The Systematic Bias of Ingestible Core Temperature Sensors Requires a Correction by Linear Regression

Andrew P. Hunt[1]*, Aaron J. E. Bach[1], David N. Borg[1], Joseph T. Costello[2] and Ian B. Stewart[1]

[1] Faculty of Health, School of Exercise and Nutrition Sciences, Institute of Health and Biomedical Innovation, Queensland University of Technology, Brisbane, QLD, Australia, [2] Extreme Environments Laboratory, Department of Sport and Exercise Science, University of Portsmouth, Portsmouth, UK

An accurate measure of core body temperature is critical for monitoring individuals, groups and teams undertaking physical activity in situations of high heat stress or prolonged cold exposure. This study examined the range in systematic bias of ingestible temperature sensors compared to a certified and traceable reference thermometer. A total of 119 ingestible temperature sensors were immersed in a circulated water bath at five water temperatures (TEMP A: $35.12 \pm 0.60°C$, TEMP B: $37.33 \pm 0.56°C$, TEMP C: $39.48 \pm 0.73°C$, TEMP D: $41.58 \pm 0.97°C$, and TEMP E: $43.47 \pm 1.07°C$) along with a certified traceable reference thermometer. Thirteen sensors (10.9%) demonstrated a systematic bias $> \pm 0.1°C$, of which 4 (3.3%) were $> \pm 0.5°C$. Limits of agreement (95%) indicated that systematic bias would likely fall in the range of -0.14 to $0.26°C$, highlighting that it is possible for temperatures measured between sensors to differ by more than $0.4°C$. The proportion of sensors with systematic bias $> \pm 0.1°C$ (10.9%) confirms that ingestible temperature sensors require correction to ensure their accuracy. An individualized linear correction achieved a mean systematic bias of $0.00°C$, and limits of agreement (95%) to 0.00–$0.00°C$, with 100% of sensors achieving $\pm 0.1°C$ accuracy. Alternatively, a generalized linear function (Corrected Temperature ($°C$) $= 1.00375 \times$ Sensor Temperature ($°C$) $- 0.205549$), produced as the average slope and intercept of a sub-set of 51 sensors and excluding sensors with accuracy outside $\pm 0.5°C$, reduced the systematic bias to $< \pm 0.1°C$ in 98.4% of the remaining sensors ($n = 64$). In conclusion, these data show that using an uncalibrated ingestible temperature sensor may provide inaccurate data that still appears to be statistically, physiologically, and clinically meaningful. Correction of sensor temperature to a reference thermometer by linear function eliminates this systematic bias (individualized functions) or ensures systematic bias is within $\pm 0.1°C$ in 98% of the sensors (generalized function).

Keywords: gastrointestinal temperature, reliability, validity, ingestible sensor, measurement error, heat, cold, thermoregulation

INTRODUCTION

The human body's capacity to regulate its internal temperature ensures optimal health and physiological function when exposed to a wide range of ambient environments (Kenney et al., 2004). In environments that are conducive to heat stress (high ambient temperature, humidity, radiant heat sources, and low air movement) increased sweating and skin blood flow facilitate thermoregulation. However, performing high intensity physical activities combined with wearing protective clothing can exacerbate and overwhelm the capacity to maintain thermal homeostasis (Montain et al., 1994). Such exposures are commonplace in sports [e.g., the American National Football League (Armstrong et al., 2010) and soccer (Taylor L. et al., 2014)], in occupational settings [e.g., fire fighters (Larsen et al., 2015; Walker et al., 2015) and other emergency first responders (Costello et al., 2015)], and in military operations (Hunt et al., 2016). Consequently, internal body temperatures can rise to dangerous extremes resulting in fatal heat stroke if the elevation in body temperature is left unchecked (Carter et al., 2005; Grundstein et al., 2012). Conversely, if individuals are exposed to extreme cold environments, particularly for a long duration of time, hypothermia becomes a serious risk to health and performance (Castellani and Tipton, 2015; Brazaitis et al., 2016). For this reason an accurate measure of core body temperature is critical to the development of safe guidelines for individuals, groups and teams undertaking physical activity in situations of prolonged exposure to extreme environments.

Ingestible temperature sensors have become a valuable and commonly used tool for monitoring core body temperature outside the clinical context (Taylor N. A. S. et al., 2014), particularly in athletic and occupational settings that require freedom of movement (Stewart and Hunt, 2011; Hunt et al., 2016). Comparisons of ingested temperature sensors with the clinical measure of rectal temperature have concluded that ingested sensors provide a valid measure of core body temperature during exercise in hot environments both indoors and outdoors (Casa et al., 2007; Ganio et al., 2009), as well as indoors in cold environments (Bagley et al., 2011). In addition to concluding that ingestible sensors are a valid surrogate of rectal temperature and oesophageal temperature, it has been recommended that the ingestible sensors are corrected to align with a certified thermometer (Byrne and Lim, 2007). To ensure optimal measurement accuracy, research laboratories are advised to assess the validity and reliability of temperature measurement devices by immersion in a sterile water bath at several temperatures in the physiological range (Simpson et al., 2006). A linear correction is then applied to align the measurements with a certified and traceable reference thermometer. Calibration of thermometry devices at regular intervals has been shown to improve accuracy of aging instruments and ensures their sensitivity and specificity is preserved for the diagnosis of various conditions and pathologies (Simpson et al., 2006).

Although ingestible temperature sensors are becoming increasingly utilized for both laboratory and field investigations of human thermoregulation, few researchers report performing calibrations. A brief review of several leading journals in the areas of exercise physiology, sports medicine, and applied ergonomics (including Frontiers in Physiology, Medicine and Science in Sports and Exercise, Journal of Applied Physiology, and Applied Ergonomics) revealed 52 papers published between 2006 and 2016 reported temperature from an ingested sensor as a primary outcome measure. However, only eight studies (14%) reported performing a calibration of the temperature sensors prior to ingestion (Byrne et al., 2006; Gant et al., 2006; Mermier et al., 2006; Goosey-Tolfrey et al., 2008; Wilkinson et al., 2008; Logan-Sprenger et al., 2012; Pyke et al., 2015; Stewart et al., 2017). The absence of a temperature correction procedure potentially introduces error into the measurement for two reasons. Firstly, it remains unclear if the manufacturers claimed level of accuracy is in fact achieved by the device, and secondly, the level of agreement between ingestible sensors cannot be quantified.

A pilot study has demonstrated the importance of conducting a correction procedure by exposing these sources of error on a limited sample of sensors (Hunt and Stewart, 2008). Manufacturer claims indicate an accuracy of $\pm 0.1°C$ (HQInc.[1] and VitalSense[2], which coincides with the recommended accuracy for measurement of internal body temperature (Moran and Mendal, 2002), yet one out of three ingestible sensors recorded temperature outside this range (Hunt and Stewart, 2008). Furthermore, a statistically significant discrepancy ($F = 10.818$, $p < 0.001$) was reported between the sensors (Hunt and Stewart, 2008). The pilot study recommended that corrections be performed by a linear regression developed between the ingested sensor and a certified traceable thermometer in a water bath heated to a minimum of four discrete temperatures in the physiological range. However, the study was limited by a small sample size of only three sensors. Consequently, the small sample size leaves it unclear if the one sensor that recorded temperature outside the $\pm 0.1°C$ standard was a simple anomaly, or a true representation of the proportion of sensors that might be expected in a larger sample. Therefore, the aims of this study were to examine the range in systematic bias in a large sample of ingestible temperatures sensors than previously reported and to evaluate the linear regression approach to temperature correction.

METHODS

A total of 119 ingestible temperature sensors (CorTemp, HQinc, Palmetto FL, USA) were evaluated as part of this investigation. All sensors where tested within their individual expiration dates. Each sensor was immersed in a circulated water bath (model 350, Contherm Scientific Inc., New Zealand) at five water temperatures that encompassed the physiological range of human body core temperature anticipated during physical activities. The average water bath conditions were: TEMP A: $35.12 \pm 0.60°C$, TEMP B: $37.33 \pm 0.56°C$, TEMP C: $39.48 \pm 0.73°C$, TEMP D: $41.58 \pm 0.97°C$, and TEMP E: $43.47 \pm 1.07°C$. A manufacturer calibrated and traceable thermometer

---

[1]http://www.hqinc.net/

[2]https://bmedical.com.au/product/core-body-temperature-capsule-ingestable-jonah/

(TL-1W, ThermoProbe, USA) with certified accuracy of ±0.06°C was suspended in the water bath and served as the reference measure of water temperature. Once the water temperature was considered stable (±0.05°C) for at least 5 min and the sensors had been submerged for >4 min to allow equilibration with the water temperature (Hunt and Stewart, 2008), sensor temperature was recorded at 10 s intervals and averaged over a 1 min period. The temperature variation in the water bath was <0.01°C during recording; a level of variation comparable to similar studies (Harper Smith et al., 2010).

Bland-Altman plots with 95% limits of agreement analysis, accounting for the repeated measurement conditions with each sensor, were performed to describe the agreement between the ingestible sensors and references temperatures (Bland and Altman, 2007). Repeated measures analysis of variance with pairwise comparisons was performed to evaluate any effect of water bath temperature on systematic bias, with significance accepted at α < 0.05. Sensor calibrations were performed by linear regression; Firstly as an individualized linear regression specific to each individual sensor. To examine the potential to utilize a generalized linear regression to calibrate ingestible sensors, the sensors were randomly allocated into development ($n = 52$) and verification ($n = 67$) groups, whereby the generalized linear regression was generated as the average slope and intercept of the development sensors, and applied to the verification sensors. Finally, ingestible sensors with a systematic bias > ±0.5°C were removed from the analysis ($n = 4$) and the generalized linear regression was calculated on the remaining development sensors ($n = 51$) and applied to the remaining verification sensors ($n = 64$).

RESULTS

A total of 71 (59.7%) sensors demonstrated a systematic bias greater than the measurement accuracy of the reference thermometer (±0.06°C). Thirteen sensors (10.9%) demonstrated a systematic bias > ±0.1°C, of which 4 (3.3%) were > ±0.5°C. Overall, sensors tended to overestimate the reference thermometer, with systematic bias averaging 0.06 ± 0.24°C (**Figure 1**). Limits of agreement (95%) indicated that systematic bias would likely fall in the range of −0.14 to 0.26°C, highlighting that it is possible for temperatures measured between sensors to differ by 0.40°C. However, the extremes of systematic bias of individual sensors ranged from −1.35 to 2.00°C. The systematic bias was significantly different across the range of water bath temperatures (TEMP A: 0.07 ± 0.23; TEMP B: 0.06 ± 0.24; TEMP C: 0.05 ± 0.24; TEMP D: 0.05 ± 0.24; TEMP E: 0.05 ± 0.24; F = 14.199, $P < 0.001$), with Temp A significantly different to all other conditions ($P < 0.001$).

A strong linear relationship was observed between the reference and sensor temperatures, with a correlation coefficients (r) of 1.00 consistently for each individual ingestible sensor. Due to the strength of the relationship it was accepted that linear regression was a suitable method for correction of the sensor temperature to the reference temperature. An individualized linear regression, specific to each ingestible sensor, corrected the
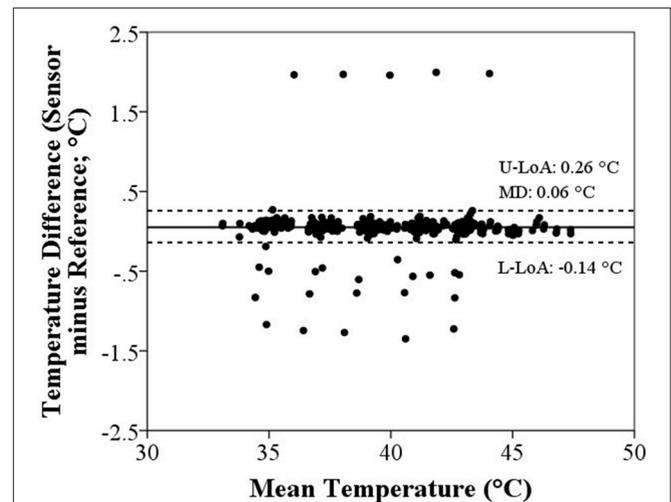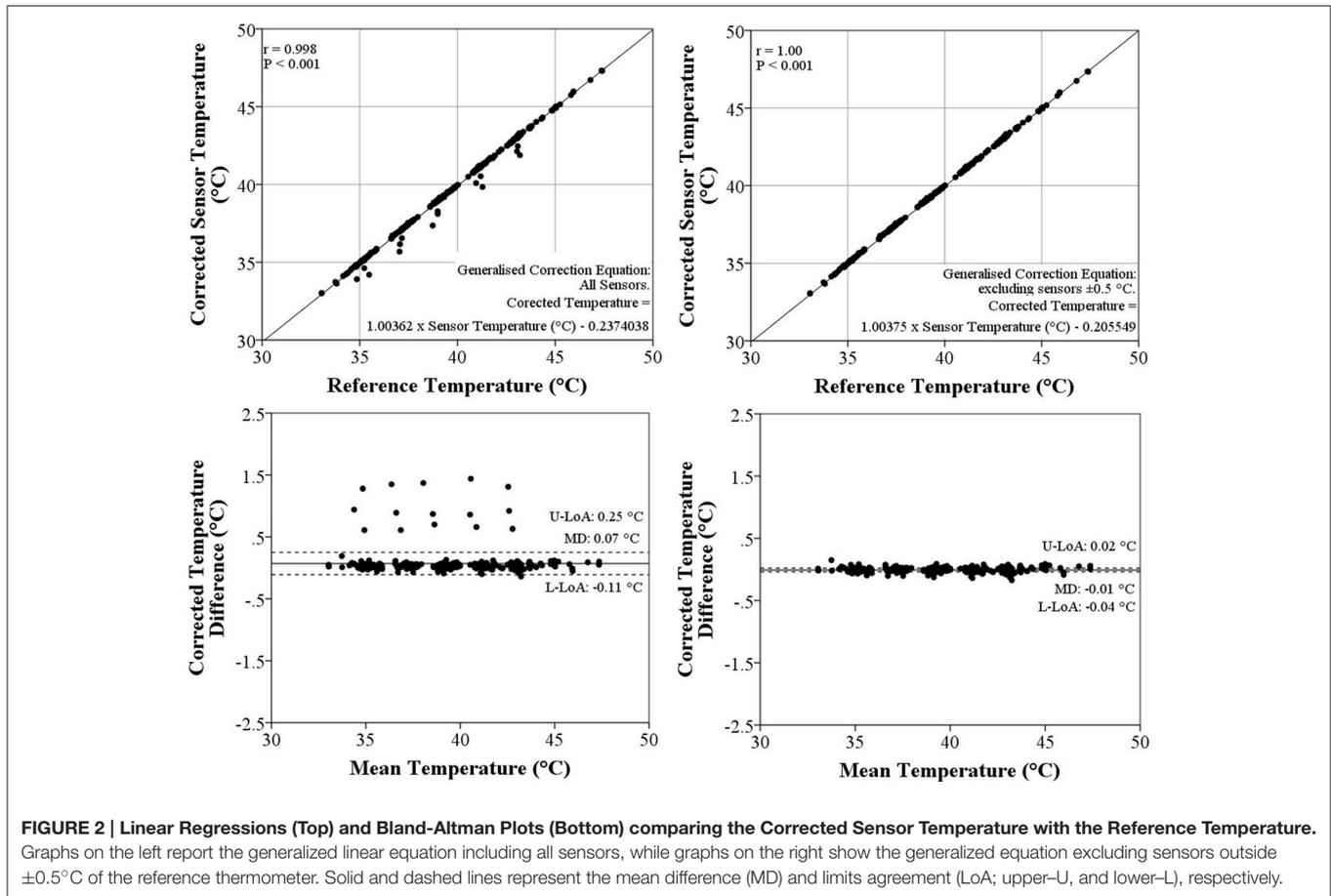

FIGURE 1 | Bland-Altman plot of the agreement between ingestible sensor and reference temperatures. Solid and dashed lines represent the mean difference (MD) and limits agreement (LoA; upper–U, and lower–L), respectively.

mean systematic bias to 0.00°C, and limits of agreement (95%) to 0.00–0.00°C. As a result, the proportion of sensors with a systematic bias < ±0.1°C rose to 100% following correction with individual regression equations.

The generalized linear regression generated from the group of development sensors (Corrected Temperature (°C) = 1.00362 × Sensor Temperature (°C) − 0.2374038) produced a mean systematic bias of 0.07 ± 0.21°C when applied to the verification sensors (**Figure 2**). The limits of agreement (95%) of the verification sensors ranged between −0.11 and 0.25 (**Figure 2**). With this generalized correction only 94.0% of the verification sensors showed a systematic bias of < ±0.1°C. However, when sensors with systematic bias > ±0.5°C were removed from the analysis, the generalized linear regression (Corrected Temperature (°C) = 1.00375 × Sensor Temperature (°C) − 0.205549) corrected systematic bias among the verification sensors to −0.01 ± 0.04°C, with limits of agreement (95%) of −0.04 to 0.02°C (**Figure 2**). Consequently the proportion of sensors with a systematic bias < ±0.1°C was 98.4%.

DISCUSSION

Overall, 89.1% ($n = 106$) of ingestible temperature sensors measured temperature in the physiological range within ±0.1°C. The remaining 11.7% ($n = 13$) displayed a wide range in systematic bias (**Figure 1**) which highlights the importance of performing a calibration procedure prior to administering the sensor for ingestion. An individual linear function for each sensor was required to eliminate the systematic bias from the reference thermometer and therefore between different sensors, however, these findings also demonstrate that an average linear function could be used to ensure error between sensors is restricted to an acceptable range. Provided that sensors with a systematic bias > ±0.5°C are identified and excluded, correcting

FIGURE 2 | Linear Regressions (Top) and Bland-Altman Plots (Bottom) comparing the Corrected Sensor Temperature with the Reference Temperature. Graphs on the left report the generalized linear equation including all sensors, while graphs on the right show the generalized equation excluding sensors outside ±0.5°C of the reference thermometer. Solid and dashed lines represent the mean difference (MD) and limits agreement (LoA; upper–U, and lower–L), respectively.

sensor temperature readings by the following equation: corrected temperature (°C) = 1.00375 × sensor temperature (°C) − 0.205549; achieved a 98.4% success rate in meeting the ±0.1°C criteria for acceptable measurement accuracy (**Figure 2**).

The proportion of sensors with a systematic bias outside ±0.1°C was smaller in the present study than previously observed in an initial pilot study of only three sensors (Hunt and Stewart, 2008). The present study has evaluated a larger sample of 119 sensors allowing greater confidence in the actual proportion of inaccurate sensors being identified at ∼11%. Furthermore, this study has revealed that a small yet important proportion of sensors have a much greater systematic bias than previously reported, in excess of ±0.5°C. These discrepancies have important practical implications for monitoring body core temperature in athletic and occupational settings. For instance, test re-test reliability indicates that intestinal temperature can vary by up to 0.34°C between repeat exercise trials (Ruddock et al., 2014). Therefore, detection of a meaningful difference in body core temperature during experimental trials must ensure that the combined error of ingestible sensors (which are one use only) is below this value. From the proportions observed in this study, it is evident that 3% of sensors may provide inaccurate temperature measurement (> ±0.5°C) if uncorrected to a reference thermometer. Consequently, in an experimental

trial of 15 participants performing two exercise conditions for example, it would be expected that at least one participant's results would be erroneous due to measurement device error instead of a treatment effect. The literature identified that utilized the ingestible sensor technology but without performing a temperature correction (40 out of the 52 studies provided enough information to determine the number of sensors used), used an average of 28 (range: 8–96) sensors per study. Proportionately, the present findings suggest that at least one participant's data would be erroneous in each of these studies. Consequently these findings highlight the necessity of performing a pre-ingestion temperature assessment of ingestible sensors in a water bath.

Correction of sensor temperature to a reference thermometer by linear regression was confirmed as an appropriate technique in the present study. The systematic bias of sensor temperature was effectively eliminated by the application of a linear function specific to each ingestible sensor. While an individual linear function was identified as the optimal correction procedure, there is also considerable time cost to researchers and practitioners in implementing this strategy. Alternatively, an average linear function has been proposed. By this method sensor temperature could be compared to a reference thermometer at a single water bath temperature in the physiological range and, if found to be within ±0.5°C, could subsequently be corrected by the average

linear function [corrected temperature (°C) = 1.00375 × sensor temperature (°C) − 0.205549)]. This technique would restrict systematic bias to the required ±0.1°C accuracy range for 98.4% of sensors.

The ingestible temperature sensors are most commonly employed for monitoring extremes of body temperature experienced in real-world situations, as such the accuracy of the temperature measurement is of paramount importance at the extremes of the physiological range. However, it has been reported that the differences between ingestible sensor and rectal temperatures tended to be greater in those who showed the highest elevations in rectal temperature (Casa et al., 2007). Similarly, in studies examining ingestible temperature sensors in a water bath, greater systematic bias has been reported as temperatures progress toward the extremes of the physiological range (Hunt and Stewart, 2008; Travers et al., 2016). In the present investigation systematic bias was significantly different at the lowest temperature condition (TEMP A: 35.12 ± 0.60°C) compared to the warmer temperature conditions. This finding further highlights the importance of ensuring accurate temperature measurement through correction before ingestion. According to the National Athletic Trainers Association Position Statement on Environmental Cold Injuries a core temperature of 35°C is an indicator of mild hypothermia (Cappaert et al., 2008). Therefore, accurate measurements of core temperature are required in the field for assessment and diagnosis of injuries associated with exposure to the environment. In addition, this finding also highlights that sensor corrections need to be based on reference temperatures measured within the physiological range. For if correction equations were to be based on temperatures outside the physiological range an incorrect linear function may be developed.

A limitation of the present study was that only one model of ingestible sensor was evaluated, yet there are several others on the commercial market. Systematic bias of other ingestible sensors such as the VitalSense and the e-Celcius, has been observed in the range of 0.18–0.34°C, respectively (Travers et al., 2016). In addition, the difference in accuracy at the extremes of the physiological range may also differ as previous evaluations have displayed sensors overestimating at low temperatures and underestimating at high temperatures (Chapon et al., 2012). Therefore, caution should be exercised before implementing the average linear function developed here to other models of ingestible sensor.

RECOMMENDED CALIBRATION METHOD

Based on the findings of the present study the following method is recommended for calibration of ingestible temperature sensors. Equipment requirements include a circulated and sterile water bath capable of maintaining a stable water temperature in the expected physiological range, as well as a certified and traceable reference thermometer with accuracy < ±0.1°C. Also, ensure the ingestible sensors are within the manufacturer's expiration date.

Individual Sensor Calibration

To develop an individual sensor calibration equation, sensor, and reference temperatures should be measured in at least four water bath temperatures. For each water bath temperature ensure:

1. Sufficient time (>4 min) is allowed for water temperature and sensor temperature to stabilize before the measurement is recorded.
2. Repeat this procedure for each water bath temperature.
3. Plot the linear relationship between sensor and reference temperatures.
4. Utilize the individualized linear equation to correct raw data from the ingestible sensor. Raw data from the sensor can either be corrected by the linear regression in a post-processing of the data after the exercise trial is completed, or could be incorporated into the investigators viewing of the data if real-time corrections are required during a trial.

Generalized Sensor Calibration

To utilize the generalized linear correction presented in this paper, sensor and reference temperature should be measured in at least one water bath temperature in the physiological range anticipated depending on the expected values for the activity to be monitored or the type of research study).

1. Sufficient time (>4 min) is allowed for water temperature and sensor temperature to stabilize before the measurement is recorded.

   a. If sensor temperature is within ±0.5°C of the reference thermometer, the following correction can be applied to the raw sensor temperature data:

   $$\text{Corrected Temperature (°C)} = 1.00375 \times \text{Sensor Temperature (°C)} - 0.205549$$

   b. If sensor temperature is outside ±0.5°C then an individual sensor calibration should be performed (see individual sensor calibration above).

The timing of sensor calibration prior to ingestion will be an important consideration in trial preparation and depend on the type of study being undertaken. For example, a study of body temperature during a mass participation endurance event would require all sensors to be calibrated in the same time window prior to the event. Alternatively, a laboratory based study with repeated trials over weeks or months requires the calibration procedure to be performed within a short time period before each trial. Performing the calibration procedure in the time window immediately prior to its ingestion will ensure battery life of the sensor is sufficient to record data for the duration of the specific trial activity being undertaken by the participant. If for practical limitations the trial activity necessitates a prolonged period between calibration and ingestion (such as long distance travel to a test location), some models of ingestible sensor can be deactivated (switched off) in the interim to preserve battery life (Hunt and Stewart, 2008).

CONCLUSION

In conclusion, this study has confirmed that ingestible temperature sensors require correction by linear regression. The optimal technique for reducing systematic bias of the sensor is the development of an individualized linear function to correct temperature data to a reference thermometer. However, application of an average linear function [corrected temperature ($^{\circ}$C) = $1.00375 \times$ sensor temperature ($^{\circ}$C) $- 0.205549$] in sensors with an initial accuracy within $\pm 0.5^{\circ}$C of the reference thermometer, was also effective at reducing systematic bias to $< \pm 0.1^{\circ}$C in 98.4% of sensors.

AUTHOR CONTRIBUTIONS

All authors made substantial contributions to the design of this study and/or acquisition, analysis, and interpretation of the data.

REFERENCES

Armstrong, L. E., Johnson, E. C., Casa, D. J., Ganio, M. S., McDermott, B. P., Yamamoto, L. M., et al. (2010). The American football uniform: uncompensable heat stress and hyperthermic exhaustion. *J. Athl. Train.* 45, 117–127. doi: 10.4085/1062-6050-45.2.117

Bagley, J. R., Judelson, D. A., Spiering, B. A., Beam, W. C., Bartolini, J. A., Washburn, B. V., et al. (2011). Validity of field expedient devices to assess core temperature during exercise in the cold. *Aviat. Space Environ. Med.* 82, 1098–1103. doi: 10.3357/ASEM.3102.2011

Bland, J. M., and Altman, D. G. (2007). Agreement between methods of measurement with multiple observations per individual. *J. Biopharm. Stat.* 17, 571–582. doi: 10.1080/10543400701329422

Brazaitis, M., Paulauskas, H., Skurvydas, A., Budde, H., Daniuseviciute, L., and Eimantas, N. (2016). Brief rewarming blunts hypothermia-induced alterations in sensation, motor drive and cognition. *Front. Physiol.* 7:592. doi: 10.3389/fphys.2016.00592

Byrne, C., Lee, J. K. W., Chew, S. A. N., Lim, C. L., and Tan, E. Y. M. (2006). Continuous thermoregulatory responses to mass-participation distance running in heat. *Med. Sci. Sports Exerc.* 38, 803–810. doi: 10.1249/01.mss.0000218134.74238.6a

Byrne, C., and Lim, C. L. (2007). The ingestible telemetric body core temperature sensor: a review of validity and exercise applications. *Br. J. Sports Med.* 41, 126–133. doi: 10.1136/bjsm.2006.026344

Cappaert, T. A., Stone, J. A., Castellani, J. W., Krause, B. A., Smith, D., and Stephens, B. A. (2008). National athletic trainers' association position statement: environmental cold injuries. *J. Athl. Train.* 43, 640–658. doi: 10.4085/1062-6050-43.6.640

Carter, III., R., Cheuvront, S. N., Williams, J. O., Kolka, M. A., Stephenson, L. A., Sawka, M. N., et al. (2005). Epidemiology of hospitalizations and deaths from heat illness in soldiers. *Med. Sci. Sports Exerc.* 37, 1338–1344. doi: 10.1249/01.mss.0000174895.19639.ed

Casa, D. J., Becker, S. M., Ganio, M. S., Brown, C. M., Yeargin, S. W., Roti, M. W., et al. (2007). Validity of Devices That Assess Body Temperature During Outdoor Exercise in the Heat. *J. Athl. Train.* 42, 333–342.

Castellani, J. W., and Tipton, M. J. (2015). Cold stress effects on exposure tolerance and exercise performance. *Comprhen. Physiol.* 6, 443–469. doi: 10.1002/cphy.c140081

Chapon, P. A., Gauthier, A., Bulla, J., and Moussay, S. (2012). Calibration and performance assessment of a temperature sensor prototype using a 1-point calibration procedure. *Rev. Sci. Instrum.* 83:114907. doi: 10.1063/1.4767244

Costello, J. T., Stewart, K. L., and Stewart, I. B. (2015). Inside the 'Hurt Locker': the combined effects of explosive ordnance disposal and chemical protective clothing on physiological tolerance time in extreme environments. *Ann. Occup. Hyg.* 59, 922–931. doi: 10.1093/annhyg/mev029

Ganio, M. S., Brown, C. M., Casa, D. J., Becker, S. M., Yeargin, S. W., McDermott, B. P., et al. (2009). Validity and reliability of devices that assess body temperature during indoor exercise in the heat. *J. Athl. Train* 44, 124–35. doi: 10.4085/1062-6050-44.2.124

Gant, N., Atkinson, G., and Williams, C. (2006). The validity and reliability of intestinal temperature during intermittent running. *Med. Sci. Sports Exerc.* 38, 1926–1931. doi: 10.1249/01.mss.0000233800.69776.ef

Goosey-Tolfrey, V., Swainson, M., Boyd, C., Atkinson, G., and Tolfrey, K. (2008). The effectiveness of hand cooling at reducing exercise-induced hyperthermia and improving distance-race performance in wheelchair and able-bodied athletes. *J. Appl. Physiol.* 105, 37–43. doi: 10.1152/japplphysiol.01084.2007

Grundstein, A. J., Ramseyer, C., Zhao, F., Pesses, J. L., Akers, P., Qureshi, A., et al. (2012). A retrospective analysis of American football hyperthermia deaths in the United States. *Int. J. Biometeorol* 56, 11–20. doi: 10.1007/s00484-010-0391-4

Harper Smith, A. D., Crabtree, D. R., Bilzon, J. L. J., and Walsh, N. P. (2010). The validity of wireless iButtons® and thermistors for human skin temperature measurement. *Physiol. Meas.* 31, 95–114. doi: 10.1088/0967-3334/31/1/007

Hunt, A. P., Billing, D. C., Patterson, M. J., and Caldwell, J. N. (2016). Heat strain during military training activities: the dilemma of balancing force protection and operational capability. *Temperature* 3, 307–317. doi: 10.1080/23328940.2016.1156801

Hunt, A. P., and Stewart, I. B. (2008). Calibration of an ingestible temperature sensor. *Physiol. Meas.* 29:N71. doi: 10.1088/0967-3334/29/11/N01

Kenney, W., DeGroot, D., and Holowatz, L. (2004). Extremes of human heat tolerance: life at the precipice of thermoregulatory failure. *J. Therm. Biol.* 29, 479–485. doi: 10.1016/j.jtherbio.2004.08.017

Larsen, B., Snow, R., Williams–Bell, M., and Aisbett, B. (2015). Simulated firefighting task performance and physiology under very hot conditions. *Front. Physiol.* 6:322. doi: 10.3389/fphys.2015.00322

Logan-Sprenger, H. M., Heigenhauser, G. J. F., Killian, K. J., and Spriet, L. L. (2012). Effects of dehydration during cycling on skeletal muscle metabolism in females. *Med. Sci. Sports Exerc.* 44, 1949–1957. doi: 10.1249/MSS.0b013e31825abc7c

Mermier, C. M., Schneider, S. M., Gurney, A. B., Weingart, H. M., and Wilmerding, M. V. (2006). Effect of whole-body cooling in patients with myasthenia gravis. *Med. Sci. Sports Exerc.* 38, 13–20. doi: 10.1249/01.mss.0000180887.33650.0f

Montain, S. J., Sawka, M. N., Cadarette, B. S., Quigley, M. D., and McKay, J. M. (1994). Physiological tolerance to uncompensable heat stress: effects of exercise intensity, protective clothing, and climate. *J. Appl. Physiol.* 77, 216–222.

Moran, D. S., and Mendal, L. (2002). Core temperature measurement: methods and current insights. *Sports Med.* 32, 879–885. doi: 10.2165/00007256-200232140-00001

Pyke, A. J., Costello, J. T., and Stewart, I. B. (2015). Heat strain evaluation of overt and covert body armour in a hot and humid environment. *Appl. Ergon.* 47, 11–15. doi: 10.1016/j.apergo.2014.08.016

Ruddock, A. D., Tew, G. A., and Purvis, A. J. (2014). Reliability of intestinal temperature using an ingestible telemetry pill system during exercise in a hot environment. *J. Strength Cond. Res.* 28, 861–869. doi: 10.1519/JSC.0b013e3182aa5dd0

Simpson, R., Machin, G., McEvoy, H., and Rusby, R. (2006). Traceability and calibration in temperature measurement: a clinical necessity. *J. Med. Eng. Technol.* 30, 212–217. doi: 10.1080/03091900600711530

Stewart, I. B., and Hunt, A. P. (2011). Negligible heat strain in armored vehicle officers wearing personal body armor. *J. Occup. Med. Toxicol.* 6, 22–27. doi: 10.1186/1745-6673-6-22

Stewart, I., Dias, B., Borg, D., Bach, A., Feigl, B., and Costello, J. (2017). Intraocular pressure is a poor predictor of hydration status following intermittent exercise in the heat. *Front. Physiol.* 8:36. doi: 10.3389/fphys.2017.00036

Taylor, L., Fitch, N., Castle, P., Watkins, S., Aldous, J., Sculthorpe, N., et al. (2014). Exposure to hot and cold environmental conditions does not affect the decision making ability of soccer referees following an intermittent sprint protocol. *Front. Physiol.* 5:185. doi: 10.3389/fphys.2014.00185

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

Travers, G. J. S., Nichols, D. S., Farooq, A., Racinais, S., and Périard, J. D. (2016). Validation of an ingestible temperature data logging and telemetry system during exercise in the heat. *Temperature* 3, 208–219. doi: 10.1080/23328940.2016.1171281

Walker, A., Argus, C., Driller, M., and Rattray, B. (2015). Repeat work bouts increase thermal strain for Australian firefighters working in the heat. *Int. J. Occup. Environ. Health* 21, 285–293. doi: 10.1179/2049396715Y.0000000006

Wilkinson, D. M., Carter, J. M., Richmond, V. L., Blacker, S. D., and Rayson, M. P. (2008). The effect of cool water ingestion on gastrointestinal pill temperature. *Med. Sci. Sports Exerc.* 40, 523–528. doi: 10.1249/MSS.0b013e31815cc43e