>re</sub> and T<sub>au</sub> pre-exercise across the 3 studies may indicate a small difference even at normal body temperature ranges. As shown in Table 2, the CIs of both Newsham et al2 and Ganio et al12 at baseline crossed zero into negative ranges, reflecting no difference, whereas the baseline point estimate and CI of Casa et al11 suggested a difference. When the baseline measures were corrected for variance, the pooled weighted MDs resulted in a difference of 0.27°C (95% CI = 0.15°C, 0.39°C), with T<sub>re</sub> higher than T<sub>au</sub>. Clinically, this difference is small and insignificant when differentiating between safe levels of hyperthermia.

**Measurements During Exercise**

As exercise duration increased, both T<sub>re</sub> and T<sub>au</sub> increased. Exercise durations as short as 15 minutes14 were associated with an MD of 0.80°C between T<sub>re</sub> and T<sub>au</sub>. When the exercise duration reached 60 minutes or longer, the MD ranged from 0.55°C to 1.70°C. Five studies11,12,14,17,18 reported positive MDs, indicating that T<sub>re</sub> was higher than T<sub>au</sub> (Figure 2). Casa et al14 demonstrated the highest MD in their field study at 90 minutes of exercise (1.70°C, 95% CI = 1.45°C, 1.95°C). Two groups reported negative MDs: Hansen et al17 (−0.30°C, 95% CI = −0.63°C, 0.03°C) in their 30°C ambient heat chamber laboratory study at 120 minutes of exercise and Newsham et al2 (−0.33°C, 95% CI = −0.67°C, 0.01°C) after 36.5 minutes of exercise. Pooled
results for the weighted measures during exercise indicated that $T_{re}$ was 0.96°C (95% CI = 0.84°C, 1.08°C) higher than $T_{au}$ (Figure 2).

### Measurements Postexercise

Six groups reported postexercise temperatures.$^{2,4,11,12,16,19}$ As the rest period increased, the reported $T_{re}$ and $T_{au}$ both gradually decreased (Figure 3). Armstrong et al.$^4$ demonstrated the highest MD in their field study immediately after an 11.5-km race (3.90°C, 95% CI = 2.60°C, 5.20°C). Pooled results for the postexercise measures indicated that $T_{re}$ was 0.71°C (95% CI = 0.65°C, 0.78°C) higher than $T_{au}$ (Figure 3). Furthermore, data were extrapolated to reflect the pooled results for MD between $T_{re}$ and $T_{au}$ measurements at different $T_{re}$ ranges: <38.0°C, 38.0°C to 38.99°C, and >39°C. As $T_{re}$ measures increased, the magnitude of difference between the techniques also increased, with an MD of 0.59°C (95% CI = 0.53°C, 0.65°C) when $T_{re}$ <38.0°C; 0.79°C (95% CI = 0.72°C, 0.86°C) when $T_{re}$ was between 38.0°C and 38.99°C; and 1.72°C (95% CI = 1.54°C, 1.91°C) when $T_{re}$ > 39.0°C (Figure 4).

### DISCUSSION

The results from our pooled data for the systematic review and meta-analysis indicate that $T_{re}$ was consistently greater than $T_{au}$ when $T_{core}$ was measured in hyperthermic, exercising individuals. In 2 of the included studies,$^{17}$ where $T_{au}$ was greater than $T_{re}$ during exercise, possible reasons as to why $T_{au}$ was greater than $T_{re}$ measures could be selective cooling of the CNS by the hypothalamus, a lag in $T_{re}$ due to decreased blood flow to the rectum during exercise, or the type of sensor used in the specific aural device (eg, infrared, thermistor, or thermocouple). The data from 1 study$^{17}$ support the idea that $T_{re}$ continues to rise after exercise stops and that the $T_{re}$ device may not have given accurate readings due to decreased blood flow.

As $T_{re}$ increases with exercise, the discrepancy between $T_{au}$ and $T_{re}$ increases. Measuring $T_{re}$ is the method recommended by the National Athletic Trainers’ Association and the American College of Sports Medicine for assessing $T_{core}$. Clinicians should be aware of the critical difference in temperature magnitude between these measurement techniques when assessing $T_{core}$ as determined by $T_{re}$ assessment in hyperthermic individuals during or postexercise. Based on the results of the meta-analysis, as $T_{core}$ increased based on $T_{re}$, the difference between $T_{re}$ and $T_{au}$ also increased (Figure 4). In other words, the more hyperthermic an exercising individual was, the greater the discrepancy between the measures. This finding is of particular importance for those clinicians who use $T_{au}$ devices and believe that adding 1°C or any other correction factor can account for the difference in $T_{core}$ as determined by rectal thermometry. This misconception may result in a crucial error in the care of a patient with exercise-induced hyperthermia, especially at higher $T_{core}$ levels.
At rest, an individual's core temperature should be less than the common resting temperature of 38.0°C. This was true at all baseline measurements, and the difference between $T_{re}$ and $T_{au}$ was approximately 0.5°C. This difference was consistent with previous systematic reviews\textsuperscript{24} that examined $T_{re}$ and $T_{au}$ in individuals at rest. Thus, clinicians may consider $T_{au}$ an accurate measure when compared with $T_{re}$ during the general medical assessment of nonexercising patients. Our meta-analysis specifically examined studies of hyperthermic exercising individuals, and we determined that as an individual's temperature increased to the 38°C to 38.99°C range via rectal thermometry, the difference between $T_{re}$ and $T_{au}$ was approximately 0.8°C. The greatest difference between devices was identified when an individual's $T_{re}$ was greater than 39°C, with a difference of 1.7°C. If an individual has a core temperature of 40°C as assessed by rectal thermometry, an aural device should theoretically display a temperature of 38.3°C. This point is of vital importance to medical personnel evaluating an individual who might have hyperthermia and not a potentially fatal case of EHS. Furthermore, an individual may present with an initial lucid period, which is why accurate $T_{core}$ assessment is so necessary.\textsuperscript{7,10} If the temperature difference between devices increases as an individual’s core temperature increases, the clinical consequences are very concerning once the true $T_{core}$ is greater than 40°C.\textsuperscript{5,7,10} This possibility cannot be examined via clinical studies because institutional review boards understand the danger of a $T_{core}$ at that high level.

### Core Body-Temperature Assessment Devices

As assessed by medical professionals, true $T_{core}$ often refers to pulmonary artery or esophageal temperature. Devices to obtain these measurements are considered the criterion standard, yet their clinical use is limited for the athletic trainer in the on-field assessment of a patient with a heat-related illness. Based on this practical restriction, other devices (eg, rectal, ingestible, aural, oral, and infrared thermometry) were developed. These devices only partially addressed the invasiveness concern, but their inherent limitation is that they measure only the anatomical region in which they are placed. Some have demonstrated more validity than others,\textsuperscript{11,12,34,36} yet even with measurement differences of more than 1°C during exercise, their use continues due to their simple operation and reduced cost.

We recognize that $T_{re}$ is a measure from the rectum and $T_{au}$ from the aural canal and that differences exist based on anatomical location. Within the literature, the measurements have yet to be compared in a systematic fashion in athletes exercising in the heat, which is key to the practicing clinician.

### Pros and Cons of $T_{re}$ and $T_{au}$

Previous researchers\textsuperscript{2,4,15,24,25,29,33} have stated the advantages and disadvantages to $T_{re}$ and $T_{au}$. Considered invasive by some, $T_{re}$ creates a privacy concern, and patient cooperation may be lacking.\textsuperscript{2} In 1 report,\textsuperscript{25} $T_{re}$ had a temperature lag during rapid changes in core temperature.
Figure 3. Mean difference (±SD) between rectal ($T_{re}$) and aural ($T_{au}$) temperatures after exercise in hyperthermic individuals. Pooled weighted mean difference reflects all included articles. * Pooled weighted mean ($T_{re} - T_{au}$) = 0.71°C.

Figure 4. Pooled weighted mean difference between rectal ($T_{re}$) and aural ($T_{au}$) temperatures at various temperature ranges based on rectal temperatures for studies conducted in a climate chamber (indoor), field setting (outdoor), and all studies combined.
due to the lack of blood supply to the rectum. When compared with pulmonary artery measurement,\textsuperscript{25,26} \( T_{re} \) accurately predicts \( T_{core} \) and may be a better predictor of EHS due to its ability to detect systemic heat stress and potential organ damage. Rectal temperature has also been described as a valid device to measure \( T_{core} \).\textsuperscript{27–29} Research has supported \( T_{au} \) as providing an accurate assessment of an individual’s \( T_{core} \).\textsuperscript{30,31} Advantages included its ease of use and noninvasive nature and patient compliance. However, these authors examined the devices in nonhyperthermic exercising individuals.\textsuperscript{30,31} Most of the research was conducted in pediatric and hospital patients, whose temperatures were usually lower or indicated a slight fever.\textsuperscript{16,32} In addition, environmental factors such as fanning one’s face, sweating rates, metabolic rates, and exposure to wind or colder air can influence the device’s accuracy.\textsuperscript{16,28} Also, \( T_{au} \) underestimates \( T_{core} \) and does not reflect changes during exercise.\textsuperscript{33}

**The Importance of a Valid \( T_{core} \) Measurement**

It is imperative to use an accurate and reliable method to assess \( T_{core} \) in hyperthermic individuals suspected of having EHS. To diagnose an individual with EHS, 2 specific criteria must be met: \( T_{core} \geq 40^\circ C \) and CNS dysfunction. An improper assessment of \( T_{core} \) can lead to misdiagnosis of EHS and possible fatality. A vital distinguishing criterion between nonfatal heat exhaustion and the medical emergency of EHS is \( T_{core} \geq 40^\circ C \). The position statements of both the National Athletic Trainers’ Association\textsuperscript{7} and American College of Sports Medicine\textsuperscript{10} recommend obtaining \( T_{re} \) when assessing \( T_{core} \).

The validity of devices that assess \( T_{au} \) as an estimate of \( T_{core} \) must be questioned. In exercising, hyperthermic individuals, \( T_{re} \) was consistently higher than \( T_{au} \). Clinicians should be aware of discrepancies when assessing \( T_{core} \) in hyperthermic individuals during or postexercise with various devices. The \( T_{au} \) devices are not in direct contact with the tympanic membrane; therefore, they may not assess true \( T_{core} \) but \( T_{au} \). Although some device manufacturers take this information into consideration and add a correction factor, those corrections are not clinically validated when assessing temperatures near 40°C. In addition, the correction factor differs with each device. Using an aural device and adding a correction value does not always provide a valid \( T_{core} \) measurement. Furthermore, manufacturers that recommend correction factors do not adjust for the growing disparity between \( T_{au} \) and \( T_{re} \) as \( T_{core} \) increases. Aural devices vary in their technological designs, which may be a source of inaccuracy when measuring \( T_{core} \) in hyperthermic patients. If reliable alternative aural thermometers that are in direct contact with the tympanic membrane (without causing damage) are developed, tympanic core temperature may accurately assess elevated \( T_{core} \). Future research in the validity and reliability of \( T_{au} \) devices is needed.

The validity of alternative \( T_{core} \) measurements has been tested in hyperthermic patients.\textsuperscript{11,12} The ingestible telemetric pill shows promise in the assessment of \( T_{core} \) due to its noninvasive nature and continuous monitoring abilities.\textsuperscript{11,12,34–39} but 3 aspects of this device must be considered. First, the CorTemp (HQ, Inc, Palmetto, FL) ingestible \( T_{core} \) sensor costs approximately $30.00, and the CorTemp data recorder device costs $2000.00. Although many levels (eg, youth, high school, and some college) of sport programs may not be able to accommodate the cost of ingestible pills, the athletes’ health and safety is a priority, and these devices should be used whenever possible. Second, the ingestible pill can take 1 to 8 hours to enter the lower gastrointestinal tract, and at that time, it must be in a location where neither food nor liquid is present.\textsuperscript{11} However, recent original research\textsuperscript{40} has shown no difference in pill effectiveness when ingested at 24 hours and 40 minutes pre-exercise. Third, younger patients and those who have difficulty swallowing pills can find ingesting the device challenging. The ingestible pill can be used as a suppository, but then invasiveness becomes a concern. However, it may be difficult to use the ingestible pill in a patient with an emergent heat-related illness.

**Limitations**

A limitation of the included studies was that different \( T_{au} \) devices were used: Thermoscan Pro 1 tympanic thermometer (Braun, South Boston, MA) by Newsham et al,\textsuperscript{2} Thermoscan ExacTemp (model IRT 4520; Braun) by Casa et al\textsuperscript{11} and Ganio et al,\textsuperscript{12} First Temp (model 2000A; Intelligent Medical Systems, Inc, Carlsbad, CA) by Hansen et al,\textsuperscript{16,18} Genius 300A (Intelligent Medical Systems, Inc) by Roth et al,\textsuperscript{19} Ottotemp (model 3000; Exgeren Corporation, Newton, MA) by Armstrong et al,\textsuperscript{4} and YSI-402 (YSI Inc, Yellow Springs, OH) by Coso et al.\textsuperscript{14} One group\textsuperscript{17} did not report the device used. Different \( T_{re} \) devices were also used: Blanketrol Hypo-Hyper Temperature Control Unit (Cincinnati Sub-Zero, Cincinnati, OH) by Newsham et al,\textsuperscript{2} YSI-401 (YSI Inc) by 5 groups,\textsuperscript{4,11,12,14,18} and IVAC model 2080A (IVAC Corporation, Naperville, IL) by Roth et al.\textsuperscript{19} Hansen et al\textsuperscript{17} omitted the model and manufacturer of their rectal device. Six groups\textsuperscript{11,12,14,16,17,19} specifically stated that their rectal devices had been calibrated before data collection, but the remaining authors did not mention calibration of the devices. We did not include a threshold for the reported reliability estimates of each measurement in the included studies.

Another limitation to the study was the use of the 10-point PEDro scale to determine the internal validity of our articles. A score of 10 signifies a high-quality study. All of our articles achieved scores of 5 or 6, which are not ideal scores for study quality. The PEDro scale is often used to determine the internal validity for randomized controlled trials. The articles included in this study had low scores due to the absence of blinding and random allocation. However, all the studies were observational, and it is extremely difficult to blind a patient in such a study from the interventions. The PEDro scale may not be the ideal reflection of study quality for these articles, but we are unaware of a better scoring tool that has been validated in the literature.

The participants in the study may also be considered a limitation, but they should not affect the validity of a device used to track temperature changes. Participants ranged from the physically active within the general population\textsuperscript{2,12} to marathon runners\textsuperscript{4,16} to moderately trained college students.\textsuperscript{14} This wide range is associated with a great deal of variance in exercise training and heat acclimation. Being well trained or acclimated to the heat.
allows the body to more efficiently dissipate heat and cool itself, which may have affected the data. Coso et al.\textsuperscript{14} openly stated that their participants were not heat acclimated.

Study design and mode of exercise also varied among studies. Casa et al.\textsuperscript{11} and Hansen et al.\textsuperscript{16} performed observational field studies that required participants to run at various speeds for 65 to 180 minutes, whereas Ganio et al.\textsuperscript{12} Hansen et al.\textsuperscript{17,18} and Newsham et al.\textsuperscript{2} performed observational laboratory studies within an environmental heat chamber. Hansen et al. studied members of the Australian Army for 100 minutes of exercise in 1 study\textsuperscript{17} and healthy adults in another.\textsuperscript{18} Newsham et al.\textsuperscript{2} investigated physically active individuals on a stair climber at self-selected intensities for an average of 36.5 minutes, and Ganio et al.\textsuperscript{12} tested physically active individuals on a treadmill for 90 minutes. Roth et al.\textsuperscript{19} and Armstrong et al.\textsuperscript{4} measured $T_{au}$ and $T_{re}$ in runners at the end of a road race. Armstrong et al.\textsuperscript{4} assessed core temperature after an 11.5-km foot race in participants who displayed CNS dysfunction and were suspected of having EHS. Roth et al.\textsuperscript{19} tested patients admitted to the field hospital at the end of a marathon. Although we note these discrepancies in the modes of exercise, we primarily compared $T_{au}$ and $T_{re}$. In order to properly perform this comparison, we could not exclude the studies based on mode of exercise.

The $T_{core}$ of the hyperthermic individuals also differed. For the purpose of this review, participants needed to reach a mean $T_{core}$ of 38°C to be considered hyperthermic. In some studies,\textsuperscript{2,11,12,16,14,17–19} a mean of 38°C was reached; another group\textsuperscript{4} recorded values as high as 41.7°C. Safety precautions played a role in limiting the surpassing of a $T_{core}$ of 40°C, as in the Casa et al.\textsuperscript{11} and Ganio et al.\textsuperscript{12} studies, whose institutional review board required participants to cease exercise if they reached a $T_{core}$ of 40°C.\textsuperscript{11,12} The advantage of the investigations by Armstrong et al.\textsuperscript{4} and Roth et al.\textsuperscript{19} was that the data were collected in runners immediately after the completion of a road race and, thus, the $T_{core}$ was greater than in the other studies. Notably, Armstrong et al.\textsuperscript{4} also found the greatest mean difference (3.9°C) between devices immediately after the race.

The fact that we drew our conclusions from 9 studies may be considered a limitation. The 28 total data points used may be fewer than in previous meta-analyses, but they were the only values available at the time of the search. More data on $T_{au}$ and $T_{re}$ may be needed in the future to support or strengthen the conclusions found in this meta-analysis.

**CONCLUSIONS**

In conclusion, $T_{re}$ measures were consistently greater than $T_{au}$ measures when assessing $T_{core}$ in hyperthermic, exercising individuals. As determined by $T_{re}$, $T_{au}$ appeared to underestimate $T_{core}$. Clinicians should be aware of this critical difference in temperature magnitude between these measurement techniques when assessing $T_{core}$ in hyperthermic individuals during or after exercise. If the device used does not provide an accurate, valid, and precise measurement of an individual’s $T_{core}$, the diagnosis of EHS may be delayed. To date, a minimal amount of research has compared $T_{re}$ with $T_{au}$ in hyperthermic, exercising individuals. Based on the data we examined, we strongly support using $T_{re}$ as the temperature assessment method for hyperthermic, exercising individuals.

**ACKNOWLEDGMENTS**

We thank Ronald N. Roth, MD; Katherine R. Newsham, PhD, ATC; and Lawrence E. Armstrong, PhD, FACSM, for their interest in and support of this project by devoting time and effort to look for additional raw data, effect sizes, and other information in their investigations to help improve and strengthen this research article.

**REFERENCES**


Address correspondence to Jay Hertel, PhD, ATC, FNATA, FACSM, Kinesiology Program, University of Virginia, 210 Emmet Street South, PO Box 400407, Charlottesville, VA 22904-4407. Address e-mail to jhertel@virginia.edu.

Figure 3 was incorrect. Please see the correct Figure 3 below. We regret the error.

**Figure 3.** Mean difference (±SD) between rectal (Tre) and aural (Tau) temperatures after exercise in hyperthermic individuals. Pooled weighted mean difference reflects all included articles. * Pooled weighted mean (Tre − Tau) = 0.71°C.